

# Open Research Online

---

The Open University's repository of research publications and other research outputs

## Automatic Feature Set Selection for Merging Image Segmentation Results Using Fuzzy Clustering

Conference or Workshop Item

How to cite:

Ali, Ameer; Dooley, Laurence S. and Karmakar, Gour (2005). Automatic Feature Set Selection for Merging Image Segmentation Results Using Fuzzy Clustering. In: 8th International Conference on Computer and Information Technology (ICCIT '05), 28-30 Dec 2005, Dhaka, Bangladesh.

For guidance on citations see [FAQs](#).

© [not recorded]

Version: [not recorded]

Link(s) to article on publisher's website:  
<http://www.iccitbd.net/>

---

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.

---

[oro.open.ac.uk](http://oro.open.ac.uk)

# AUTOMATIC FEATURE SET SELECTION FOR MERGING IMAGE SEGMENTATION RESULTS USING FUZZY CLUSTERING

*M. Ameer Ali, Laurence S Dooley and Gour C Karmakar*

Gippsland School of Information Technology, Monash University, Australia

Email: {Ameer.Ali, Laurence.Dooley and Gour.Karmakar }@infotech.monash.edu.au

## ABSTRACT

The image segmentation performance of clustering algorithms is highly dependent on the features used and the type of objects contained in the image, which limits the generalization ability of such algorithms. As a consequence, a *fuzzy image segmentation using suppressed fuzzy c-means clustering* (FSSC) algorithm was proposed that merged the initially segmented regions produced by a fuzzy clustering algorithm, using two different feature sets each comprising two features from pixel location, pixel intensity and a combination of both, which considered objects with similar surface variations (SSV), the arbitrariness of fuzzy c-means (FCM) algorithm using pixel location and the connectedness property of objects. The feature set selection for the initial segmentation in the merging technique was however, inaccurate because it did not consider all possible feature set combinations and also manually defined the threshold used to identify objects having SSV. To overcome these limitations, a new *automatic feature set selection for merging image segmentation results using fuzzy clustering* (AFMSF) algorithm is proposed, which considers the best feature set selection and also calculates the threshold based upon human visual perception. Both qualitative and quantitative analysis prove the superiority of AFMSF algorithm compared with other clustering techniques including FSSC, FCM, possibilistic c-means (PCM) and SFCM, for different image types.

**Keyword:** Image Segmentation, Connectedness.

## 1. INTRODUCTION

Object-based image segmentation is a very challenging task because of the inherent large number of objects and variations amongst them, which ultimately limits the segmentation process performance for each object within the general framework [1], [2]. Clustering algorithms [3]-[9] use many different feature types<sup>1</sup>, such as brightness (pixel *intensity* of a gray-scale image) and geometric information (pixel *location*). The segmentation performance of clustering algorithm is highly dependent on the type of objects and the feature used, which raises an open question about which types of feature produce the best results for which type of image and hence limit the generalisation of a clustering algorithm [3]. For instance, objects having similar pixel intensities in an image cannot be separated well by FCM

[3], PCM [5] and SFCM [10] by considering only their PI. They may however be able to segment well, by exploiting PL information or a combination (CIL) of PL and PI. Similarly, clustering cannot segment asymmetrically oriented adjacent regions having different intensities by only considering PL, but may well be able to do so by considering PI or CIL. It had been reported [1] that even clustering algorithms that use both features do not always produce the expected results for all images. This was the motivation of merging the separate initial segmented regions produced by a clustering algorithm using different feature sets for final segmentation. To address these issues, Ameer et al. [1] proposed an algorithm namely *fuzzy image segmentation using suppressed fuzzy c-means clustering* (FSSC) which considered connectivity, objects having similar surface variations (SSV) in an image and the arbitrariness of FCM using PL. In the FSSC algorithm, the objects having SSV are segmented by SFCM using PL, because SFCM is insensitive to the fuzzy factor used. It also emphasizes the highest membership value, while suppressing all the others [10] so providing better segmented results than FCM using PL for objects with SSV. For merging objects having dissimilar surface variations (DSV), FSSC considered only CIL and PL; and PI and PL feature sets, but it did not take account of CIL and PI combination for initial segmentation. Also it did not consider the overlapping regions for merging strategy. To address these issues, this paper presents a new algorithm called *automatic feature set selection for merging image segmentation results using fuzzy clustering* (AFMSF) by considering all the combination of feature sets and the minimum overlapping regions generated when initial segmented results are merged. In this paper, a perceptual threshold is employed to detect those objects having SSV and for this reason, the new AFMSF algorithm is a more generalized technique which produces improved segmented results compared with existing algorithms. Embedded within AFMSF are two constituent algorithms, which handle the *merging initial segmented regions* (MISR) and the *separation of objects having similar and dissimilar surface variations* (SOSDS). The paper includes a full numerical analysis of FCM, PCM and SFCM for all feature sets and the proposed AFMSF algorithm, in addition to FSSC using the objective segmentation evaluation methods in [1].

The remainder of the paper is organized as follows: The SFCM algorithm that segments objects having SSV is detailed in Section 2 and issues relating to the identification of SSV are discussed fully in Section 3. The merging technique is analysed in Section 4 with a detailed description of the proposed AFMSF algorithm presented in Section 5. Section 6 provides a comprehensive qualitative and quantitative performance analysis

---

<sup>1</sup> The terminologies *PL*, *PI* and *CIL* refer respectively to *pixel location*, *pixel intensity* and a *combination* of pixel intensity and normalized pixel location.

of the segmentation results of the new algorithm, while some conclusions are provided in Section 7.

## 2. THE SFCM ALGORITHM

Wei et al. [11] originally proposed an algorithm called *rival checked fuzzy c-means clustering algorithm* (RCFCM) which magnified the largest membership value and suppressed the second largest membership value  $\mu$ . Since RCFCM gives priority to the biggest and the second biggest membership values of  $\mu$ ,  $\mu$  is dependent on the value of the parameter  $\alpha$ , where ( $0 \leq \alpha \leq 1$ ). This leads to distortion of the original order of the membership values when the selection of  $\alpha$  is unsuitable. To address this issue, [10] proposed a *suppressed fuzzy c-means clustering algorithm* (SFCM) which gives the highest priority to the biggest membership and suppresses the rest. Assume  $X_j$  is a datum; if the membership value of  $X_j$  belonging to cluster  $p$  is the biggest cluster, the value is noted as  $\mu_{pj}$ . The modified membership value is then defined as:-

$$\mu_{pj} = 1 - \alpha \sum_{\substack{i=1 \\ i \neq p}}^c \mu_{ij} = 1 - \alpha + \alpha \mu_{pj} \quad (1)$$

$$\mu_{ij} = \alpha \mu_{ij}, \quad i \neq p \quad (2)$$

where  $0 \leq \alpha \leq 1$  and  $c$  is the number of clusters. This modification will not disturb the original order and overcomes the drawback in RCFCM. When  $\alpha = 0$ , the algorithm provides hard clustering, while when  $\alpha = 1$  the result is exactly the same as FCM.

## 3. IDENTIFICATION OF OBJECTS HAVING SIMILAR SURFACE VARIATIONS (SSV)

To effectively segment objects having SSV, it is important to identify similar and dissimilar object surfaces in an image, though this is a challenging task [1]. Two possible cases exist by which surfaces may be considered as similar with respect to brightness perspective. These are that the surfaces have: (i) similar intensity and (ii) possess SSV. In case (i), FCM using CIL produces similar results to FCM using PL. For the later case, when objects have SSV i.e., objects with repeated patterns of bright and dark pixels, FCM using CIL is unable to separate them [1]. In this case one cluster covers the whole area of these objects. This motivated the use of FCM using CIL to determine the type of objects surface variations, such that when FCM using CIL cannot separate a group of objects, all these objects have SSV.

## 4. THE MERGING TECHNIQUE

As mentioned in Section 1, there are a huge number of objects and myriad variations among them. For this reason, no single feature or combination of them is suitable for segmenting every object in an image. This was the rationale behind independently merging the segmented results produced by FCM using PL, PI and CIL in [1]. In the *merging initial segmented regions* (MISR) algorithm [1], misclassified pixels are distributed using 8-connected objects and 8-connectivity property to the

corresponding merged pair. The reason for using 8-connectivity is to ensure all weak object connections are considered. All the steps involved in the MISR algorithm are given in *Algorithm 1*. Note that  $R^I$ ,  $R^L$  and  $R^C$  represent the initial segmented regions produced by FCM separately using PI, PL and CIL respectively and  $\mathfrak{R}$  stands for the number of clusters (regions).

*Algorithm 1: Merging initial segmented regions* (MISR)

**Precondition:** A selected pair of the initial segmented regions  $R^I$ ,  $R^L$  and  $R^C$ ;  $\mathfrak{R}$ , connectivity.

**Post-condition:** The segmented regions  $R$ .

1. Determine similar regions.
2. Merge these similar regions.
3. Calculate the overlap between the two merging regions and remove overlapping pixels from them.
4. IF (connectivity) THEN distribute 8-connected objects of the overlap to merging regions using 8-connectivity.
5. Redistribute any remaining overlapping pixels by a clustering algorithm using CIL.

## 5. THEORETICAL MODELING OF THE PROPOSED ALGORITHM

As mentioned in Section 1, the FSSC algorithm [1] does not consider the combination of all feature sets and the amount of overlap during merger. This paper presents a new algorithm called *automatic feature set selection for merging image segmentation results using fuzzy clustering* (AFMSF) which considers the selection of the feature sets based on the overlapping regions from all combination of feature sets (CIL, PI, PL) and calculates a threshold  $T_{\max}$  considering human visual perception.

Since any image may contain different objects having SSV and DSV, for object based image segmentation any clustering algorithm that is able to identify them in the segmentation process has the potential to provide superior segmented results. Once this is achieved, SSV objects are segmented by SFCM using PL as mentioned in Section 1, while DSV object segmentation requires selecting the most appropriate pair of feature sets from PI, PL and CIL. Similar regions of the initially segmented regions are then merged and the overlap for each merging region pair is calculated and removed. The overlapping pixels are then distributed between this pair of merging regions. Before detailing the AFMSF algorithm, the following sections examine firstly the issue of separating objects having SSV and DSV and secondly, selecting the best feature set.

### 5.1 The Separation Of Objects Having Similar And Dissimilar Surface Variations Algorithm

Any strategy able to identify and distinguish objects having SSV and DSV undoubtedly affords the potential of improved image segmentation. A description of one such strategy is given in Algorithm 2 called the *separation of objects having similar and dissimilar surface variations* (SOSDS) algorithm. As discussed in Section 3, FCM using CIL is used to determine whether an

object has either SSV or DSV. To locate SSV regions, the area  $A^{R_i^C}$  of the segmented region  $R_i^C$  is calculated using a convex hull:-

$$A^{R_i^C} = \text{Area}(\text{Convexhull}(R_i^C)) \quad (3)$$

*Algorithm 2:* Separation of objects having similar and dissimilar surface variations (SOSDS)

**Precondition:** Initially segmented regions  $R^C$  and  $\mathfrak{R}$  .  
**Post condition:** Objects having SSV ( $R_i^M$ ) and DSV ( $R^D$ ) .

1. Set  $M = 1$  and  $k = 1$  .
2. Form region  $R_k^M$  by combining  $R_i^C$  and  $R_j^C$  .
3. Calculate areas  $A^{R_k^M}$ ,  $A^{R_i^C}$  and  $A^{R_j^C}$  using (3).
4. Find the maximum area  $A_i^C$  of  $(M + 1)$  regions in  $R_k^M$  .
5. IF  $\frac{|A^{R_k^M} - A_i^C|}{A_i^C} \leq T_{\max}$  THEN two regions  $R_i^C$  and  $R_j^C$  have SSV and  $M$  is incremented.  
Repeat Steps 2-5 forming  $R_k^M$  by merging  $R_k^M$  and another region (which has not already been merged) from  $R^C$
6. IF  $(M \geq 2)$  THEN increment  $k$  and GOTO Step 1
7. Separate region  $R^D$  which has  $D$  objects with DSV from  $R^C$  .

where  $\text{Area}(\bullet)$  and  $\text{Convexhull}(\bullet)$  are respectively the area and vertices of the convex hull of a segmented region. To identify objects having SSV, two regions are merged (Step 2) to form  $R_k^M$  and then calculate the area  $A^{R_k^M}$ ,  $A^{R_i^C}$ ,  $A^{R_j^C}$  of regions  $R_k^M$ ,  $R_i^C$  and  $R_j^C$  respectively using (3) in Step 3, where  $M$  is the number of merged regions in  $R_k^M$ ,  $2 \leq M \leq \mathfrak{R}$ ,  $1 \leq k \leq \lfloor \frac{\mathfrak{R}}{2} \rfloor$  and  $1 \leq i, j \leq \mathfrak{R}$ . The difference between the area of the largest merged region of the  $k^{\text{th}}$  merging region and  $A^{R_k^M}$  (the area of the  $k^{\text{th}}$  merging region) is a measure of shape distortion (Step 5), because the merging region  $R_k^M$  always contains the largest merged region. If this distortion measure is less than  $0.5dB$ , the human eye will not perceive a change in shape and the segmentation algorithm cannot separate the objects i.e., the region has SSV. A perceptual threshold  $T_{\max}$  is now introduced into the SSV identification process which incorporates both shape distortion and human perception. The maximum value of  $T_{\max}$  is calculated as follows:-

$$0.5 = 20 \log \frac{\text{Area of Merging Region}}{\text{Area of the Largest Merged Region}}$$

$$\Rightarrow \text{Area of Merging Region} - \text{Area of the Largest Merged Region} = 0.059 \times \text{Area of the Largest Merged Region}$$

$$\Rightarrow \frac{\text{Area of Merging Region} - \text{Area of the Largest Merged Region}}{\text{Area of the Largest Merged Region}} = 0.059$$

so this threshold is bounded  $T_{\max} = 0.059$  (4)

Finally, objects having SSV are separated from those with DSV which are represented by region  $R^D$ , where  $D$ ,  $0 \leq D \leq \mathfrak{R}$  is the number of DSV objects. Those clusters not merged are treated as clusters containing objects having DSV.

## 5.2 Selection Strategy for Feature Sets

To segment objects having DSV, an appropriate feature set needs to be determined as any clustering algorithm separately using PL, PI and CIL will be unable to segment such objects [1]. Since for  $\mathfrak{R}$  regions, the degree of arbitrariness of FCM using PL increases by  $O(\mathfrak{R}^2)$  [1], so in proposing a feature set selection strategy, two scenarios are considered: i)  $\mathfrak{R} > 2$  ii)  $\mathfrak{R} = 2$ . In the former, FCM using CIL provides comparatively better results than using only PL because of the arbitrariness of FCM using PL. Thus selecting CIL and PI as the feature set is fully justified for the initial segmentation in the MISR algorithm. For  $\mathfrak{R} = 2$  however, it is necessary to choose a pair of feature sets from PI, PL and CIL.

To select the best set, the amount of overlap between pairs of merged regions representing the misclassified pixels of these regions is exploited. The less the degree of overlap, the lower the distribution time complexity and the more visually distinctive are the objects. Conversely, ambiguity during the distribution of overlapping pixels increases proportionally, so the risk of misclassification is proportional to any overlap. To minimize this risk, emphasis is given to the pair of feature sets that produce the minimal overlap. The approach used in this paper to calculate the overlap between a pair of merged regions is formalised in the following lemma.

**Lemma 1:** The amount of overlap between a pair of merging regions is directly proportional to the acute angle value between the decision boundaries of the initial segmented regions, separately produced using the two selected feature sets.

**Proof:** Let  $L_1$  and  $L_2$  be two non-parallel decision boundaries for FCM using PL and CIL respectively and  $\theta_1$  be the acute angle between them (Figure 1 (c)). The two segmented regions yielded by these two decision boundaries are shown in Figure 1 (a) ( $R_1^L$  and  $R_2^L$ ) and Figure 1 (b) ( $R_1^C$  and  $R_2^C$ ) respectively.

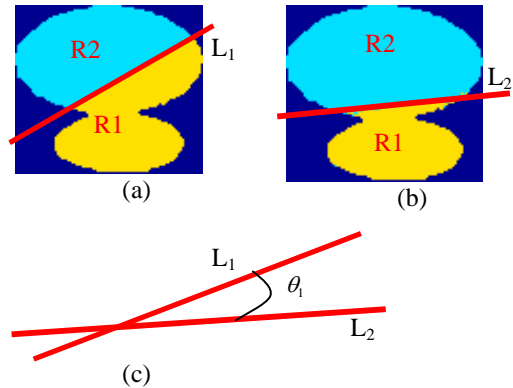


Figure 1: Examples of the initial segmented results produced by FCM using a two-region synthetic image, (a) Only PL, (b) Only CIL, (c) Angle between the two decision boundaries.

The merging of  $R_i^L$  with its similar region  $R_j^C$  produces a merged region  $R_K^M$ . The overlapping area between a pair of merging regions  $R_K^M$  and  $R_{K+1}^M$  for initially produced region  $R_i^L$  considering PL is defined as:-

$$A^{R_{k,k+1}^M} = \sum_{i=1}^2 \rho_i \theta_i * A^{R_i^L} / \pi = \theta_1 / \pi \times (\rho_1 * A^{R_1^L} + \rho_2 * A^{R_2^L}) \quad (5)$$

where  $\rho_i$  is a matching factor used in calculating the exact area of  $R_i^L$  subtended by  $\theta_1$ . Since  $A^{R_1^L}$  and  $A^{R_2^L}$  are constant for an image and specific feature set assuming  $\rho_1$  and  $\rho_2$  are constant, then:-

$$A^{R_{k,k+1}^M} \propto \theta_1 \quad (6)$$

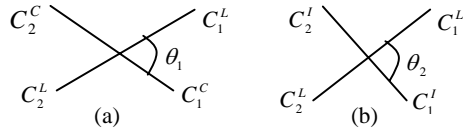


Figure 2: Acute angle between two decision boundaries produced by FCM separately using (a) PL and CIL, (b) PL and PI.

In the best case,  $L_1$  and  $L_2$  are equal so  $\theta_1 = 0$  and the overlapping region  $A^{R_{k,k+1}^M} = 0$ . For the average case  $\theta_1 = \pi/4$ , while the worst case  $\theta_1 = \pi/2$  which corresponds to maximum overlap  $A^{R_{k,k+1}^M} = 1/2 f$ , where  $f$  is the foreground of objects.

Since the maximum overlap is effectively half the foreground, it represents the average cluster size. ■

As previously mentioned, misclassification increases with overlap, so to limit to some extent, the arbitrariness effect of FCM using PL in selecting the feature sets, a constraint is applied to  $\theta_1$  by comparing it with its average value i.e.  $\pi/4$ . To select the best set, the following two cases are now considered,  $\theta_1 > \pi/4$  and  $\theta_1 \leq \pi/4$ .

**Lemma 2:** For any acute angle  $\theta_1 > \pi/4$  between the decision boundaries of FCM separately using CIL and PL, the feature sets CIL and PI are used in the MISR algorithm (*Algorithm 1*).

**Proof:** For  $\theta_1 > \pi/4$ , CIL strongly dominates PL in the segmentation process and has a high misclassification risk (Lemma 1) when merging. It is because of the existence of two objects with quite different brightness values that PI outweighs PL. The feature set CIL and PI will thus produce less overlapping regions. ■

For the case  $\theta_1 \leq \pi/4$ , to minimise misclassification the feature sets are selected based on the minimum value of the angle between the corresponding decision boundaries as follows:-

$$\text{feature sets} = \begin{cases} \text{CIL, PL} & \text{if } \theta_1 \text{ is minimum} \\ \text{PL, PI} & \text{if } \theta_2 \text{ is minimum} \\ \text{CIL, PI} & \text{if } \theta_3 \text{ is minimum} \end{cases} \quad (7)$$

where  $\theta_1$  = angle between the decision boundaries for FCM using only CIL and PL;  $\theta_2$  = angle for FCM using only PL and PI;  $\theta_3$  = angle for FCM using only CIL and PI.

### 5.3 The Automatic Feature Set Selection For Merging Image Segmentation Results using Fuzzy Clustering Algorithm

Having described both the MISR (*Algorithm 1*) and SOSDS (*Algorithm 2*) algorithms, these are embedded in the *automatic feature set selection for merging image segmentation results using fuzzy clustering* (AFMSF) algorithm detailed in *Algorithm 3*. Steps 1 and 2 respectively segment the foreground  $f$  and separate those objects in the image having SSV and DSV, using the SOSDS algorithm. FCM using CIL is applied since as alluded in Section 3,  $R^C$  is used to determine whether objects have either SSV or DSV. For those objects with SSV, SFCM using PL is used for segmentation (Step 3), while those objects with DSV are segmented by the MISR algorithm since they are visually distinct in terms of pixel intensity. Step 4 considers whether the connectivity feature is to be employed within the MISR algorithm. For two regions, if  $\theta_1 > \pi/4$  then CIL and PI (Lemma 2) are used in MISR. In this case, however it can be intuitively argued that connectivity should not be applied because each region has a distinct PI, one or more may possess a similar PI to another region that is actually connected to it. In such circumstances, to eliminate the possibility of misclassification, connectivity is not applied. For all other scenarios, the feature set selection strategy (Section 5.2) with connectivity is exploited in the MISR algorithm due to the potential impact of PL over PI in CIL.

**Algorithm 3:** *Automatic feature sets selection for merging image segmentation results using fuzzy clustering* (AFMSF)

**Precondition:** The foreground region  $f$  to be segmented,  $\mathfrak{R}$ ,  $\theta_1$  and  $\theta_2$ .

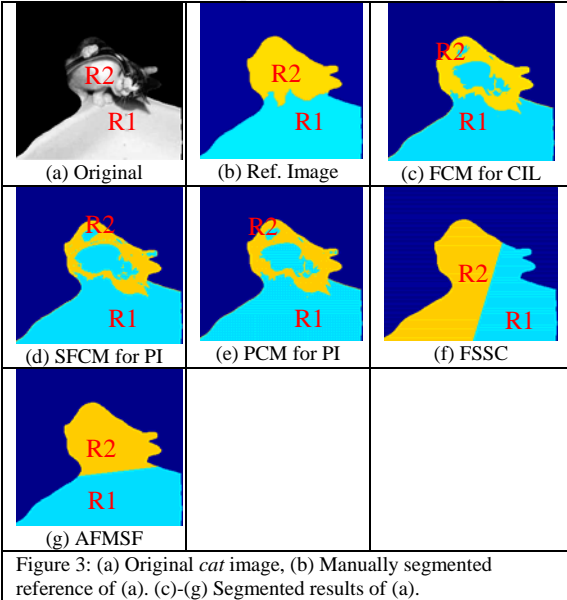
**Post condition:** The final segmented regions  $R$ .

1. Segment  $f$  by FCM using CIL into  $\mathfrak{R}$  regions represented by  $R^C$ .
2. Find  $R_k^M$  and  $R^D$  using **SOSDS** for  $R^C$ .
3. IF ( $k \geq 1$ ) THEN FOR  $i = 1, \dots, k$   
Segment  $R_i^M$  into  $M$  regions by SFCM using PL.
4. IF ( $D \geq 2$ ) THEN  
Connectivity=TRUE  
IF  $D = 2$  THEN  
IF ( $\theta_1 > \pi/4$ ) THEN  
Connectivity=FALSE  
Segment  $R^D$  into  $D$  regions using **MISR** for  $R^I$  and  $R^C$  (Lemma 2).  
ELSE  
Select feature sets using (7).  
Segment  $R^D$  into  $D$  regions by **MISR**.  
ELSE  
Segment  $R^D$  by MISR using  $R^I$  and  $R^C$ .

## 6. EXPERIMENTAL RESULTS

The new AFMSF, FSSC, fuzzy c-means (FCM) [3], PCM [5] and SFCM [10] algorithms were implemented using Matlab 6.1 (The Mathworks Inc.). The feature sets: PI, PL, and CIL were used for FCM, PCM and SFCM. A total of 148 different natural and synthetic gray-scale images were randomly selected for the experimental analysis, comprising up to 5 different regions (objects) having various degrees of surface variation, (obtained from IMSI<sup>2</sup>, own collection, and the Internet). To segment only the foreground objects in an image, the background was manually removed by setting it to zero. Any zero-valued foreground object pixels were replaced by 1, which had no effect upon visual perception and avoided the possibility of foreground pixels merging with the background. PL in the form of the  $x, y$  coordinates were normalized within the range [0, 255] in order to constrain them to the same range of PI for 8-bit gray-scale images. The perceptual threshold was set  $T_{\max} = 0.05$  as discussed in Section 5.1.

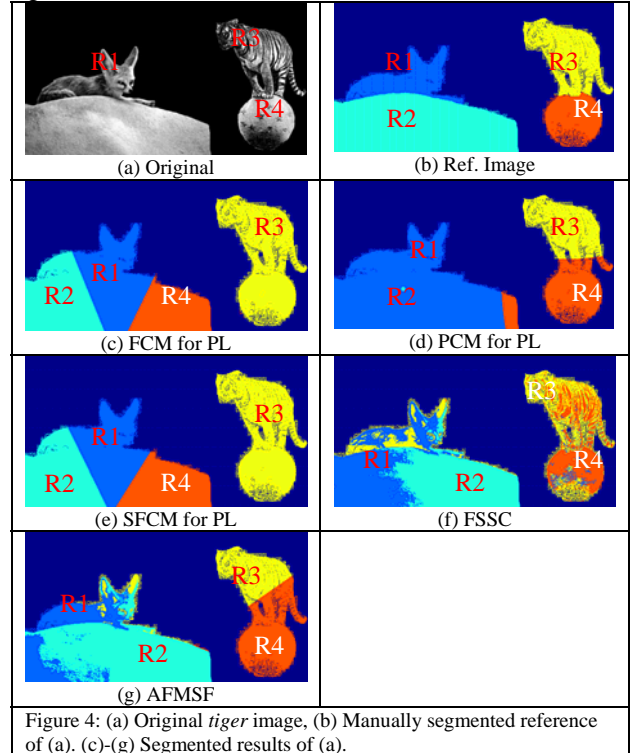
To quantitatively appraise the performance of all the various fuzzy clustering algorithms, the efficient objective segmentation evaluation method, *discrepancy based on the number of misclassified pixels* [1] was used. Two types of error, namely Type I,  $errorI_i$  and Type II,  $errorII_i$  are computed, the former being the percentage error of all  $i^{th}$  region pixels misclassified into other regions, while the latter is the error percentage of all region pixels misclassified into  $i^{th}$  region. Representative samples of the manually segmented reference regions together with their original images are shown in Figures 3(a)-3(b) and 4(a)-4(b). To provide a better visual interpretation



of the segmented results, both the reference and segmented regions are displayed using different colours rather than their original gray-scale intensities. Note that due to space limitations, only the best segmentation results of FCM, PCM and SFCM

with related feature and FSSC are presented as a comparison with AFMSF in Figures 3 and 4.

The experiments were performed upon the *cat* image shown in Figure 3(a) which has two different objects: the basin ( $R_1$ ) and the cat ( $R_2$ ). The three best segmented results for FCM, PCM, and SFCM are taken for each algorithm with the results of FSSC and the proposed AFMSF shown in Figures 3 (c)-(g). If the segmented results in Figure 3 (c)-(f) are compared with the manually segmented reference regions in Figure 3 (b), it is clear that a considerable number of pixels of  $R_2$  are misclassified into region  $R_1$  since both of the regions have similar white pixels. The segmented results in Figure 3(f) show that FSSC arbitrarily divided the *cat* image, for which both  $R_1$  and  $R_2$  regions contain lot of misclassified pixels. Regions  $R_1$  and  $R_2$  are correctly classified by the proposed AFMSF algorithm shown in Figure 3(g) with a minimum error, because the correct feature sets were selected for the segmentation process. The numerical results (average of Type I and Type II errors) of FCM, PCM and SFCM for all feature sets (PI, PL and CIL), FSSC with the proposed AFMSF algorithm for the *cat* image are shown in Table 1, which reveals that the average percentage error of AFMSF algorithm is 3.6%, while the second best average percentage error 19% was achieved by FCM for CIL. This confirms the superiority of AFMSF algorithm over FCM, PCM and SFCM using any of the three feature sets and FSSC algorithms.



The second sample image used in the experiment was *tiger* image shown in Figure 4(a) having four different regions: the rabbit ( $R_1$ ), the rock1 ( $R_2$ ), the tiger ( $R_3$ ) and the rock2 ( $R_4$ ) where both  $R_1$  and  $R_2$  regions appear to have DSV while  $R_3$

<sup>2</sup> IMSI's Master Photo Collection, 1895 Francisco Blvd. East, San Rafael, CA 94901-5506, USA.

and  $R_4$  have SSV. The best segmented results of FCM, PCM and SFCM with FSSC and the proposed AFMSF algorithms are shown in Figure 4(c)-(g). If the segmented results of FCM using PL and SFCM using PL in Figures 4 (c) and (e) respectively are compared with the reference regions shown in Figure 4(b), it is clear that  $R_4$  is not separated from  $R_3$  regions, while the other regions also contain a number of misclassified pixels. This is because the image contains two objects having SSV, and the other objects have DSV for which the clustering algorithms such as FCM and SFCM using PL arbitrarily divide the objects. PCM using PL approximately segments  $R_3$  and  $R_4$  well but are unable to segment  $R_1$  and  $R_2$  as shown in Figure 4(d). The segmented results shown Figure 4(f) produced by FSSC algorithm contain a large number of misclassified pixels in all the regions because of inaccurate selection of feature sets. The new AFMSF algorithm in contrast, separated all the four regions with a minimum number of misclassified pixels as shown in Figure 4(g) because of selecting the most appropriate feature sets for initial segmentation for objects having DSV ( $R_1$  and  $R_2$ ) and then merging the initial segmented results considering connectivity and the use of SFCM using PL for objects having SSV ( $R_3$  and  $R_4$ ). Table 1 shows that the AFMSF algorithm produces a minimum average error of 15%, while the next best average error performance was 22.3% produced by PCM using PL, so endorsing the superior segmentation performance achieved by AFMSF.

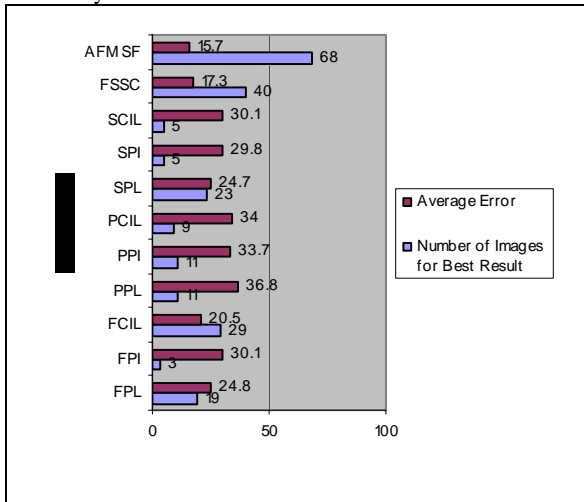


Figure 5: Performance analysis with best number of images and average error

To assess the generalization capability, robustness and effectiveness of the proposed algorithm, experiments were conducted on 148 real and synthetic images including multiple

regions. Out of 148 test images, AFMSF algorithm produced superior results for 68 images and minimum overall average error (15.7%) for all images shown in Figure 5 which also exhibits the overall superiority of the proposed algorithm.

## 7. CONCLUSIONS

This paper has introduced a new algorithm called *automatic feature sets selection for merging image segmentation results using fuzzy clustering* (AFMSF) which has proven superior segmented performance compared to FSSC, FCM, PCM and SFCM both qualitatively and quantitatively. The automatically selection of feature sets improves the segmentation performance where two objects have DSV. For the case where an image contain all the objects which have SSV, AFMSF produces exactly the same results as FSSC because they use the same strategy to segment objects having SSV. A perceptual threshold is applied in the AFMSF algorithm to assist in the segmentation by considering shape distortion.

## 8. REFERENCES

- [1] M. Ameer Ali, G C Karmakar and L S Dooley, "Fuzzy Image Segmentation using Suppressed Fuzzy C-Means Clustering (SFCM)", International Conference on Computer and Information Technology (ICCIT 04), pp.363-368, 2004.
- [2] I. Gath, and A. B. Geva, "Unsupervised Optimal Fuzzy Clustering," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 2(7), pp. 773-781, 1989.
- [3] J.C. Bezdek, "Pattern Recognition with fuzzy objective function algorithm," (New York: Plenum Press, 1981).
- [4] Jian Yu, Houkuan Huang and Shengfeng Tian, "An Efficient Optimality Test for the Fuzzy C-Means Algorithm," IEEE Intern. Con. on Fuzzy Systems, pp. 98-103, 2002.
- [5] R. Krishnapuram and J. M. Keller, "A Possibilistic Approach to Clustering," Fuzzy Systems, Vol. 2(2), pp. 98-110, 1993.
- [6] Carl G. Looney, "Interactive clustering and merging with a new fuzzy expected value," Pattern Recognition Letters, Vol. 35(11), pp. 2413-2423, 2002.
- [7] Gerardo Beni and Xiaomin Liu, "A least biased fuzzy clustering method," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 16(9), pp. 954-960, 1994.
- [8] Z. Chi, H. Yan, and T. Pham, "Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition," (Singapore: World Scientific Publishing Co. Pte. Ltd. 1996).
- [9] Frank Hoppner and Frank Klawonn, "A new approach to fuzzy partitioning," IEEE, 2001(C), pp. 1415-1424.
- [10] Jiu-Lun Fan, Wen-Zhi Zhen and Wei-Xin Xie, "Suppressed fuzzy c-means clustering algorithm," Pattern Recognition Letters, Vol. 24, pp. 1607-1612, 2003.
- [11] Wei L. M. and Xie W. X., "Rival checked fuzzy c-means algorithm", Acta Electronica Sinica, China, 2000, Vol. 28(7), pp. 63-66.

Table 1: Average error percentages and the number of images for the superior results of different algorithms

Image\ Average error	Algorithms										
	FCM			PCM			SFCM			FSSC	ASAM
	PL	PI	CIL	PL	PI	CIL	PL	PI	CIL		
<i>Cat</i>	35	24.9	19	49.9	21.6	49.5	35.9	24.6	24.6	35.9	3.6
<i>Tiger</i>	31.1	41.3	33	22.3	48.5	49.7	30.8	41.3	41.3	25.4	15