Firefly Algorithm Based Band Selection for Spectral Unmixing in Hyperspectral Images

G. Veera Senthil Kumar [1], S. Vasuki [2], M. Rashini [3]
Assistant Professor, ECE Dept, Vcet, Madurai, India[1]
Professor And Head, ECE Dept, Vcet, Madurai, India[2]
Student, ECE Dept, Vcet, Madurai, India[3]

***Abstract*— In remote sensing, hyperspectral imaging technique finds widespread application because of its high spectral resolution. The advantage of high spectral resolution leads to the problem of high dimensionality. The high dimensionality of hyperspectral image leads to increase in computational complexity in their analysis. Hence, Dimensionality Reduction (DR) is essential in all hyperspectral image analysis. Dimensionality reduction can be achieved by either transform based method or band selection method. This paper presents a novel approach of dimensionality reduction by means of band selection using firefly algorithm in Hyperspectral images (HSI). Firefly algorithm provides a comprehensive excellent result compared to other regular searching approaches. In band selection, there is an issue of determining the number of bands to be selected, which is resolved by the Virtual Dimensionality (VD) estimation. Spectral unmixing consists of endmember extraction followed by abundance fraction estimation. In this work, Simplex Growing Algorithm (SGA) is adopted for endmember extraction and Fully Constrained Least Squares (FCLS) is used for abundance fraction estimation. Finally, the performance of the proposed method is evaluated using Root Mean Square Error (RMSE) as a measure and compared with N-FINDR algorithm. Experimental results prove that the proposed method of Firefly based band selection with SGA outperforms the N-FINDR algorithm.**

Index Terms* — *Band selection (BS), Virtual Dimensionality (VD), Firefly Algorithm, SGA, FCLS

I.INTRODUCTION

Hyperspectral sensors can acquire number of adjoining bands over a wide electromagnetic spectrum for each pixel. The affluent spectral information allows for different materials with sophisticated spectral divergence, but it essentially leads to the "nuisance of dimensionality." Dimensionality reduction (DR) has been frequently used to tackle this issue [1]. Though, these methods usually change the substantial significance of original bands the channels in the low-dimensional space does not communicate to each bands. Band selection (BS) methods are used to select a section of the original bands which might be preferred when the physical meaning of bands required for maintenance [2], [3], [4]. Feature selection and extraction are the two methods involved for dimensionality reduction. The unique dimension of the hyperspectral image has to be reduced without any loss of information. Therefore, feature extraction method can be used for band selection. VD estimation provides reliable result of number of endmembers and the number of bands to be selected [5]. In this paper, firefly algorithm is proposed for band selection because firefly can find an overall optimal result more efficiently compared with other searching strategies.

SGA algorithm is recently developed simplex volume based method for finding endmembers in hyperspectral image. It overcomes the issues of random initialization and output inconsistency in N-FINDR algorithm. FCLS based abundance fraction estimation has been used because it considers both non-negative and sum-to-one constraints.

The rest of the paper is organized as follows: VD estimation is discussed in section II. Firefly algorithm based band selection is explained in section III. SGA is explained in section IV. The proposed methodology is elucidated in section V. Experimental results are discussed in section VI and conclusion is given in section VII.

II.VD ESTIMATION

VD estimation is a reliable measure for finding number of distinct spectral signatures available in hyperspectral image [5]. Eigen threshold based method, referred to as the Harsanyi-Farrand-Chang (HFC) method, has been previously developed in to find the number of endmembers in AVIRIS data. It is based on Neyman-Pearson detection theory and the number of bands are found by testing the number of failures for all bands, for a given false alarm rate PF. Noise-Whitened HFC (NWHFC) method [3] is an alternative for estimating VD accurately because HFC does not have noise-whitening process.

III. FIREFLY ALGORITHM BASED BAND SELECTION

The Firefly algorithm was developed by Dr.Xin She yangat Cambridge University stimulated by the matingor flashing behavior of fireflies. The Firefly Algorithm (FA) is one of the latest and best artificial intelligence algorithms being developed. The Firefly algorithm has proved to be much simpler in both concept and implementation. Fireflies or lightning bugs belong to a family of insects [6] which are capable to provide natural light to attract a mate or prey. There are about two thousand firefly species which generate short and rhythmic flashes. These flashes may often appears to be in an unique pattern and make an amazing sight in the tropical areas during summer. When distance (r) increases, the intensity (I) of flashes decreases. And therefore, most of the fireflies can communicate only up to several meters. In this algorithm, the flashing light is formulated in such a way that it is allied with the objective funtion to survive optimization. The steps of a Firefly algorithm are as follows.

Step 1) *Objective function:* The light intensity I(r) varies according to the inverse square law.

$$I\,(\,r\,) = \frac{I_s}{r^2} \qquad (1)$$

where I(r)-intensity at the source.

r-observers distance from source.

If we take absorption coefficient γ into account, the light intensity I varies with the square of distance r.

$$I = I_0\, e^{-\gamma r^2} \qquad (2)$$

Step 2) *Generate initial population of fireflies:* Initial population of fireflies (say n)

$$X_{t+1} = x_t + \beta_0 e^{-\gamma r^2} + \alpha\varepsilon \qquad (3)$$

where $\beta_0 e^{-\gamma r^2}$ - attraction
αε - randomization

Step 3) *Determine the light intensity Ii at xi via f(xi):* Determine the light intensities of each of the fireflies to find the brightness of each fireflies.

$$I = I_0\, e^{-\gamma r^2} \qquad (4)$$

Step 4) *Calculate the attractiveness of fireflies:*

$$\beta = \beta_0 e^{-\gamma r^2} \qquad (5)$$

Step 5) *Movement of less brightness fireflies towards n brighter one:* The movement of the firefly I is attracted to another more attracted firefly j is determined by,

$$x_i = x_i + \beta_0 e^{-\gamma r_{i,j}}\,(xj - xi) + \alpha\varepsilon \quad (6)$$

Step 6) *Find the current best:* Update light intensities of the fireflies and rank the fireflies, after ranking of the fireflies, find the current best solution.

IV. SIMPLEX GROWING ALGORITHM

SGA is a sequential algorithm which is used to find a simplex with the maximum volume every time a new vertex is being added. Since an endmember is an idealized pure signature, it is not essentially an image pixel. On the other hand, in a real image experiments, an endmember is directly extracted from the data [7][8]. Therefore, when an image occurs as a pixel, it is referred to as endmember pixel in this paper. Here, we represent a new algorithm known as Simplex Growing Algorithm (SGA), for endmember extraction to determine a set of desired endmembers by increasing a sequence of simplexes. It starts with two vertices and develop a simplex by increasing its vertices in one at a time. This algorithm is done when the number of vertices reaches the number of endmembers *p*, that can be estimated by the VD using a method introduced by Harsanyi, Farrand, and Chang, referred to as HFC method in [9][10], which do not require noise estimation. In order to select an appropriate pixel as its initial endmember pixel, a selection process for determining the first endmember pixel is introduced for this purpose and described as follows.

First Endmember Selection Process

Step 1) Randomly generate a target pixel, denoted by **t**.
Step 2) Find a pixel $\mathbf{e}_1$ that yields the maximum of absolute determinant of the matrix, $\left|\det\begin{bmatrix}1 & 1\\ t & r\end{bmatrix}\right|$ overall sample vectors r,i.e

$$e_1 = \arg\left\{max_r\left[\left|\det\begin{bmatrix}1 & 1\\ t & r\end{bmatrix}\right|\right]\right\} \qquad (7)$$

where principal components analysis (PCA) or MNF is required to reduce the original data dimensionality *L* to the dimension 2 to find the maximum.

A different target pixel **t** may result in a different $\mathbf{e}_1$. The experiments shows that the generated $\mathbf{e}_1$ is always a pixel consists of either a maximum or a minimum value in the first component of Dimensionality Reduction (DR) transform. So, there is no final set of endmembers in the target pixel **t**. Moreover, according to our extensive experiments, the generated $\mathbf{e}_1$ finally becomes one of the final generated endmembers. This shows that the final set of endmembers generated by the SGA is always same and consistent.

Considering the above first endmember selection process as a preprocessing for initialization, the SGA can be described in detail as follows.

Simplex Growing Algorithm

Step 1) Initialization:

a) Use the VD to estimate the number of endmembers to be generated *p*; and

b) Use the $\mathbf{e}_1$ found by the first endmember selection process as the desired initial endmember pixel and set *n* = 1.

Step 2) At $n \geq 1$ and for each sample vector r, we calculate $V\,(e_1, \ldots, e_n, r)$ defined by

$$V(\,e_1, \ldots, e_n, r\,) = \frac{\left|\det\begin{bmatrix}1 & 1\cdots 1 & 1\\ e_1 & e_2\cdots e_n & r\end{bmatrix}\right|}{n!} \qquad (8)$$

which is the volume of the simplex specified by vertices $\mathbf{e}_1$, $\mathbf{e}_2, \ldots, \mathbf{e}_n$, **r**, denoted by

$S\,(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \mathbf{r})$. Since the matrix $\begin{bmatrix}1 & 1\cdots 1 & 1\\ e_1 & e_2\cdots e_n & r\end{bmatrix}$ in (8) is not necessarily a square matrix, a DR technique such as PCA or MNF is needed to reduce the original data dimensionality *L* to the dimension *n*.

Step 3) Find e_{n+1} that yields the maximum of (8), i.e.,

$$e_{n+1} = \arg\{max_r[\,V\,(e_1, \ldots, e_n, r)]\} \qquad (9)$$

Step 4) *Stopping rule:* If $n < p$, then $n \leftarrow n + 1$ and go step 2). Otherwise, the final set of $\{e_1, e_2, \ldots, e_p\}$ is the desired *p* endmembers.

V. PROPOSED METHODOLOGY

In this proposed method, the hyperspectral image is first read and Virtual Dimensionality (VD) is estimated to find the number of bands to be selected from hyperspectral imagecube. Band selection is a general approach to minimize the data dimensionality of the hyperspectral imagery. It extracts several bands of significance in some sense by considering the advantage of high spectral correlation. Band selection can be viewed as an optimization problem. Many optimization tools have been used for band selection. Firefly algorithm has been deployed as searching algorithm for selecting optimal bands in this work. Then SGA is employed for endmember extraction to determine the set of desired endmembers by increasing the sequence of simplexes. It starts with two vertices and starts to rise a simplex by increasing its vertices one at a time. Abundance fraction of endmembers is estimated using FCLS method. Finally, the performance of the proposed method is evaluated using RMSE as a measure.

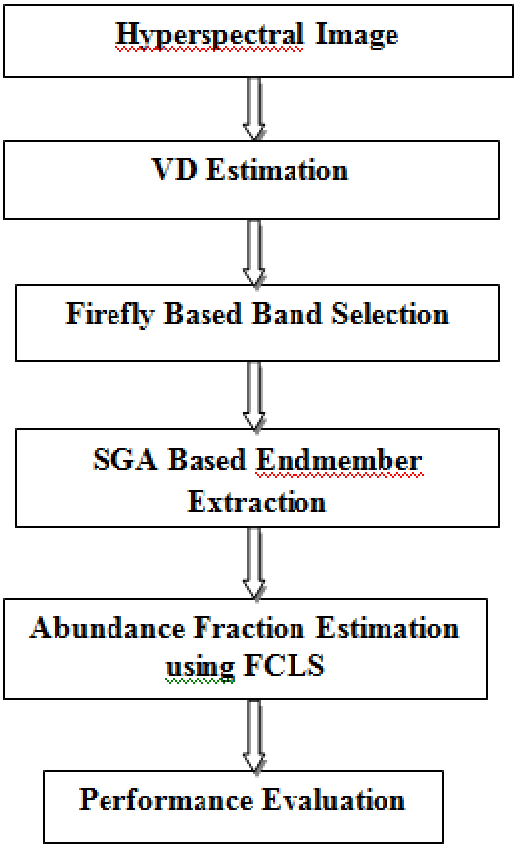


Fig 1. Block Diagram of proposed method

VI. EXPERIMENTAL RESULTS

In this work, Jasper ridge image, a popular hyperspectral data has been used. It is having size 512 x 614 pixels ranging from 0.38 to 2.5 μm wavelength, as shown in Fig 2. A sub image of size 100X100 has been selected from the original image in order to reduce the computational complexity. T he endmembers present in the image are Water, Tree, Soil and Road. The groundtruth of the endmembers are shown in Fig.3.

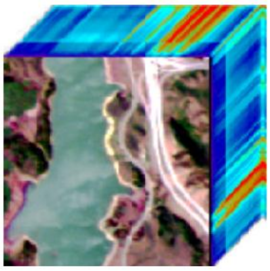


Fig 2. Jasper Ridge

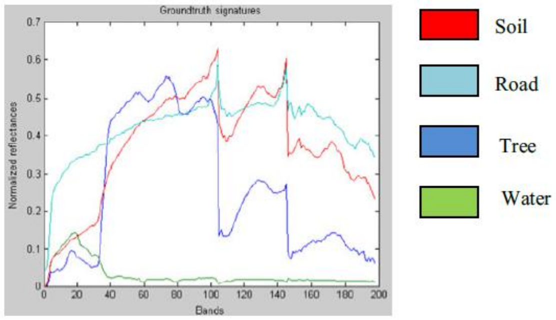


Fig 3. Groundtruth of the Endmembers

VD Estimation with different false alarm rate for Jasper ridge image is shown in Table I.

Table I. VD Estimation with different false alarm rate for Jasper ridge image

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

The bands selected for Jasper ridge image is shown in Table II.

Table II. Selected bands for Jasper ridge

Feature	Band Selection
Variance	13,10,100,152,185,21,105,87,2

The endmembers which are extracted using SGA is shown in Fig 4.

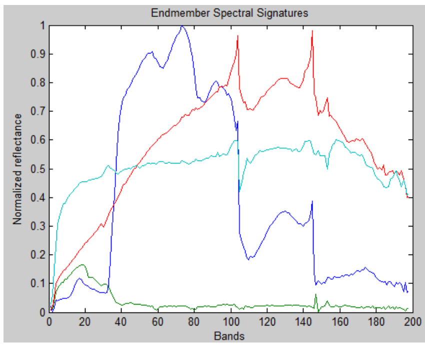
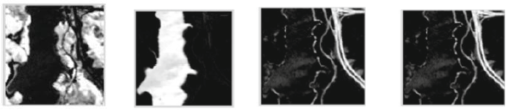
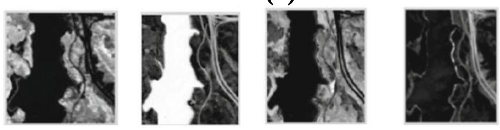


Fig 4. Endmember Extraction Using SGA

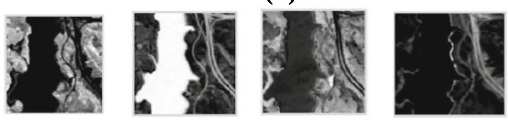
Then, the abundance fraction of endmembers are estimated using FCLS and the results are shown in Fig 5.



(a)



(b)



(c)

Fig 5.a) Groundtruth Abundance Map of Endmember
b) Abundance Map of Endmember using SGA
c) Abundance Map of Endmember using N-FINDR

The Comparison of performance of a proposed method of firefly based band selection with SGA and N-FINDR using RMSE is given in Table III.

Table III. Comparison of performance of proposed clustering based band selection with SGA and NFINDR using RMSE

Endmembers	RMSE(%)	
	Firefly based band selection with N-FINDR	Firefly based band selection with SGA
Tree	15.64	26.02
Water	21.71	7.15
Soil	10.01	11.66
Road	10.56	10.17
Average	14.48	13.75

where

$$RMSE\ (z, \hat{z}) = \left(\frac{1}{N}\ \|z - \hat{z}\|^2\right)^{1/2}$$

where N is the number of pixels in the image, $\hat{z}$ is the estimated abundance map and z is the corresponding groundtruth.

VII. CONCLUSION

Firefly algorithm based band selection method is adopted in this work for selection of optimal bands from a high dimensionality hyperspectral image cube. VD estimation is used to resolve the issue of determining the number of bands to be selected. SGA algorithm is adopted in endmember extraction. It has been proved from the experimental results that firefly based band selection with SGA provides superior performance than existing N-FINDR algorithm. It has also been observed from the experimental results that the proposed firefly algorithm based band selection with SGA produces lower average RMSE by 1% compared to that of N-FINDR algorithm for Jasper Ridge Dataset. Hence, the proposed firefly based band selection with SGA yields superior performance in spectral unmixing.

REFERENCES

[1] Muhammad Sohaib, Ihsan-Ul-Haq, and Qaiser Mushtaq “Dimensional Reduction of Hyperspectral Image Data Using Band Clustering and Selection Based on Statistical Characteristics ofBand Images” International Journal of Computer and Communication Engineering, Vol. 2, No.2, March 2013

[2] M. Ghamary Asl , B. Mojaradi, “Virtual Dimensionality Estimation In Hyperspectral Imagery Based On Unsupervised Feature Selection”, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume III-7, 2016 XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic.

[3] Chang, C.-I, Xiong, W., Liu, W., Chang, M.-Li, Wu, C.-C., and Chen, C. C.-C., 2010b, “Linear Spectral Mixture Analysis Based Approaches to Estimation of Virtual Dimensionality in Hyperspectral Imagery”, IEEE Transactions on Geoscience and Remote Sensing, 47(7), pp. 3960–3979.

[4] Ghamary Asl, M., Mobasheri, M. R., Mojaradi, B., 2014b, “Unsupervised Feature Selection using Geometrical Measures in Prototype Space for Hyperspectral Imagery” IEEE Transactions on Geoscience and Remote Sensing,52(7), pp. 3774-3787.

[5] Veera Senthil Kumar G & Vasuki S, “Clustering based band selection for endmember extraction using simplex growing algorithm in hyperspectral images” An International Journal ISSN 1380-7501 Multimed Tools Appl DOI 10.1007/s11042-016-3420-4

[6] Adil Hashmi, Nishant Goel, Shurti Goel, Divya Gupta,”Firefly Algorithm for Unconstrained Optimization. IOSR Journal of Computer Engineering (IOSR-JCE) Jun. 2013,

[7] M.Rashini, G.Veera Senthil Kumar, S.Vasuki, “Analysis of Band Selection Algorithms for Endmember Extraction in Hyperspectral Images”, IOSR Journal of Electronics and Communication Engineering (IOSR-JECE) e-ISSN: 2278-2834,p- ISSN: 2278-8735.Volume 11, Issue 6, Ver. IV (Nov.-Dec .2016), PP 25-32

[8] Chang C-I, Du Q (2004), “Estimation of number of spectrally distinct signal sources in hyperspectral imagery.” IEEE Trans Geosci Remote Sens 42(3):608–619

[9] Kang Sun, Xiurui Geng, and Luyan Ji, “Exemplar Component Analysis: A Fast Band Selection Method for Hyperspectral Imagery,” IEEE Geoscience And Remote Sensing Letters, vol. 12, no. 5, may 2015.

[10] Q. Du and H. Yang, “Similarity-based unsupervised band selection for hyperspectral image analysis,” IEEE Geosci. Remote Sens. Lett., vol. 5, no. 4, pp. 564–568, Oct. 2008.

[11] K. Sun, X. Geng, L. Ji, and Y. Lu, “A new band selection method for hyperspectral image based on data quality,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 7, no. 6, pp. 2697-2703,Jun. 2014.

[12] C. Sheffield, “Selecting band combinations from multispectral data,”Photogramm. Eng. Remote Sens., vol. 51, no. 6, pp. 681–687, 1985.

[13] M.E. Winter, “N-FINDR: An Algorithm For Fast Autonomous Spectral Endmember Determination In Hyperspectral Data,” Imaging Spectrometry V, Proc. SPIE 3753, pp. 266-277, 1999.

[14] D. Landgrebe, “Hyperspectral image data analysis,” IEEE Signal Process.Mag., vol. 19, no. 1, pp. 17–28, Jan. 2002.

[15] R. Huang and M. He, “Band selection based on feature weighting for classification of hyperspectral imagery,” IEEE Geosci. Remote Sens.Lett., vol. 2, no. 2, pp. 156–159, Apr. 2005