

## 基于代数余子式的 N-FINDR 快速端元提取算法

李琳<sup>\*①②</sup> 孟令博<sup>①②</sup> 孙康<sup>①②</sup> 赵永超<sup>①</sup>

<sup>①</sup>(中国科学院电子学研究所 北京 100190)

<sup>②</sup>(中国科学院大学 北京 100049)

**摘要:** 基于高光谱图像特征空间几何分布的端元提取方法通常可分为投影类算法和单形体体积最大类算法, 通常前者精度不好, 后者计算复杂度较高。该文提出一种基于代数余子式的快速 N-FINDR 端元提取算法(FCA), 该算法融合了投影类算法速度快和单形体体积最大类算法精度高的优势, 利用像元投影到端元矩阵元素的代数余子式构成的向量上的方法, 寻找最大体积的单形体。此外, 该算法在端元搜索方面较为灵活, 每次迭代都可用纯度更高的像元代替已有端元, 因此能保证用该端元确定的单形体, 可以将特征空间中全部像元包含在内。仿真和实际高光谱数据实验结果表明, 该文算法在精准提取出端元的同时, 收敛速度非常快。

**关键词:** 图像处理; 高光谱; 端元提取; 单形体; 体积最大; 代数余子式; 投影

中图分类号: TP751.1

文献标识码: A

文章编号: 1009-5896(2015)05-1128-07

DOI: 10.11999/JEIT140923

## A Fast N-FINDR Algorithm Based on Cofactor of a Determinant

Li Lin<sup>①②</sup> Meng Ling-bo<sup>①②</sup> Sun Kang<sup>①②</sup> Zhao Yong-chao<sup>①</sup>

<sup>①</sup>(Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China)

<sup>②</sup>(University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** Endmember extraction methods based on geometric distribution of hyperspectral images usually divide into projection algorithm and the maximum volume formula for simplex, which the former has lower computational complexity and the latter has better precision. A Fast endmember extraction method based on Cofactor of a determinant Algorithm (FCA) is proposed. The algorithm combines the two kinds of algorithms, and which means it has a high speed and accuracy performance for endmember extraction. FCA finds the max volume of simplex by making pixels project to vectors, which are composed of the cofactors of elements in endmember determinant. Besides, FCA is flexible in endmember search, for it can use higher purity pixels to replace the endmembers extracted in the last iteration, which ensures that all the endmembers extracted by FCA are the vertices of simplex. The theoretical analysis and experiments on both simulated and real hyperspectral data demonstrate that the proposed algorithm is a fast and accurate algorithm for endmember extraction.

**Key words:** Image processing; Hyperspectral; Endmember extraction; Simplex; Maximum volume; Cofactor; Projection

### 1 引言

高光谱遥感图像空间分辨率通常相对较低, 再加上自然界地物的复杂性、多样性, 故混合像元普遍存在于高光谱遥感图像中。而混合像元的存在正是像元级遥感分类和面积测量精度难以达到使用要求的主要原因。因此, 光谱解混是分析高光谱图像的关键所在<sup>[1]</sup>。在光谱解混过程中, 端元提取算法可以提取出图像中只包含一种特征地物的纯像元。

常见的端元提取算法通常是在线性混合模型成立条件下<sup>[2,3]</sup>, 基于高光谱图像特征空间几何分布的。

该几何分布是指图像中的像元分布在以端元为顶点的单形体结构中<sup>[4]</sup>。目前, 绝大多数端元提取算法都致力于寻找单形体的顶点, 具体可分为基于特征空间的投影算法和单形体体积最大算法。其中投影算法有: 纯像元指数算法(Pixel Purity Index, PPI)<sup>[5]</sup>, 顶点成分分析算法(Vertex Component Analysis, VCA)<sup>[6]</sup>, 正交子空间投影算法(Orthogonal Subspace Projection, OSP)<sup>[7]</sup>等。这类算法通常将像元投影到一组向量上, 位于线段端点的像元是纯像元的可能性最大。其优点是复杂度较低, 缺点是投影向量的选择具有一定随机性, 端元提取的结果很大程度上依赖于向量的选择, 因此精度不高。而

2014-07-15 收到, 2014-11-28 改回

\*通信作者: 李琳 linli\_iceas8@163.com

另一类基于单形体体积最大的算法则精度较好。如 N-FINDR 算法<sup>[8]</sup>，通过寻找具有最大体积的单形体来自动获取图像中的所有端元，但是其算法复杂度很高，因而应用受到一定限制。故有很多算法对其改进，如单形体增长算法 (Simplex Growing Algorithm, SGA)<sup>[9]</sup>和基于 Householder 变换的体积最大算法 (Maximum Volume based on Householder Transformation, MVHT)<sup>[10]</sup>等。尽管 SGA 和 MVHT 算法在计算复杂度上低于 N-FINDR，但是其寻找端元所用的是贪婪算法，即从第 1 个初始端元出发，逐步一个个地确定后面的端元，因而这种搜索方法总是做出在当前看来最好的选择，从局部而非整体考虑结果。此方法缺点在于一旦端元入选后就不能将其删除，故容易陷入局部极值。而体积最大算法 (New Volume Formula for Simplex, NVFS)<sup>[11]</sup>虽然在端元搜索方面和 N-FINDR 一样，即：已寻找到的端元仍可被其他像元替代，但是其计算复杂度比 N-FINDR 还高，因此应用也受到了一定限制。

由此可知，投影算法虽然复杂度较低，但精度不高；单形体体积最大算法虽然精度较好，但算法复杂度较高。所以在此背景下，本文提出一种基于代数余子式的 N-FINDR 快速端元提取算法 (A Fast endmember extraction method based on Cofactor of a determinant Algorithm, FCA)。该算法是对 N-FINDR 的改进，有如下几点创新和优势：首先，FCA 本质思想是寻找最大体积的单形体，故克服了投影算法精度低的缺点；其次，FCA 可将图像全部像元投影到一个端元矩阵元素产生的代数余子式的向量上，故具有投影算法复杂度低的优势；最后，在端元搜索方面，FCA 采用随机搜索算法，所选择的端元在计算过程中可随时被其他纯度更高的像元替代，故克服了 SGA 和 MVHT 采用的贪婪搜索算法的缺点，减少了提取出异常极值的情况。综上，FCA 兼备投影算法和单形体体积最大算法的优点，可以高精度高速度地提取出纯像元。

## 2 线性混合模型和 N-FINDR 算法

### 2.1 线性混合模型

线性光谱混合模型 (Linear Mixture Model, LMM) 假设：太阳入射辐射只与一种地物表面发生作用，物体间没有相互作用。故图像中任意一个像元光谱  $\mathbf{r}$  可以表示为

$$\mathbf{r} = \sum_{i=1}^p c_i \mathbf{e}_i + \mathbf{n} = \mathbf{E}\mathbf{c} + \mathbf{n} \quad (1)$$

$$\sum_{i=1}^p c_i = 1, \quad 0 < c_i < 1 \quad (2)$$

其中  $\mathbf{e}_i$  为端元光谱， $p$  为图像中端元个数， $c_i$  为像元光谱  $\mathbf{r}$  中每个端元所占的丰度， $\mathbf{n}$  为误差和噪声。当图像满足线性混合模型，且不考虑噪声影响时，图像中的所有像元都被包含在一个以端元为顶点的单形体中<sup>[12]</sup>。如图 1 所示，2 个波段 3 个端元的图像像元分布在特征空间中以端元为顶点的三角形中。故可通过寻找构成特征空间中最大单形体体积的顶点来提取端元。N-FINDR 即是基于这种思想的端元提取算法。

### 2.2 N-FINDR 算法

N-FINDR 通过计算式(3)和式(4)来提取一组端元。

$$\mathbf{E} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_p \end{bmatrix} \quad (3)$$

$$V(\mathbf{E}) = \frac{1}{(N-1)!} \text{abs}(|\mathbf{E}|) \quad (4)$$

其中  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p$  是图像中任意选取的一组像元光谱， $|\cdot|$  是行列式运算符号。由于式(4)需计算行列式，所以  $\mathbf{E}$  必为方阵，故  $\mathbf{e}_i$  的维数必为  $p-1$ ， $i=1, 2, \dots, p$ 。当像元光谱维数不满足此条件时，需要用降维 (Dimension Reduction, DR) 方法来降维，常用的算法有 MNF, PCA, ICA 等，本文均采用 PCA 算法。N-FINDR 实现方式有两种：(1) 用任意一个像元光谱依次代替  $\mathbf{e}_1$  到  $\mathbf{e}_p$ ，计算其体积，然后再对图像中每个像元重复这个循环过程，进而找到最大体积单形体对应的端元光谱。(2) 依次用每个像元光谱代替  $\mathbf{e}_1$  计算相应的体积，然后对  $\mathbf{e}_2$  到  $\mathbf{e}_p$  循环这个过程，找到最大体积单形体对应的端元光谱。这两种方式的端元提取结果较为接近，因此都被广泛采用。本文提出的 FCA 算法属于 N-FINDR 的第 2 种实现方式。

## 3 FCA 算法描述

FCA 算法是建立在线性混合模型成立条件下，高光谱图像特征空间中单形体体积最大的一类方法。本质上和 N-FINDR 原理一样，但 N-FINDR

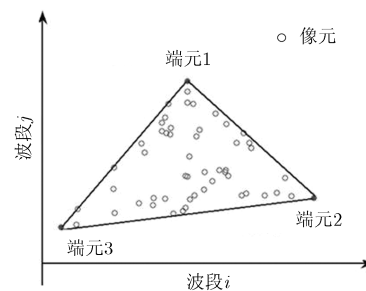


图 1 2 维特征空间下含有 3 个端元的图像像元三角形结构(图中圆圈“o”代表像元)

需要大量运算,尤其是很多矩阵运算需要大量迭代,时间复杂度高,不利于工程实际中对高光谱图像的大量处理。本文将图像所有像元投影到以端元矩阵元素的代数余子式构成的向量上的方法对 N-FINDR 第 2 种实现方式进行改进,使算法在保持端元提取精度的同时,时间复杂度大大降低。

### 3.1 算法原理及步骤

设  $\mathbf{R} = \{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$  为图像像元矩阵,  $\mathbf{E} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p\}$  为端元矩阵,以确定第 1 个端元  $\mathbf{e}_1$  为例阐述原理。应用 N-FINDR 算法时,要在保持  $\mathbf{e}_2$  到  $\mathbf{e}_p$  不变的情况下,将  $\mathbf{r}_i (i = 1, 2, \dots, N)$  依次替代  $\mathbf{e}_1$  求方阵  $\mathbf{E}$  的最大体积  $V(\mathbf{E})$ 。根据式(4)可知,体积越大,  $\det(\mathbf{E})$  的绝对值越大,故求体积最大简化为求方阵  $\mathbf{E}$  的行列式的绝对值最大,而  $\mathbf{E}$  中  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p\}$  一直保持不变,却被重复计算  $N - 1$  次,因此算法复杂度很高。

应用元素与其代数余子式乘积之和求行列式  $\det(\mathbf{E})$  的方法可以降低算法复杂度。即

$$\det(\mathbf{E}) = \sum_{i=1}^p e_{i,1} \times (-1)^{i+1} \times M(e_{i,1}) \quad (5)$$

其中  $(-1)^{i+1} \times M(e_{i,1})$  为元素  $e_{i,1}$  的代数余子式,  $M(e_{i,1})$  为元素  $e_{i,1}$  的余子式,即

$$M(e_{i,1}) = \begin{vmatrix} \cancel{e_{1,1}} & e_{1,2} & \dots & e_{1,p} \\ \dots & \dots & \dots & \dots \\ \cancel{e_{i,1}} & \cancel{e_{i,2}} & \dots & \cancel{e_{i,p}} \\ \dots & \dots & \dots & \dots \\ \cancel{e_{p,1}} & e_{p,2} & \dots & e_{p,p} \end{vmatrix} \quad (6)$$

由式(6)可知,元素  $e_{i,1}$  的代数余子式的值与  $e_{i,1}$  的具体值无关,只与  $e_{i,1}$  在方阵中的位置有关。因此,用不同像元代替端元  $\mathbf{e}_1$ ,相同位置的元素对应的代数余子式相同。故可以得到端元  $\mathbf{e}_1$  的  $p$  个元素所在位置对应的代数余子式向量为

$$\mathbf{Me}_1 = \{(-1)^{1+1} M(e_{1,1}), (-1)^{1+2} M(e_{2,1}), \dots, (-1)^{1+p} M(e_{p,1})\} \quad (7)$$

由式(6)可知,  $\mathbf{Me}_1$  对于端元  $\mathbf{e}_1$ ,即第 1 列的所有元素而言是常量,故可将图像的  $N$  个像元投影到这一向量上。

由图 2 可知,像元投影点在  $\mathbf{Me}_1$  上的距离即为方阵  $\mathbf{E}$  行列式的绝对值,即  $|\det(\mathbf{E})|$ 。故投影在  $\mathbf{Me}_1$  向量端点的像元能使  $|\det(\mathbf{E})|$  最大,进而使式(4)的体积最大,因此也最有可能是端元向量  $\mathbf{e}_1$ 。由此可知,对于方阵  $\mathbf{E}$  中不变的部分  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p\}$ , FCA 只计算了一次,相比于 N-FINDR 的  $N - 1$  次,计算复杂度大大降低。FCA 可重复确定  $\mathbf{e}_1$  的过程来确定

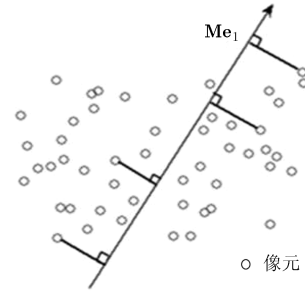


图 2 图像像元在  $\mathbf{Me}_1$  上的投影(图中圆圈“o”代表像元)

$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p$ 。可见, FCA 虽本质是单形体体积最大的方法,但在实现过程中又应用了投影类方法的思想,故兼具两类方法优点,既保证了端元提取速率又保证了精度。

此外, FCA 的终止条件和 N-FINDR 相同,即:当两次迭代得到的最大体积相同时,算法结束。这种终止条件保证了在前一次迭代中确定的端元,在下次迭代过程中都有可能被其他纯度更高的像元替代,避免了如 SGA 和 MVHT 算法在确定一个端元后不可替换的缺点,保证了一直从整体角度考虑问题。

### 3.2 FCA 算法步骤

步骤 1 根据虚拟维度(Virtual Dimensionality, VD)<sup>[13]</sup>计算出图像中的端元个数  $p$ 。

步骤 2 初始化方阵  $\mathbf{E}$

(1)用 PCA 等降维方法将原图像光谱维降到  $p - 1$  维;

(2)任取降维后图像中  $p$  个像元的光谱作为  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p$ ,代入式(3)中矩阵  $\mathbf{E}$  的第 2 行至第  $p$  行,同时矩阵  $\mathbf{E}$  的第 1 行置 1,此时矩阵  $\mathbf{E}$  为  $p$  阶方阵;

(3)设  $V_{\max} = 0$ ; 迭代次数  $it = 0$ 。

步骤 3 提取端元

(1)方阵  $\mathbf{E}$  去掉第  $j$  列,即第  $j$  个端元光谱向量,得到矩阵  $\mathbf{Ma}$ ,  $j = 1, 2, \dots, p$ ;

(2)去掉矩阵  $\mathbf{Ma}$  的第  $i$  行,即可得到原方阵  $\mathbf{E}$  第  $i$  行第  $j$  列的元素所对应的代数余子式的方阵,设该方阵为  $\mathbf{Mb}$ , 阶数为  $p - 1$ ,  $i = 1, 2, \dots, p$ ;

(3)计算第  $i$  行第  $j$  列元素所对应的代数余子式:  $(-1)^{(i+j)} \times \det(\mathbf{Mb})$ ;

(4)分别计算出第  $j$  列每个元素所对应的代数余子式的值,得到一个含有  $p$  个代数余子式值的行向量;

(5)用该代数余子式的行向量乘以像元矩阵,即可得到  $N$  个  $p$  阶方阵行列式的值。 $N$  为图像像元个数;

(6)用  $N$  个行列式中绝对值最大的值对应的像

元向量代替矩阵  $\mathbf{E}$  的第  $j$  列，直到  $j = p$  时结束；

(7)根据式(4)计算出此时的体积  $V(\mathbf{E})$ ；

(8)如果  $V(\mathbf{E}) > V_{\max}$ ，则令  $V_{\max} = V(\mathbf{E})$ ， $it = it + 1$ ；否则退出循环，算法结束。

### 4 实验结果与讨论

#### 4.1 仿真数据实验

仿真数据由已知特定的 5 种端元光谱和相应的丰度矩阵构成。端元光谱均从 USGS 光谱库中得到，分别为：AlumiteGDS83, CalciteWS272, Desert\_VarnishGDS141, KaoliniteCM9, NontroniteGDS41, 其波段数为 210。丰度矩阵由 Dirichlet<sup>[5]</sup>分布构成，满足丰度非负与和为 1 约束<sup>[14]</sup>。将端元矩阵与丰度矩阵相乘，并加上不同强度的高斯白噪声，即得到了本实验所用的仿真数据。

试验中参与比较的 5 种算法为：N-FINDR, SGA, MVHT, NVFS, FCA, 均为基于单形体体积最大的端元提取算法。在端元提取的基础上，利用全约束最小二乘法(Full Constraint Least Square, FCLS)<sup>[15]</sup>进行丰度估计，最后，通过分析算法执行时间和端元提取以及丰度估计的精度综合评价 5 种算法的性能。

#### 实验 1 算法复杂度分析

算法的浮点操作次数(Floating-point operations, Flops)常作为衡量其复杂度的标准之一。5 种算法理论上的时间复杂度如下：SGA 复杂度为  $N \times \sum_{n=2}^p n^{\eta[9]}$ ，MVHT 复杂度为  $2N \times (\sum_{n=2}^{p-1} 4p - 4n + 3) = N \times (4p^2 - 6p - 4)^{[10]}$ ，NVFS 复杂度为  $\sigma N \times (p^{\eta+1} + p^3 L - 2p^2 L + pL)^{[11]}$ ，N-FINDR 复杂度为  $\sigma p^{\eta+1} N^{[8]}$ ，而 FCA 的复杂度为  $\sigma p(p-1)N$ 。5 种算法中只有 NVFS 不需要降维，对于其他 4 种算法本文使用 PCA 降维，PCA 复杂度为  $(L^2 N / 2) + (p-1)LN$ 。以上公式中  $2.3 \leq \eta \leq 2.9$ ， $3 \leq \sigma \leq 5$ ， $L$  是波段数， $N$  是像元个数， $p$  是端元数。

如图3(a)和图3(b)，当像元数  $N$  和端元数  $p$  分别增加时，NVFS复杂度最高，FCA比N-FINDR复杂度低了很多，特别是当  $p$  逐渐增加时，N-FINDR和 FCA的算法复杂度差距会进一步加大。此外，在  $p$  和  $N$  增加的两种情况下，FCA的算法复杂度几乎和 MVHT相同，都是最低的。但是实验2和实验3将证明，在实际运算中FCA所用时间要短于MVHT，同时由于MVHT采用的贪婪搜索算法更容易陷入局部极值，而FCA采用的是随机搜索算法，因而FCA往往可以获得较高的端元提取精度。

#### 实验 2 算法时间复杂度实验

该组实验分别改变了端元个数和像元个数，以测试分析 5 种算法运算时间。时间测试的硬件环境为：Intel(R)Core(TM)i5-3470, CPU3.20 GHz, Windows 7, MATLAB R2010a。

(1)改变仿真数据端元个数，以测试不同算法的时间复杂度与端元个数的关系。端元个数  $p$  从 2 增加到 10，均是从 USGS 矿物光谱库中得到，像元个数  $N$  为  $500^2$ ，SNR 为 20 dB。

(2)改变仿真数据像元个数，以测试不同算法时间复杂度与像元个数的关系。像元个数  $N$  为  $100^2, 200^2, 300^2, 400^2, 500^2, 600^2, 700^2$ ，端元个数  $p$  为 8，均是从 USGS 矿物光谱库中得到，SNR 为 20 dB。

从图 4(a)和图 4(b)中可看出随着端元个数和像元个数增加，5 种算法的运行时间都在增加，其中 FCA 时间最短，几乎是 N-FINDR 的百分之一。可见，应用投影思想的 FCA 的时间复杂度要大大优于其他 4 种算法。值得注意的是，实验 1 中 MVHT 与 FCA 复杂度几乎相同，而本实验中 FCA 运行时间却短于 MVHT，这主要是因为应用 MATLAB 实现这两种算法时，MVHT 需要用到较多的循环语句实现，而 FCA 的投影过程使用矩阵运算就可实现。在 MATLAB 软件中，循环语句的解释时间要长于矩阵运算，故 FCA 在实际中运算速度快于 MVHT。

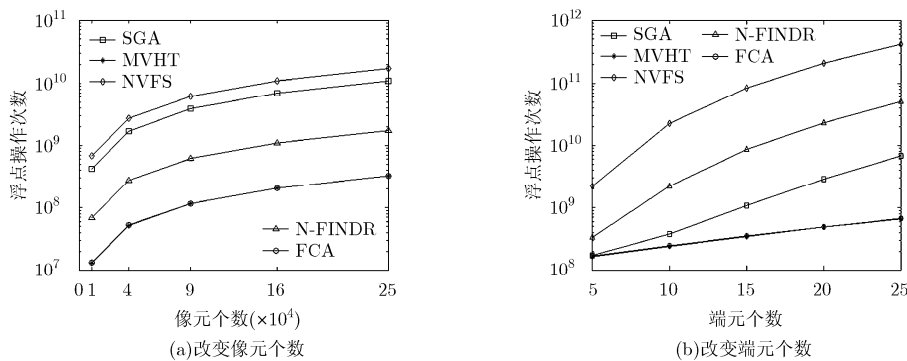


图 3 各种算法的运算复杂度

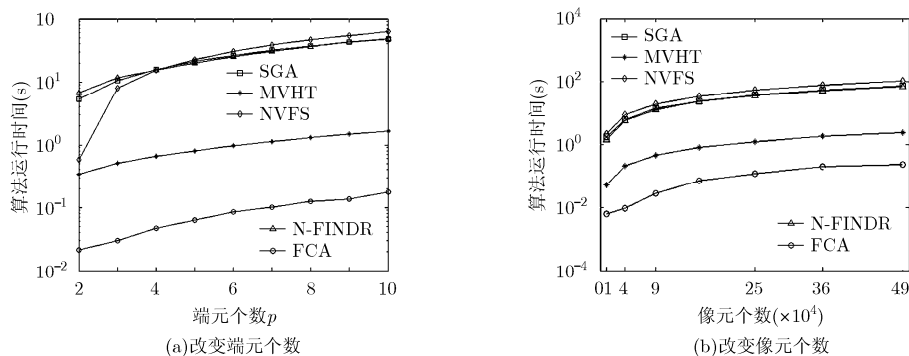


图 4 不同情况下算法运行时间

### 实验 3 算法抗噪声实验

本实验选取光谱信息散度(Spectral Information Divergence, SID)<sup>[16]</sup>和均方根误差(Root Mean Square Error, RMSE)<sup>[3]</sup>两个指标来衡量不同噪声强度下,各算法提取端元的性能。SID 用以衡量提取的端元和原端元的相似程度,取其均值作为性能评价的标准,即  $\overline{SID} = (1/p) \sum_{k=1}^p SID_k$ 。RMSE 用以衡量原图像和反混图像的相似程度,取其均值作为性能评价标准,即  $\overline{RMSE} = (1/N) \sum_{i=1}^N RMSE_i$ 。SID 和 RMSE 的值越低,所比较的两个对象越相似。

实验中,高斯白噪声的强度 SNR 分别为:10 dB, 15 dB, 20 dB, 25 dB, 35 dB, 40 dB。端元个数  $p$  为 5, 像元个数  $N$  为  $200^2$ 。为避免误差,对仿真数据进行 50 次实验,取其均值作为最终结果。

表 1 不同噪声强度下算法 RMSE( $\times 10^{-2}$ )性能比较

算法	SNR(dB)					
	40	35	25	20	15	10
SGA	2.94	2.98	3.63	4.70	7.52	10.26
NVFS	<b>0.36</b>	<b>0.63</b>	1.94	3.32	5.81	10.19
MVHT	0.37	<b>0.63</b>	1.99	3.34	5.75	9.73
N-FINDR	0.37	0.64	2.08	3.30	5.73	10.16
FCA	<b>0.36</b>	<b>0.63</b>	<b>1.92</b>	<b>3.19</b>	<b>5.49</b>	<b>9.57</b>

表 2 不同噪声强度下算法 SID( $\times 10^{-3}$ )性能比较

算法	SNR(dB)					
	40	35	25	20	15	10
SGA	5.92	7.21	28.81	49.20	82.88	123.17
NVFS	0.90	3.18	16.39	32.51	61.74	133.57
MVHT	0.90	3.18	16.45	31.85	61.03	111.28
N-FINDR	<b>0.87</b>	<b>2.94</b>	14.53	<b>30.68</b>	<b>48.50</b>	119.14
FCA	<b>0.87</b>	<b>2.94</b>	<b>14.89</b>	<b>30.68</b>	<b>48.50</b>	<b>102.31</b>

如表 1 和表 2 所示,随着 SNR 降低, SID 和 RMSE 逐渐增加,即 5 种算法的精度都在下降。同时,在不同信噪比条件下, FCA 的 SID 和 RMSE 值均低于 MVHT 和 SGA,与 N-FINDR 相近。说明 SGA 和 MVHT 所采用的逐一确定端元,并且确定后不再更改的方法不能很好地从整体把握单形体体积最大的思想,所选择的点有可能为异常极值点。而以 FCA 为代表的算法,所选择的端元在运算过程中,随时可被能构成更大体积的像元代替,更能从整体考虑特征空间中像元的分布情况,更具有灵活性,因此算法精度也更好。

### 4.2 实际数据实验

本节使用的数据是美国 Nevada 州南部沙漠地区的 Cuprite 数据。它是由机载可见光及红外成像光谱仪(Airborne Visible/ Infrared Imaging Spectrometer, AVIRIS)拍摄于 1997 年 6 月 19 日。该数据共有 224 个波段,波长范围为 370 nm 到 2500 nm,光谱分辨率为 10 nm。截取原始数据中一块大小为  $250 \times 191$  比较典型的数据用于算法实现<sup>[6]</sup>。其中波段 1~2, 104~113, 148~167, 221~224 共 36 个波段因为信噪比太低或为水吸收波段而被移除,余下的 188 个波段用于算法验证。虚拟维度 VD 估计的端元个数及相应的虚警率  $p_f$  如表 3 所示,虚警率  $p_f$  越低,被错判为纯像元的可能性越低。但依据该地区地物的实际分布<sup>[17]</sup>,和相关文献资料<sup>[10,18]</sup>,以及基于不遗漏端元的考虑,本实验将端元个数设定为 18。分别用 SGA, MVHT, NVFS, N-FINDR, FCA 这 5 种算法提取出该图像的 18 个端元进行丰度反演并重新混合图像,图 5 为用 FCA 提取的该地区几种典型矿物端元的丰度反演结果。利用丰度反演的结果初步确定相应地物,并与 USGS 矿物光谱库比较,结果如表 4。表中数值表示光谱相似度,是光谱角度匹配(Spectral Angle Mapper, SAM)和光谱特征拟合(Spectral Feature Fitting, SFF)的加

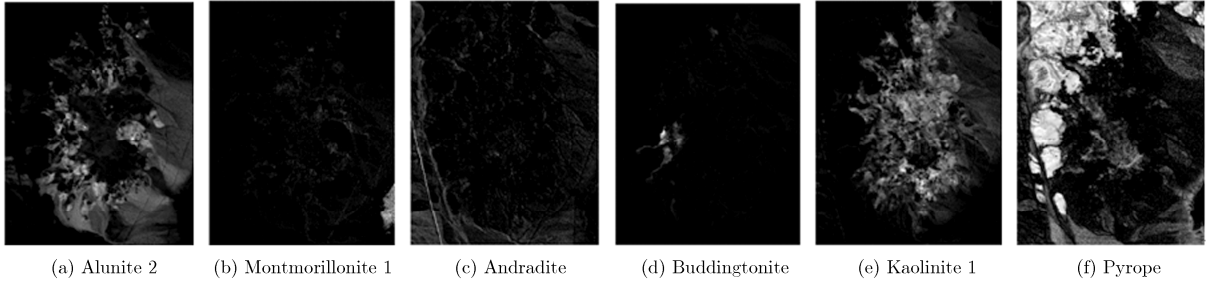


图 5 FCA 提取的几种典型矿物端元的丰度反演结果

权平均值，权重分别为 0.6 和 0.4。此数值越接近 1，说明两个光谱曲线越相似。表 4 的最后给出了 5 种算法求得的光谱相似度的均方根。

表 3 虚拟维度 VD 估计的端元个数及相应的虚警率

估计的端元数目 $p$	14	16	18	20	23
虚警率 $p_f$	$10^{-6}$	$10^{-5}$	$10^{-4}$	$10^{-3}$	$10^{-2}$

表 4 不同算法提取出的端元光谱与 USGS 光谱库中光谱的相似度比较

光谱库中矿物光谱	FCA	MVHT	NVFS	SGA	N-FINDR
Alunite 1	<b>0.801</b>	<b>0.801</b>	<b>0.801</b>	<b>0.801</b>	<b>0.801</b>
Muscovite 1	<b>0.813</b>	/	<b>0.813</b>	/	0.811
Chalcedony	0.561	0.602	0.543	0.614	<b>0.637</b>
Nontronite 1	<b>0.785</b>	0.760	0.784	0.784	0.773
Montmorillonite 1	0.874	0.860	<b>0.887</b>	/	0.868
Andradite	<b>0.721</b>	<b>0.721</b>	<b>0.721</b>	<b>0.721</b>	<b>0.721</b>
Nontronite 2	<b>0.806</b>	0.801	0.786	0.801	0.802
Nontronite 3	0.846	0.843	0.847	0.837	<b>0.848</b>
Muscovite 2	<b>0.824</b>	0.754	/	0.795	0.804
Buddingtonite	0.857	0.856	0.856	<b>0.869</b>	/
Montmorillonite 2	<b>0.859</b>	0.839	0.815	/	0.840
Kaolinite 1	<b>0.828</b>	0.816	0.807	0.807	0.817
Kaolinite 2	<b>0.815</b>	/	0.765	0.814	0.783
Dumortierite 1	0.606	0.632	0.632	<b>0.654</b>	0.603
Pyrope	<b>0.858</b>	0.849	/	0.801	0.849
Quartz	<b>0.804</b>	0.785	0.802	0.793	/
Alunite 2	0.731	0.743	<b>0.761</b>	0.701	0.722
Dumortierite 2	0.571	0.536	0.573	0.522	<b>0.602</b>
均方根	<b>0.782</b>	0.768	0.768	0.760	0.772

表 5 算法运行时间(s)

算法	FCA	NVFS	MVHT	SGA	N-FINDR
时间	<b>0.57</b>	59.82	1.43	46.69	52.63

表 4 中的数据表明用 FCA 算法提取出的端元光谱与 USGS 光谱库中光谱较为相似，其结果优于其他几种算法。同时表 5 列出的算法运行时间表明，FCA 算法运行速度最快，是 N-FINDR 的百倍多。综上，实际数据表明，本论文提出的 FCA 算法相对于其他相似算法而言，具有较好的端元提取精度和较快的算法运行时间，可以快速准确地实现高光谱图像端元提取。

## 5 结束语

本文提出一种基于代数余子式的 N-FINDR 快速端元提取算法，简称 FCA。该算法通过寻找在特征空间中最大体积的单形体来提取端元。在计算过程中，FCA 可得到一个代数余子式构成的向量，并能将所有像元投影到这一向量上，从而将复杂度极高的行列式计算转化为复杂度较低的内积(投影)运算，进而得到具有最大体积的单形体。由于 FCA 将循环迭代计算体积的方法改进为计算投影，故 FCA 具有投影类端元提取算法计算复杂度低的优势。同时，在端元搜索方面，不同于 SGA 和 MVHT 所采用的逐一确定端元并且端元不可更改的方法，FCA 所确定的端元都可在下一次迭代过程中被纯度更高的像元替换，故降低了提取出异常极值点的概率，并且保证所确定的端元都是特征空间中由全部像元构成的最大体积的单形体顶点。这也是 FCA 算法保持良好精度的原因。综上，经过仿真和真实实验验证，FCA 算法可以在计算复杂度很低的情况下保持很好的端元提取精度，这对于大数据量的高光谱图像端元提取的快速处理很有意义。

## 参考文献

- [1] Sun Kang, Geng Xiu-rui, Wang Pan-shi, *et al.*. A fast endmember extraction algorithm based on gram determinant[J]. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(6): 1124-1128.
- [2] Ji Lu-yan, Geng Xiu-rui, Yu Kai, *et al.*. A new non-negative matrix factorization method based on barycentric

- coordinates for endmember extraction in hyperspectral remote sensing[J]. *International Journal of Remote Sensing*, 2013, 34(19): 6577–6586.
- [3] Deng Cheng-bin and Wu Chang-shan. A spatially adaptive spectral mixture analysis for mapping subpixel urban impervious surface distribution[J]. *Remote Sensing of Environment*, 2013, 133(1): 62–70.
- [4] Beauchemin M. A method based on spatial and spectral information to reduce the solution space in endmember extraction algorithms[J]. *Remote Sensing Letters*, 2014, 5(5): 471–480.
- [5] Boardman J, Kruse F, and Green R. Mapping target signatures via partial unmixing of AVIRIS data[C]. Proceedings of the Summaries of the VI Jet Propulsion Laboratory Airborne Earth Science Workshop, Pasadena, USA, 1995: 23–26.
- [6] Nascimento P and Bioucas D. Vertex component analysis: a fast algorithm to unmix hyperspectral data[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, 43(4): 898–910.
- [7] Chang Chein-I. Hyperspectral Imaging: Techniques for Spectral Detection and Classification[M]. New York: Kluwer Academic/ Plenum Publishers, 2003: 73–88.
- [8] Winter M. N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data[C]. Proceedings of the SPIE Imaging Spectrometry V, San Diego, USA, 1999: 266–275.
- [9] Chang Chein-I, Wu Chao-cheng, Liu Wei-min, *et al.* A new growing method for simplex-based endmember extraction algorithm[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44(10): 2804–2819.
- [10] Liu Jun-min and Zhang Jiang-she. A new maximum simplex volume method based on householder transformation for endmember extraction[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(1): 104–118.
- [11] Geng Xiu-rui, Zhao Yong-chao, Wang Fu-xiang, *et al.* A new volume formula for a simplex and its application to endmember extraction for hyperspectral image analysis[J]. *International Journal of Remote Sensing*, 2010, 31(4): 1027–1035.
- [12] Shi Chen and Wang Le. Incorporation spatial information in spectral unmixing: a review[J]. *Remote Sensing of Environment*, 2014, 149(2): 70–87.
- [13] Chang Chein-I and Du Qian. Estimation of number of spectrally distinct signal sources in hyperspectral imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, 42(3): 608–619.
- [14] Zhang Xian-da. Matrix Analysis and Applications[M]. Beijing: Tsinghua University Press, 2004: 70–81.
- [15] Heinz D and Chang Chein-I. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2001, 39(3): 529–545.
- [16] Chang Chein-I. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis[J]. *IEEE Transactions on Information Theory*, 2000, 46(5): 1927–1932.
- [17] Roger N, Clark R, Gregg A, *et al.* Evolution in imaging spectroscopy analysis and sensor signal-to-noise: an examination of how far we have come[C]. Proceedings of the Summaries of the 6th Annual JPL Airborne Geoscience Workshop, Pasadena, USA, 1996: 4–8.
- [18] Geng Xiu-rui, Xiao Zheng-qing, Ji Lu-yan, *et al.* A Gaussian elimination based fast endmember extraction algorithm for hyperspectral imagery[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2013, 7(9): 211–218.
- 李琳: 女, 1989年生, 博士生, 研究方向为高/多光谱数据优化和混合像元分析.
- 孟令博: 女, 1989年生, 博士生, 研究方向为高光谱遥感图像处理和水光谱学.
- 孙康: 男, 1988年生, 博士生, 研究方向为高光谱特征提取与高性能计算.
- 赵永超: 男, 1971年生, 研究员, 研究方向为高光谱遥感图像去噪、大气纠正和辐射定标.