SIMULATION OF AN ALGORITHM FOR THE SPECTRAL IDENTIFICATION OF SPACE TARGET MATERIALS

Qing-Bo Li [*], Qi Wang, Shao-Lin Shi

School of Instrumentation and Optoelectronic Engineering, Precision Opto-Mechatronics Technology Key Laboratory of Education Ministry, Beihang University, Xueyuan Road No. 37, Haidian District, Beijing, 100191, China; e-mail: qbleebuaa@buaa.edu.cn

A novel extraction and purification method for identifying pure materials in space targets is presented in this paper. Simulation experiments are carried out using data from the United States Geological Survey and laboratory sample material. Experimental results show that the proposed method can successfully improve the unmixing of sparse spectra. This algorithm should also be suitable for other spectral analysis applications.

Keywords: *spectral Information, pure material, spectral unmixing.*

МОДЕЛИРОВАНИЕ АЛГОРИТМА СПЕКТРАЛЬНОЙ ИДЕНТИФИКАЦИИ МАТЕРИАЛОВ КОСМИЧЕСКИХ ОБЪЕКТОВ

Q.-B. Li [*], Q. Wang, Sh.-L. Shi

УДК 543.42

Школа приборостроения и оптоэлектронной техники, Университет Бейхан, Хайдянь, Пекин, 100191, Китай; e-mail: qbleebuaa@buaa.edu.cn

(Поступила 15 мая 2018)

Предложен метод выделения и подготовки для идентификации чистых веществ в космических мишенях. Модельные эксперименты проведены с использованием данных Геологической службы США и лабораторных проб. Показано, что предложенный метод может улучшить несмешивание разреженных спектров и быть полезен для других приложений спектрального анализа.

Ключевые слова: *спектральная информация, чистое вещество, спектральное несмешивание.*

Introduction. With the current international rate of growth of the aerospace industry, methods for identifying the materials of space targets are becoming increasingly important. Spectral analysis techniques can be used for detecting pure materials and have the advantages of being non-contact, efficient, and non-destructive [1]. In recent years, a method based on the spectral linear hybrid model has been split into three categories. The first type is the pixel purity index (PPI) algorithm, which is based on the geometric convex set [2]. In this case, the original data are processed with dimensionality reduction. Randomly generated test vectors passing through the dataset are used as one-dimensional coordinate axes, and the number of times each pixel is projected onto a one-dimensional test vector coordinate is denoted by the PPI index [3–5]. However, a disadvantage of the PPI algorithm is that the selection of the initial test vector has a considerable influence on the result [6]. The second option is the N-finder algorithm. This method calculates the volume of a single body to determine the location of the pure material. It requires randomly selecting pixels from the datasets as initial candidate pure materials and then calculates the volume of a single body enclosed by these pixels. Then, the candidate pixels are replaced by the observed pixels from the image, and the volume is recalculated [7]. The candidates are repeatedly substituted, compared, and replaced until the largest body is found. The disadvantage of this algorithm is that the result is affected by the selected initial pure material. The third method is vertex component analysis (VCA). This algorithm is based on the geometric description of the linear spectral mixture model, which takes the pixel with the largest projection length after projection

as the first pure material [8, 9]. Then, a direction orthogonal to the end member is used as the projection direction in the second iteration. After calculating the orthogonal projection, all the pure material can be identified. The pure material extraction is implemented with extremely low time complexity. The disadvantage of VCA is that the initial selection has a considerable impact on the results [10].

In recent years, some difficulties have presented themselves in the detection of space targets. The main difficulty is that the spectral resolution of ground-based spectrometers is low, resulting in a relatively small number of space target bands. Less effective information obtained from fewer bands has a certain impact on the accuracy of the results. Similar problems, such as low precision, small quantities of information, and low concentrations as a result of chemical measurement conditions, are also encountered in the chemical field. In addition, owing to the uniform material distributions of the space targets, the endmember extraction accuracy of the traditional algorithm will be reduced. To solve these problems, a pure material purification and optimization method named the endmember extraction purification method (EEPM) is presented in this paper.

Calculation. Two experiments are described in this paper to validate the new method described below. The pure material spectrum used in the first experiment was obtained from space target surface material in the laboratory. Data from the database of the United States Geological Survey was adopted to carry out the second simulation experiment.

Method of EEPM. The method we present, EEPM, works by increasing purity to optimize results after extracting the pure materials and improving the accuracy of the mixed pixel decomposition. The EEPM algorithm is implemented based on simplex identification via the split augmented Lagrange algorithm. EEPM optimizes purification after calculating the initial pure material and concentrations. Typically, each pixel in a hyperspectral image can be approximately considered as a linear combination of pure materials:

$$Y = EZ + \varepsilon, \tag{1}$$

$$\sum_{i=1}^{N} z_i = 1, \quad 0 \le z_i \le 1, \tag{2}$$

where $Y = [y_1, y_2, ..., y_N]$ is an $L \times N$ matrix in which each column is a mixed material vector, N is the number of the mixed material spectra, L is the number of wavelengths; $E = (e_1, e_2, ..., e_p) \in R^{L \times p}$ represents the pure material matrix, e_i ($i = 1, 2, ..., p$) is a pure material vector, $Z = [z_1, z_2, ..., z_N]$ is the concentration of each pure material, and ε is the error term.

This article assumes that the total number of pure materials is known (denoted by p). Thus, for a concentration matrix Z, which has $1_p^T Z = 1_N$ on the right hand side of the equation, the same constraint as $Y^T(YY^T)^{-1}$, $1_p^T M^{-1} Y = 1_N^T$ can be obtained. Order $a^T \equiv 1_N Y^T (YY^T)^{-1}$, $1_p^T M^{-1} = a^T$ can be obtained. The concept of a small simplex method is applied here to obtain the pure material spectrum. It is assumed that the mixed spectral data will be in a simplex, and that the vertices of the simplex correspond to the pure pixel spectrum. We hope to find the smallest volume of the simplex that can include all the mixed spectral data, and the vertices of this simplex will then provide the pure material spectrum. The volume of a simplex that contains the mixed spectral data can be approximated by the determinant value of the mixed spectral matrix, or det(M). To calculate det(M), we need M to perform dimensionality reduction processing to make $L = p$. We do not describe this process in detail here:

$$M^* = \arg\min |\det(M)|, \tag{3}$$

$$M^{-1} Y \ge 0, \; 1_p^T M^{-1} = a^T. \tag{4}$$

Let $Q \equiv M^{-1}$ such that $\det(Q) = 1/\det(M)$. Then, Eq. (3) can be converted to

$$Q^* = \arg\min_{Q} - \log\left|\det(Q)\right|, \tag{5}$$

$$QY \ge 0, \quad 1_p^T Q = a_n^T. \tag{6}$$

Equation (5) can be solved by turning it into an optimization problem:

$$Q^* = \arg\min_{Q} - \log\left|\det(Q)\right| + \lambda \left\|QY\right\|_h, \tag{7}$$

$$1_p^T Q = a^T. \tag{8}$$

For the optimization problem in Eq. (7), we can solve the problem of the variable augmented splitting Lagrange algorithm and obtain the initial pure material extraction results. The result obtained using this method is not a real pure material but a linear combination of real pure materials. Therefore, we assume that the initial extraction of the pure material is $M = (m_1, m_2, ..., m_p)$:

$$m_1 = A_1^{\,1} E_1 + A_2^{\,1} E_2 + ... + A_p^{\,1} E_p,\ m_2 = A_1^{\,2} E_1 + A_2^{\,2} E_2 + ... + A_p^{\,2} E_p,\ ...\ m_p = A_1^{\,p} E_1 + A_2^{\,p} E_2 + ... + A_p^{\,p} E_p, \quad (9)$$

where (E_1, E_2, ..., E_p) is the real pure material, and ($A_1^{\,b}$, $A_2^{\,b}$, ..., $A_p^{\,b}$) is the actual pure material scale factor that corresponds to the extracted pure material m_b. In the extracted pure material, the largest proportionality coefficient of a particular pure material is called the main component of the extracted pure material. Moreover, we assume that the main component of the pure material m_b is E_b. The extracted pure material is purified and optimized:

$$\begin{aligned} m_1' &= \frac{c}{c-(d_2+d_3+...+d_p)}(m_1 - m_2 d_2 / c - m_3 d_3 / c - ... - m_p d_p / c), \\ m_2' &= \frac{c}{c-(d_1+d_3+...+d_p)}(m_2 - m_1 d_1 / c - m_3 d_3 / c - ... - m_p d_p / c), \\ &... \\ m_p' &= \frac{c}{c-(d_1+d_2+...+d_{p-1})}(m_p - m_1 d_1 / c - m_2 d_2 / c - ... - m_{p-1} d_{p-1} / c), \end{aligned} \quad (10)$$

where m_b' is the first b pure material after purification and optimization, d is the termination coefficient whose initial value is 1, and c is a regulatory factor, $c > 1$. We use m_1' as an example to explain the algorithm. Assuming that the space target has three materials, we can express m_1 as follows:

$$m_1' = \frac{c}{c-(d_2+d_3)}\left(m_1 - m_2 d_2 / c - m_3 d_3 / c\right) \quad (11)$$

and $m_1 = A_1^1 E_1 + A_2^1 E_2 + A_3^1 E_3$. Substituting m_1 into Eq. (11), we obtain

$$\begin{aligned} m_1' &= \frac{c}{c-(d_2+d_3)}\left[\left(A_1^1 - \frac{d_2}{c}A_1^2 - \frac{d_3}{c}A_1^3\right)E_1 + \left(A_2^1 - \frac{d_2}{c}A_2^2 - \frac{d_3}{c}A_2^3\right)E_2 + \left(A_3^1 - \frac{d_2}{c}A_3^2 - \frac{d_3}{c}A_3^3\right)E_3\right] = \\ &= \left(\frac{c}{c-(d_2+d_3)}A_1^1 - \frac{d_2}{c-(d_2+d_3)}A_1^2 - \frac{d_3}{c-(d_2+d_3)}A_1^3\right)E_1 + \left(\frac{c}{c-(d_2+d_3)}A_2^1 - \frac{d_2}{c-(d_2+d_3)}A_2^2 - \frac{d_3}{c-(d_2+d_3)}A_2^3\right)E_2 + \\ &+ \left(\frac{c}{c-(d_2+d_3)}A_3^1 - \frac{d_2}{c-(d_2+d_3)}A_3^2 - \frac{d_3}{c-(d_2+d_3)}A_3^3\right)E_3. \end{aligned} \quad (12)$$

Because $A_1^{\,1} > A_1^{\,2}$ and $A_1^{\,1} > A_1^{\,3}$, the coefficient ($cA_1^{\,1}/[c - (d_2 + d_3)] - A_1^{\,2} d_2/[c - (d_2 + d_3)] - A_1^{\,3} d_3/[c - (d_2 + + d_3)]$) is bigger than the coefficient $A_1^{\,1}$. This result shows that the proportion of real pure material in the pure material identified by the purification algorithm increases. Through optimization of the pure material extraction process, EEPM can highlight the main elements of pure material m_1 and simultaneously reduce the proportion of minor components in pure material m_1. Other pure materials obtained by the algorithms (m_1, m_2, ..., m_p) can also be subjected to similar operations. Then, the original pure materials are replaced by the optimized pure material, and a new pure material matrix is obtained. The new concentration matrix is obtained through the fully constrained least squares method, and the reconstructed spectral data can be obtained. The reconstructed spectral data are compared with the real spectral data, and the reconstruction error is a root-mean-square-error (RMSE). If the error is reduced, the algorithm will be terminated and the coefficient $d + 1$ continues the process until the reconstruction error becomes large. The algorithm flow is as follows:

1. Input spectral data Y and pure material p.

2. The optimization problem is solved using the variable splitting augmented Lagrange algorithm:

$$Q^* = \arg\min_Q -\log\left|\det Q\right| + \lambda\left\|QY\right\|_h,$$

where $1_p^T Q = a^T$, $Q \equiv M^{-1}$, $a^T \equiv 1_n Y^T (YY^T)^{-1}$.

3. The extracted pure materials are iteratively purified and optimized, and a new pure material matrix is obtained:

$$M_1 = \frac{c}{c-(d_2+d_3+...+d_p)}(m_1 - m_2 d_2 / c - m_3 d_3 / c - ... - m_p d_p / c).$$

The initial values of (d_1, d_2, ... d_p) are 1.

4. The new concentration matrix is obtained through the fully constrained least squares method, and the reconstructed spectral data can be obtained. The obtained spectral data are compared with the real spectral data, and the reconstruction error is the RMSE.

5. If the reconstruction error is less than that on the previous iteration, the corresponding d_2 will be increased by 1. The preceding operations are repeated until the reconstruction error increases. Then d_2 is fixed and d_3 is increased by 1 and so forth until all the pure material operations are complete.

6. Output the optimized pure material, end.

Results and discussion. *Simulation experiment with space target spectral data.* The pure material spectrum used in the first experiment is a spectrum of space target surface material obtained in the laboratory. The spectrum measured is shown in Fig. 1a. Figure 1b shows the spectral simulation curve of the pure material with continuous variation of concentrations under 40 dB noise. The three measured pure material spectra are multiplied by the consecutive change concentrations to obtain a mixed spectral matrix and 40 dB of noise is added. The concentration herein uses linear variation to simulate the slow rotation of the target, with a maximum concentration of no more than 0.7 to ensure the absence of pure pixels in the spectral matrix. The EEPM, VCA, PPI, and SISAL algorithms are used to obtain the pure material extraction results and concentration contrast, as shown in Fig. 1c. The pure material optimization extracted by the EEPM algorithm is seen to be effective. Moreover, the result of the purification coincides with the real pure material and is obviously improved compared with SISAL. The reconstruction error associated with the pure material concentration extracted from SISAL and the reconstructed error after optimization are 3×10^{-3} and 1×10^{-4}, respectively. From Fig. 2, it is evident that the trend of the pure material concentration extracted by the EEPM algorithm is close to the true value.

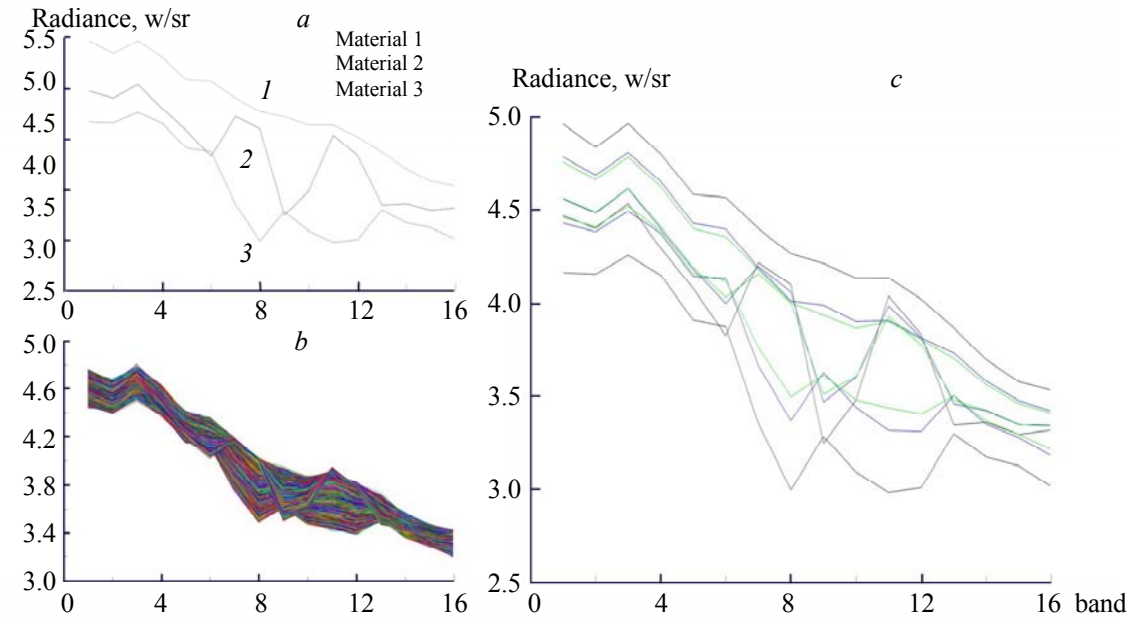

Fig. 1. Surface spectrum of the space target material (a), simulated spectral curve (b), and comparison of the pure material extraction results (c).

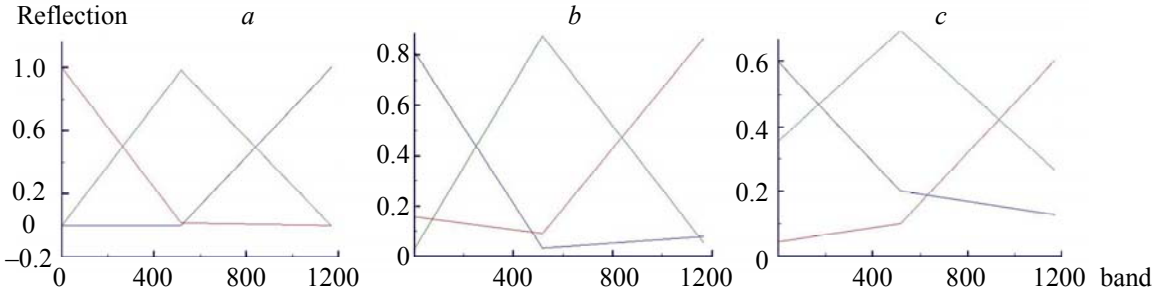

Fig. 2. Concentration from SISAL (a), optimized concentration from EEPM (b), and true concentration (c).

TABLE 1. Experimental Results for Different Noise Conditions

Method	RMSE		
	SNR = 40 dB	SNR = 30 dB	SNR = 20 dB
VCA	0.0030	0.0047	0.0103
PPI	0.0555	0.0643	0.0870
SISAL	0.0030	0.0037	0.0084
EEPM	0.0001	0.0004	0.0081

Table 1 compares the experimental data from the VCA, PPI, SISAL, and EEPM algorithms under different noise conditions. In the experiment, the same data were used for spectral unmixing under 40, 30, and 20 dB noise conditions, and the reconstruction error of each method was compared with the true value. Table 1 shows that the unmixing result of the EEPM algorithm is accurate under the same noise conditions. Moreover, the result has good noise immunity and high mixing precision at low SNR.

Simulation experiment with USGS spectral data. Under the conditions in which the spectral wavenumber is small and the number of pure materials increases, the spectra of pure materials 3, 4, and 5 are selected from the database of the United States Geological Survey (USGS) to verify the spectral unmixing effect of the EEPM algorithm. The wavelength range is 370–2500 nm, with a total of 224 wavelengths. The selected pure material spectra are shown in Fig. 3a.

Materials 3, 4, and 5 are selected as the pure materials, and the spectral matrix is combined with the continuously changing concentration. We take these three pure materials as an example. The experimental results for pure materials 4 and 5 are shown in Table 2. Figure 3b shows the mixed spectrum of pure material 3 and a spectral curve with 40 dB noise.

Figure 3c shows the three pure material spectra at 224 wavelengths and the extraction results achieved using the SISAL and EEPM methods. The results provided by EEPM are clearly closer to the true values compared with those from SISAL. The reconstruction error is reduced from the original 5.5×10^{-3} to 4.2×10^{-3}, while the concentration changes are nearly the same as the true values, as shown in Fig. 4.

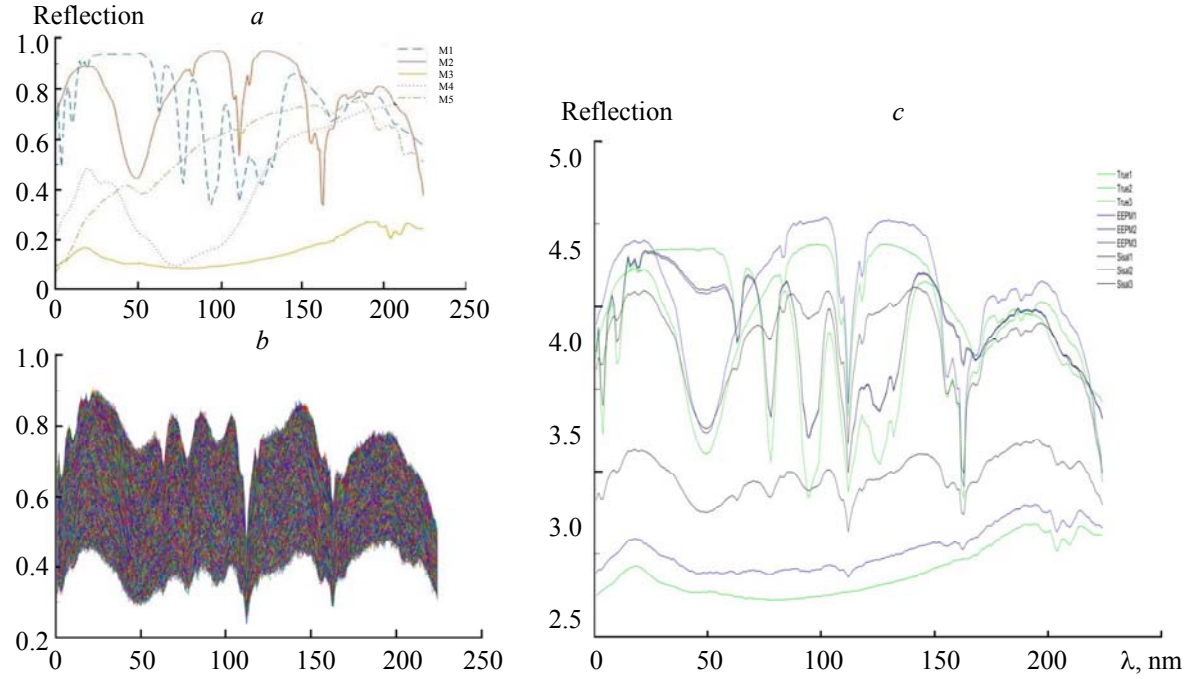

Fig. 3. Spectral curve from the USGS database (a), synthetic spectrum curve (b), and comparison of three pure material extraction results (c).

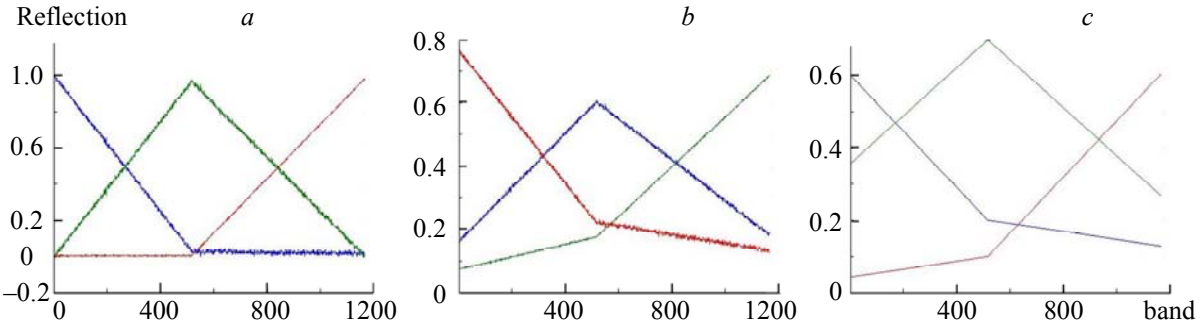

Fig. 4. Concentration from SISAL (a), optimized concentration from EEPM (b), and true concentration (c).

Table 2 is a comparison of the results from the different pure material extraction algorithms under 40 dB noise. Table 2 shows that the reconstruction error of the EEPM algorithm is at a minimum, and the results are clearly better than those of the other algorithms. The EEPM algorithm is also applicable in multiple waves, and the EEPM unmixing effect improves with an increase in the number of pure materials.

The actual observation of a space target is limited by the conditions. We can only obtain the mixed spectral data of the space target material at a certain angle for a continuous time, which means that the pure material concentration in the mixed spectral data is continuously changed. Therefore, the observed concentrations are continuous in a small range, which is not feasible for the main current algorithm based on the volume growth of the simplex. Besides, the traditional algorithms of the minimum simplex need the uniform distribution of the concentration of pure materials in the mixed spectral data. The data we get in the actual measurements do not satisfy the above conditions. Thus, the difference between the measured pure materials and the true values will be large if the traditional minimum simplex methods are applied. In this case, the extracted pure material is further refined and optimized to effectively improve the accuracy of the unmixing results.

TABLE 2. Extraction Results of Different Pure Material 3, 4, 5

Method	RMSE		
Pure material	3	4	5
VCA	0.0041	0.0047	0.0050
PPI	0.0475	0.0322	0.0477
SISAL	0.0034	0.0055	0.0045
EEPM	0.0029	0.0042	0.0032

Compared with the commonly used pure material extraction algorithms, the experimental results show that the advantages of the proposed algorithm are as follows. First, the algorithm overcomes the disadvantage in which the simplex growth algorithm cannot correctly extract the pure material under low information. In the case where the concentration is not uniformly distributed, we cannot get the pure material spectrum from the smallest simplex volume method because the pure material spectral data do not correspond to the simplex vertices. The result approximates the real pure material as well as possible using the purification method. The concentration of unmixing is more accurate and the spectral wavelength range is further extended. As shown in the first simulation experiment, the algorithm performs unmixing well when the spectral wavenumber is small, and the concentration is not significantly changed. Second, in the pure material optimization process, the existing spectral unmixing method, which has high versatility and a wide range of applications, can be further optimized and purified. As shown in the second simulation experiment, the algorithm works well in the case of multi-wavelength multi-terminal elements, and the method is simple, easy to implement, highly accurate, and robust.

Conclusion. A pure material extraction method is improved, and simulation experiments of pure material extraction are conducted. Two groups of simulation experiments show that the algorithm adopts the purification method and optimizes the results of the initial pure material extraction. Moreover, the accuracy of the results is further improved by highlighting its main components and reducing minor components. This method can provide accurate results for spectral unmixing with low information and low noise ratios. The

algorithm is not limited to the volume growth method of the simplex. Under the condition that the spectral data are not a single shape, this algorithm still has certain versatility and optimization effects on the existing spectral unmixing methods. Therefore, the algorithm can considerably improve the reliability of species and concentration measurements from the spectra of space targets. The algorithm has broad prospects in the field of spatial object recognition.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (Grant No. 61575015).

REFERENCES

1. S. M Li, W. Niu, X. Ma, *Defence Technol.*, **30**, No. 3, 6–13 (2009).
2. Q. B. Li, K. J. Wu, Q. S. Gao, *Spectrosc. Spectral Anal.*, **36**, No. 12, 4067–4071 (2016).
3. J. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, J. Chanussot, *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.*, **5**, No. 2, 354–379 (2012).
4. X. L. Jin, Z. J. Tang, C. H. Sui, *J. Space Sci*., **34**, No. 1, 95–103 (2014).
5. R. Marrero, S. Lopez, Gustavo M. Callicó, M. A. Veganzones, A. Plaza, J. Chanussot, R. Sarmiento, *IEEE Transact. Geosci. Remote Sens.*, **53**, No. 7, 3772–3790 (2015).
6. J. Boardman, *Int. Geosci. Remote Sensing Symb.*, **4**, No. 4, 2369–2371 (1994).
7. M. E. Winter, *Image Spectrom.*, **3753**, 266–277 (1999).
8. J. M. Bioucas-Dias, J. M. P. Nascimento, *IEEE Trans. Geosci. Remote Sens.*, **46**, No. 8, 2435–2445 (2008).
9. R. Guerra, L. Santos, S. Lopez, R. Sarmiento, *IEEE Trans. Geosci. Remote Sens.*, **4**, No. 12, 6752–6765 (2015).
10. K. Jorgensen, J. Africano, G. Stansbery, *Int. Symposium on Optical Science and Technology*, 237–244 (2001).