

doi: 10.3788/gzxb20154410.1011002

一种改进的快速 N-FINDR 端元提取算法

赵春晖, 郭蕴霆

(哈尔滨工程大学 信息与通信工程学院, 哈尔滨 150001)

摘 要:为了解决传统 N-FINDR 算法计算量大,提取结果对噪声和初始端元选取敏感,且容易将异常点作为端元而造成误提取的问题,提出一种改进的快速 N-FINDR 端元提取算法.该方法通过光谱距离提取并去除高光谱图像中的冗余信息,减少 N-FINDR 提取端元的搜索范围,平滑噪声影响,并自适应剔除异常点,通过最大化光谱距离选取 N-FINDR 的初始端元,避免了随机选择的盲目性.采用合成数据和真实高光谱数据进行仿真分析并与现有算法进行对比,结果表明,本文算法能在噪声与奇异点干扰下正确提取端元,其提取效率和鲁棒性均优于现有算法.

关键词:遥感; 端元提取; N-FINDR; 高光谱图像; 光谱距离; 快速算法; 抗噪声

中图分类号: TP751.1

文献标识码: A

文章编号: 1004-4213(2015)10-1011002-9

An Improved Fast N-FINDR Endmember Extraction Algorithm

ZHAO Chun-hui, GUO Yun-ting

(College of Information and Communication Engineering, Harbin Engineering University,
Heilongjiang 150001, China)

Abstract: Traditional N-FINDR algorithm suffers from complicated calculations, and is sensitive to noise and initial endmembers, resulting in wrong extraction by outliers. In order to solve these problems, an improved fast N-FINDR endmember extraction algorithm was proposed, which gets rid of the redundant information and reduces the search area of commonly used N-FINDR algorithm by spectral distance, then smoothes the spectral noise and gets rid of the outliers adaptively. In addition, the initial endmembers of N-FINDR are selected by maximizing the spectral distance, avoiding the blindness of random selection. The synthetic data and real hyperspectral data were used for simulation analysis, and the proposed method was compared with existing algorithms. The experiment results show that the proposed method is able to extract the endmembers correctly in the noise, and has higher extraction efficiency and robustness than existing algorithms.

Key words: Remote sensing; Endmember extraction; N-FINDR; Hyperspectral image; Fast algorithm; Spectral distance measure; Anti-noise

OCIS Codes: 110.4234; 100.4145; 070.4790

0 引言

随着成像技术和图像处理技术的发展,高光谱图像因其极高的光谱分辨率得到广泛应用,例如地物类别分析.然而,由于高光谱传感器较低的空间分辨率,混合像元在高光谱图像中的普遍存在,增加了对高光谱图像进行精确地物分析和定位的难度.为了解决这一问题,端元提取技术被提出.在高光谱图像中,理想

化的纯净地物光谱数据被定义为“端元”^[1],即高光谱图像中可以表征某一类纯净地物的光谱数据.提取一幅高光谱图像中端元的技术被称为端元提取技术,它是对高光谱图像进行进一步分析的前提,其有效与否直接关系到高光谱图像信息获取的成败.因此,对端元提取算法的研究一直是高光谱图像处理领域的研究热点之一.目前,已有多种有效的端元提取算法被相继提出,例如 N-FINDR^[2-3]、顶点成分分析(Vertex

基金项目:国家自然科学基金(No. 61405041),黑龙江省自然科学基金重点资助项目(No. ZD201216)和哈尔滨市优秀学科带头人基金(No. RC2013XK009003)资助

第一作者:赵春晖(1965-),男,教授,博士,主要研究方向为高光谱图像处理. Email: zhaochunhui@hrbeu.edu.cn

收稿日期:2015-04-02; **录用日期:**2015-06-08

<http://www.photon.ac.cn>

Component Analysis, VCA)^[4-5]、单体生长算法 (Simplex Growing Algorithm, SGA)^[6-7]、正交子空间投影 (Orthogonal Subspace Projection, OSP)^[8]、像素纯度索引 (Pure Pixel Index, PPI)^[9-10] 和自动形态学端元提取 (Automated Morphological Endmember Extraction, AMEE)^[11-12] 等。其中 N-FINDR-基于高光光谱图像中的像元在高维空间中的几何结构,即所有像元在高维空间构成一凸体点集,通过计算凸体体积的最大来提取端元,因其需求参数少,提取效果好而受到广泛关注。然而,该算法的搜索空间通常包括整幅图像中的所有像元,在多次迭代以求解最大体积的过程中必然出现重复计算,影响算法效率^[13];此外,凸体体积的计算易受到噪声以及异常点干扰,易造成端元的误提取。

近年来,已有不少文献针对 N-FINDR 提出各种改进,文献[14]利用卡方分布分离出较少的候选端元集合,减少算法的计算量;文献[15]利用距离计算代替体积计算,并用 PPI 选择初始端元集以减少收敛的迭代次数;文献[16]采用遍历过程是否存在顶点像素更替作为停机准则,并引入支持向量机,以增加计算量为代价优化端元提取结果。然而,这些方法都未能解决噪声和奇异点对端元提取的消极影响。本文在深入分析高光光谱图像的数据特征以及 N-FINDR 算法原理的基础上,提出一种改进的 N-FINDR 端元提取算法。利用光谱距离及像元光谱特征减少候选端元集个数,消除噪声及奇异点干扰,并通过最大化光谱距离选取 N-FINDR 初始端元,提高算法运算效率。

1 N-FINDR 端元提取算法

给定一幅高光光谱图像,其所有的像元光谱在高维空间构成凸体点集,纯像元(即端元)位于凸体的顶点,混合像元(非端元)依据线性光谱混合模型位于凸体表面或凸体内部。在该凸体点集中,由凸体顶点(端元)组成的凸体体积最大,因此,端元提取在 N-FINDR 中被转化为求取凸体顶点的问题。假设 e_i 为第 i 个像元光谱, n 个像元构成像元矩阵 $\mathbf{A}_n = (e_1, e_2, \dots, e_n)$,它们张成的单形体体积为

$$V(\mathbf{A}_n) = \frac{1}{(n-1)!} \text{abs}(|\mathbf{E}|) \quad (1)$$

$$\mathbf{E} = \begin{bmatrix} 1 \\ \mathbf{A}_n \end{bmatrix} \quad (2)$$

式中 V 表示这 n 个像元构成的体积, $\text{abs}(\cdot)$ 为绝对值运算, $|\cdot|$ 为行列式运算。

由于体积公式中存在行列式运算,即要求 \mathbf{E} 为方阵,像元矩阵 \mathbf{A}_n 的大小被限制为 $(n-1) \times n$ 。因此,原高光光谱图像的波段数必须降维至 $(n-1)$ 维。这不仅对降维算法提出了较高的要求,也在一定程度上降低了

端元提取的精度。因此,文献[17]提出了另一种凸体体积计算公式来代替,即

$$V(\mathbf{A}_n) = \frac{1}{n!} \sqrt{|\mathbf{A}_n^T \mathbf{A}_n|} \quad (3)$$

式中, $\mathbf{A}_n^T \mathbf{A}_n$ 为方阵,故使用式(3)计算凸体体积时无需对高光光谱数据进行降维。

N-FINDR 算法通过遍历所有构成凸体的像元组合,将 \mathbf{A}_n 中的像元迭代替换为能使凸体体积最大的新像元,以完成端元提取过程。具体迭代步骤为

1) 随机选取 n 个像元,利用式(3)计算凸体体积 $V(e_1, e_2, \dots, e_n)$ 。

2) 从高光光谱图像中选取任意一个像元 r ,分别计算凸体体积 $V(r, e_2, \dots, e_n)$ 、 $V(e_1, r, \dots, e_n)$ 、 \dots 、 $V(e_1, e_2, \dots, r)$,若这 n 个体积的最大值大于原体积 $V(e_1, e_2, \dots, e_n)$,则将相应的 e_i 替换为 r ,其他像元不变;否则更换像元 r 。

3) 重复过程 2),直到不再发生体积更迭,此时的像元 $e_i (i=1, 2, \dots, n)$ 即为所提取的端元。

2 改进的 N-FINDR 算法

2.1 光谱距离

对端元提取而言,所提取端元的准确性以及覆盖性至关重要,它应尽可能包括高光光谱数据中存在的所有端元并尽可能接近理想状态下地物的真实光谱。然而,在实际观测的高光谱数据中,由于光照,大气,噪声等客观条件的影响,光谱数据会发生不同程度的失真,该失真程度通常通过光谱角 (Spectral Angle Mapping, SAM) 和欧式距离 (Euclidean Distance, ED) 来衡量,其定义分别为

$$\text{SAM}(\mathbf{x}, \mathbf{y}) = \cos^{-1} \left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \right) \quad (4)$$

$$\text{ED}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_p \quad (5)$$

式中, $\langle \mathbf{x}, \mathbf{y} \rangle$ 是两个向量 \mathbf{x} 和 \mathbf{y} 的内积, $\|\cdot\|_p$ 表示向量的 p -范数。

通常, SAM 计算的是光谱向量间的夹角,表征地物的类别差异; ED 计算的是光谱在幅值上的距离,表征的是地物对不同波段反射强度的差异。但需要注意的是,同种地物的光谱夹角为 0,但在某些情况下可能存在较大的幅值差异。因此,加权求和的办法并不适用于计算光谱间的距离。本文综合 SAM 和 ED 的特点,以乘积形式生成光谱距离 (Spectral Distance, SD) 来衡量光谱间差异,即

$$\text{SD}(\mathbf{x}, \mathbf{y}) = \text{SAM}(\mathbf{x}, \mathbf{y}) \cdot \text{ED}(\mathbf{x}, \mathbf{y})^\lambda \quad (6)$$

式中, λ 是一个常数,用于控制 SAM 和 ED 的权重,为了均衡考虑光谱角和欧氏距离在光谱距离中的权重,设置 $\lambda=1$ 。

2.2 基于光谱距离的候选端元选取

一般情况下,受客观条件影响导致的光谱失真,其

光谱距离会控制在一定的阈值范围内,且该失真范围内的光谱可以被正确匹配和识别,这也是对端元提取算法所提取的端元的要求.然而,传统的 N-FINDR 算法极端追求凸体体积最大,并未考虑噪声以及奇异点的影响,这会影响到端元提取的正确性.此外,在高光谱图像中,存在大量的光谱冗余信息,传统 N-FINDR 算法并未对这些信息进行有效去除,不仅造成算法的效率低下,也影响了算法的鲁棒性.设一幅由 3 个端元构成的高光谱图像,其二维空间点集示意图如图 1.

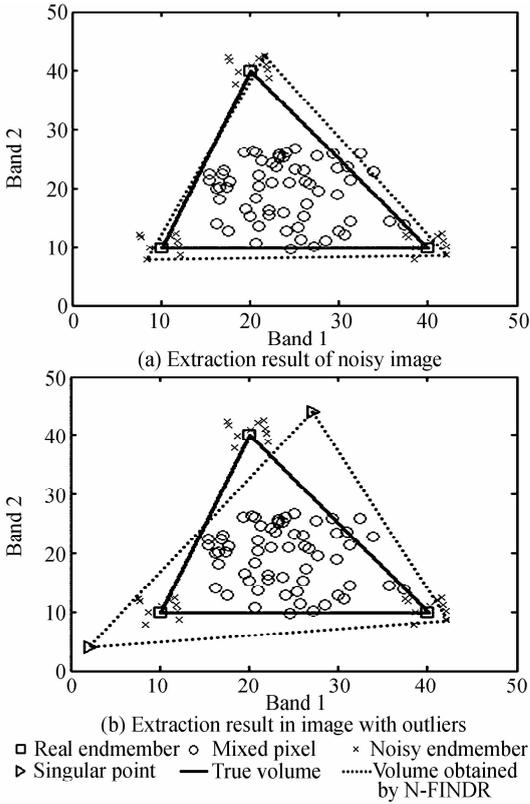


图 1 N-FINDR 算法端元提取示意图

Fig. 1 Diagram of endmember extraction via N-FINDR

在图 1 中可以看出,由于噪声的影响,高光谱数据中的混合像元和纯像元(端元)都偏离了原有的空间位置.基于求解最大凸体体积的 N-FINDR 算法所提取的端元,往往是被噪声影响最明显、与真实端元偏离最大的像元,奇异值存在的情况下,提取的端元完全与真实端元无关,这不符合端元提取的基本要求.此外,混合像元与纯像元在空间内表现出较高的聚集性,N-FINDR 算法对此类在高维空间内临近的像元点进行大量的重复计算,严重影响了算法的效率.

因此,本文提出基于光谱距离 SD 的候选端元选取算法,算法步骤为:

- 1) 给定高光谱数据 $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$, N 为高光谱数据所含有的像元个数.
- 2) 选取 \mathbf{X} 中的一个像元 x_i , 计算 \mathbf{X} 中的其他像元与像元 x_i 的光谱距离 SD, 提取 SD 小于阈值 t_{SD} 的像元, 标记为 P_i , 即

$$P_i = \{x | SD(x, x_i) < t_{SD}, x_i \in \mathbf{X}\} \quad (7)$$

3) 若像元集 P_i 中的像元个数小于 t_{NUM} , 则认为 P_i 中的端元为奇异点, 将其从 \mathbf{X} 中删除; 反之, 将这些端元光谱取平均(标记为 e)以平滑噪声或失真, 并将 e 选取为候选端元加入候选端元集(标记为 \mathbf{E}).

$$e = \text{avg}(P_i) \quad (8)$$

$$\begin{cases} \mathbf{E} = \mathbf{E} \oplus e & \text{num}(P_i) > t_{NUM} \\ \mathbf{E} = \mathbf{E} & \text{num}(P_i) \leq t_{NUM} \end{cases} \quad (9)$$

$$\mathbf{X} = \mathbf{X} \ominus P_i \quad (10)$$

式中, 运算符 \oplus 、 \ominus 分别定义为从集合中添加、删除元素, $\text{avg}(\cdot)$ 表示计算集合元素的平均值, $\text{num}(\cdot)$ 表示元素个数统计.

4) 重复步骤 2), 3), 直至 \mathbf{X} 为空, 此时 \mathbf{E} 中像元为候选端元.

注意, 可以对提取的候选端元集 \mathbf{E} 进行若干次迭代, 以获取更稳定的结果. 经过上述步骤, 候选端元集 \mathbf{E} 中的像素具有特点:

- 1) 大量冗余的同种地物像元被去除, 生成具有代表性的地物光谱, 像元个数相比原有的高光谱数据大大减少, 显著提高后续 N-FINDR 算法的运算效率.
- 2) 包括奇异点在内的异常端元被剔除, 保证了 N-FINDR 提取结果的鲁棒性.
- 3) 由于大量的同类端元被加权平均, 不同的光照条件导致的失真以及噪声影响被平滑, 保证了提取端元的纯净性.

经过本文算法处理后的简化空间点集示意图如图 2.

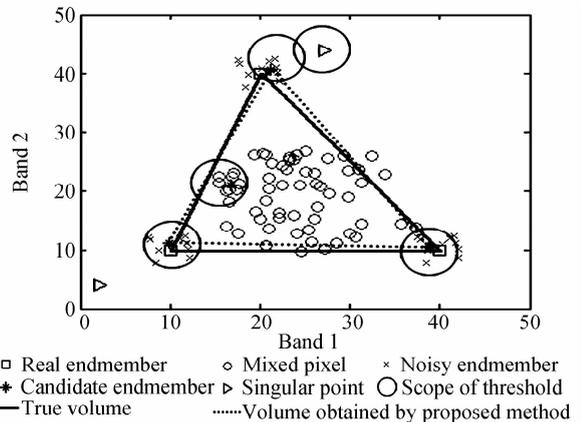


图 2 本文算法端元提取示意图

Fig. 2 Diagram of endmember extraction via proposed method

从图 2 中可以看出, 对高光谱数据中的每个像元, 给定的光谱距离阈值 t_{SD} 可以限定一个空间范围, 该范围中的所有像元被它们取平均后的候选端元点所代替, 减少了待提取端元的像元个数, 使得被噪声污染的像元恢复到其原有的空间位置; 若该范围中的像元点的个数小于 t_{NUM} (如图 2 中的奇异点), 则认为这些像元为奇异点, 直接删除该部分像元以排除奇异点干扰.

因此,该算法能较好地平滑噪声并去除奇异点,正确有效地提取端元.

2.3 基于光谱距离最大化的端元初始化算法

在 N-FINDR 算法中,端元的初始化是极为重要的步骤之一,初始端元的随机选取容易造成提取结果的不一致,使得算法无法收敛至全局最优解.此外,初始端元的随机选取还会造成迭代次数的增加,增加算法的处理时间. N-FINDR 的原理是寻找使凸体体积最大的端元,高维空间中的体积最大可以简化为端元间的光谱距离之和的最大,基于该原理和光谱距离 SD,提出改进的 N-FINDR 端元初始化算法,步骤为:

1) 获得候选端元集 E , 并确定高光谱图像中的端元个数 n , 设置初始端元集 E_0 为空.

2) 从 E 中选择光谱能量最小的像元 m , 计算其余像元与 m 的光谱距离, 选择与 m 的光谱距离最大的像元加入初始端元集 E_0 , 即

$$m = \arg \min_x (\|x\|_2), \quad x \in E \quad (11)$$

$$E_0 = E_0 \oplus \arg \max_x (\|x - m\|_2), \quad x \in E \quad (12)$$

3) 依次计算 E 中每个像元 x 与 E_0 中所有像元的光谱距离, 将光谱距离之和最大的像元加入 E_0 , 即

$$E_0 = E_0 \oplus \arg \max_x (\sum_i \|x - e_i\|_2), \quad x \in E, e_i \in E_0 \quad (13)$$

4) 重复步骤 3), 使 E_0 中的像元个数为 n . 此时 E_0 中的像元为 N-FINDR 算法的初始端元.

由于光谱距离的计算复杂度远小于体积计算的复杂度, 且距离计算是凸体体积计算的线性优化, 因此该端元初始化算法能次优地选取使凸体体积最大的端元, 显著提高 N-FINDR 算法的计算效率.

2.4 改进的 N-FINDR 端元提取算法

改进 N-FINDR 端元提取算法流程如图 3. 该算法在候选端元集的选择, 以及 N-FINDR 算法迭代的初始端元选取两个方面减少 N-FINDR 的计算量, 提高了算法的端元提取效率. 并通过阈值过程, 解决了传统 N-FINDR 算法对噪声敏感的问题, 能在各种噪声环境下较为准确的提取端元.

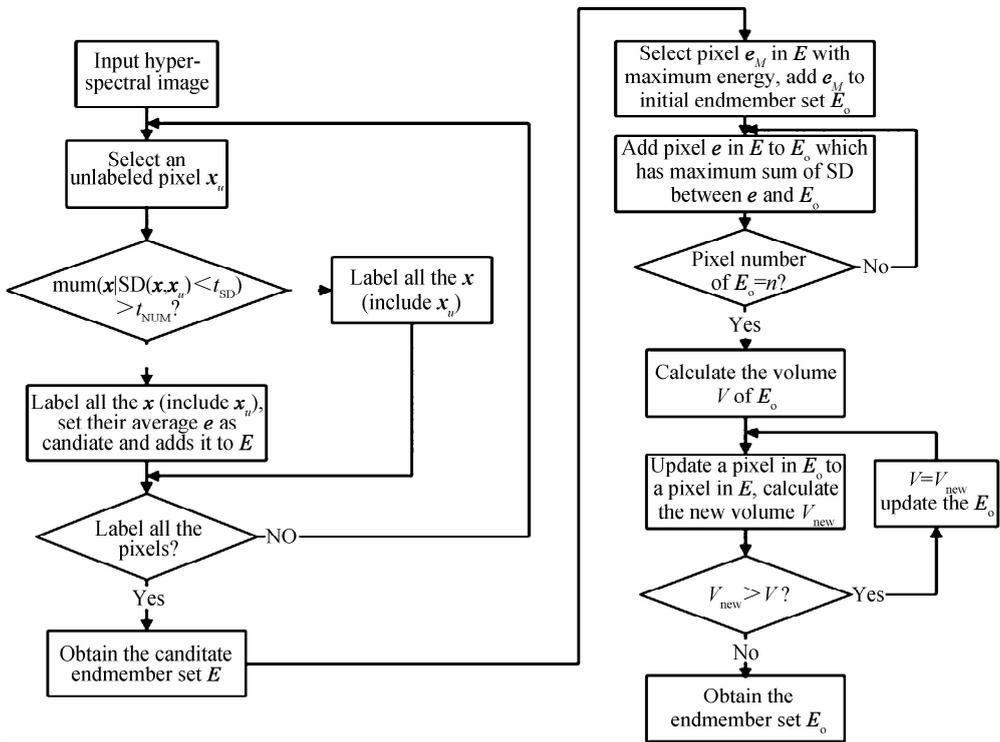


图 3 改进的 N-FINDR 端元提取算法流程图

Fig. 3 Flow chart of the improved N-FINDR endmember extraction algorithm

3 阈值选取

算法最关键的环节在于光谱距离阈值 t_{SD} 的选取, 既要保证尽可能多地将同种地物的光谱选进端元集 P_i , 又要避免混入异种地物像元和奇异点. 此外, 阈值的设置与算法的计算效率息息相关, 过小的阈值会导致算法对噪声的鲁棒性缺失, 无法有效地合并同类地

物光谱而导致处理时间急剧增加, 在阈值为 0 的时候, 本文算法退化成传统 N-FINDR 算法; 而过大的阈值会使得提取的候选端元个数过少, 光谱代表性和纯度不足, 出现混合光谱导致提取端元失败. 因此, 为了更好地确定算法阈值, 通过一组实验测试噪声对光谱距离的影响, 作为设定阈值的依据. 选取美国地质调查局矿物光谱库中的 2 种矿物光谱和 AVIRIS 高光谱数据中

的 2 种植被光谱,将这 4 种地物光谱进行归一化处理后再再次加噪(控制噪声强度,使得信噪比分别为 20 dB、30 dB、40 dB),并对噪声导致的光谱距离 SD 取平均,统计结果见表 1.

表 1 各地物加噪时的平均光谱失真

Table 1 Spectral distortion of different objects in noise

Objects	Mineral 1	Mineral 2	Plant 1	Plant 2
SD(20dB)	9.95×10^{-3}	9.94×10^{-3}	9.94×10^{-3}	9.94×10^{-3}
SD(30dB)	9.98×10^{-4}	9.97×10^{-4}	9.98×10^{-4}	9.96×10^{-4}
SD(40dB)	9.98×10^{-5}	9.98×10^{-5}	9.96×10^{-5}	9.98×10^{-5}

如表 1,同一信噪比下,不同的地物光谱受噪声引起的光谱失真趋于稳定,本文使用的光谱距离 SD 与信噪比近似成线性关系,即在信噪比一定的条件下,噪声引起的光谱距离仅与光谱能量有关,与地物类别无关.遥感获取的高光谱图像数据的信噪比通常在 20~30 dB 之间,由于噪声的随机性较强,多次实验结果表明,较好的距离阈值应设置为相应信噪比下的平均光谱失真的 2~3 倍,精确取值应根据高光谱图像的复杂程度而定.

此外,对于奇异点个数的判别阈值 t_{NUM} ,可设置为常数,也可依据总像元个数设置一个较小的比例.若需要进一步剔除占据比例较小的混合像元,可提高阈值比例,以获取更好的提取效果.

4 仿真实验分析

将本文算法定义为 DN-FINDR,为了验证 DN-FINDR 的有效性,将它与传统 N-FINDR 方法以及文献[13]提出的 MN-FINDR 算法分别在零噪声环境、高斯白噪声环境、高斯白噪声+奇异点干扰环境中进行耗时和端元提取结果的对比.为获得监督评价,采用合成数据进行测算.该合成数据由第 3 节选取的四种矿物光谱合成,图像大小为 160×160 ,有 224 个波段,如图 4(a),图像的四个角为四种纯净的矿物光谱,中心的十字区域为比例渐变的四种矿物的混合像元.实验环

境为 Intel Core i7-2630QM CPU 2.1 GHz, RAM 4 GB.

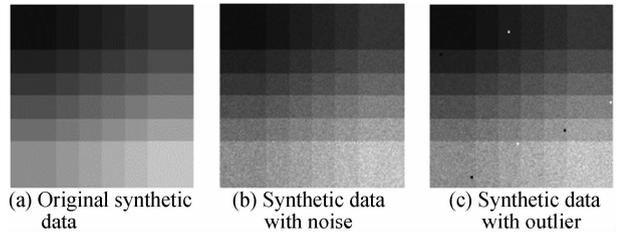


图 4 合成数据

Fig. 4 Synthetic data

4.1 零噪声实验分析

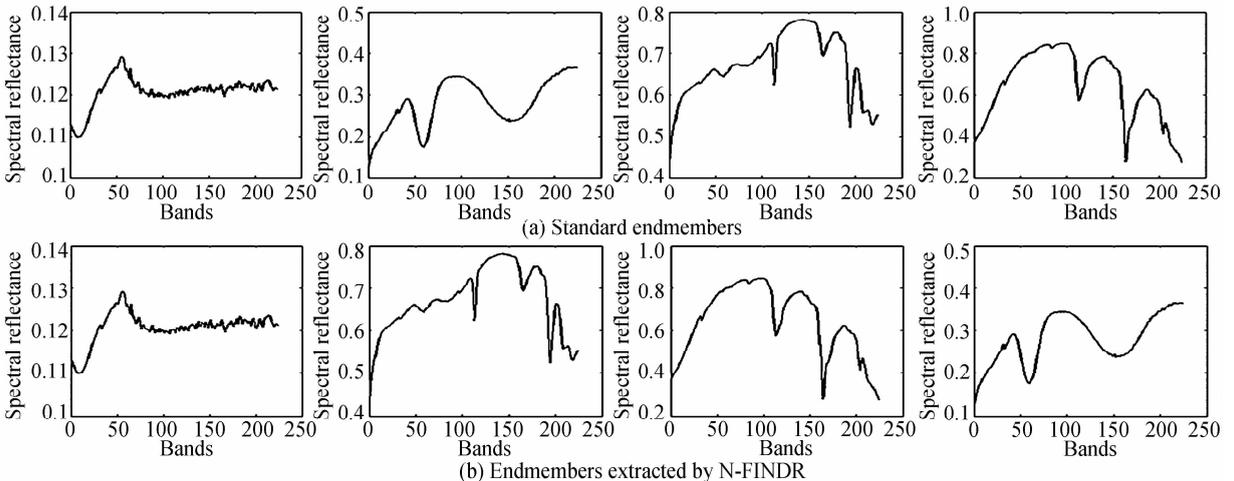
实验中,合成数据被直接用于端元提取,设置光谱距离阈值为 10^{-4} ,端元提取的结果如图 5(忽略光谱排序),耗时对比见表 2.

表 2 算法耗时对比

Table 2 Comparison of the processing time

Algorithm	Time of pre-treatment/s	Number of candidates	Extraction time/s	Total time/s
N-FINDR	-	25600	431.44	431.44
MN-FINDR	5.41	1600	28.46	33.87
DN-FINDR	13.78	34	0.56	14.34

图 5 中,三种方法提取的端元光谱与标准光谱完全重合,说明在零噪声条件下,三种算法都可以准确提取高光谱图像中的端元.然而,从表 2 中可以看出,经过 MN-FINDR 和 DN-FINDR 处理后,候选端元的个数从 25600 个分别降至 1600 个和 34 个,这是由于两种算法的选取原理不同导致的. MN-FINDR 通过卡方分布阈值化剔除混合像元,保留下的是所有可能为纯像元的像元;而 DN-FINDR 是在光谱距离的阈值范围内保留每一类像元的代表(包括纯像元与混合像元).因此,两种算法提取所得的候选端元个数有较大偏差, DN-FINDR 算法提取的候选端元冗余信息更少,效率更高.



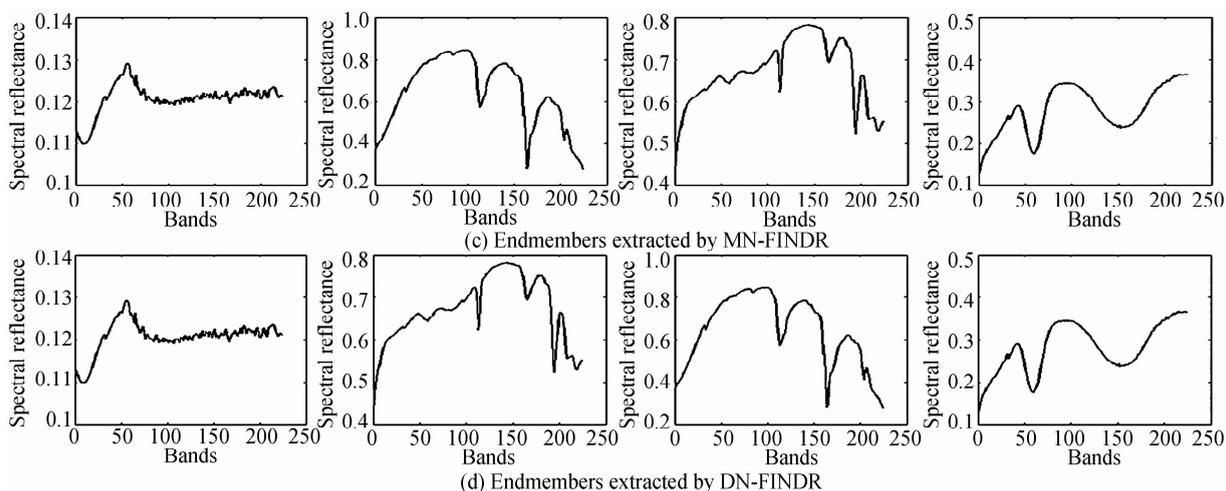


图5 端元提取结果对比

Fig. 5 Comparison of the endmembers extraction results

4.2 高斯白噪声实验分析

将高斯白噪声添加进合成数据,控制噪声强度使得信噪比为 25 dB,含噪高光谱数据如图 4(b),设置光谱距离阈值为 0.01,端元提取的结果如图 6,耗时对比见表 3.

表3 算法耗时对比

Table 3 Comparison of the processing time

Algorithm	Time of pre-treatment/s	Number of candidates	Extraction time/s	Total time/s
N-FINDR	-	25600	420.41	420.41
MN-FINDR	5.24	1600	27.33	32.57
DN-FINDR	10.83	36	0.59	11.42

从图 6 中可以看出,N-FINDR 算法在进行加噪高光谱数据的端元提取时,由于噪声的添加反而可以增加凸体体积,故提取出的端元光谱往往受到噪声的严重污染,出现相当程度的失真.在 MN-FINDR 中,由于算法只剔除混合像元,最后依旧在被噪声污染的纯像元中提取端元,因此该算法提取的端元同样含有较大的噪声.而在 DN-FINDR 的提取结果中,候选端元是光谱距离阈值内的所有像元的平滑结果,噪声被大幅抑制,可以较精确地匹配标准光谱.表 3 相对于表 2 无明显变化,说明噪声的加入并不会影响各算法的耗时.

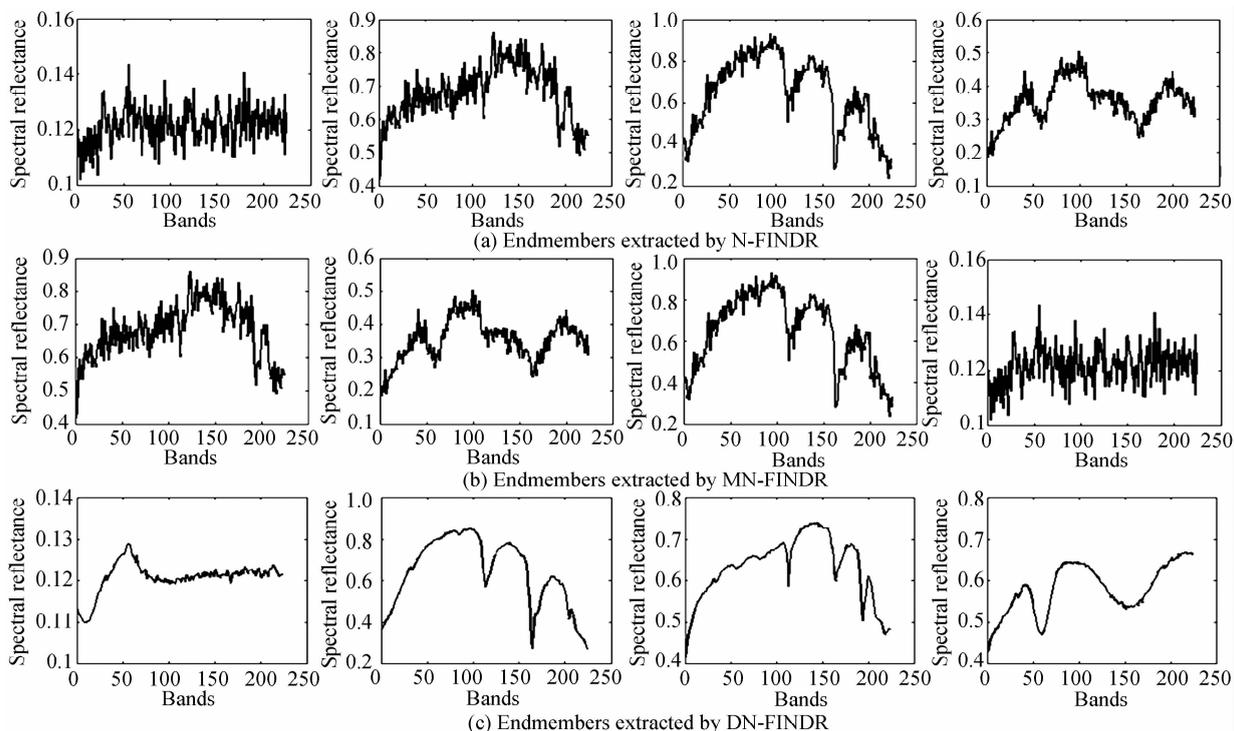


图6 端元提取结果对比

Fig. 6 Comparison of the endmembers extraction results

4.3 高斯噪声+奇异点干扰实验分析

高斯白噪声添加进合成数据,控制噪声强度使得信噪比为 25dB,并随机添加 6 个奇异点作为干扰,含噪高光谱数据如图 4(c),设置光谱角阈值为 0.015,端元提取的结果如图 7,耗时对比见表 4.

表 4 算法耗时对比

Table 4 Comparison of the processing time

Algorithm	Time of pre-treatment/s	Number of candidates	Extraction time/s	Total time/s
N-FINDR	-	25600	424.65	424.65
MN-FINDR	5.18	1605	29.01	34.19
DN-FINDR	5.52	14	0.24	5.76

从图 7 中可以看出,N-FINDR 算法对含有奇异点干扰的高光谱图像无能为力,错误地将奇异点作为端元

元提取出来,造成正确端元的缺失.在 MN-FINDR 算法中,奇异点的卡方分布与上侧分位点的距离通常大于纯像元的分布,因此奇异点造成的干扰在 MN-FINDR 算法中同样无法避免.而 DN-FINDR 通过个数阈值的判定能够排除奇异点的干扰,提取出正确、纯净的端元.表 4 中,N-FINDR 与 MN-FINDR 算法的耗时依旧没有明显变化,而 DN-FINDR 由于阈值的增加(相比实验 4.2 提高了 0.005),使得阈值范围的像元数目增加,筛选出的候选端元个数减少,致使算法效率得到一定程度的提高,反映了阈值选取对于算法效率和效果的影响.然而,阈值选取并不是越大越好,应根据噪声强度以及高光谱数据的复杂度综合选取.

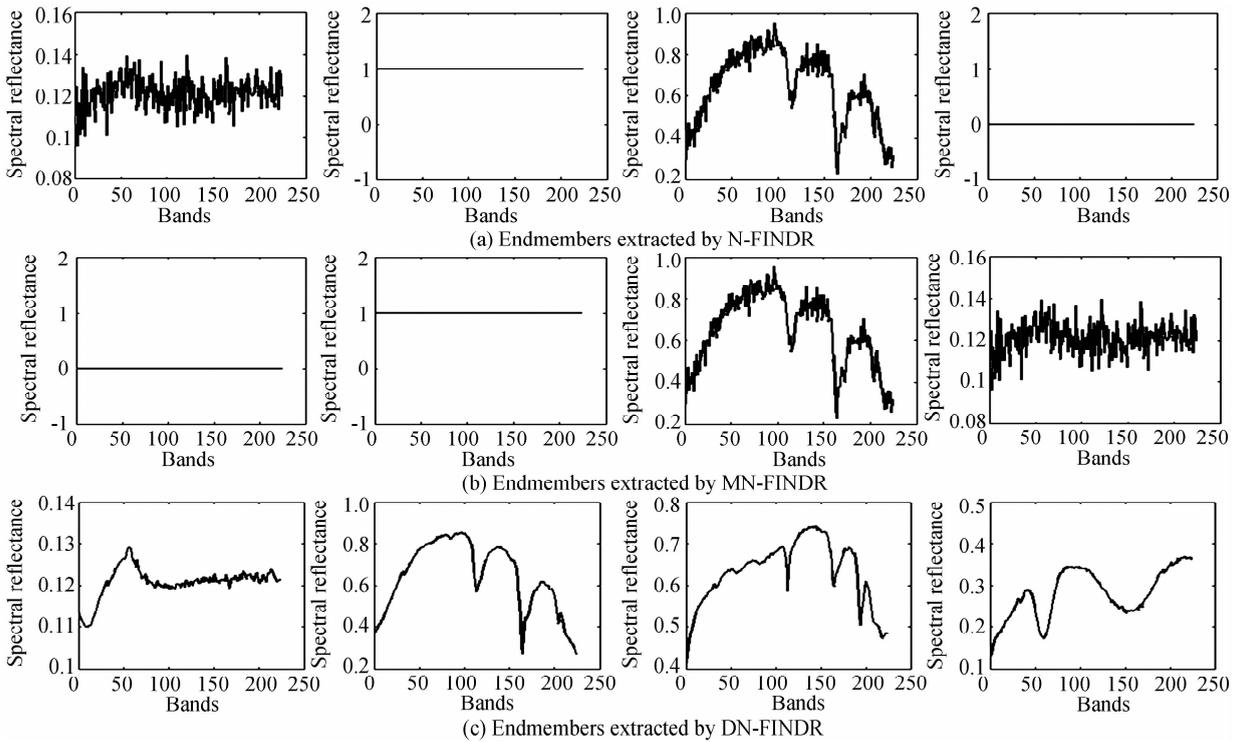


图 7 端元提取结果对比

Fig. 7 Comparison of the endmembers extraction results

4.4 真实高光谱数据实验分析

采用 AVIRIS 采集的高光谱图像对各算法的端元提取性能进行研究.图像包括 145×145 个像元,200 个波段,16 种已标记地物类别以及其他未标记地物类别.该图像的第 20 波段以及其他地物分布如图 8.图 8 (b)中的黑色部分为未标记地物,灰色部分根据灰度值的不同代表不同的已标记地物.在该高光谱图像中,由于地物种类皆为农作物,其光谱较为类似,无法采用视觉评价分析端元提取的性能,故采用 SAM 和 ED 来定量评价各算法的提取结果.

实验选取图像中光谱特征较明显的 9 种地物共 6416 个像元作为待提取端元的高光谱数据,分别标记

为地物 1~9.表 5~7 给出各方法端元提取结果的比较.

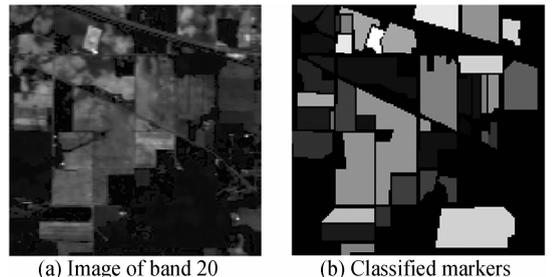


图 8 AVIRIS 高光谱数据

Fig. 8 Hyperspectral image from AVIRIS

表 5 N-FINDR 算法提取结果

Table 5 Endmember(EM) extraction result via N-FINDR algorithm

N-FINDR	EM1	EM 2	EM 3	EM 4	EM 5	EM 6	EM 7	EM 8	EM 9	Processing time/s	Missing Object(Ob)
Category	Ob 1	Ob 9	Ob 7	Ob 9	Ob 1	Ob 7	Ob 5	Ob 9	Ob 6		
SAM	0.0114	0.1756	0.0694	0.0595	0.0661	0.0164	0.0922	0.0502	0.0371	203.22	Ob 2, 3, 4, 8
ED	0.0597	2.1846	0.9281	1.0010	0.9994	0.1780	0.8992	0.3475	0.3078		

表 6 MN-FINDR 算法提取结果

Table 6 Endmember extraction result via MN-FINDR algorithm

MN-FINDR	EM1	EM 2	EM 3	EM 4	EM 5	EM 6	EM 7	EM 8	EM 9	Processing time/s	Missing Object
Category	Ob 9	Ob 1	Ob 7	Ob 1	Ob 7	Ob 9	Ob 9	Ob 7	Ob 5		
SAM	0.1756	0.0661	0.0694	0.0114	0.0164	0.0502	0.0595	0.0164	0.0922	56.87	Ob 2, 3, 4, 8
ED	2.1846	0.9994	0.9281	0.0597	0.1780	0.3475	1.0010	0.1780	0.8992		

表 7 DN-FINDR 算法提取结果

Table 7 Endmember extraction result via DN-FINDR algorithm

DN-FINDR	EM1	EM 2	EM 3	EM 4	EM 5	EM 6	EM 7	EM 8	EM 9	Processing time/s	Missing Object
Category	Ob 1	Ob 6	Ob 9	Ob 3	Ob 8	Ob 5	Ob 7	Ob 6	Ob 4		
SAM	0.0217	0.0547	0.0250	0.0946	0.0489	0.0392	0.0119	0.1135	0.0187	9.57	Ob 2
ED	0.2252	0.6707	0.1528	0.9578	0.2302	0.5051	0.1263	0.6932	0.3063		

在表 5~7 的结果对比中,可以得出与模拟数据实验相似的结果. N-FINDR 提取的端元 2 与 MN-FINDR 提取的端元 1 在光谱数据中为奇异点,它在光谱距离上与地物 9 的光谱最相似,即便如此,两者间的光谱角和欧式距离仍相差极大.此外, N-FINDR 提取的端元 3、5、7 以及 MN-FINDR 提取的端元 2、3、9 都受噪声影响较为严重,导致其光谱角和欧氏距离与真实光谱有较大差异,将造成地物识别的困难.此外,由于该高光谱数据中地物的光谱相似性较高,这两种算法存在严重的重复提取现象,地物 7 被提取了两次,地物 9 被提取三次,地物 1 被提取两次,端元提取的结果并不理想.在 DN-FINDR 端元提取结果中,地物 2 未能被正确提取,在提取的端元 4、8 中,与地物的匹配结果也并不理想,这是由于在真实的遥感拍摄中,容易出现较大面积的干扰,这些干扰由于影响像元较多,无法被算法的阈值处理所去除,进而影响了端元提取的结果.除此之外, DN-FINDR 算法能较好地提取出其余 7 种地物,并在光谱角和光谱距离上都更为接近标准光谱.此外, DN-FINDR 算法的耗时分别是 N-FINDR 和 MN-FINDR 算法的 1/20 和 1/6,端元提取效率较高.综上所述,本文算法在提取端元的精度、覆盖度和效率上都较另外两种算法更佳.

5 结论

本文综合利用高光谱图像中丰富的光谱信息,以及常见噪声和干扰的特点,提出了一种改进的快速 N-FINDR 端元提取算法,通过光谱距离阈值减少候选端元个数并平滑噪声,通过个数阈值剔除奇异点的干扰.

为了进一步提高算法效率,本文还提出一种基于最大化光谱间距离的初始端元选取算法.实验表明该算法克服了现有的 N-FINDR 算法对噪声敏感的缺点,较大程度上消除噪声及奇异点的影响,提高端元光谱的纯度和正确率,并明显提高算法运算速度.本文提出的两种改进方法可以结合使用,也可单独应用于其他端元提取算法,具备一定的应用价值.

参考文献

- [1] SCHOWENGERDT A R. Remote sensing: models and methods for image processing [M]. San Diego: Academic Press, 1997: 120-121.
- [2] WINTER E M. N-finder: an algorithm for fast autonomous spectral endmember determination in hyperspectral data[C]. SPIE: Conference on Imaging Spectrometry V, 1999, **3753**: 266-275.
- [3] PLAZA A, CHANG Chein-I. An improved N-FINDR algorithm in implementation. Algorithms and Technologies for Multispectral[C], Hyperspectral and Ultraspectral Imagery XI, 2005, **5806**: 298-306.
- [4] NASCIMENTO J M P, DIAS J M B. Vertex component analysis: a fast algorithm to unmix hyperspectral data[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, **43**(4): 898-910.
- [5] FANG Ling-jiang, NIAN Yong-jian, LEI Shu-tao, et al. Endmembers extraction for hyperspectral images based on vertex component analysis[J]. *Ship Electronic Engineering*, 2014, **34**(8): 154-157.
方凌江, 粘永健, 雷树涛, 等. 基于顶点成分分析的高光谱图像端元提取算法[J]. *舰船电子工程*, 2014, **34**(8): 154-157.
- [6] CHANG Chein-i, WU Chao-cheng, LIU Wei-min, et al. A new growing method for simplex-based endmember extraction algorithm[J]. *IEEE Transactions Geosciences and Remote Sensing*, 2006, **44**(10): 2804-1819.
- [7] WANG Li-jiao, LI Xiao-run, ZHAO Liao-ying. Fast

- implement of the simplex growing algorithm for endmember extraction[J]. *Acta Optica Sinica*, 2014, **34**(11): 1-7.
- 王丽姣, 厉小润, 赵辽英. 快速实现基于单形体体积生长的端元提取算法[J]. *光学学报*, 2014, **34**(11): 1-7.
- [8] LI Na, ZHAO Hui-jie. Endmember extraction method based on morphology and orthogonal subspace projection [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2010, **36**(12): 1457-1460.
- 李娜, 赵慧洁. 基于形态学与正交子空间投影的端元提取方法[J]. *北京航空航天大学学报*, 2010, **36**(12): 1457-1460.
- [9] CUI Jian-tao, WANG Jing, LI Xiao-run, *et al.* Endmember extraction algorithm based on spatial pixel purity index[J]. *Journal of Zhejiang University (Engineering Science)*, 2013, **47**(9): 1524-1530.
- 崔建涛, 王晶, 厉小润, 等. 基于空间像素纯度指数的端元提取算法[J]. *浙江大学学报*, 2013, **47**(9): 1524-1530.
- [10] SETOAIN J, PRIETO M, TENLLADO C, *et al.* Parallel morphological endmember extraction using commodity graphics hardware [J]. *Geoscience and Remote Sensing Letters, IEEE*, 2007, **4**(3): 441-445.
- [11] WANG Ying, LIANG Nan, GUO Lei. A hyperspectral remote sensing image endmember extraction algorithm based on modified extended-morphological operator [J]. *Acta Photonica Sinica*, 2012, **41**(6): 672-677.
- 王瀛, 梁楠, 郭雷. 一种基于修正扩展形态学算子的高光谱遥感图像端元提取算法[J]. *光子学报*, 2012, **41**(6): 672-677.
- [12] BOARDMAN J W, KRUSE F A, GREEN R O. Mapping target signatures via partial unmixing of AVIRIS data: in Summaries [C]. Fifth JPL Airborne Earth Science Workshop. 1995, 23-26.
- [13] XIONG Wei, CHANG Chein-i, WU Chao-cheng, *et al.* Fast algorithms to implement N-FINDR for Hyperspectral endmember extraction[C]. Proceedings of the SPIE, 2010, **76951Q**: 1-12.
- [14] DING Hai-yong, SHI Wen-zhong. Fast N-FINDR algorithm for endmember extraction based on chi-square distribution [J]. *Journal of Remote Sensing*, 2013, **17**(1): 122-129.
- [15] WANG Li-guo, ZHANG Ye. Speed-up for N-FINDR algorithm[J]. *Journal of Harbin Institute of Technology*, 2008, **15**(1): 141-144.
- [16] ZHAO Chun-hui, QI Bin, WANG Yu-lei. An improved N-FINDR hyperspectral endmember extraction algorithm[J]. *Journal of Electronics & Information Technology*, 2012, **34**(2): 499-503.
- 赵春晖, 齐滨, 王玉磊. 一种改进的 N-FINDR 高光谱端元提取算法[J], *电子与信息学报*, 2012, **34**(2): 499-503.
- [17] 耿修瑞, 赵永超, 周冠华. 一种利用单形体体积自动提取高光谱图像端元的算法[J]. *自然科学进展*, 2006, **16**(9): 1196-1200.
- GENG Xiu-ru, ZHAO Yong-chao, ZHOU Guan-hua. An automatic extraction algorithm of hyperspectral image endmember using single volume form [J]. *Progress in Natural Science*, 2006, **16**(9): 1196-1200.