ISPRS Journal of Photogrammetry and Remote Sensing 126 (2017) 108–119

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Endmember extraction from hyperspectral image based on discrete firefly algorithm (EE-DFA)

Chengye Zhang [a], Qiming Qin [a,*], Tianyuan Zhang [a], Yuanheng Sun [a], Chao Chen [b,c]

[a] Institute of Remote Sensing and Geographic Information System, School of Earth and Space Sciences, Peking University, Beijing 100871, China
[b] Marine Science and Technology College, Zhejiang Ocean University, Zhoushan, Zhejiang 316022, China
[c] Marine Acoustics and Remote Sensing Laboratory, Zhejiang Ocean University, Zhoushan, Zhejiang 316022, China

ARTICLE INFO

Article history:
Received 7 July 2016
Received in revised form 22 January 2017
Accepted 12 February 2017

Keywords:
Endmember extraction
Discrete firefly algorithm (DFA)
Hyperspectral image
Remote sensing

ABSTRACT

This study proposed a novel method to extract endmembers from hyperspectral image based on discrete firefly algorithm (EE-DFA). Endmembers are the input of many spectral unmixing algorithms. Hence, in this paper, endmember extraction from hyperspectral image is regarded as a combinational optimization problem to get best spectral unmixing results, which can be solved by the discrete firefly algorithm. Two series of experiments were conducted on the synthetic hyperspectral datasets with different SNR and the AVIRIS Cuprite dataset, respectively. The experimental results were compared with the endmembers extracted by four popular methods: the sequential maximum angle convex cone (SMACC), N-FINDR, Vertex Component Analysis (VCA), and Minimum Volume Constrained Nonnegative Matrix Factorization (MVC-NMF). What's more, the effect of the parameters in the proposed method was tested on both synthetic hyperspectral datasets and AVIRIS Cuprite dataset, and the recommended parameters setting was proposed. The results in this study demonstrated that the proposed EE-DFA method showed better performance than the existing popular methods. Moreover, EE-DFA is robust under different SNR conditions.

1. Introduction

Hyperspectral remote sensing acquires image covering hundreds of narrow contiguous spectral bands, which provides a contiguous spectrum for each pixel (Zhang et al., 2015). However, due to the relative low spatial resolution of hyperspectral image, it is inevitable that several materials are presented together in one pixel in natural environment. Thus, it has triggered a popular research interest to determine the abundance fraction of each material in a mixed pixel from hyperspectral image, which called spectral unmixing. The first step for solving this problem (spectral unmixing) is usually endmember extraction (EE), which is to determine the idealized pure spectral signature of each material. In fact, an endmember extracted from a real hyperspectral image is generally still a mixed pixel that contains high fraction of one kind of material (Filippi and Archibald, 2009). In computer science and mathematics, discrete optimization is a topic on finding a best object from a finite set of objects. In detail, discrete optimization problem is to find a combination of discrete variables from all feasible solutions to optimize the objective function. What's more, endmember extraction from a hyperspectral image is to find some pixels and combine them as pure spectral signatures of the materials in the image to get the best accuracy of spectral unmixing results. In this sense, endmember extraction from hyperspectral image can be transferred into a typical discrete optimization problem that is of particular importance for spectral unmixing of hyperspectral image.

For recent more than two decades, many methods for endmember extraction from hyperspectral image have been proposed. Typical algorithms developed for endmember extraction include pixel purity index (PPI) (Chang and Plaza, 2006; Qin et al., 2016), N-FINDR (Winter, 1999), iterative error analysis (IEA) (Neville, 1999), vertex component analysis (VCA) (Nascimento and Dias, 2005), the sequential maximum angle convex cone (SMACC) (Gruninger et al., 2004), the optical real-time adaptive spectral identification system (ORASIS) (Bowles et al., 1998), convex cone analysis (CCA) (Ifarraguerri and Chang, 1999), automated morphological endmember extraction (AMEE) (Plaza et al., 2002), iterated constrained endmembers (ICE) (Berman et al., 2004), the single individual evolutionary (SIE) strategy (Graña et al., 2004). In addi-

* Corresponding author.
E-mail addresses: zhangchengye@pku.edu.cn (C. Zhang), qmqin@pku.edu.cn, qmqinpku@163.com (Q. Qin).

http://dx.doi.org/10.1016/j.isprsjprs.2017.02.005

tion, some new methods for endmember extraction were also proposed in recent years (Chang et al., 2006; Miao and Qi, 2007; Plaza et al., 2010; Zare and Gader, 2010; Chang et al., 2010a; Bing et al., 2011a, 2011b; Chan et al., 2013; Geng et al., 2013; Zhang and Xie, 2014). In fact, currently, although there are a variety of methods for endmember extraction from hyperspectral image, these methods have respective shortcomings. Comparative analysis and discussions on the existing endmember extraction algorithms have been conducted by several scholars (Filippi and Archibald, 2009; Bing et al., 2011a; Plaza et al., 2004). In this paper, disadvantages of several typical algorithms are displayed as examples. For instance, the PPI algorithm is a typical manual/interactive endmember extraction method, which is slow and artificially influenced (Filippi and Archibald, 2009). In the N-FINDR algorithm, it is generally inevitable to reduce the dimensions of the hyperspectral dataset (Winter, 1999). What's more, many algorithms including the N-FINDR algorithm, VCA algorithm and SMACC algorithm, are based on a simple geometrical precondition that the endmembers are the vertexes of a convex simplex in spectral space (Winter, 1999; Nascimento and Dias, 2005; Gruninger et al., 2004). However, due to some nonlinear factors and the noise in a real hyperspectral image, this precondition is not reliable (Nascimento and Dias, 2005). Thus, the existing algorithms are still limited and it is necessary to propose new algorithms with particular advantages to conduct endmember extraction from hyperspectral image and provide more optional methods for endmember extraction for scholars in the area of hyperspectral remote sensing.

Firefly algorithm (FA) is a novel nature-inspired algorithm, which is inspired by social behavior of fireflies and was presented by Gilang and Suyanto (2011). It has been proven that FA is an efficient method to solve optimization problem and is better than both particle swarm optimization (PSO) algorithm and genetic algorithm (GA) in the respects of both efficiency and success rate (Yang, 2009; Sayadi et al., 2010). Moreover, Reference Sayadi et al. (2010) has demonstrated that discrete firefly algorithm (DFA) performed better than the ant colony optimization (ACO) algorithm in solving discrete optimization problems. Hence, DFA provides a better way for solving discrete optimization problems including endmember extraction from hyperspectral image, than other popular swarm intelligence algorithms, such as PSO, GA and ACO.

This paper proposes an endmember extraction method from hyperspectral image based on discrete firefly algorithm (EE-DFA). To the best of our knowledge, there is little information available in published literature about the application of firefly algorithm on endmember extraction. To realize the endmember extraction, we redefined the objective function and several variables in FA, and proposed a new discretization method of the firefly algorithm.

2. Firefly algorithm (FA)

A brief introduction to FA is presented here, with the purpose to define necessary concepts to explain the EE-DFA. More details of FA can be found in the reference Yang (2008). Nature-inspired algorithms are among the most important and powerful methods for optimization problems. In the nature, most firefly species produce short and rhythmic flashes that is to act as a signal system to attract other fireflies. Therefore, FA obeys the following rules: (1) All fireflies are attractive to other fireflies regardless of their sex; (2) the attractiveness of a firefly is determined by its brightness, which means that for any two flashing fireflies the less bright one will move toward the brighter one. If a particular firefly is the brightest in the firefly population, it will move randomly; and (3) the brightness of a firefly is determined by the objective function. For a minimization problem in this study, the brightness of a firefly should be inversely proportional to the value of the objective function (Sayadi et al., 2010).

2.1. Distance

In FA, the distance r_{ij} of two fireflies $\mathbf{x_i}$ and $\mathbf{x_j}$ can be determined as follows.

$$r_{ij} = \sqrt{\sum_{k=1}^{n} (x_{ik} - x_{jk})^2} \tag{1}$$

where x_{ik} is the k-th component of the i-th firefly $\mathbf{x_i}$ and n is the number of components of the firefly $\mathbf{x_i}$. Here we give an explanation on the word 'component'. 'component' is used to express a dimension of a vector. For example, if there is a vector $\boldsymbol{V} = (a, b, c, d, e, f)$, we could say that the vector $\boldsymbol{V}$ possesses 6 components and "d" is the 4th component of the vector $\boldsymbol{V}$.

2.2. Attractiveness

The attractiveness β_{ij} between two fireflies $\mathbf{x_i}$ and $\mathbf{x_j}$ is a monotonically decreasing function. For instance, it can be the following form.

$$\beta_{ij} = \beta_0 e^{-\gamma r_{ij}^2} \tag{2}$$

where r_{ij} is the distance between two fireflies $\mathbf{x_i}$ and $\mathbf{x_j}$, β_0 is the attractiveness at $r_{ij} = 0$ and γ is a fixed light absorption coefficient.

2.3. Movement

The movement of the i-th firefly $\mathbf{x_i}$, when it is attracted by another brighter one ($\mathbf{x_j}$), can be determined as follows.

$$x_{ik} = x_{ik} + \beta_0 e^{-\gamma r_{ij}^2}(x_{jk} - x_{ik}) + \alpha(\text{rand} - 0.5) \tag{3}$$

where the second term is for the attraction between two fireflies, and the third term is randomization with α being the randomization step and "rand" is a random number uniformly distributed in [0,1]. Eq. (3) takes x_{ik} as an example to illustrate the calculation method of the components of $\mathbf{x_i}$ after one movement. All the components of $\mathbf{x_i}$ should be calculated according to Eq. (3) in one movement of $\mathbf{x_i}$.

To apply the FA on endmember extraction, we redefined the inputs, outputs, objective function, the meaning of fireflies, and proposed discretization method.

3. Endmember extraction based on discrete firefly algorithm (EE-DFA)

3.1. Definition of the parameters and variables

In this section, the meaning of all the parameters and variables in EE-DFA is presented.

m: Number of endmembers that will be extracted.
n: Number of candidate endmembers and also the number of components of the firefly $\mathbf{x_i}$.
N: Number of the pixels in one spectral band.
L: Number of spectral bands in hyperspectral dataset.
t: $t = 0$ means the i-th firefly is the current best solution in all fireflies. Otherwise, the i-th firefly is not the current best solution in all fireflies.
$\mathbf{p_i} = (p_{i1}, p_{i2}, p_{i3}, \ldots, p_{iL})$, $i = 1, 2, 3, \ldots, N$: The spectral reflectance of the i-th pixel.
$\mathbf{p}'_i = (p'_{i1}, p'_{i2}, p'_{i3}, \ldots, p'_{iL})$, $i = 1, 2, 3, \ldots, N$: The approximate estimation of $\mathbf{p_i}$.
$\mathbf{e_j} = (e_{j1}, e_{j2}, e_{j3}, \ldots, e_{jL})$, $j = 1, 2, 3, \ldots, m$: The spectral reflectance of the j-th endmember.