Frontiers of Information Technology & Electronic Engineering
www.zju.edu.cn/jzus; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Fast implementation of kernel simplex volume analysis based on modified Cholesky factorization for endmember extraction*

Jing LI[1], Xiao-run LI[†‡1], Li-jiao WANG[1], Liao-ying ZHAO[2]

([1]*College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China*)

([2]*Institute of Computer Application Technology, Hangzhou Dianzi University, Hangzhou 310018, China*)

[†]E-mail: lxr@zju.edu.cn

Received July 31, 2015; Revision accepted Oct. 16, 2015; Crosschecked Feb. 23, 2016

Abstract: Endmember extraction is a key step in the hyperspectral image analysis process. The kernel new simplex growing algorithm (KNSGA), recently developed as a nonlinear alternative to the simplex growing algorithm (SGA), has proven a promising endmember extraction technique. However, KNSGA still suffers from two issues limiting its application. First, its random initialization leads to inconsistency in final results; second, excessive computation is caused by the iterations of a simplex volume calculation. To solve the first issue, the spatial pixel purity index (SPPI) method is used in this study to extract the first endmember, eliminating the initialization dependence. A novel approach tackles the second issue by initially using a modified Cholesky factorization to decompose the volume matrix into triangular matrices, in order to avoid directly computing the determinant tautologically in the simplex volume formula. Theoretical analysis and experiments on both simulated and real spectral data demonstrate that the proposed algorithm significantly reduces computational complexity, and runs faster than the original algorithm.

Key words: Endmember extraction, Modified Cholesky factorization, Spatial pixel purity index (SPPI), New simplex growing algorithm (NSGA), Kernel new simplex growing algorithm (KNSGA)

http://dx.doi.org/10.1631/FITEE.1500244 **CLC number:** TP75

1 Introduction

Endmember extraction has become increasingly significant in remote sensing image analysis due to the wide existence of mixed pixels. Extraction is a process of selecting a set of pure signature spectra (called endmembers) present in a remotely sensed hyperspectral scene (Schowengerdt, 1997). A variety of endmember extraction algorithms (EEAs), such as the pixel purity index (PPI) (Boardman *et al.*, 1995), the N-finder algorithm (N-FINDR) (Zhao *et al.*, 2012), vertex component analysis (VCA) (Nascimento and Bioucas-Dias, 2005), the automatic target generation process (ATGP) (Ren and Chang, 2003), and the orthogonal based algorithm (OBA) (Tao *et al*., 2009), have been proposed and developed over the last decade. Among these, the simplex growing algorithm (SGA) (Chang *et al*., 2006), which is based on maximum simplex volume theory, has grown popular due to its full automaticity and high efficiency. In conjunction with the new simplex volume formula introduced in previous research (Geng *et al*., 2010), the new simplex growing algorithm (NSGA) and the kernel new simplex growing algorithm (KNSGA) (Zhao *et al*., 2014) were developed as linear and nonlinear alternatives to SGA, respectively.

However, the algorithms developed still suffer from two main issues that limit their application. One is random initialization, which creates inconsistency in the final results, and the other is the excessive computation due to the iterations of simplex volume calculation.

‡ Corresponding author

* Project supported by the Zhejiang Provincial Natural Science Foundation of China (Nos. LY13F020044 and LZ14F030004) and the National Natural Science Foundation of China (No. 61571170)

ORCID: Jing LI, http://orcid.org/0000-0001-8436-1193; Xiao-run LI, http://orcid.org/0000-0001-7611-845X

To solve the first problem, in this paper we integrate the spatial pixel purity index (SPPI) addressed in Cui *et al*. (2013) to find the first endmember for KNSGA. By applying SPPI to the extraction of the first endmember, the process forms a fixed endmember and solves the initialization problem.

As for the second issue, many algorithms aimed at speeding up the computing process have been proposed recently, among which several methods reduce the computation complexity of N-FINDR to some degree (Xiong *et al*., 2011; Dowler *et al*., 2013). Liu and Zhang (2012) proposed a new maximum simplex volume method based on householder transformation (HT), which contributes to lower computational complexity. Wang *et al*. (2013) presented a simple distance measure to replace simplex volume evaluation to speed up the original N-FINDR and SGA. Because a pixel is an endmember when it owns a maximum value in any spectral band of the hyperspectral scene, researchers were able to form a fast extraction algorithm based on the theory of Gaussian elimination, in which fewer bands are required along with the increasing number of extracted endmembers (Geng *et al*., 2013). Another fast endmember extraction algorithm, called the fast gram determinant based algorithm (FGDA), was introduced by Sun *et al*. (2014) using a recursion rule of simplex volume calculation. Xia *et al*. (2012) deduced a new method which uses triangular factorization to simplify simplex volume calculations, thus improving the extraction efficiency.

In this paper, a new, rapid method that takes advantage of modified Cholesky factorization (Gill and Murray, 1974; Gill *et al*., 1981) is presented to transform simplex volume computation into triangular matrix decomposition. The proposed method uses triangular factorization to avoid directly calculating the matrix determinant, thereby reducing the computational complexity and improving the searching efficiency of the algorithm.

2 Impoved KNSGA

This section reviews KNSGA, and then introduces the SPPI that finds the first endmember for KNSGA.

2.1 Kernel new simplex growing algorithm

KNSGA functions by first mapping the original input data in χ into a potentially associated high dimensional feature space $\mathcal{F}$ (Zhao *et al*., 2014),

$$\phi: \boldsymbol{r} \in \chi \to \phi(\boldsymbol{r}) \in \mathcal{F}, \tag{1}$$

and then finds a simplex with the maximum volume in the feature space for a given value of p. To avoid determining the nonlinear mapping feature ϕ, the kernel functions, defined as

$$k(\boldsymbol{r}_i, \boldsymbol{r}_j) = <\phi(\boldsymbol{r}_i) \cdot \phi(\boldsymbol{r}_j)>, \tag{2}$$

are used to calculate the simplex volume, where $\phi(\boldsymbol{r}_i)$ and $\phi(\boldsymbol{r}_j)$ are the mappings of pixel vectors $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$, respectively, $<>$ represents the dot product, and k is a kernel function in terms of the original data.

Considering a hyperspectral datum of N distinct pixels over L spectral bands, the detailed procedure of KNSGA is described in Algorithm 1.

Algorithm 1 Outline for KNSGA

Input: data matrix $\boldsymbol{X}=[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N] \in \mathbb{R}^{L\times N}$ and $\boldsymbol{r} \in \boldsymbol{X}$.

Step 1: initialization.

1a) Let p be the number of endmembers to be generated and use virtual dimension VD (Chang and Du, 2004) to estimate p.

1b) Randomly generate a target pixel, denoted by $\boldsymbol{t}$. Search for the first endmember by $\boldsymbol{e}_1=\arg\{\max_{\boldsymbol{r}}(k(\boldsymbol{r}, \boldsymbol{r})+k(\boldsymbol{t}, \boldsymbol{t})-2k(\boldsymbol{r}, \boldsymbol{t}))\}$, and set i=1.

1c) For each sample $\boldsymbol{r}$, calculate $V(\phi(\boldsymbol{e}_1), \phi(\boldsymbol{r}))=k(\boldsymbol{r}, \boldsymbol{r})+k(\boldsymbol{e}_1, \boldsymbol{e}_1)-2k(\boldsymbol{r}, \boldsymbol{e}_1)$ and find the pixel that satisfies $\boldsymbol{e}_2=\arg\{\max_{\boldsymbol{r}}[V(\phi(\boldsymbol{e}_1), \phi(\boldsymbol{r}))]\}$, and set i=2.

Step 2: start iteration.

2a) Set

$$\boldsymbol{A}_i^{\phi} = [\phi(\boldsymbol{e}_2)-\phi(\boldsymbol{e}_1), \cdots, \phi(\boldsymbol{e}_i)-\phi(\boldsymbol{e}_1), \phi(\boldsymbol{r})-\phi(\boldsymbol{e}_1)].$$

For each sample $\boldsymbol{r}$, calculate

$$V(\phi(\boldsymbol{e}_1), \cdots, \phi(\boldsymbol{e}_i), \phi(\boldsymbol{r})) = \frac{1}{i!}\sqrt{|\det((\boldsymbol{A}_i^{\phi})^{\mathrm{T}}\boldsymbol{A}_i^{\phi})|}.$$

Next, find the pixel that satisfies $\boldsymbol{e}_i=\arg\{\max_{\boldsymbol{r}}[V(\phi(\boldsymbol{e}_1), \ldots, \phi(\boldsymbol{e}_i), \phi(\boldsymbol{r}))]\}$.

2b) If $i<p-1$, then $i=i+1$, and move to step 2a; otherwise, terminate.

Step 3: output the results $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_p\}$.

Experimentation has shown that the KNSGA results vary occasionally for different initializations (Zhao *et al*., 2014). More appropriate methods are still required to more effectively determine the first endmember in KNSGA.

2.2 Determining the first endmember in KNSGA using SPPI

SPPI (Cui *et al*., 2013), based on the space pixel purity index, uses the differences between pure pixels and mixed pixels in a local area to reduce the influence of any spectral changes or abnormal pixels on endmember extraction.

In general, spectral angle distance (SAD) and Euclidean distance (ED) (Cui *et al*., 2013) are used to measure spectral similarity; however, in SPPI, the following mixing distance measure is used to evaluate spectral similarities and differences:

$$M(\boldsymbol{x},\boldsymbol{y})=\alpha A_{\mathrm{sad}}(\boldsymbol{x},\boldsymbol{y})+(1-\alpha)D_{\mathrm{emd}}(\boldsymbol{x},\boldsymbol{y}), \quad (3)$$

where $M(\boldsymbol{x},\boldsymbol{y})$ represents the mixing distance measure of pixels $\boldsymbol{x}$ and $\boldsymbol{y}$, $A_{\mathrm{sad}}(\boldsymbol{x},\boldsymbol{y})$ and $D_{\mathrm{emd}}(\boldsymbol{x},\boldsymbol{y})$ denote the SAD and EMD values of pixels $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively, and the regularization parameter $\alpha\in[0,1]$.

For each pixel $\boldsymbol{x}(i,j)\in\mathbb{R}^L$, the corresponding spatial pixel purity index can be defined as follows:

$$P_{\mathrm{sppi}}(\boldsymbol{x}(i,j))=\max_{(s,t)} M(\boldsymbol{x}(i,j),\boldsymbol{x}(s,t)),\ \ (s,t)\in K, \quad (4)$$

where (i,j) are the space coordinates of the pixel, K is the neighboring window centering on pixel $\boldsymbol{x}(i,j)$, and (s,t) are the space coordinates of the pixel in the neighboring window centering on K.

As described in Cui *et al*. (2013), the smaller the SPPI value, the higher the purity of a pixel; thus, the pixel with the smallest SPPI can be extracted as the first endmember in KNSGA to avoid selecting the first pixel randomly.

2.3 Fast implementation of KNSGA

This section studies the fast implementation of KNSGA using modified Cholesky factorization.

2.3.1 Modified Cholesky factorization

The Cholesky factorization of matrix $\boldsymbol{M}$ is written as follows (Golub and van Loan, 1996):

$$\boldsymbol{M}=\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}, \quad (5)$$

where $\boldsymbol{M}$ is a positive definite symmetric matrix, and $\boldsymbol{L}$ is a lower triangular matrix with strictly positive diagonal entries. According to mathematical theory, the diagonal elements of $\boldsymbol{L}$ can be obtained by $l_{i,i}=\left(m_{i,i}-\sum_{k=1}^{i-1}l_{i,k}^2\right)^{1/2}$, where $l_{i,i}$ and $m_{i,i}$ indicate the diagonal elements of matrix $\boldsymbol{L}$ and matrix $\boldsymbol{M}$ respectively, and $l_{i,j}$ represents the element in the ith row and jth column of matrix $\boldsymbol{L}$. This requires square root calculation, which results in large calculation complexity. To circumvent square root calculation, in this paper we use modified Cholesky factorization to simplify the calculation. The modified Cholesky factorization can be written as follows (Gill and Murray, 1974; Gill *et al*., 1981):

$$\boldsymbol{M}=\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}}, \quad (6)$$

where $\boldsymbol{L}$ is a lower triangular matrix with diagonal elements equal to one, and $\boldsymbol{D}$ is a diagonal matrix with strictly positive diagonal entries. This is implemented as follows:

$$l_{i,j}=\frac{1}{d_j}\left(m_{i,j}-\sum_{k=1}^{j-1}l_{i,k}d_k l_{j,k}\right), \quad (7)$$

$$d_j=m_{j,j}-\sum_{k=1}^{j-1}l_{j,k}d_k l_{j,k},\ \ i>j, \quad (8)$$

where d_j marks the diagonal element of matrix $\boldsymbol{D}$, and $m_{i,j}$ and $l_{i,j}$ represent elements in the ith row and jth column of matrix $\boldsymbol{M}$ and matrix $\boldsymbol{L}$, respectively.

2.3.2 Implementing KNSGA based on Cholesky factorization

The time cost of KNSGA is devoted to the calculation of the determining matrix $(\boldsymbol{A}_{p-1}^{\phi})^{\mathrm{T}}\boldsymbol{A}_{p-1}^{\phi}$. Because $(\boldsymbol{A}_{p-1}^{\phi})^{\mathrm{T}}\boldsymbol{A}_{p-1}^{\phi}$ is a positive definite symmetric matrix, it can be decomposed by modified Cholesky factorization $(\boldsymbol{A}_{p-1}^{\phi})^{\mathrm{T}}\boldsymbol{A}_{p-1}^{\phi}=\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}}$, as described in Eq. (6). For convenience, the diagonal elements of matrix $\boldsymbol{D}$ are put here into the corresponding diagonal positions of matrix $\boldsymbol{L}$ in actual calculation. Next, the new simplex volume proposed by Geng *et al*. (2010) can be further expressed as

$$V=\frac{1}{(p-1)!}|\det((\boldsymbol{A}_{p-1}^{\phi})^{\mathrm{T}}\boldsymbol{A}_{p-1}^{\phi})|^{1/2}$$
$$=\frac{1}{(p-1)!}|\det(\boldsymbol{LDL}^{\mathrm{T}})|^{1/2} \quad (9)$$
$$=\frac{1}{(p-1)!}|d_1||d_2|\cdots|d_{p-1}|,$$

where p is the number of endmembers to be generated and d_j (j=1, 2, …, p−1) is obtained using Eqs. (7) and (8). According to Eq. (9), the computation of the complicated determinant is transformed into simple multiplication, reducing the calculation complexity and saving time. Maximizing the simplex volume is, clearly, equivalent to maximizing $|d_j|$.

Based on the above analysis, the progress of the fast kernel new simplex growing algorithm (FKNSGACF) can be described as follows:

Algorithm 2 Outline for FKNSGACF

Input: matrix $\boldsymbol{X}=[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N]\in\mathbb{R}^{L\times N}$.

Step 1: initialization.

1a) Use VD to estimate p.

1b) For each pixel $\boldsymbol{x}_n$, calculate the spatial pixel purity index defined in Cui *et al*. (2013), $\boldsymbol{e}_1=\arg[\min_{x_n}(P_{\mathrm{sppi}})]$.

1c) Find the pixel farthest from $\boldsymbol{e}_1$, set it as endmember $\boldsymbol{e}_2$, and set its position as id(1). For each sample pixel $\boldsymbol{x}_n$, calculate

$$d_n^1=k(\boldsymbol{x}_n,\boldsymbol{x}_n)+k(\boldsymbol{e}_1,\boldsymbol{e}_1)-2k(\boldsymbol{x}_n,\boldsymbol{e}_1),\ n=1,2,\cdots,N,$$
$$\boldsymbol{e}_2=\arg[\max_{x_n}(d_n^1)],\ \mathrm{id}(1)=\arg[\max_n(d_n^1)].$$

1d) Set iteration index m=2.

Step 2: start iteration.

2a) For each sample vector $\boldsymbol{x}_n$, calculate

$$l_n^m=\frac{1}{d_{\mathrm{id}(m)}^m}[\phi(\boldsymbol{x}_n)-\phi(\boldsymbol{e}_1)]^{\mathrm{T}}[\phi(\boldsymbol{e}_m)-\phi(\boldsymbol{e}_1)]$$
$$-\frac{1}{d_{\mathrm{id}(m)}^m}\sum_{k=1}^{m-1}l_n^k d_{\mathrm{id}(k)}^k l_{\mathrm{id}(m)}^k.$$

2b) Calculate $d_n^{m+1}=d_n^m-l_n^m d_{\mathrm{id}(m)}^m l_n^m$. Find the pixel that satisfies the largest d_n^{m+1}, and then calculate

$$\boldsymbol{e}_{m+1}=\arg[\max_{x_n}(d_n^{m+1})],$$
$$\mathrm{id}(m+1)=\arg[\max_n(d_n^{m+1})].$$

2c) If m<p−1, then m=m+1 and move to step 2a; otherwise, terminate.

Step 3: output the results $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_p\}$.

Note that fast KNSGA is called FKNSGACF here to ensure consistency with the name of the fast new growing algorithm based on Cholesky factorization (FNSGACF) (Wang *et al*., 2014).

In KNSGA, endmembers are extracted one by one and their simplex volumes must be calculated repeatedly in iteration until all endmembers are extracted. FKNSGACF, however, needs only to compute the simplex volume after all the endmembers are extracted, which considerably reduces the computation complexity compared to KNSGA. From the mathematical view, generally the computational complexity of KNSGA, in which LU decomposition (also imposed in the Matlab built-in function 'det') is used to calculate the determinant of a square matrix, is about $O(Np^3+N^2p^3)$. The computational complexity of the proposed FKNSGACF, in which modified Cholesky factorization is used for determinant calculation, is nearly $O(Np^3+N^2)$. So, theoretically FKNSGACF offers a more time-saving alternative.

3 Experiments with simulated data

This section describes a set of experiments in which simulated hyperspectral data was used to facilitate comprehensive analysis of the proposed method. All experiments were conducted on the same configured PC. Algorithms were programmed in Matlab 7.0.

The endmembers and their corresponding abundances were known in simulated data. Spectral signatures were selected from the USGS Digital Spectral Library (http://speclab.cr.usgs.gov) to generate the simulated data, and mixed nonlinearly using the bilinear spectrum mixing model described in Zhao *et al*. (2014). To evaluate accuracy, the mean spectra angle distance (SAD), a criterion commonly used to measure the similarity of the extracted endmembers and true endmembers, was used in a similar way to previous studies (Ren and Chang, 2003; Miao and Qi, 2007; Nascimento and Bioucas-Dias, 2008). The value of simplex volumes is an additional criterion that is used to evaluate results obtained by different methods. Here, the polynomial kernel defined as $k(\boldsymbol{r}_i,\boldsymbol{r}_j)=(a\boldsymbol{r}_i^{\mathrm{T}}\boldsymbol{r}_j+c)^b$ was used, where a, b, and c are kernel parameters set as b=8/9, c=1, and a the reciprocal of the square of the maximum value of the image data.

Experiment 1 This experiment was designed to determine whether the final results of endmember extraction algorithms depend on their initializations. Each algorithm was implemented separately 10 times. The algorithms in the following experiments with simulated data were also implemented separately 10 times, which will not be repeated hereinafter. In this experiment, simulated data was generated by five endmembers, with 189 bands, a fixed 30 dB signal-to-noise ratio (SNR), and an image size of 64×64. Fig. 1 shows the endmember extraction results, where the markers with different colors indicate the ground truth of five pure pixels and the extracted endmembers are marked with white Arabic numerals of 1–5. Specifically, Figs. 1a–1c show the endmember extraction results that appeared the most and the second most frequently among the 10 implementations of KNSGA, plus the only result obtained by the improved KNSGA with SPPI (temporarily called KNSGA (SPPI)), and demonstrate that KNSGA (SPPI) solves the initialization problem. Table 1 lists the processing time, the average SADs, and simplex volumes that correspond to the results shown in Fig. 1, where KNSGA (most) and KNSGA (second most) represent the results that appeared the most and second most frequently, respectively.

Table 1 shows that the two results obtained by KNSGA were similar, as the average SADs and volumes were almost the same. KNSGA (SPPI) was shown to perform better, as the final endmembers extracted had smaller SADs and larger simplex volume values than the results obtained by the original KNSGA. Use of SPPI during initialization helps find the first endmember correctly; thus, the KNSGA with SPPI eliminates the influence of initialization.

Experiment 2 This experiment was designed to detail the relationship between the performance of two algorithms and different SNR values. Here, five endmembers were selected from the USGS Digital Spectral Library to generate a synthetic image of size 64×64. Additionally, six levels of SNRs, namely, 10, 20, 30, 40, 50, and 60 dB, were added to the simulated data, respectively. The corresponding simplex volumes and SADs for both methods are listed in Table 2.

As shown in Table 2, generally the SADs and simplex volumes for both methods experienced decrease with the decrease of SNR and the new algorithm performed better with larger volumes and smaller SADs.

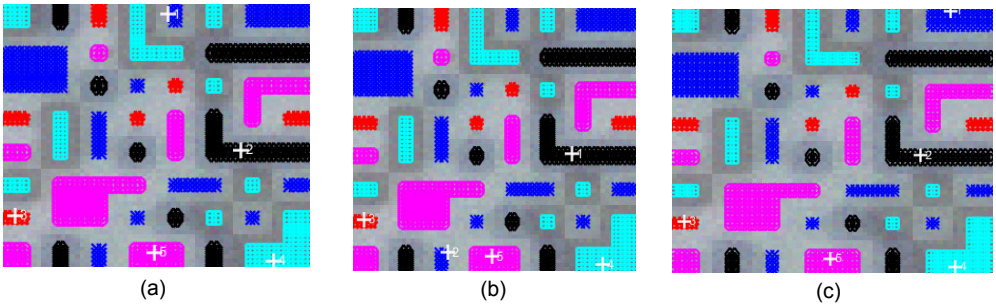
Table 1 Computation times, average SADs and simplex volumes obtained by KNSGA and KNSGA (SPPI)

Algorithm	Time (s)	Simplex volume	Average SAD
KNSGA (most)	1.342	1.520	0.042
KNSGA (second most)	1.351	1.508	0.043
KNSGA (SPPI)	1.340	1.749	0.036

Table 2 Simplex volumes and average SADs obtained by two methods with respect to different SNRs

SNR (dB)	Simplex volume		Average SAD	
	KNSGA	FKNSGACF	KNSGA	FKNSGACF
10	2.04	2.16	0.067	0.066
20	1.37	1.37	0.048	0.045
30	1.32	1.32	0.038	0.038
40	1.28	1.28	0.035	0.035
50	1.18	1.25	0.037	0.033
60	1.23	1.23	0.032	0.032

Experiment 3 This experiment was conducted to detail the relationship between the performance of endmember extraction algorithms and the number of endmembers they extracted. Simulated data was generated with a fixed 30 dB SNR and an image size of 64×64, and the estimated number of endmembers



(a) (b) (c)

Fig. 1 Endmember extraction results: (a) result of KNSGA that appeared the most frequently; (b) result of KNSGA that appeared the second most frequently; (c) only result of KNSGA (SPPI)

ranged from 6 to 18.

Table 3 displays the average SADs and simplex volumes for both methods. The results of the fast implementation algorithm were occasionally better than those of the original algorithm with larger volumes and smaller SADs, which can potentially be attributed to the use of SPPI in initialization. At the same time, the performances of all algorithms deteriorated with increase in the estimated number of endmembers. Table 4 shows the running time for both algorithms, indicating various degrees of reduction in computation time when the fast method is adopted. Another notable conclusion suggested by the results in Table 4 is that as the estimated number of endmembers increases, the speedup ratio increases.

Table 3 Simplex volumes and average SADs obtained by the two methods with respect to the number of endmembers

Number of endmembers	Simplex volume		Average SAD	
	KNSGA	FKNSGACF	KNSGA	FKNSGACF
6	3.3×10^{-1}	3.7×10^{-1}	0.039	0.037
9	2.3×10^{-3}	2.2×10^{-3}	0.052	0.052
12	8.1×10^{-7}	8.2×10^{-7}	0.051	0.051
15	1.1×10^{-10}	1.2×10^{-10}	0.072	0.071
18	1.8×10^{-14}	1.8×10^{-14}	0.071	0.071

Table 4 Time consumed by two methods with respect to the number of endmembers

Number of endmembers	Time (s)		Speedup
	KNSGA	FKNSGACF	
6	1.78	1.47	1.21×
9	3.65	2.01	1.82×
12	6.17	2.85	2.16×
15	9.37	3.60	2.60×
18	13.06	4.36	3.00×

Experiment 4 This experiment was conducted to detail the performance of the new algorithm for images with various numbers of pixels. Five simulated images with different pixel sizes ranging from 64×64 to 256×256 were used, with the SNR fixed at 30 dB and five endmembers extracted. The corresponding SADs and the simplex volumes are provided in Table 5. A comparison of corresponding time consumptions between these two methods is shown in Table 6.

The new algorithm, FKNSGACF, performed better than KNSGA when it came to similarly sized pixels, as shown in Table 5, due to its larger simplex volume values and smaller SADs (except for sizes 144×144 and 196×196). The results of the two algorithms changed slightly alongside the increased number of pixels. Additionally, according to Table 6, the speedup ratios picked up less variation with respect to the image size, and the new fast computational method saved computing time for data with a speed increase of a factor 2, which can play a major role in saving time when the image size was very large.

Table 5 Simplex volumes and average SADs obtained by two methods with respect to various pixel sizes

Pixel size	Simplex volume		Average SAD	
	KNSGA	FKNSGACF	KNSGA	FKNSGACF
64×64	1.50	1.75	0.044	0.036
100×100	1.57	1.85	0.043	0.038
144×144	1.60	1.58	0.041	0.043
196×196	1.63	1.57	0.042	0.045
256×256	1.68	1.92	0.042	0.037

Table 6 Time consumed by two methods with respect to various pixel sizes

Pixel size	Time (s)		Speedup
	KNSGA	FKNSGACF	
64×64	2.26	1.24	1.82×
100×100	3.50	2.01	1.74×
144×144	7.13	4.05	1.76×
196×196	12.84	6.94	1.85×
256×256	18.47	9.62	1.92×

4 Experiments on real data

The real hyperspectral Cuprite image data (NASA, 1997) obtained by the airborne visible/infrared imaging spectrometer (AVIRIS) was used to further investigate the performance of FKNSGACF. The original image consists of 350×350 pixels and 224 bands with a 0.4- to 2.5-μm spectral range. Fig. 2 shows the first band of the image. Prior to analysis, bands 1–3, 105–115, and 150–170 were removed due to water absorption and low SNR, resulting in a remaining total of 189 spectral bands. Here, simplex volume was used to compare the accuracies of extracted endmembers, and the polynomial kernel was used in experiments where parameter a was set to 100 divided by the square of the maximum reflectance

data in the whole image, b was set to 8/9, and c was set to 1.

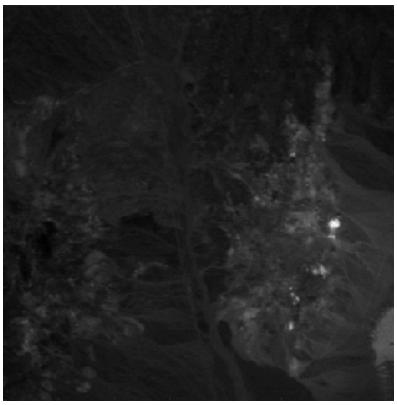


Fig. 2 The first band of the Cuprite airborne visible/ infrared imaging spectrometer (AVIRIS) image

Tables 7 and 8 provide simplex volumes and running time of the two different methods for the Cuprite data, where the number of endmembers varied from 6 to 22 and the results of KNSGA were average values after running 10 times. The traditional KNSGA and its fast implementation, FKNSGACF, extracted the same number of endmembers, and the values of simplex volumes were the same. As expected, the fast implementation algorithm greatly reduced the computing time for the Cuprite data, and the speedup ratio for FKNSGACF increased along with increase in the number of pixels to be extracted.

Table 7 Simplex volumes obtained by two methods for Cuprite data with respect to the number of endmembers

Number of endmembers	Simplex volume	
	KNSGA	FKNSGACF
6	4.33×10^{18}	4.33×10^{18}
10	1.67×10^{29}	1.67×10^{29}
14	1.83×10^{38}	1.83×10^{38}
18	1.29×10^{46}	1.29×10^{46}
22	8.33×10^{52}	8.33×10^{52}

Table 8 Time consumed by two methods for Cuprite data with respect to the number of endmembers

Number of endmembers	Time (s)		Speedup
	KNSGA	FKNSGACF	
6	52.8	49.2	1.07×
10	135.5	83.5	1.62×
14	273.4	151.9	1.80×
18	415.4	180.2	2.31×
22	607.7	236.5	2.57×

5 Conclusions

In this paper we propose a new endmember extraction algorithm called FKNSGACF, a fast implementation of KNSGA. The new algorithm incorporates the advantages of SPPI to resolve the problem inherent to initialization. The proposed method also uses modified Cholesky factorization to reduce the computational complexity by decomposing the volume matrix into a product of the triangular matrix without repeatedly calculating the matrix determinant. Experimentation on both simulated and real data sets indicates that the proposed algorithm achieves ideal performance in reducing computation time and acquiring larger simplex volumes.

References

Boardman, J.W., Kruse, F.A., Green, R.O., 1995. Mapping target signatures via partial unmixing of AVIRIS data. JPL Airborne Earth Science Workshop, p.23-26.

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. http://dx.doi.org/10.1109/TGRS.2003.819189

Chang, C.I., Wu, C., Liu, W., *et al*., 2006. A new growing method for simplex-based endmember extraction algorithms. *IEEE Trans. Geosci. Remote Sens.*, **44**(10):2804-2819. http://dx.doi.org/10.1109/TGRS.2006.881803

Cui, J.T., Wang, J., Li, X.R., *et al*., 2013. Endmember extraction algorithm based on spatial pixel purity index. *J. Zhejiang Univ. (Eng. Sci.)*, **47**(9):1517-1523 (in Chinese). http://dx.doi.org/10.3785/j.issn.1008-973X.2013.09.002

Dowler, S.W., Takashima, R., Andrews, M., 2013. Reducing the complexity of the N-FINDR algorithm for hyperspectral image analysis. *IEEE Trans. Image Process.*, **22**(7):2835-2848. http://dx.doi.org/10.1109/TIP.2012.2219546

Geng, X.R., Zhao, Y.C., Wang, F.X., *et al.*, 2010. A new volume formula for a simplex and its application to endmember extraction for hyperspectral image analysis. *Int. J. Remote Sens*., **31**(4):1027-1035. http://dx.doi.org/10.1080/01431160903154283

Geng, X.R., Xiao, Z.Q., Ji, L.Y., *et al*., 2013. A Gaussian elimination based fast endmember extraction algorithm for hyperspectral imagery. *ISPRS J. Photogr. Remote Sens*., **79**(5):211-218. http://dx.doi.org/10.1016/j.isprsjprs.2013.02.020

Gill, P.E., Murray, W., 1974. Newton-type method for unconstrained and linearly constrained optimization. *Math. Programm.*, **7**(1):311-350. http://dx.doi.org/10.1007/BF01585529

Gill, P.E., Murray, W., Wright, M.H., 1981. Practical Optimization. Academic Press, London.

Golub, G.H., van Loan, C.F., 1996. Matrix Computations. The John Hopkins University Press, Baltimore, Mariland.

Liu, J.M., Zhang, J.S., 2012. A new maximum simplex volume method based on householder transformation for endmember extraction. *IEEE Trans. Geosci. Remote Sens*., **50**(1):104-118. http://dx.doi.org/10.1109/TGRS.2011.2158829

Miao, L., Qi, H., 2007. Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. *IEEE Trans. Geosci. Remote Sens*., **45**(3):765-777. http://dx.doi.org/10.1109/TGRS.2006.888466

NASA, 1997. NASA AVIRIS Data. Available from http://aviris.jpl.nasa.gov.

Nascimento, J.M.P., Bioucas-Dias, J.M., 2005. Vertex component analysis: a fast algorithm to unmix hyperspectral data. *IEEE Trans. Geosci. Remote Sens*., **43**(4):898-910. http://dx.doi.org/10.1109/TGRS.2005.844293

Nascimento, J.M.P., Bioucas-Dias, J.M., 2008. New developments on VCA unmixing algorithm. *SPIE*, **7109**: 71090F. http://dx.doi.org/10.1117/12.799838

Ren, H., Chang, C.I., 2003. Automatic spectral target recognition in hyperspectral imagery. *IEEE Trans. Aerosp. Electron. Syst.*, **39**(4):1232-1249. http://dx.doi.org/10.1109/TAES.2003.1261124

Schowengerdt, R.A., 1997. Remote Sensing: Models and Methods for Image Processing. Academic Press, New York.

Sun, K., Geng, X., Wang, P., 2014. A fast endmember extraction algorithm based on gram determinant. *IEEE Geosci. Remote Sens. Lett.*, **11**(6):1124-1128. http://dx.doi.org/10.1109/LGRS.2013.2288093

Tao, X., Wang, B., Zhang, L., 2009. Orthogonal bases approach for decomposition of mixed pixels for hyperspectral imagery. *IEEE Geosci. Remote Sens. Lett*., **6**(2): 219-223. http://dx.doi.org/10.1109/LGRS.2008.2010529

Wang, L., Wei, F., Liu, D., 2013. Fast implementation of maximum simplex volume-based endmember extraction in original hyperspectral data space. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens*., **6**(2):516-521. http://dx.doi.org/10.1109/JSTARS.2012.2234439

Wang, L.J., Li, X.R., Zhao, L.Y., 2014. Fast implement of the simplex growing algorithm for endmember extraction. *Acta Opt. Sin.*, **34**(11):1128001 (in Chinese). http://dx.doi.org/10.3788/AOS201434.1128001

Xia, W., Pu, H.Y., Wang, B., *et al*., 2012. Triangular factorization-based simplex algorithms for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.*, **50**(11): 4420-4440. http://dx.doi.org/10.1109/TGRS.2012.2195185

Xiong, W., Chang, C.I., Wu, C.C., 2011. Fast algorithms to implement N-FINDR for hyperspectral endmember extraction. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens*., **4**(3):545-564. http://dx.doi.org/10.1109/JSTARS.2011.2119466

Zhao, C.H., Qi, B., Wang, Y.L., 2012. An improved N-FINDR hyperspectral endmember extraction algorithm. *J. Electron. Inform. Technol*., **34**(2):499-503 (in Chinese).

Zhao, L.Y., Zheng, J.P., Li, X.R., *et al*., 2014. Kernel simplex growing algorithm based on a new simplex volume formula for hyperspectral endmember extraction. *J. Appl. Remote Sens*., **8**(1):083594. http://dx.doi.org/10.1117/1.JRS.8.083594