

文章编号: 1001 - 9014 (2009) 06 - 0476 - 05

一种新的基于 Fisher 判别的混合像元分解算法： 室内控制实验结果分析

陈学泓，王胜强，陈晋*，沈妙根，朱孝林

(北京师范大学资源学院 地表过程与资源生态国家重点实验室(筹), 北京 100875)

摘要:混合像元分解技术(Spectral Mixture Analysis, SMA)是遥感图像处理的重要手段之一。传统方法假设每个端元具有稳定的光谱特征,然而端元内光谱差异普遍存在,这将导致混合像元分解精度的降低。针对该问题,提出了基于 Fisher 判别(Fisher Discriminant Analysis, FDA)的混合像元分解算法。Fisher 判别对光谱各波段进行线性组合,使得转换后的光谱值分离度最大,即端元内的光谱差异较小而端元间的光谱差异较大。利用转换后的光谱对混合像元进行分解可以最大程度地减少端元内光谱差异对分解结果的影响。利用该方法对室内控制实验的模拟混合像元光谱进行分解,并与过去提出的几种混合像元分解技术进行比较,结果显示新方法相比传统方法在分解精度上有相当程度的提高。

关键词:混合像元分解;端元内光谱差异;Fisher 判别;室内控制实验
中图分类号: TP751.1 **文献标识码:** A

NEW ALGORITHM FOR SPECTRAL MIXTURE ANALYSIS BASED ON FISHER DISCRIMINANT ANALYSIS: EVIDENCE FROM LABORATORY EXPERIMENT

CHEN Xue-Hong, WANG Sheng-Qiang, CHEN Jin*, SHEN Miao-Gen, ZHU Xiao-Lin

(State Key Laboratory of Earth Surface Processes and Resources Ecology, College of Resources
Science & Technology, Beijing Normal University, Beijing 100875, China)

Abstract: Spectral mixture analysis (SMA) is one of the most important methods in remote sensing image processing. Traditional SMA assumes a constant spectral signature for each endmember. However, the endmember spectral variability commonly exists, which leads to the lower accuracy of pixel unmixing. In order to solve this problem, a novel unmixing method based on Fisher discriminant analysis (FDA) was developed. FDA aimed to find a linear combination of the spectral bands for getting the largest separation degree among the endmember spectra, i.e. small variability of spectra inside one endmember group but a large difference of spectra among endmember groups. Mixture pixel was unmixed by using transformed spectra, as a result, the adverse impact caused by endmember spectral variability on unmixing accuracy could be diminished to a large extent. A laboratory experiment was designed to obtain a group of mixed spectra with endmember spectral variability. The measured spectra were used to test the performance of the new method and the traditional SMA methods. The comparison results suggest that the new method outperforms the traditional methods with considerable improvement of unmixing accuracy.

Key words: spectral mixture analysis; endmember spectral variability; Fisher discriminant analysis; laboratory experiment

引言

混合像元问题一直是遥感科学研究中的热点问题。在过去的三十年里,国内外已发展了不少混合像元分解模型,包括线性混合模型、神经网络、模糊模

型等^[1-3],并在很多领域得到成功的应用。在这些模型中,线性混合模型是最简单、应用最广泛的混合像元模型。在线性模型中,混合光谱等于端元光谱与端元面积比例的线性组合。如果存在 m 个端元, L 个波段,线性模型的数学表达式如下:

收稿日期: 2008 - 10 - 15, 修回日期: 2009 - 05 - 27

Received date: 2008 - 10 - 15, revised date: 2009 - 05 - 27

基金项目: 本研究得到国家高技术研究发展计划(863计划)(2006AA12Z103)及教育部新世纪人才计划资助。

作者简介: 陈学泓(1985-),男,福建泉州人,硕士研究生,现从事资源环境遥感应用研究。

* 通讯作者: 陈晋(1967-),博士,教授, Email: chenjin@ires.cn

$$r = \sum_{i=1}^m f_i x_i + w = XF + w \quad (1)$$

其中, r 是 $L \times 1$ 的向量, 表示混合光谱; X 是 $L \times m$ 矩阵, 表示端元光谱; F 是 $m \times 1$ 向量, 表示各端元的比例; w 是 $L \times 1$ 向量, 表示观测误差. 为使模型具有物理意义, 需要受到两个约束条件限制: 一是端元比例之和为 1, 即 $\sum_{i=1}^m f_i = 1$; 二是所有的端元比例都大于 0, 即 $f_i > 0, i = 1, \dots, m$. 通常, 利用最小二乘法或约束最小二乘法求得各端元的比例.

混合像元分解技术的一个重要假设是: 在一个给定的图像里, 地表由少数的几种地物 (端元) 组成, 并且这些地物具有相对稳定的光谱特征. 然而, 由于同物异谱现象的存在, 一个端元的光谱并非恒定的值, 这就是端元内光谱差异现象. 一般的混合像元分解算法对整幅图像采用相同的端元光谱, 这种现象的存在会导致分解结果的误差. 目前, 解决该问题的方法可以分为三类: (1) 多端元方法对每一类地物选取多个端元光谱参与混合像元分解, 典型的方法包括 MESMA (Multiple Endmember Spectral Mixture Analysis) 方法^[4]和端元束方法^[5]; (2) 光谱变换法对光谱进行特定的变换后再作混合像元分解, 常见的变换方式有标准化^[6,7]及微分法^[8]; (3) 贝叶斯混合光谱分析方法 (Bayesian Spectral Mixture Analysis, BSMA) 用概率密度函数来表示端元内的光谱差异, 进而根据贝叶斯理论推导出最似然的端元百分比^[9].

上述方法的提出虽然推动了端元内光谱差异的混合像元分解问题的解决, 但第一、三类方法存在求解过程复杂、耗时过长等方面问题, 而第二类方法则存在物理意义不明确、普适性差的问题. 针对上述需求与问题, 我们认为端元内光谱差异并非毫无规律, 可以寻找到一个由各波段的线性组合而成的特征空间中的方向, 使得在这样一个方向上, 端元内的光谱差异较小而端元间的光谱差异较大, 在这个方向上对混合像元进行分解可以显著地减少端元内光谱差异对分解结果的影响. 基于该思路, 本文提出一种基于 Fisher判别的混合像元分解算法, 能够在最大程度上减少端元内光谱差异对分解结果的影响. 为对新算法进行检验, 我们设计了室内控制光谱实验.

1 算法原理

1.1 多类 Fisher判别

Fisher判别分析是一种监督分类方法, 其主要

思想是将多元观测值进行线性组合来建立新的判别量, 使得新判别量的组间方差与组内方差的比值达到最大^[10].

假设有 m 个类别的训练样本 x_1, x_2, \dots, x_m . x_i 为 $L \times n_i$ 的矩阵, 表示了第 i 类端元有 n_i 个训练样本. 首先构造组间样本矩阵, 表示各端元平均光谱的协方差矩阵, 即端元间的光谱差异

$$B = \sum_{i=1}^m (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T \quad (2)$$

其中, \bar{x}_i 表示第 i 类端元的样本均值, \bar{x} 表示“总平均向量”,

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \quad (3)$$

$$\bar{x} = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}}{m} \quad (4)$$

再构造组内样本矩阵为:

$$W = \sum_{i=1}^m (n_i - 1) S_i \quad (5)$$

其中, S_i 表示各类端元内样本光谱的协方差矩阵, 如式 (6):

$$S_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)^T \quad (6)$$

组内样本矩阵表示了各个端元各自的协方差矩阵之和, 用于构造端元内的光谱差异. 考虑线性组合

$$y = Ax \quad (7)$$

其中, A 是 $1 \times L$ 的矩阵, 表示对原光谱的某种线性组合操作. 则经过矩阵 A 变换后的分离度为:

$$\frac{ABA^T}{AWA^T} = \frac{\sum_{i=1}^m (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T A^T}{\sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)^T A^T} \quad (8)$$

式 (8) 中分子表示组间方差, 即端元间的光谱差异; 分母表示组内方差, 即端元内的光谱差异; 分离度表示了两者的比值, 因此分离度越大意味着端元间的光谱差异越大, 端元内的光谱差异越小. 使得分离度最大的线性组合称为第一判别量, 分离度次大的线性组合为第二判别量, 以此类推可获得 L 个判别量. 每个判别量的变换矩阵对应于矩阵 $W^{-1}B$ 的特征向量, 而每个判别量的分离度对应于 $W^{-1}B$ 的特征值. 分开 m 个类别至少需要 $m - 1$ 个判别量, 因此通常选择前 $m - 1$ 个判别量用于判别分析.

1.2 基于 Fisher判别的混合像元分解方法

由于线性混合像元模型在线性变换后依然成

立,因此 Fisher判别量可以直接应用于混合像元分解.以 m 个类别为例,根据 Fisher判别可提取各类端元平均光谱的前 $m - 1$ 个判别量:

$$\begin{aligned} y_{1i} &= A_1 x_i \\ y_{2i} &= A_2 x_i \\ &\dots \\ y_{(m-1)i} &= A_{m-1} x_i \\ i &= 1, \dots, m \end{aligned} \quad (9)$$

对混合光谱作同样的变换

$$\begin{aligned} r_1 &= A_1 r \\ r_2 &= A_2 r \\ &\dots \\ r_{m-1} &= A_{m-1} r \end{aligned} \quad (10)$$

由于变换为线性,变换后的光谱仍然满足线性混合模型,有

$$r = \sum_{i=1}^m f_i y_i + w \quad (11)$$

其中, $r = (r, r, \dots, r_{m-1})^T$ 表示变换后的混合光谱, $y_i = (y_{1i}, y_{2i}, \dots, y_{(m-1)i})^T$, $i = 1, \dots, m$ 表示变换后的端元光谱.加上约束条件 $\sum_{i=1}^m f_i = 1$,即可求解方程组(式 11).由于前 $m - 1$ 个判别量有较大的分离度,可以在很大程度上减少端元内光谱差异对混合像元分解产生的影响.基于 Fisher判别的混合像元分解方法的具体步骤如下:

- (1) 等量选取各类地物的纯光谱作为训练样本;
- (2) 利用 Fisher判别方法获得前 $m - 1$ 个分离度最大的判别量对应的变换矩阵,同时对端元光谱和混合光谱作变换;
- (3) 利用变换后的光谱构建新的方程组并利用最小二乘法求解.

2 室内控制实验

2.1 实验设计

为检验新算法的精度,本研究设计了一个室内控制实验来模拟遥感图像中的混合像元.模拟的三端元包括植被、土壤和枯叶,将三种材料按预定的比例平铺于一个圆柱形容器中(图 1(a)),圆柱形容器的内壁和底部均被涂成黑色.实验装置如图 1(b)所示,实验光源由两盏卤素灯提供,卤素灯位于容器正上方两侧,互成 45° 角.光谱由手提式分光辐射度计(MS-720, EKO Instrument Co. Ltd, Japan)测量,波长范围为 $350 \sim 1050\text{nm}$,光谱分辨率为 10nm ,视场角为 25° .手提式分光辐射度计位于容器正上方

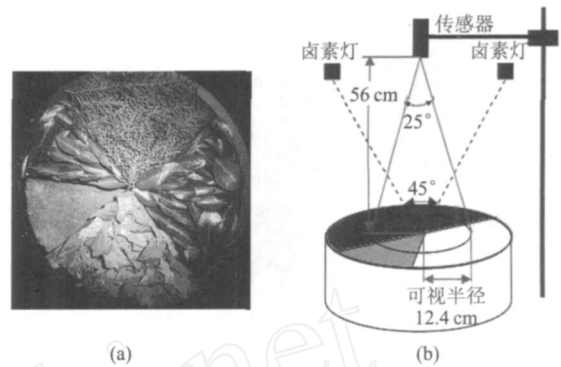


图 1 (a)模拟混合像元 (b)实验装置

Fig 1 (a) simulated mixed pixels (b) the instrument for experiment

56cm处,根据视场角可知其可视半径为 12.4cm.一块标准白板用于辐射定标.由于光谱质量问题,本实验只采用 $500 \sim 1000\text{nm}$ 波段的数据.

2.2 端元光谱测量

本实验模拟了三类端元,端元光谱如图 2所示:

- (1) 植被端元:植被端元由野外采集的两种植被叶片(阔叶和针叶)模拟,并通过混合不同比例的阔叶和针叶来实现端元内光谱差异.本次实验测量了五条植被端元光谱,分别对应不同比例的阔叶和针叶.阔叶的比例从 0,按步长 0.25 递增至 1,同时针叶的比例从 1,按步长 0.25 递减至 0.如图 2所示,不同的植被光谱在近红外区相差很大.
- (2) 土壤端元:土壤端元由野外采集的土壤模拟,并通过改变湿度构造端元内光谱差异.同样地,本次实验测量了五条土壤端元的光谱,分别对应于不同的土壤湿度.土壤湿度由烘干称重法获得,质量百分比分别为 1%、11%、20%、27%及 33%.如图 2所示,不同湿度的土壤光谱形状相似,但在亮度上有很大差异.

(3) 枯叶端元:枯叶端元由野外采集的枯叶模拟,并通过改变摆放方式构造端元内光谱差异.同样本次实验测量了五条枯叶端元的光谱,由于枯叶光谱的端元内光谱差异只是因为其几何结构的变化导致的,因此端元内光谱差异较不明显.

2.3 混合光谱测量

将三类端元按各种比例混合,构造模拟混合像元,并测量光谱.本研究共模拟 49 个不同盖度的混合像元,并测得混合光谱 49 条.其中不同模拟混合像元的端元光谱并不相同,通过不同类型叶片的混合,不同湿度土壤及不同的枯叶摆放方式构造了端元内光谱差异.

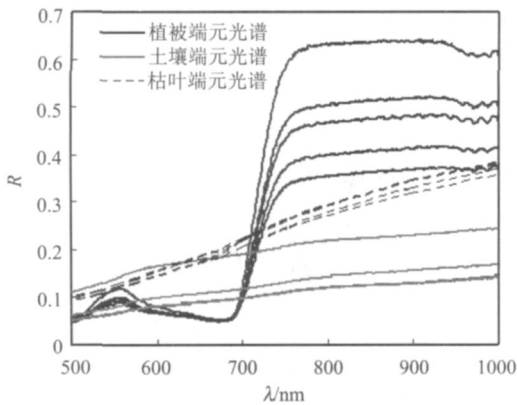


图 2 端元光谱曲线
Fig 2 Spectra of the three endmembers

3 算法检验

本文用上述实验中得到的实测混合光谱和端元比例来检验新算法的精度. 将 15 个端元光谱作为训练样本用以作 Fisher判别分析, 并利用前两个判别量对应的变换矩阵对混合光谱及端元光谱作变换. 由于高光谱数据波段间存在极大的相关性, 我们采用主成分变换对原始光谱进行变换并提取前四个主成分 (方差解释量 99.99%). 图 3 (a) 显示了由第一、二主成分构成的特征空间中端元光谱与混合光谱的分布, 可以明显看到在主成分空间中, 端元内光谱差异十分明显, 同类地物的光谱很分散. 而在 Fisher判别量构成的特征空间中 (图 3b), 同类端元的光谱很集中, 端元内光谱差异得到明显改善.

本文利用普通最小二乘法分别对主成分变换后的光谱 (提取前四个主成分) 和 Fisher判别变换后的光谱进行分解, 同时为了与其它现有方法比较, 采用标准化混合光谱分析 (Normalized SMA)^[6]、微分法^[9]及全约束最小二乘法 (FCLS)^[11]对原始混合光谱进行分解, 并将分解结果与新方法的求解结果进行了比较. 在应用普通最小二乘法时, 将求解结果中的负值取为 0, 再对其余的正值作归一化, 以使得求解结果满足两个约束条件. 评价指标采用均方根误差 (RMSE), 表达式如式 (12) 所示:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{f}_i - f_i)^2}{N}}, \quad (12)$$

其中, \hat{f}_i 表示估计的端元比例, f_i 表示真实的端元比例, N 表示样本量. 各种方法的求解精度如表 1 所示. 可以看出, 相对于其它方法, 基于 Fisher判别的分解

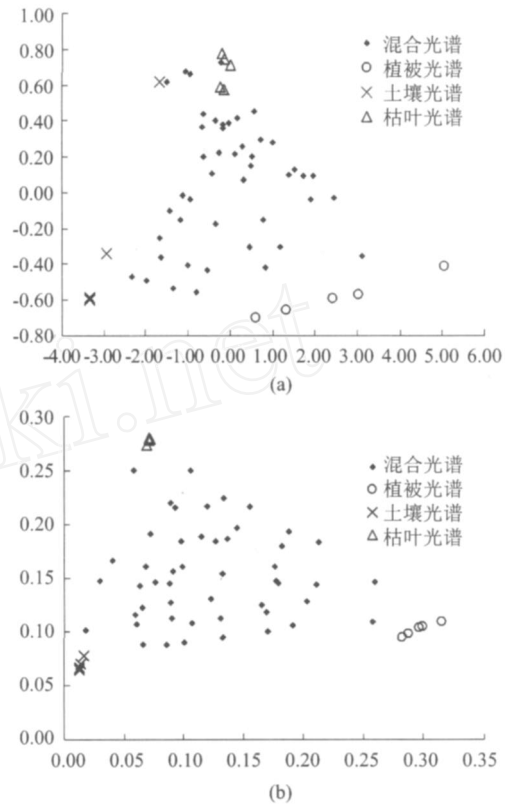


图 3 特征空间中的端元与混合像元分布 (a)主成分空间 (b) Fisher判别量空间
Fig 3 Scatter plot of endmember and mixed pixels in the feature space (a) feature space of the first two PC components (b) feature space of the first two Fisher discriminating components

算法具有明显的优势. 尽管 FCLS能够求得同时满足两个约束条件的最优解, 但是稳定的端元光谱依然是其重要前提. 在端元光谱可变的情况下, FCLS并不能改善求解结果. 标准化 SMA确实能够压缩端元内的光谱差异, 但是对光谱进行标准化后, 线性混合模型已经得不到满足, 求解得到的比例并非真实比例^[12]. 微分法能够克服光谱整体平移的一类端元内光谱差异 (如土壤), 但是对于更复杂的端元光谱变化 (如植被), 微分法并不能有效克服, 而且对光谱进行微分后往往会减小端元间的光谱差异, 从而影响求解结果. 对于本实验的数据, FCLS, 标准化 SMA及微分法并没有改善求解结果. 图 4表示了主成分变换后光谱的普通最小二乘法与新方法求解结果的对比. 不难发现新方法的求解精度有较大幅度的提高.

表 1 不同方法对模拟像元的分解精度比较
Table 1 Comparison of unmixing accuracy by different methods

端元	新方法	最小二乘法 (主成分)	标准化 SMA	微分法	FCLS
植被	0.037	0.052	0.095	0.056	0.098

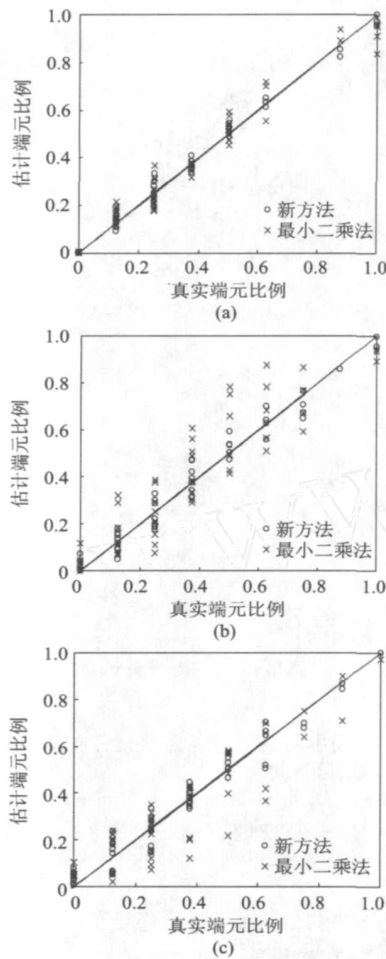


图 4 新方法与最小二乘法的分解结果比较 (a)植被 (b)土壤 (c)枯叶

Fig 4 Comparison of fractions estimated by the new method and least square method for PCA-possessed spectra (a) vegetation (b) soil (c) litter

端元	新方法	最小二乘法 (主成分)	标准化 SMA	微分法	FCLS
土壤	0.043	0.108	0.091	0.098	0.147
枯叶	0.048	0.101	0.098	0.095	0.109

4 结论

针对混合像元分解技术中存在的端元内光谱差异问题,本文提出了基于 Fisher判别的混合像元分解算法. Fisher判别法通过对光谱波段进行线性组合,使得变换后的端元光谱的分离度最大. 经过 Fisher变换后的前面几个判别量的端元内光谱差异很小,同时端元间光谱差异很大,这使得混合像元分解的假设在最大程度上得到满足,从而提高了混合像元分解精度. 而过去提出的几种方法尽管原理各异,但都使用了原始光谱的绝大部分信息. 但是原始

光谱中包含了相当一部分端元内光谱差异的信息,而这部分信息会给混合像元分解带来很大的混淆作用. 对室内控制实验实测光谱的分解结果表明,新方法确实能减少端元内光谱差异的影响,在分解精度上优于过去提出的几种混合像元分解技术. 但是在真实的遥感影像中,同物异谱现象可能更加复杂,新方法的有效性还需要进一步的检验,这将是我们下一步的工作.

REFERENCES

- [1] Adams J B, Smith M O, Johnson P E. Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander I site[J]. *Journal of Geophysical Research Atmospheres*, 1986, **91** (B8): 8098—8112
- [2] Atkinson PM, Cutler M E J, Lewis H. Mapping sub-pixel proportional land cover with AVHRR imagery[J]. *International Journal of Remote Sensing*, 1997, **18** (4): 917—935.
- [3] Marsh S E, Suitzer P, Kowalik W S, et al. Resolving the percentage of component terrains within single resolution elements[J]. *Photogrammetric Engineering and Remote Sensing*, 1980, **46** (8): 1079—1086
- [4] Roberts D A, Gardner M, Church R, et al. Mapping chaparral in the santa monica mountains using multiple endmember spectral mixture models[J]. *Remote Sensing of Environment*, 1998, **65** (3): 267—279.
- [5] Bateson C A, Asner G P, Wessman C A. Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2000, **38** (2): 1083—1094.
- [6] Wu C S. Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery[J]. *Remote Sensing of Environment*, 2004, **93** (4): 480—492
- [7] Garcia-Haro F J, Sommer S, Kemper T. A new tool for variable multiple endmember spectral mixture analysis (VMESMA) [J]. *International Journal of Remote Sensing*, 2005, **26** (10): 2135—2162
- [8] Asner G P, Lobell D B. A biogeophysical approach for automated SWIR unmixing of soils and vegetation[J]. *Remote Sensing of Environment*, 2000, **74** (1): 99—112
- [9] Song C H. Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability [J]. *Remote Sensing of Environment*, 2005, **95** (2): 248—263.
- [10] Johnson R A, Wichern D W. *Applied Multivariate Statistical Analysis*[M]. New Jersey: Prentice-Hall, 1982: 496—498
- [11] Heinz D C, Chang C 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.
- [12] Zhang J, Rivard B, Sánchez-Azofeifa A. Spectral unmixing of normalized reflectance data for the deconvolution of lichen and rock mixtures[J]. *Remote Sensing of Environment*, 2005, **95** (1): 57—66