

Artificial Bee Colony Algorithm Based Endmember Extraction in Hyperspectral Images

G. Veera Senthil Kumar ¹, A. Annie Persis ², K. Aneesa Sulthana ², K. Kavya ²

Department of Electronics and Communication Engineering,
Velammal college of Engineering and Technology, Madurai-625009

Abstract--Hyperspectral imaging, like other spectral imaging, collects and processes information across the electromagnetic spectrum. The hyperspectral remote sensing images usually consist of mixed pixels. Spectral unmixing is the procedure by which the measured spectrum of a mixed pixel is decomposed into a collection of constituent spectra, or endmembers, and a set of corresponding fractions, or abundances, that indicate the proportion of each endmember present in the pixel. Since hyperspectral image suffers from high dimensionality, Dimensionality Reduction (DR) is employed in this work. In the proposed work, spectral unmixing of images is done using Artificial Bee Colony (ABC) algorithm. This method improves the robustness which was a major drawback in the existing methods. The ABC approach possesses several advantages over other swarm intelligence algorithms in the context of continuous-domain optimization. The design requirements for the objective function are also flexible. Therefore, the proposed approach in this paper improves upon the models used in geometric methods to achieve greater robustness. Experiments are carried out using real hyperspectral dataset and the performance of the proposed method is evaluated using spectral angle distance.

Keywords-- Hyperspectral images, Dimensionality Reduction, Endmember extraction, Spectral unmixing, Artificial Bee Colony Algorithm.

I. INTRODUCTION

Hyperspectral imaging collects and processes information like other spectral imaging from across the electromagnetic spectrum. The main objective of the hyperspectral imaging is to obtain the spectrum for each pixel in the image, with the purpose to find objects and to identify materials. Hyperspectral imaging is commonly referred to as spectral imaging or spectral analysis. The distinction between hyper and multispectral imaging is sometimes based on the number of bands. The hyperspectral image consists of much narrower bands (10-20 nm) and could have hundreds of thousands of bands which uses an imaging spectrometer whereas the multispectral image has 3-10 wide bands. Hence hyperspectral images provide abundant information. However, it suffers from the problem of high dimensionality. Hence dimensionality reduction is an essential preprocess in hyperspectral image analysis. The problem of dimensionality reduction can be resolved by two methods namely feature extraction and feature selection. Feature selection, also called as band selection, preserves the originality of the data in comparison with feature

extraction method. In band selection, there is an issue of determination of number of bands to be selected. It is resolved by the concept of Virtual Dimensionality (VD) [1] because VD estimate provides the reliable estimation of number of bands.

Hyperspectral images generally have mixed pixels. Hence spectral unmixing is an essential step in hyperspectral image analysis. Spectral unmixing consists of two process namely endmember extraction and abundance fraction estimation[2]. Endmember is defined as the spectral signature of a pure pixel. Linear Spectral Mixture Model(LSMM) is widely used in hyperspectral image analysis. LSMM based endmember extraction methods are classified into geometric methods, statistical methods, sparsity-based methods, spatial contextual information based methods. Typical geometric methods include the Pixel Purity Index (PPI) [3], N-FINDR [4], Vertex Component Analysis (VCA) [5], Simplex Growing Algorithm (SGA) [6], Alternating Volume Maximization (AVMAX) [7], Successive Volume Maximization (SVMAX) [7], Minimum Volume Simplex Analysis (MVSA) [8], the Minimum Volume Enclosing Simplex (MVES) approach [9], the Robust MVES (RMVES) approach [9], and Minimum Volume Constrained Nonnegative Matrix Factorization (MVC-NMF) [10]. Swarm based algorithms have been developed rapidly and have significant advantages for solving optimization problems. Some swarm based algorithms are Ant Colony Optimization (ACO) [11] [15], Particle Swarm Optimization (PSO) [12], Firefly Algorithm (FA) [13], Artificial Bee Colony (ABC) algorithm [14] and Discrete PSO (DPSO) [16] for endmember extraction. The reason for choosing ABC algorithm over other swarm based algorithm is its fast coverage, few control parameters and better exploration and exploitation[18].

In the rest of the paper, the Virtual Dimensionality is first reviewed in section II. Band selection is explained in section III. Artificial Bee Colony algorithm for endmember extraction is presented in section IV. The proposed methodology is explained in section V. Experimental results are discussed in section VI. Conclusion is given in section VII.

II. VIRTUAL DIMENSIONALITY

The Virtual Dimensionality (VD) estimates the number of spectrally distinct signatures in the given data. It also estimates the number of bands to be selected for band selection process[22]. The Harsanyi-Farrand-Chang (HFC) method, also called as Eigen threshold based method, is used to estimate VD. The VD estimation by HFC method is based on Neyman-Pearson detection theory. For a given false alarm rate P_F , the number of bands are calculated by testing the number of failures for all bands.

III. BAND SELECTION

Band Selection is one of the Dimensionality Reduction (DR) techniques. Many algorithms for band selection are found in literature. Clustering based band selection is a simple and effective method. A cluster is a collection of similar objects between them and they are dissimilar to the other clusters. Clustering based band selection identifies the bands with high information and groups them based on minimum variance within the cluster and maximum variance between them[17]. The clustering problem can be resolved by K-MEANS which is a simple, unsupervised, robust, fast algorithm and gives better results. For each clusters the algorithm starts with K number of centroids. The K group's results when the data points in the data set are associated to nearest centroid. The new centroids are calculated and the process is repeated until the centroids are fixed.

K-MEANS Clustering Algorithm:

1. Randomly select cluster centres and calculate distance between them and the data points.
2. Assign the data points to the cluster centre which has minimum distance.
3. Form new cluster by recalculating new centroids.
4. Repeat (2) and (3) until the centroids are fixed.

IV. ARTIFICIAL BEE COLONY ALGORITHM:

Artificial Bee Colony (ABC) defined by Dervis Karaboga in 2005, is a swarm based Meta heuristic algorithm based on the foraging behavior of the bees. This algorithm is used to solve the optimization problem and has many advantages over other algorithms like few control parameters, fast coverage, good exploration and exploitation. The forage selection of bees consists of three essential components and they are food sources, employed bees and unemployed bees. The unemployed bees are further divided into onlooker bees and scout bees [20]. The solution depends on how the bees search for the best food source. The best food source depends on the richness and amount or concentration of the nectar. In artificial bee colony algorithm the swarm of bees search for the best food sources (solution). First they consider the optimization problem, and then it is converted to the best parameter vector which in turn minimizes the objective

function. Now the bees randomly find initial solution vectors and then it improves the solution by neglecting the poor solutions. The three groups of bees are assigned with a task. The employed bee are associated with a particular food source, the onlooker bee decides a food source by watching the waggle dance of employed bees and the scout bees searches the food source randomly. The ABC algorithm can be classified into four phase and they are Initial phase, Employed Bee phase, Onlooker Bee phase, Scout Bee phase. In the initial phase, all the vectors of the population of food sources are initialized and control parameters are set by scout bees. Here each food source is a solution vector to the optimization problem and each vector holds variables, which are to be optimized to minimize the objective function. In the employed bee phase, the bees have the previous solution in mind and searches for a better solution and if it finds a better solution it produces a modification on the solution (position) in its memory and forgets the old one. In the onlooker bee phase, it first listens to the waggle dance (probability of the food source) and selects the best solution. In the scout phase, the employed bees whose solution cannot be improved for a long time are called as scout bees. These bees randomly checks for food source and if it is better the onlooker bees consider this to be the best solution [19].

Pseudocode:

1. Perform virtual dimensionality estimation and Dimensionality reduction on hyperspectral images.
2. The input image (dimensionally reduced hyperspectral image) is given as $\{\tilde{r}_i\}_{i=1}^N$. The parameters are the number of endmembers is M, the number of employed bees is N_e , the number of onlooker bees N_o , the maximum number of iterations i_{max}
3. Initialize the food source $\{x_i\}_{i=1}^{N_e}$ for employed bees.
4. Repeat
5. Employed bee phase
 - Search for new food sources in the neighborhood based on the equation

$$x'_{ij} = x_{ij} + \varphi (x_{ij} - x_{kj}) \quad (1)$$
 - Objective function calculation based on the formula

$$\text{Minf}(E) = V(\{\tilde{e}_j\}_{j=1}^M) + \mu_R \text{RMSE}(\{\tilde{r}_i\}_{i=1}^N, \{\tilde{e}_j\}_{j=1}^M) \quad (2)$$
 - Fitness calculation based on

$$f_{it_i} = 1/f(x_i) \quad (3)$$
 - Select the better food source

6. Onlooker bee phase

- Calculate the probability $\{p_i\}_{j=1}^{N_e}$ for each food source chosen by the employed bees
- Repeat step 5 to select the better food source

7. Update the optimal solution x_{best} with the final food source.

8. Scout bee phase

- Randomly select food source
- If the food source is good repeat step 5 and update x_{best}

9. Process until the maximum iteration is reached.

10. The extracted endmember will be the output $[x_{best}]$.

11. End.

V. PROPOSED METHODOLOGY:

In the proposed work, water absorption and low SNR bands are first removed from the hyperspectral image. VD is then estimated using HFC method to estimate the number of bands. K-means clustering is applied to select the most informative bands from the hyperspectral image cube. The flow diagram of the proposed work has been depicted in Fig.1

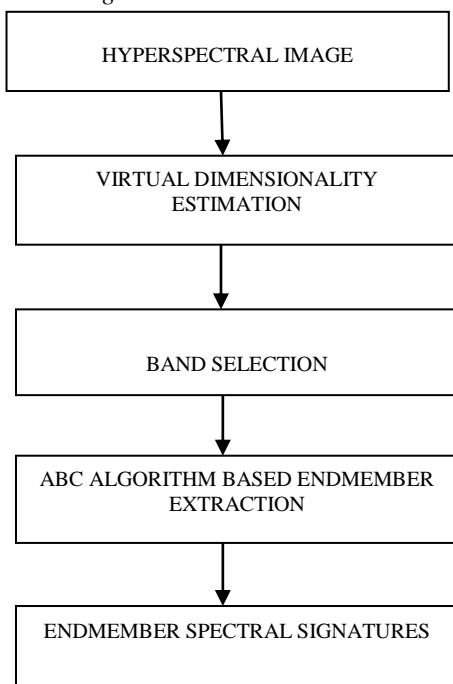


Fig.1 Block Diagram of Proposed Methodology

In the clustering method, the variance of all the bands is calculated and the bands are clustered based on the variance. The bands having maximum variance from each cluster are selected as the most informative bands. The variance of each band is calculated using the formula:

$$d_i = \frac{1}{MN} \sum_{k=1}^{MN} (b_k - \bar{B}_i) \tag{4}$$

Endmembers are then extracted from the selected bands using ABC algorithm. The ABC algorithm converges in a fast manner and has better exploration and exploitation. At last the obtained results are compared with the groundtruth using spectral angle distance.

VI. EXPERIMENTAL RESULTS:

In this experiment, a 224-band JasperRidge image as shown in Fig. 2 having size 512 x 614, pixels ranging from 0.38 to 2.5µm, is used.



Fig.2 Jasper Ridge

The spectral resolution of Jasper Ridge is up to 9.46nm. A sub image of 100 x100 pixels is considered in this experimentation because it is too complex to get the ground truth for the original image. After removing the low SNR channels 1--3, 108--112, 154--166 and 220--224, we have 198 channels. There are four endmembers in this data Road, Soil, Water and Tree. The groundtruth of Jasper Ridge is shown in Fig.3.

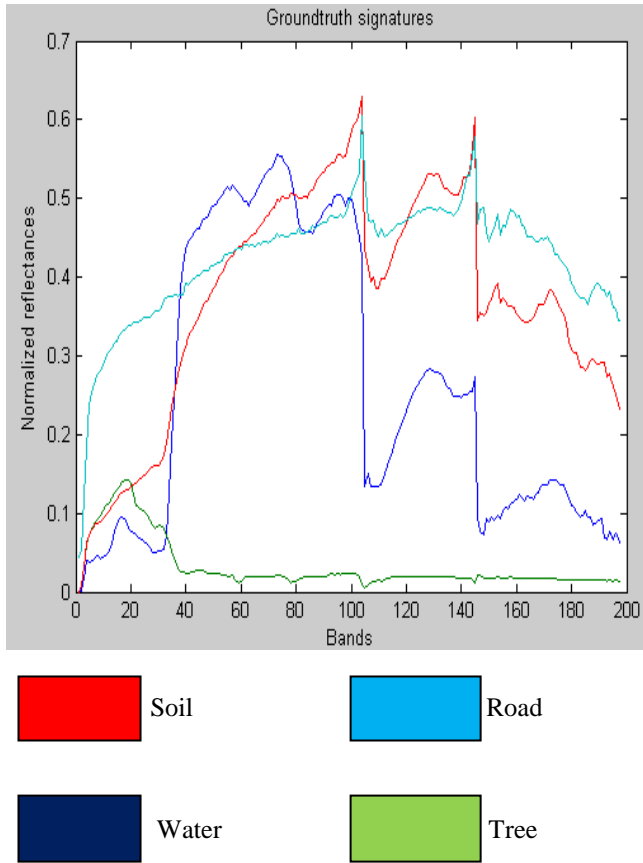


Fig.3 Groundtruth spectral signatures of endmembers

The number of bands required is estimated using the concept of VD by HFC method with false-alarm probability set to $P_f=10^{-5}$, beyond which the value remains constant at 9. The VD estimation for different false-alarm rate (PF) is shown in Table I. Then variance is calculated for 198 bands. Then the bands are clustered using K-means algorithm based on the variance. The number of clusters is equal to the number of endmembers. In each cluster bands having maximum variance value are selected and considered to be the informative bands for subsequent process. The selected bands are tabulated in Table II.

TABLE I. VD estimation using HFC method for different false-alarm rate

P_f	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
VD	21	17	12	10	9

TABLE II. Selected bands for Jasper ridge

Features	Selected Bands
Variance	34,178,113,119, 135,48, 53, 67, 104

In endmember extraction phase, the endmembers are extracted using Artificial Bee Colony algorithm and their performance is evaluated on selected bands. The spectral signatures are found and they are plotted in Fig. 3.

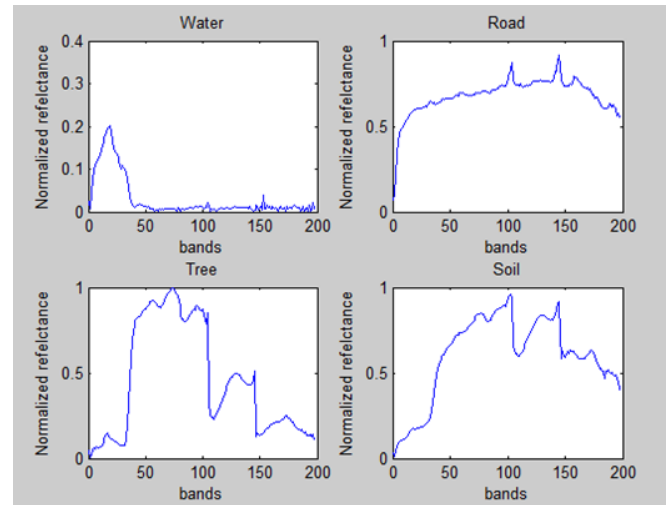


Fig.3 Spectral signatures of endmembers using Artificial Bee Colony algorithm

Spectral Angle Distance (SAD), a benchmark metric has been used to assess the performance of the proposed algorithm. The SAD measures the spectral similarity that exists between the estimated endmember using EEA and its groundtruth. Lower the value of this measure indicates better the performance of the algorithm.

$$SAD(m_k, \hat{m}_k) = \arccos\left(\frac{m_k^T \hat{m}_k}{\|m_k\| \|\hat{m}_k\|}\right) \quad (5)$$

Where \hat{m}_k is the k^{th} estimated endmember and m_k is the corresponding groundtruth

TABLE III. Comparison of the proposed ABC algorithm with N-FINDR algorithm

Endmembers	N-FINDR algorithm	Proposed ABC algorithm
Soil	11.14	10.88
Tree	15.59	12.56
Water	46.89	35.22
Road	10.69	9.77
Average	21.08	17.10

The performance comparison of ABC algorithm with N-FINDR is tabulated in Table III. The obtained values for each endmembers are different and the average error percentage obtained for the proposed ABC algorithm is lower. It is clearly shown that there is a considerable improvement in endmember determination by ABC algorithm.

VII. CONCLUSION:

In this work, the number of distinct spectral signatures is calculated using Virtual Dimensionality and then the bands with high information are selected using K-MEANS clustering. Artificial Bee Colony algorithm is used for endmember extraction and it is carried out with selected bands. ABC provides superior performance in endmember extraction than N-FINDR algorithm because it suffers from the problem of inconsistency in providing final endmembers. Hence, our proposed algorithm is well suited for real time applications. The future scope of this work is to develop an efficient hardware design for implementing our proposed algorithm in real world applications.

VIII. REFERENCES

- [1] Chang C-I, Wang S (2006), "Constrained Band Selection For Hyperspectral Imagery," *IEEE Trans Geosci Remote Sens* 44(6):1575–1585.
- [2] Keshava, N, Mustard, J.F., "Spectral Unmixing," *IEEE Signal Process. Mag.* 2002, 19, 44–57.
- [3] Boardman, J.W.; Kruse, F.A.; Green, "R.O. Mapping Target Signatures Via Partial Unmixing Of Aviris Data," In Proceedings of the Fifth Annual JPL Airborne Earth Science Workshop, Pasadena, California, 23–26 January 1995.
- [4] Winter, M.E., "N-Findr: An Algorithm For Fast Autonomous Spectral End-Member Determination In Hyperspectral Data," *SPIE Proc.* 1999, 3753, 266–275.
- [5] Nascimento, J.M.; Dias, J.M., "Vertex Component Analysis: A Fast Algorithm To Unmixhyperspectral Data," *IEEE Trans. Geosci. Remote Sens.* 2005, 43, 898–910.
- [6] Chang, C.; Wu, C.; Liu, "W. A New Growing Method For Simplex-Based Endmember Extraction Algorithm," *IEEE Trans. Geosci. Remote Sens.* 2006, 44, 2804–2819.
- [7] Chan, T.; Ma, W.; Ambikapathi, A., "A Simplex Volume Maximization Framework For Hyperspectral Endmember Extraction," *IEEE Trans. Geosci. Remote Sens.* 2011, 49, 4177–4193.
- [8] Jun, L.; José, M.B., "Minimum Volume Simplex Analysis: A Fast Algorithm To Unmixhyperspectral Data," *IEEE Trans. Geosci. Remote Sens.* 2008, 3, 250–253.
- [9] Ambikapathi, A.; Chan, T.; Ma, W., "Chance-Constrained Robust Minimum-Volume Enclosing Simplex Algorithm For Hyperspectralunmixing," *IEEE Trans. Geosci. Remote Sens.* 2011, 49, 4194–4209.
- [10] Miao, L.; Qi, H., "Endmember Extraction From Highly Mixed Data Using Minimum Volume Constrained Nonnegative Matrix Factorization," *IEEE Trans. Geosci. Remote Sens.* 2007, 45, 765–777.
- [11] Dorigo, M.; Maniezzo, V.; Colomi, A., "Ant System: Optimization By A Colony Of Cooperating Agents," *IEEE Trans. Syst. Man Cybern.B Cybern.* 1996, 26, 29–41.
- [12] Eberhart, R.; Kennedy, J., "A New Optimizer Using Particle Swarm Theory," In Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 4–6 October 1995.
- [13] Yang, X., "Nature-Inspired Metaheuristic Algorithms," *Luniver Press: Frome, UK,* 2008.
- [14] Karaboga, D.; Basturk, B., "A Powerful And Efficient Algorithm For Numerical Function Optimization: Artificial Bee Colony (ABC) Algorithm," *J. Glob. Optim.* 2007, 39, 459–471.
- [15] Zhang, B.; Sun, X.; Gao, L., "Endmember Extraction Of Hyperspectral Remote Sensing Images Based On The Ant Colony Optimization (ACO) Algorithm," *IEEE Trans Geosci. Remote Sens.* 2011, 49, 2635–2646.
- [16] Zhang, B.; Sun, X.; Gao, L., "Endmember Extraction Of Hyperspectral Remote Sensing Images Based On The Discrete Particle Swarm Optimization Algorithm," *IEEE Trans Geosci. Remote Sens.* 2011, 49, 4173–4176.
- [17] Martinez-Usó A, Pla F, Sotoca JM, Garcia-Sevilla P, "Clustering-Based Hyperspectral Band Selection Using Information Measures," *IEEE Trans Geosci Remote Sens* 45(12):4158–4171.
- [18] Karaboga, D.; Basturk, B., "Artificial Bee Colony (ABC) Optimization Algorithm For Solving Constrained Optimization Problems," In *Foundations of Fuzzy Logic and Soft Computing*; Springer: Berlin, Germany, 2007; pp. 789–798.
- [19] Xu Sun, Lina Yang, Bing Zhang, "An Endmember Extraction Method Based On Artificial Bee Colony Algorithms For Hyperspectral Remote Sensing Images,"
- [20] Karaboga, D.; Basturk, B., "A Powerful And Efficient Algorithm For Numerical Function Optimization: Artificial Bee Colony (ABC) Algorithm," *J. Glob. Optim.* 2007, 39, 459–471
- [21] Daniela Baran, NicolaeApostolescu, "A Virtual Dimensionality Method for Hyperspectral Imagery,"
- [22] B. Ramakrishna, J. Wang, A. Plaza, and C.-I Chang, "Spectral/Spatial Hyperspectral Image Compression In Conjunction With Virtual Dimensionality," presented at the SPIE Symp. Defense and Security Conf. Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI, vol. 5806, Orlando, FL, Mar./Apr. 28–1, 2005.