International Conference on "Emerging Trends in Computer Engineering, Science and Information Technology”- 2015
Special Issue of International Journal of Electronics, Communication & Soft Computing Science and Engineering, ISSN: 2277-9477

Particle Swarm Optimization (PSO) based approach for Classification of Remote Sensing Images

Geeta R. Gupta **Prof. S. M. Kamalapur**

Abstract- **Dimensionality reduction is a major task in remote sensing images. Feature selection is applied for performing dimensionality reduction. It selects the spectral features(i.e. Bands) and find a feature subset that preserves the semantics of the hyperspectral image. Based on particle swarm optimization (PSO), this paper proposes multi-objective functions for selecting the spectral feature subsets for classification. The multi-objective function select feature subsets based on Jeffries Matusita(JM) distance and classifier(i.e. SVM). This paper performs optimal band selection and dimensionality reduction of hyperspectral imagery. The goal of the system is to perform spectral feature selection using particle swarm optimization (PSO) based multi-objective function. The system implements multi-objective functions which performs spectral feature selection (i.e. most informative bands) from the hyperspectral image dataset. These selected features are further used for evaluating the overall classification accuracy.**

***Keywords* — Dimension Reduction; Hyperspectral Imaging; Image Spectroscopy; Particle Swarm Optimization(PSO); Remote Sensing; Spectral features**

I. INTRODUCTION

Remote sensing images are high resolution images having huge dimensions of data. It obtains information about objects over the surface of the earth without physically coming into contact with them. Observations are made by making use of sensors. Various applications of remote sensing are the effects of change in climate on glaciers and Arctic and Antarctic regions, monitoring deforestation in areas such as the Amazon Basin [1]. Remote sensing data consists of its spatial and spectral resolutions: Spatial resolution is the angular difference between two objects that can be identified by the sensor. Spectral resolution is the dimension and number of specific wavelength intervals in the electromagnetic spectrum. Due to narrow bandwidths in certain regions of the spectrum various features are easily differentiated. Image spectroscopy or hyperspectral analysis is the process of scanning hundreds of closely spaced and very narrow spectral bands as shown in Fig. 1. Hyperspectral data is stored in 3D Data cube form as shown in Fig. 2.

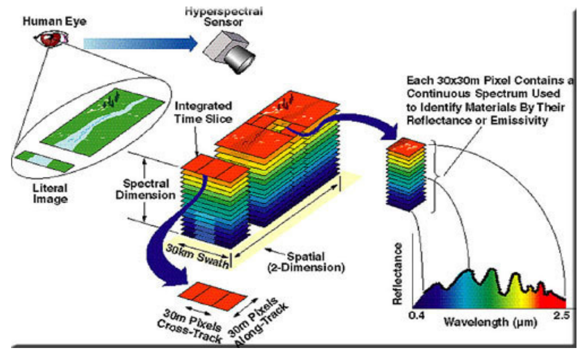

Fig. 1. Imaging Spectroscopy[6]

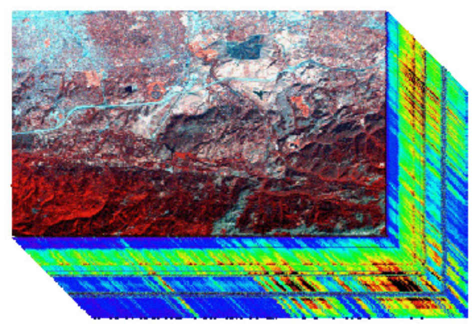

Fig. 2. Hyperspectral datacube[6]

However, the large number of features causes the problem of “the curse of dimensionality”, which is a major problem in classification. All of the features are not useful for classification. Redundant and irrelevant features may increase the classification error rate. Feature selection (FS) can reduce the number of features by eliminating redundant and irrelevant features, and thus it results in increased classification performance and/or increased efficiency[5].

For the applications of hyperspectral image classification, it seems that the high dimensionality of hyperspectral data should increase the abilities in classifying land use/cover types. But, as the dimensionality increases with the number of bands, the number of training samples required for training a specific classifier should be increased as well. So, the limitation in number of training samples leads to the peaking phenomenon or Hughes phenomenon [1, 2]. To tackle with the hyperspectral image Classification problem many supervised and unsupervised classification have been addressed.

This paper is organized as follows: Section I gives a brief introduction of proposed system. Section II includes the literature survey. Section III describes the Block diagram and implementation of proposed system. Section IV includes the results and discussion . Finally in Section V, we conclude with the summary of this paper.

II. Literature Survey

Remote sensing images are categorized into three types which are Multispectral Image, Superspectral Image and Hyperspectral Image. They differ based on the number of spectral bands. The proposed system focuses on hyperspectral images which consists of hundred or more contiguous spectral bands forming a three-dimensional (one spectral and two spatial dimension) image cube. Dimensionality reduction has been widely used in hyperspectral image analysis to reduce data volume and redundancy[4]. **Supervised classification** requires training data in order to build an predictive model. The training data set is obtained by registering the hyperspectral imagery with ground measurements. Supervised methods are computationally intensive than unsupervised methods due to an arbitrarily high model complexity and an iterative nature of model formation[5]. Another requirement is that the number of examples in a training set should be larger than the number of attributes (i.e. bands). This requirement might be hard to meet.

In general, supervised classification, can be classified into two approaches : **Filter approach -** The filter approach is performed independently from any classification algorithm, so before the classification process begins unwanted features are filtered or omitted. **Wrapper approach -** It optimizes the classification accuracy of the desired classifier by selecting feature subsets. In supervised Classification, Land cover classes are defined. Sufficient reference data are available and this data are used as training samples [11]. Most commonly used supervised classification approaches are maximum likelihood, neural network and decision tree. Supervised Classification is one of the tasks carried out by Intelligent System.

III. Implementation Details

This section contain details of proposed system i.e. block diagram and implementation of proposed system.

A. *Block Diagram*

The overall Block diagram of the system is shown in the Fig. 3.

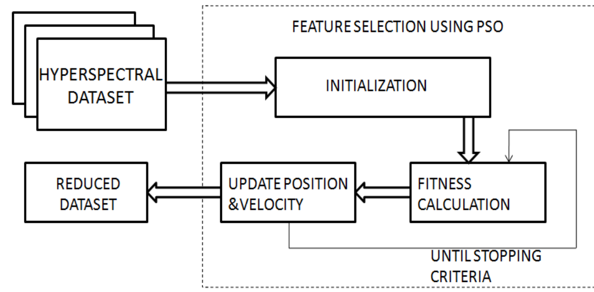

Fig. 3. Block diagram of proposed system

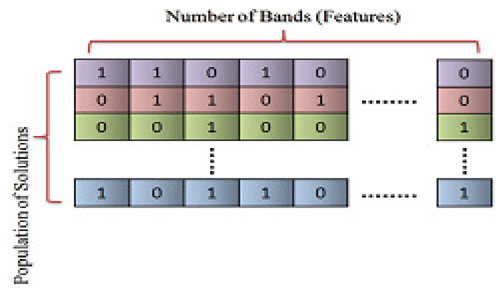

Fig. 4. Initialization of the population of solutions

1. Hyperspectral Dataset : Hyperpectral data which is to be optimized(reduced). It contains information of hyperspectral image and its ground truth image.

2. Feature Selection : Particle Swarm Optimization (PSO) based multi-objective function approach is used for selecting the most informative bands/features from the original bands/features.

   Following are the steps for feature selection using multi-objective PSO:

   i) Initialization : All the particles/solutions of PSO are initialized with a string with length equal to the number of hyperspectral bands. For binary PSO, an n bit string represents the bands where, value 1 means inclusion of band and 0 represents its absence as shown in Fig.4.

   ii) Fitness calculation : Based on the multi-objective function(i.e. JM distance and SVM classifier) the fitness value for each particle is calculated.

   iii) Update position and velocity : After selecting the best neighborhood particle (i.e. the particle having maximum fitness value), update the position and velocity of each particle.

   Finally a particle having maximum fitness value amongst all the particles is selected and reduced dataset is obtained.

B. Implementation of Proposed System

Input to the system is hyperspectral images consisting of various bands. This dataset is preprocessed i.e. all the atmospheric noise is removed. Multi-objective function is used for selecting the feature subsets. The multi-objective function used are Jeffries Matusita(JM) Distance and classifier(SVM). Optimization will be done based on PSO approach by using the multi-objective function (i.e. JM distance and SVM) defined in (1).

$$JM_{ij} = \sqrt{2(1 - e^{-\alpha})} \quad (1)$$

where,

$$\alpha = \frac{1}{8}(\mu_i - \mu_j)^T\left(\frac{C_i + C_j}{2}\right)^{-1}(\mu_i - \mu_j) + \frac{1}{2}\ln\left(\frac{|(C_i + C_j)/2|}{\sqrt{|C_i| \times |C_j|}}\right) \quad (2)$$

i and j are two signatures (classes) being compared, C_i is the covariance matrix of signature i, μ_i is the mean vector of signature i, $|C_i|$ is the determinant of C_i. For multiclass following averaging function is used as shown in (3):

$$d_{ave} = \sum_{i=1}^{C}\sum_{j=1}^{C} p(\omega_i)p(\omega_k)J_{ij} \quad (3)$$

where $p(\omega_i)$,$p(\omega_j)$ are class prior probabilities.

Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a technique used to explore the search space of a given problem to find the parameters required to maximize a particular objective. This technique, first described by James Kennedy and Russell C. Eberhart in 1995 [3], originates from two separate concepts: the idea of swarm intelligence based on the observation of swarming habits of animals (such as birds and fish); and the field of evolutionary computation. Particle Swarm Optimization (PSO) based problem solving approach is an evolutionary algorithm that computes the fittest solution out of the candidate solutions using the local and global best information. This algorithm examines all possible ways to solve the problem and will pick the best solution.

A solution is chosen from a swarm of possible solutions accordingly and is checked for optimality of fitness value. Thus after each unfit solution is encountered the process is back tracked and the swarm is updated accordingly. Finally we obtain the optimized solution. Further, this optimized solution (i.e. reduced features) can be combined to be used with classifier in order to improve the classification quality of the dataset.

In this paper, Binary Particle Swarm Optimization(BPSO) based multi-objective feature selection approach is used for reducing the number of bands in hyperspectral data and is discussed here.

Input: S is the set of all bands with n-dimensions i.e. {$b_1, b_2, ..., b_n$}, swarm size (p)

Output: The system will generate the dataset with reduced number of dimensions say {$b'_1, b'_2, ..., b'_p$} where $p < n$

The basic steps of the algorithm for reducing dimensions are listed below:

1. Assume p bands are to be selected. Randomly initialize M particles x_{id} , and each particle includes p indices of the bands to be selected.

2. Evaluate the multi-objective function(JM distance and SVM classifier) for each particle, and determine the local and global optimal solution p_{id} and p_{gd} respectively.

3. Update all the particles.
   Based on the global best particle updated all other particles.
   (a) Update particles velocity by using(2):
   $$v_i[t+1] = \omega \times v_i[t] + c_1 \times r_1[t] \times (p_{id}[t] - present[t]) + c_2 \times r_2[t] \times (p_{gd}[t] - present[t]) \quad (2)$$
   (b) Update particles position by using (3),(4):
   $$x_{id} = \begin{cases} 1, & \text{if } rand() < s(v_{id}) \\ 0, & otherwise \end{cases} \quad (3)$$
   where
   $$s(v_{id}) = \frac{1}{1 + e^{-v_{id}}} \quad (4)$$
   where rand() is a random number ranging from [0,1].

4. If the algorithm is converged, then stop; otherwise, go to step 2.

5. The particle yielding the global optimum solution p_{gd} is the final result and overall classification accuracy is evaluated.

BPSO preserves the fundamental concept of the PSO algorithm. Hence, for feature selection BPSO is used as it includes discrete values i.e. [0,1].

IV. Results & Discussion

This section describes about the results and hyperspectral image datasets used by the system for experimentation purpose. AVIRIS (Airborne Visible/Infrared Imaging Spectrometer), ROSIS and HYDICE are the well known sensors from which hyperspectral images are captured. So, this system uses AVIRIS dataset to reduce its high dimensional data to low dimension space and to evaluate the classification accuracy. Overall accuracy(OA) as defined in (5) is adopted as assessment criteria.

$$O^c = \frac{\sum_{i=1}^{k} n_{ij}}{|T|} \qquad (5)$$

where, nij is element of error matrix and |T| is the number of pixel used for testing.

A. AVIRIS

Hyperspectral image used in the system was acquired by the AVIRIS sensor over the Indian Pines region in

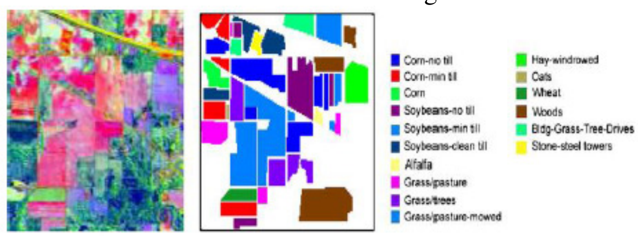

Fig. 5. (a) False-color composition of the AVIRIS Indian Pines scene. (b) Reference map containing (right) 16 mutually exclusive land-cover classes.

Northwestern Indiana in 1992. The image measured 145 × 145 samples was acquired over a mixed agricultural/forest area. The scene comprises 220 spectral channels. There are 16 land-cover classes as shown in Fig. 5(b).

B. Experimental Setup

In experiment, BPSO based multi-objective feature selection approach is applied on the original dataset to reduce its dimensions. Algorithm includes initialization, in this random initialization of a population of particles is done as shown in table I. For each particle fitness is calculated by using multi-objective function as shown in table II & III and best neighborhood particle is selected. Based on the best neighborhood particle the velocity and position of each particle is updated. The algorithm continues until the best neighborhood particle amongst all the particle is selected or maximum number of iteration is reached.

TABLE I. INITIALIZATION OF PARTICLES

Particles	B1	B2	B3	B4	.....	B200
1	1	0	0	1	.....	1
2	1	0	1	0	.....	0
3	0	1	0	0	.....	0
4	1	0	0	0	......	1
.	.	.	.	.	......	.
.	.	.	.	.	......	.
.	.	.	.	.	......	.
.	.	.	.	.	......	.

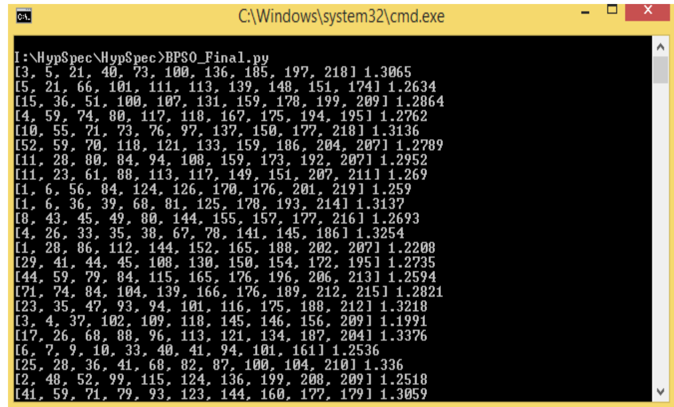

Fig. 5. Fitness Calculation of each particle using JM-Distance

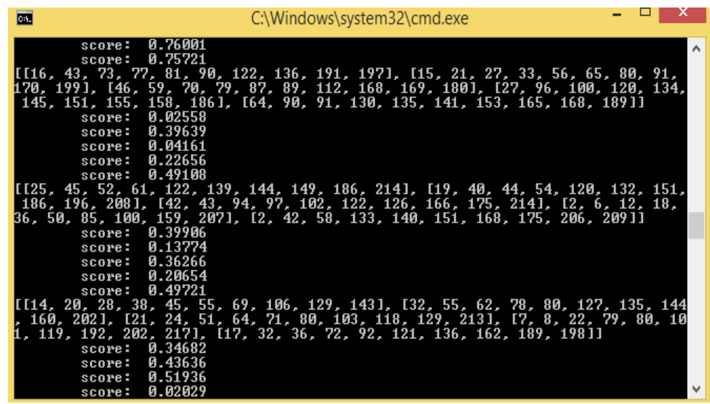

Fig. 6. Fitness Calculation of each particle using SVM

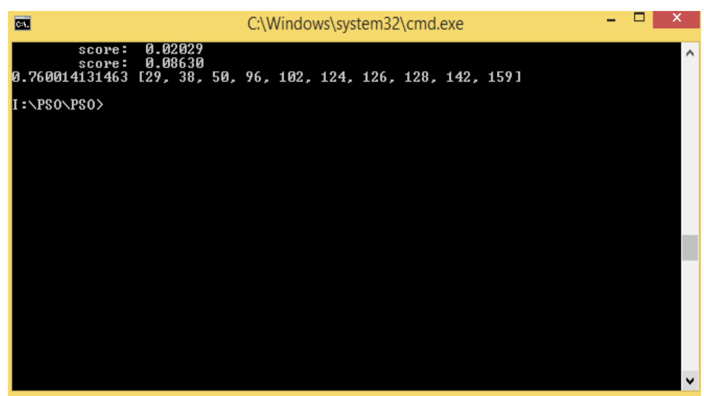

Fig. 7. Final solution obtained using SVM

CONCLUSION

Hyperspectral images have high spatial coverage and large number of spectral features. Handling this huge dimensionality of data becomes a vital problem. So Feature selection(FS) has become an important issue in hyperspectral images. The objective is to optimize the hyperspectral dataset with minimum number of spectral features (i.e. spectral bands) without any loss of semantic information (i.e. informative bands). Hence, a BPSO-based multi-objective system is introduced to select the optimal number of bands for hyperspectral dimensionality reduction.

Based on this, the system implements the BPSO based multi-objective feature selection algorithm which can select the most informative bands with optimized set of spectral features that yields increased classification performance. In this way, the original hyperspectral dataset has been reduced containing less number of spectral bands and having improved classification accuracy.

ACKNOWLEDGMENT

I would like to express my special thank to all those people who have helped me to complete this work. I am very grateful to my guide, Prof. S. M. Kamalapur, for insightful comments. She has continuously helped and encouraged me in my work.

REFERENCES

[1] Hongjun Su, Member, IEEE, Qian Du, Senior Member, IEEE, Genshe Chen, and Peijun Du, Senior Member, IEEE, "Optimized Hyperspectral Band Selection Using Particle Swarm Optimization", IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, VOL. 7, NO. 6, JUNE 2014.

[2] Jun Li, Prashanth Reddy Marpu, Antonio Plaza, Jose M.Bioucas-Dias, and Jon Atli Benediktsson, "Generalized Composite Kernel Framework for Hyperspectral Image Classification", IEEE transactions. on geoscience and remote sensing, vol. 51, no. 9, September 2013.

[3] H. Yang, Q. Du, and G. Chen, "Particle swarm optimization-based hyperspectral dimensionality reduction for urban land cover classification", IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 5, no. 2, Apr. 2012.

[4] H. Yang, Q. Du, H. Su, and Y. Sheng, "An efficient method for supervised hyperspectral band selection", IEEE Geosci. Remote Sens. Lett., vol. 8, no. 1, Jan. 2011.

[5] N. Keshava, "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries", IEEE Trans. Geosci. Remote Sens., vol. 42, no. 7, Jul. 2004.

[6] Lu .D and Weng.Q, (2007), "A Survey of Image Classification methods and techniques for improving classification Performance", International Journal of Remote Sensing, 28, 5.

[7] I. C. Trelea, "The particle swarm optimization algorithm: Convergence analysis and parameter selection", Inf. Process. Lett., vol. 85, no. 6, 2003.

[8] S. V. Stehman, "Estimating the kappa coefficient and its variance under stratified random sampling", Photogramm. Eng. Remote Sens., vol. 62, no. 4, 1996.

[9] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swam theory", in Proc. 6th Int. Symp. Micromach. Human Sci., 1995.

[10] Q. Du and H. Yang, "Similarity-based unsupervised band selection for hyperspectral image analysis", IEEE Geosci. Remote Sens. Lett., vol. 5, no. 4, Oct. 2008.

[11] H. Su, H. Yang, Q. Du, and Y. Sheng, "Semi-supervised band clustering for dimensionality reduction of hyperspectral imagery", IEEE Geosci. Remote Sens. Lett., vol. 8, no. 6, Nov. 2011.

[12] S. Jia, Z. Ji, Y. Qian, and L. Shen, "Unsupervised band selection for hyperspectral imagery classification without manual band removal", IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 5, no. 2, Apr. 2012.

[13] S. T. Monteiro and Y. Kosugi, "A particle swarm optimization-based approach for hyperspectral band selection", in Proc. IEEE Congr. Evol. Comput., 2007.

[14] Yao.F and Qian,Y.(2009), "Band Selection based Gaussian processes for Hyperspectral Remote Sensing Images Classification", IEEE, ICIP.

[15] Melgani.F and L. Bruzzone,(2004), "Classification of hyperspectral remote sensing images with support vector machines", IEEE transaction on geo science and remote sensing, 42.

[16] Mercier.,G.and Lennon,M. "SVM for hyperspectral image classification with spectral- based kernels", IEEE, 6.

[17] Schott.J.R, Lee.K, Raqueno. R.V, Hoffmann G.D, Healey.G, (2003), A Subpixel target detection technique based on the invariant approach, AVIRIS, AVIRIS workshop, Pasadena.

[18] Gualteri, J. A., Chettri, S. R., Cromp, R. F., & Johnson, L. F. (1999). "Support vector machines applied to AVIRIS data". Summaries of the Airborne Earth Science Workshop.

[19] Bruzzone, L., Chi, M., & Maeconcini, M. (2006). "A novel transductive SVM for the semisupervised classification of remote sensing images". IEEE Transactions on Geoscience and Remote Sensing.

[20] B. Chakraborty. Genetic algorithm with fuzzy fitness function for feature selection. In IEEE International Symposium on Industrial Electronics (ISIE'02), volume 1, pages 315– 319, 2002.

[21] B. Chakraborty. Feature subset selection by particle swarm optimization with fuzzy fitness function. In 3rd International Conference on Intelligent System and Knowledge Engineering (ISKE'08), volume 1, pages 1038–1042, 2008.

AUTHOR'S PROFILE

Geeta R. Gupta
Received the BE Degree from Pune University, in 2013. Her research interests include Machine learning, Image processing, Swarm intelligence.

Prof. S. M. Kamalapur
Associate Professor, Department of Computer Engineering, KKWIEER, Nashik. Her research interests include Machine learning, Swarm intelligence, Pattern Recognition.