Multiobjective Genetic Clustering for Pixel Classification in Remote Sensing Imagery

Sanghamitra Bandyopadhyay, *Senior Member, IEEE,* Ujjwal Maulik, *Senior Member, IEEE,*
and Anirban Mukhopadhyay

Abstract— **An important approach for unupervised landcover classification in remote sensing images is the clustering of pixels in the spectral domain into several fuzzy partitions. In this article, a multiobjective optimization algorithm is utilized to tackle the problem of fuzzy partitioning where a number of fuzzy cluster validity indices are simultaneously optimized. The resultant set of near-Pareto-optimal solutions contains a number of non-dominated solutions, which the user can judge relatively and pick up the most promising one according to the problem requirements. Real-coded encoding of the cluster centers is used for this purpose. Results demonstrating the effectiveness of the proposed technique are provided for numeric remote sensing data described in terms of feature vectors. Different landcover regions in remote sensing imagery have also been classified using the proposed technique to establish its efficiency.**

Index Terms— **Fuzzy clustering, genetic algorithm, remote sensing imagery, pixel classification, multiobjective optimization, Pareto-optimal, cluster validity measures.**

I. INTRODUCTION

REMOTE sensing satellite images have significant applications in different areas such as climate studies, assessment of forest resources, examining marine environments etc. For remote sensing applications, classification is an important task where the pixels in the images are classified into homogeneous regions, each of which corresponds to some particular landcover type. The problem of pixel classification is often posed as clustering in the intensity space [1]. Clustering [2], [3] is a popular exploratory pattern classification technique which partitions the input space into K regions based on some similarity/dissimilarity metric where the value of K may or may not be known *a priori*. The main objective of any clustering technique is to produce a $K \times n$ partition matrix $U(X)$ of the given data set X, consisting of n patterns, $X = \{x_1, x_2, \ldots, x_n\}$. The partition matrix may be represented as $U = [u_{kj}]$, $k = 1, \ldots, K$ and $j = 1, \ldots, n$, where u_{kj} is the membership of pattern x_j to the kth cluster. For fuzzy clustering of the data, $0 < u_{kj} < 1$, i.e., u_{kj} denotes the probability of belongingness of pattern x_j to the kth cluster.

Remotely sensed multispectral images contain several spectral bands. In a satellite image, a pixel represents an area of the land space, which may not necessarily belong to a single landcover type. Hence, it is evident that a large amount of

Sanghamitra Bandyopadhyay is with the Machine Intelligence Unit, Indian Statistical Inst, Kolkata-700108, India. *Email* : *sanghami@isical.ac.in*

Ujjwal Maulik is with the Dept of Computer Sc & Engg, Jadavpur Univ., Kolkata-700032, India. *Email* : *drumaulik@jdvu.ac.in*

Anirban Mukhopadhyay is with the Dept of Computer Sc & Engg, Univ. of Kalyani, Kalyani-741235, India. *Email* : *anirbanbuba@yahoo.com*

imprecision and uncertainty can be associated with the pixels in such images. Therefore, it is natural to apply the principles of fuzzy set theory in the domain of pixel classification.

Recently, application of Genetic Algorithms (GAs) [4] in the field of pixel classification has attracted the attention of the researchers [1], [5]. These techniques use a single cluster validity measure as the fitness function to reflect the goodness of an encoded clustering. However, a single cluster validity measure is seldom equally applicable for different kinds of data sets with different characteristics. Hence it is necessary to simultaneously optimize several validity measures that can capture the different data characteristics. In order to achieve this, in this article the problem of fuzzy partitioning is posed as one of multiobjective optimizations (MOO) [6]-[9], where search is performed over a number of, often conflicting, objective functions. The final solution set contains a number of Pareto-optimal solutions, none of which can be further improved on any one objective without degrading it in another. A popular elitist real-coded multiobjective GA, Non-dominated Sorting Genetic Algorithm-II (NSGA-II) [6] is used to determine the appropriate cluster centers and the corresponding partition matrix. (Note that although NSGA-II has been used in the present article, any other multiobjective GA can be used instead; a study in this regard is in progress). The Xie-Beni (XB) index [10] and the fuzzy C-means (FCM) [11] measure (J_m) are used as the objective functions.

Clustering results are reported for remote sensing data available as both set of labelled feature vectors as well as images. Comparison with the well known FCM algorithm [11] (which directly optimizes the J_m criterion), the single objective version of the genetic clustering method that minimizes the XB index [1], and a widely used hierarchical clustering algorithm (average linkage) [3] are also provided. The comparative performance of the algorithms for the labelled remote sensing numeric data, obtained from the satellites Systeme Probatoire d'Observation de la Terre (SPOT) and LANDSAT is reported in terms of J_m and XB indices (indicating the goodness of obtained clustering), and a cluster goodness index $\mathcal{I}$ [12]. For the image data, comparison is provided in qualitative terms (i.e., visually) and in quantitative terms using index $\mathcal{I}$. Indian remote sensing (IRS) satellite image of parts of the cities of Calcutta and Mumbai, and a SPOT image of a part of the city of Calcutta have been segmented using the proposed method.

II. MULTIOBJECTIVE GENETIC ALGORITHMS

In many real world situations there may be several objectives to be optimized simultaneously in order to solve a certain

problem. The multiobjective optimization can be formally stated as following [9]: Find the vector $\overline{x}^* = [x_1^*, x_2^*, \ldots, x_v^*]^T$ of v decision variables which optimizes the objective function vector $\overline{f}(\overline{x}) = [f_1(\overline{x}), f_2(\overline{x}), \ldots, f_k(\overline{x})]^T$ satisfying some equality and inequality constraints. A decision vector $\overline{x}^*$ is called Pareto-optimal if and only if there is no $\overline{x}$ that dominates $\overline{x}^*$, i.e., there is no $\overline{x}$ such that $\forall i \in \{1, 2, \ldots, k\}, f_i(\overline{x}) \leq f_i(\overline{x}^*)$ and $\exists i \in \{1, 2, \ldots, k\}, f_i(\overline{x}) < f_i(\overline{x}^*)$. Pareto optimality usually admits a set of solutions called *non-dominated* solutions.

NSGA-II [6], Strength Pareto Evolutionary Algorithm (SPEA) [7] and SPEA2 [8] are some recently developed multiobjective elitist techniques. Of these, NSGA-II has been used in this article for developing the proposed multiobjective fuzzy clustering technique, described in the next section.

III. The Proposed Technique

The proposed NSGAII based clustering technique is described below in detail.

A. String Representation and Population Initialization

In NSGA-II based clustering, real valued chromosomes representing the coordinates of the centers of the partitions are used. For K clusters, the centers encoded in a chromosome in the initial population are randomly selected K distinct points from the data set.

B. Computing the Objectives

The Xie-Beni (XB) index [10] and J_m measure [11] are taken as the two objectives which need to be simultaneously optimized. For computing the measures, the centers $z_1, z_2, \ldots, z_K$ encoded in a chromosome are first extracted. The membership values u_{ik}, $i = 1, 2, \ldots, K$ and $k = 1, 2, \ldots, n$ are computed as follows [11]:

$$u_{ik} = \frac{1}{\sum_{j=1}^{K}\left(\frac{D(z_i, x_k)}{D(z_j, x_k)}\right)^{\frac{2}{m-1}}}, \quad (1)$$

where $D(z_i, x_k)$ is the euclidean distance between point x_k and cluster center z_i. m is the weighting coefficient. (Note that while computing u_{ik} using Eqn. 1, if $D(z_j, x_k)$ is equal to zero for some j, then u_{ik} is set to zero for all $i = 1, \ldots, K$, $i \neq j$, while u_{jk} is set equal to one.) Subsequently, the centers encoded in a chromosome are updated using the following equation [11]:

$$z_i = \frac{\sum_{k=1}^{n}(u_{ik})^m x_k}{\sum_{k=1}^{n}(u_{ik})^m}, \quad 1 \leq i \leq K, \quad (2)$$

and the cluster membership values are recomputed.

The XB index is defined as a function of the ratio of the total variation σ $(= \sum_{i=1}^{K}\sum_{k=1}^{n} u_{ik}^2 D^2(z_i, x_k))$ to the minimum separation sep $(= \min_{i \neq j}\{D^2(z_i, z_j)\})$ of the clusters, i.e.,

$$XB = \frac{\sigma(U, Z; X)}{n \ sep(Z)} = \frac{\sum_{i=1}^{K}(\sum_{k=1}^{n} u_{ik}^2 D^2(z_i, x_k))}{n(\min_{i \neq j}\{D^2(z_i, z_j)\})}. \quad (3)$$

Note that when the partitioning is compact and good, σ should be low while sep should be high, thereby yielding lower values

of the XB index. The objective is therefore to minimize the XB index for achieving proper clustering.

The J_m measure, which needs to be minimized also, is defined as [11]:

$$J_m = \sum_{j=1}^{n}\sum_{k=1}^{K} u_{kj}^m D^2(x_j, z_k), \quad 1 \leq m \leq \infty, \quad (4)$$

where m is the fuzzy exponent.

C. Genetic Operators

Crowded binary tournament selection as in NSGA-II, followed by conventional crossover and mutation are used here. The most characteristic part of NSGA-II is its elitism operation, where the non-dominated solutions among the parent and child populations are propagated to the next generation. For details on the different genetic processes, the reader may refer to [6]. The near-Pareto-optimal strings of the last generation provide the different solutions to the clustering problem.

IV. Choice of Objectives

Performance of multiobjective clustering highly depends on the choice of objectives which should be as contradictory as possible. In this article XB and J_m have been chosen as the two objectives to be optimized. From Eqn. 4 it can be noted that J_m calculates the global cluster variance, i.e., it considers the within cluster variance summed up over all the clusters. Lower value of J_m implies better clustering solution. On the other hand, XB index (Eqn. 3) is a combination of global (numerator) and local (denominator) situations. Although the numerator of XB index (σ in Eqn. 3) is similar to J_m, the denominator contains an additional term (sep) representing the separation between the two nearest clusters. Therefore, XB index is minimized by minimizing σ (or J_2), and by maximizing sep. These two terms may not attain their best values for the same partitioning when the data has complex and overlapping clusters. Since remote sensing data sets typically have such overlapping clusters (as evident from Fig. 2), considering J_m and XB (or in effect σ and sep) will provide a set of alternate partitionings of the data.

Fig. 1 shows, for the purpose of illustration, the final Pareto-optimal front (composed of non-dominated solutions) of one of the runs of the proposed algorithm for the numeric SPOT data set (described in the next section), to demonstrate the contradictory nature of J_m and XB indices.
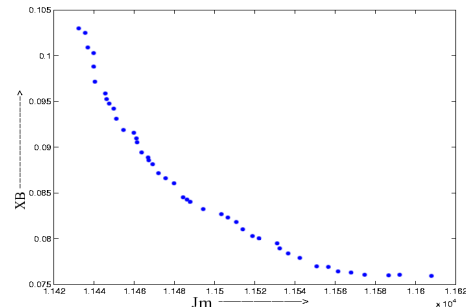

Fig. 1.   Non-dominating Pareto front for SPOT data set

V. Results for Numeric Remote Sensing Data

Two numeric satellite image data obtained from SPOT (a part of the city of Calcutta) and LANDSAT are considered here for experimenting. A numeric image data means that some pixels from several known landcover types are extracted from an image. The landcover type serves as the class label, and the intensity values (in multiple bands) serve as the different feature values. Moreover, in contrast to the normal image data, in the numeric image data, no spatial information is available as the pixel locations are not retained. The data sets are first described below.

SPOT: This is a three dimensional data set [5] (corresponding to green, red, near infrared (NIR) bands) consisting of 932 samples partitioned into seven distinct classes of turbid water (TW), pond water (PW), concrete (Concr.), vegetation (Veg), habitation (Hab), open space (OS), and roads (including bridges) (B/R). For the purpose of illustration, Fig. 2 shows the scatter plot of the data set from where it can be seen that the clusters are highly complex and overlapping in nature.
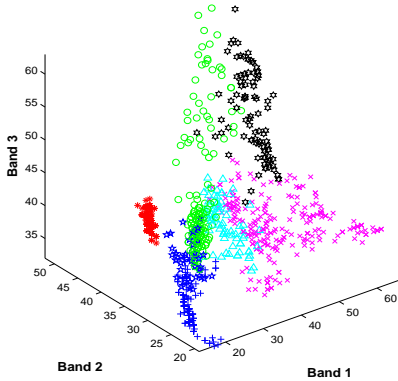

Fig. 2. Scatter plot of 932 points of SPOT image of Calcutta having 7 classes

LANDSAT: This data set has 795 samples and four bands (features) [5]: green, red, near infrared and infrared. Since the features are highly correlated, the feature space is reduced to two by using principal component analysis. Tha data set contains five classes, *viz.*, Manda Granite, Romapahari Granite, Vegetation, Black Phillite and Alluvium.

A. Performance Measure

The clustering results have been evaluated objectively, i.e., by measuring the goodness of the clusters. For this purpose, a validity index $\mathcal{I}$ [12] as well as the XB and J_m indices, described in Section III-B are used. The $\mathcal{I}$ index has been proposed recently as a measure of indicating the goodness/validity of a cluster solution. It is defined as follows:

$$\mathcal{I}(K) = \left(\frac{1}{K} \times \frac{E_1}{E_K} \times \mathcal{D}_K \right)^p, \qquad (5)$$

where $E_K = \sum_{k=1}^{K} \sum_{j=1}^{n} u_{kj} D(x_j, z_k)$ and $\mathcal{D}_K = \max_{i,j=1}^{K} \{D(z_i, z_j)\}$. The different terms are defined earlier. $\mathcal{I}$ index has been shown to provide superior performance when compared to several other validity indices [12]. In this article, we have taken $p = 2$. Larger value of $\mathcal{I}$ index implies better

solution. Note that for computing the index $\mathcal{I}$, knowledge about the true partioning of the data is not necessary.

B. Input Parameters

The parameters of the GA based clustering (both single objective and multiobjective) are as follows: population size = 50, number of generations = 100, crossover probability = 0.8, mutation probability = $\frac{1}{length\ of\ chromosome}$. Results reported in the tables are average values obtained over ten runs of the algorithms. FCM is executed for a maximum of 100 iterations, with m, the fuzzy exponent, equal to 2.0.

C. Results

Tables I and II present the performance of the proposed method for clustering the SPOT and LANDSAT numeric image data (which, as mentioned earlier, correspond to pixel values of some known landcover types) respectively. The solution providing the best value of the $\mathcal{I}$ index is selected from the set of non-dominated solutions. The J_m, XB and $\mathcal{I}$ index values, along with the percentage of correctly classified pairs ($\%CP$), are reported corresponding to this solution. For the purpose of comparison, three other clustering algorithms, *viz.*, a single objective genetic clustering optimizing the XB index only, FCM and average linkage, are considered.

TABLE I
RESULTS FOR NUMERIC SPOT DATA

Method	XB	J_m	$\mathcal{I}$	$\%CP$
Multi objective	0.0818	11535.7695	689.6587	89.12
Single objective (XB)	0.0714	11984.5623	635.3235	87.35
FCM	0.1712	11425.6113	634.3225	87.26
Average linkage	1.3476	14890.8848	517.8401	79.81

TABLE II
RESULTS FOR THE LANDSAT DATA

Method	XB	J_m	$\mathcal{I}$	$\%CP$
Multi objective	0.1148	23226.9785	38334.5092	91.18
Single objective (XB)	0.0924	23627.5443	37901.9102	88.42
FCM	0.1591	23014.8878	38242.7409	88.94
Average linkage	2.2355	24080.2842	34253.9776	84.33

The single objective genetic clustering algorithm minimizes the XB index only. Hence, as expected, it provides the best XB index value (Tables I and II). However, the J_m value reported for this algorithm (as attained by the solution with the best XB index) is quite poor. Again, since FCM optimizes the J_m index, it provides the best value for this index. The corresponding XB index value (as attained by the solution with the best J_m value) is once again quite poor. The multiobjective method is found to provide values of the XB and J_m indices that are only slightly poorer than the best values. Interestingly, in terms of $\mathcal{I}$ index, which none of the algorithms optimizes directly, the multiobjective method significantly outperforms the other methods. Also, the multiobjective scheme provides better classification accuracy in terms of the %CP scores. This signifies that in terms of the algorithm independent measure of clustering goodness (or, validity), i.e.., the $\mathcal{I}$ index, just

optimizing the XB index (as the single objective version does) or the J_m index (as the FCM does) is not good enough. A trade-off solution, provided by the multiobjective scheme, that balances both the objectives, appears to be better. The average linkage method provides the poorest scores which could be due to the highly overlapping nature of the clusters as evident from Fig. 2. Thus, this section highlights the importance of utilizing multiple criteria and optimizing them simultaneously rather than using only a single optimizing objective.

VI. Pixel Classification

Three image data sets used for the experiments are two Indian remote sensing satellite (IRS) images of the parts of the cities of Calcutta and Mumbai [1], and a SPOT satellite image of a part of the city of Calcutta [5]. Each image is of size 512×512, i.e., the size of the data set to be clustered in all the images is 262144. The IRS images consist of four bands *viz.*, blue, green, red and near infrared, whereas, the SPOT image consists of green, red and near infrared bands. To show the effectiveness of the proposed multiobjective technique quantitatively, the cluster goodness index $\mathcal{I}$ has been examined. The efficiency of multiobjective genetic clustering can also be verified visually from the clustered images and comparing them with the available ground knowledge about the landcover areas.

A. IRS image of Calcutta

Fig. 3 shows the IRS Calcutta image clustered using proposed multiobjective GA clustering scheme. From our ground knowledge, we know that the image has four classes [1]: turbid water (TW), pond water (PW), concrete (Concr.) and open space (OS). The river Hooghly cutting across the middle of the image has been classified as TW, whereas several fisheries observed towards the lower-right portion of the image are correctly identified as PW. It appears from the figure that the water class has been differentiated into TW (the river Hooghly) and PW (canal, fisheries etc.) because they differ in their spectral properties. Towards the lower right side of the image, a township, Salt Lake, has come out partially as combination of concrete and open space, which appears to be correct, since this particular region is known to have several open spaces. The canal bounding Salt Lake from the upper portion has also been correctly classified as PW. Two parallel lines observed towards the upper right hand side of the image correspond to the airstrips in the Dumdum airport, and the airstrip is classified rightly as belonging to the class concrete. Presence of some small areas of PW beside the airstrips is correct again as these correspond to the several ponds around the region. The predominance of concrete on both sides of the river, particularly towards the bottom of the image, i.e., the central part of the city is also correct.

Fig. 4 shows the Calcutta image partitioned using FCM algorithm. From the figure, it can be noted that the river Hooghly and the city region has been incorrectly classified as belonging to the same class. Another flaw is apparent that the whole Salt Lake city has been put into one class. It is evident that although some portions like the canals, parts of

airstrip, fisheries are correctly identified, a significant amount of confusion lies in the FCM clustering result.
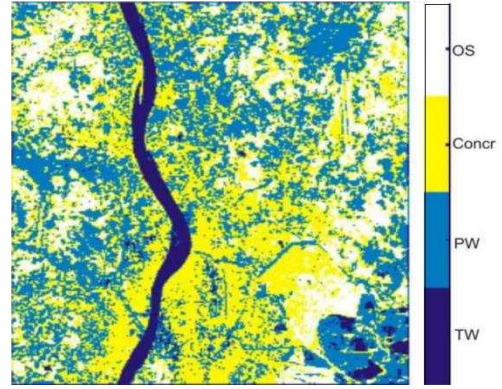

Fig. 3. Clustered IRS image of Calcutta using multiobjective GA
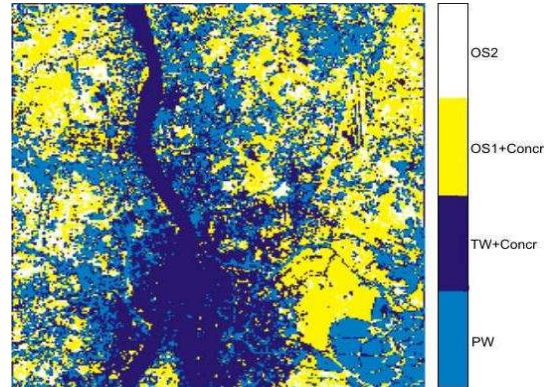

Fig. 4. Clustered IRS image of Calcutta using FCM

B. IRS image of Mumbai

Fig. 5 shows the Mumbai image segmented using proposed multiobjective technique. According to the available ground knowledge [1], the different clusters are labelled as, concrete (Concr.), open spaces (OS1 and OS2), vegetation (Veg), habitation (Hab) and turbid water (TW1 and TW2). As can be seen, the elongated city area is surrounded on three sides by the Arabian sea which is distinguished into two classes TW1 and TW2. It is evident from figure that the sea water has two distinct regions with different spectral properties. Hence the clustering result providing two partitions for this region is expected. Towards the bottom right of the image, there are several islands, including the well known Elephanta islands. The dockyard is situated on the south eastern part of Mumbai, which can be seen as a set of three finger like structure. Note that the classes habitation and concrete share common properties. The islands, dockyard, several road structures have mostly been correctly identified in the image. Within the islands, as expected, there is a high proportion of open space and vegetation. The southern part of the city, which is heavily industrialized, has been classified as primarily belonging to habitation and concrete.

Fig. 6 demonstrates the Mumbai image clustered using the FCM technique. As evident from the figure, the water of the

Arabian sea has been partitioned into three regions, rather than two as obtained earlier. The other regions appear to be classified more or less correctly for this data.
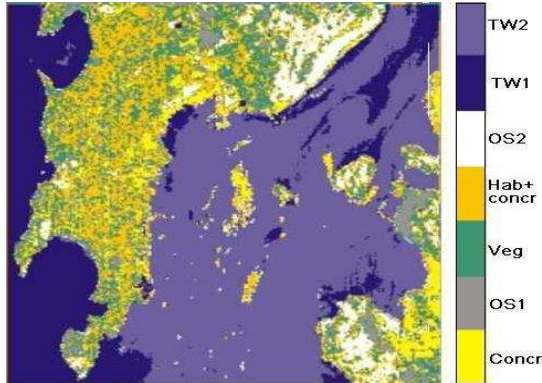

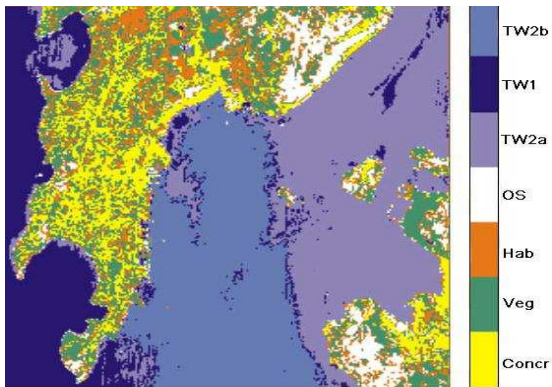Fig. 5. Clustered IRS image of Mumbai using multiobjective GA

Fig. 6. Clustered IRS image of Mumbai using FCM

C. SPOT Image of Calcutta

Fig. 7 shows the SPOT image of Calcutta clustered using proposed multiobjective scheme. The image has seven classes [5]: turbid water (TW), pond water (PW), concrete (Concr.), vegetation (Veg), habitation (Hab), open space (OS), and roads (including bridges) (B/R). The river Hooghly cuts through the image and it is correctly classified as TW. Below the river, on the left side, two distinct patches are the water bodies Khidirpore dockyard (on the right), and Garden Reach lake (on the left). Each of the water bodies is rightly classified as PW. Just on right side of those water bodies, a very thin line is noticed, starting from the river and going towards the bottom edge of the picture. This is a canal called the Talis nala and this is also correctly placed into TW class. The triangular patch on right side of Talis nala is the race course. There is a thin line starting from the top right hand corner of the picture and stretching towards the middle of it. This is the Beleghata canal (classified as TW) with a road beside it. Several roads (combination of habitation, concrete and B/R) are found at the right side of the picture mainly near the top and middle portions. Around the center of the image, a bridge (second Hooghly bridge) can be noticed across the river Hooghly. The

bridge was being constructed when the picture was taken. Another bridge, called Rabindra Setu, cuts the river near the top of the image. Both the bridges are correctly identified as a combination of B/R and concrete.
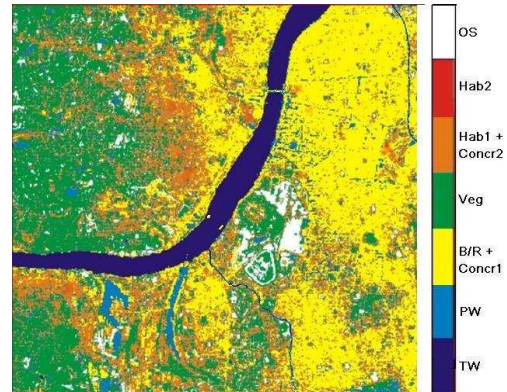

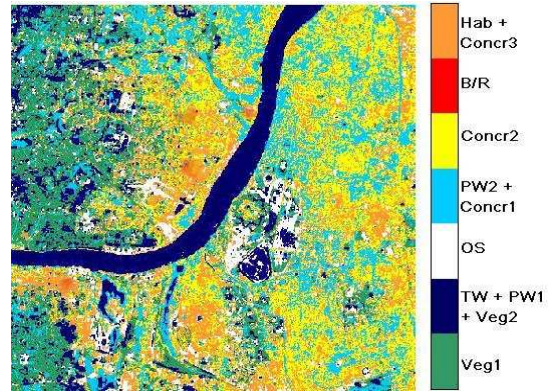Fig. 7. Clustered SPOT image of Calcutta using multiobjective GA

Fig. 8. Clustered SPOT image of Calcutta using FCM

It seems from the Fig. 7 that there are some confusions between the classes Concr and B/R and between Concr and Hab. These are expected as these classes have large amount of overlap in reality. In contrast, FCM algorithm performs poorly (Fig. 8), having a significant amount of confusion among the classes TW, PW, Veg and Concr. For example, the Khidirpore dockyard has come out as a combination of PW and Concr. Also, FCM cannot distinguish between the two water classes (TW and PW). The Talis nala is also not very prominent in this case. Also FCM fails to recognize Rabindra Setu properly.

For the purpose of illustration, the segmentation results of the race course part of the image are magnified in Fig. 9 for different clustering algorithms. It can be noticed that only the multiobjective method is able to identify a lighter outline within it properly. This line corresponds to the tracks of the race course and belongs to the class open space (OS).

Table III shows the $\mathcal{I}$ index values for all the images. The values indicate the superiority of the multiobjective technique which produces a far better value of the $\mathcal{I}$ index than those of single objective clustering optimizing XB and FCM algorithm. The results therefore demonstrate that the multiobjective fuzzy genetic clustering technique clearly outperforms the related methods for all three images.
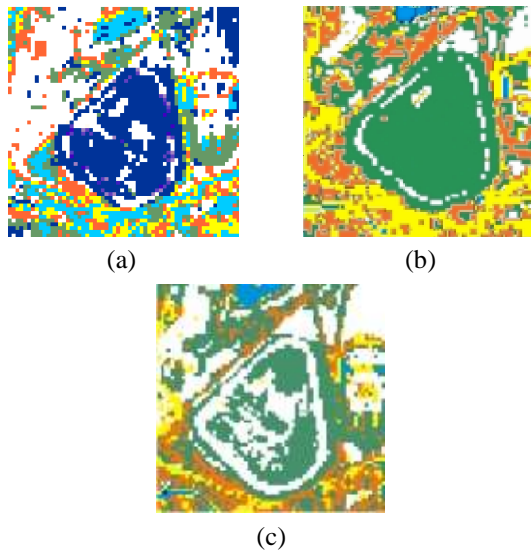
Fig. 9.   Segmented SPOT image of Calcutta (zooming the race course) using (a) FCM algorithm (b) Single objective GA minimizing XB (c) Multiobjective GA optimizing J_m and XB

TABLE III

$\mathcal{I}$ INDEX VALUES FOR CALCUTTA AND MUMBAI IMAGES FOR DIFFERENT ALGORITHMS

Method	Calcutta(IRS)	Mumbai(IRS)	Calcutta(SPOT)
Multi-objective clustering	96.2788	183.7727	97.6453
Single objective clustering (XB)	81.5934	180.4512	88.9346
FCM	31.1697	178.0322	82.2081

VII. DISCUSSION AND CONCLUSIONS

The problem of fuzzy clustering has been modelled as simultaneous optimization of two cluster validity measures, namely, Xie-Beni (XB) index and the J_m. In this regard, the well known NSGA-II algorithm has been used. Results on numeric image data sets indicate that rather than considering these indices individually, better clustering performance is obtained if both of them are optimized simultaneously. In this context, IRS satellite images of Calcutta and Mumbai and a SPOT image of Calcutta have been classified using the proposed technique and compared with single objective GA and FCM algorithms to show its effectiveness. Good performance of multiobjective genetic clustering method for such large image data sets shows that it may be motivating to use this algorithm in data mining applications also.

As a scope of further research, the technique of multiobjective optimization with other cluster validity indices needs to be studied. Moreover, new ways of comparing the performance of multiobjective solutions have to be defined. Finally, comparative study with other approaches of multiobjective optimization should be performed. The authors are currently working in these directions.

REFERENCES

[1] U. Maulik and S. Bandyopadhyay, "Fuzzy partitioning using a real-coded variable-length genetic algorithm for pixel classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 5, pp. 1075– 1081, 2003.

[2] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[3] J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles*. Reading: Addison-Wesley, 1974.

[4] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. New York: Addison-Wesley, 1989.

[5] S. Bandyopadhyay and S. K. Pal, "Pixel classification using variable string genetic algorithms with chromosome differentiation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 2, pp. 303– 308, 2001.

[6] K. Deb, *Multi-objective Optimization Using Evolutionary Algorithms*. England: John Wiley and Sons, Ltd, 2001.

[7] E. Zitzler and L. Thiele, "An evolutionary algorithm for multiobjective optimization: The Strength Pareto approach," Gloriastrasse 35, CH-8092 Zurich, Switzerland, Tech. Rep. 43, 1998.

[8] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the Strength Pareto Evolutionary Algorithm," Gloriastrasse 35, CH-8092 Zurich, Switzerland, Tech. Rep. 103, 2001.

[9] C. A. Coello Coello, "A comprehensive survey of evolutionary-based multiobjective optimization techniques," *Knowledge and Information Systems*, vol. 1, no. 3, pp. 129–156, 1999.

[10] X. L. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 841–847, 1991.

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

[12] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1650–1654, 2002.

Sanghamitra Bandyopadhyay, Associate Professor, Indian Statistical Institute, Kolkata, India, did her PhD in 1998. She has worked in Los Alamos National Laboratory, USA, University of New South Wales, Australia, University of Texas at Arlington, USA, University of Maryland, Baltimore County, USA, Fraunhofer Institute AiS, Germany and Tsinghua University, China. Dr. Bandyopadhyay is a co-author of 3 books and more than 100 papers. Her main research interests include Pattern Recognition, Data Mining, Soft Computing and Bioinformatics.

Ujjwal Maulik, Professor, Dept. of Comp. Sc. and Engg., Jadavpur Univ., Kolkata, India, did his PhD in 1997. Dr. Maulik has worked in Center for Adaptive Systems Application, New Mexico, USA, University of New South Wales, Australia, University of Texas at Arlington, USA, University of Maryland Baltimore County, USA and Fraunhofer Institute AiS, Germany. He is a co-author of 2 books and about 100 papers. His main research interests include Evolutionary Computing, Pattern Recognition, Data Mining, Bioinformatics and Distributed System.

Anirban Mukhopadhyay, faculty member, Dept. of Comp. Sc. and Engg., Univ. of Kalyani, India, received the MS in Comp. Sc. in 2004 and is currently pursuing his PhD from Jadavpur Univ., Kolkata, India. He is the recipient of the University Gold Medal and Amitava Dey Memorial Gold Medal from Jadavpur Univ. in 2004. He is a co-author of about 15 papers. His research interests include Soft and Evolutionary Computing, Data Mining, Bioinformatics and Optical Networks.