

一种基于单形体正化的高光谱数据 全约束线性解混方法

许宁^{1,2,3*}, 耿修瑞^{1,2}, 尤红建^{1,2}, 曹银贵⁴

(1. 中国科学院空间信息处理与应用系统技术重点实验室, 北京 100190;

2. 中国科学院电子学研究所, 北京 100190;

3. 中国科学院大学, 北京 100049;

4. 中国地质大学, 北京 100083)

摘要: 在端元已知情况下, 线性混合模型的非负约束最小二乘无闭式解, 需要多次迭代得收敛最优解, 时间复杂度高. 通过高光谱数据凸面几何特性分析, 指出当数据为正单形体时, 可经有限步骤快速得线性混合模型最优解. 据此提出一种单形体正化的高光谱数据全约束线性解混方法, 据已知端元进行单形体正化, 采用和为一约束求解丰度系数, 最后迭代剔除丰度负值端元得全约束解. 实验结果表明该方法可获得传统全约束解一致的丰度估计, 且效率大大提升.

关键词: 高光谱数据; 光谱解混; 端元白化; 单形体正化; 全约束最小二乘

中图分类号: TP394.1 **文献标识码:** A

A fully constrained linear unmixing method: Simplex regularization for hyperspectral imagery

XU Ning^{1,2,3*}, GENG Xiu-Rui^{1,2}, YOU Hong-Jian^{1,2}, CAO Yin-Gui⁴

(1. Key Laboratory of Technology in Geo-spatial Information Processing and Application System, IECAS, Beijing 100190, China;

2. Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China;

3. University of Chinese Academy of Sciences, Beijing 100049, China;

4. School of Land Science and Technology, China University of Geosciences, Beijing 100083, China)

Abstract: With a priori information of the known endmembers in hyperspectral image, there is no closed-form solution of Least Square (LS) method for linear mixing model under the Abundance Non-negativity Constraint (ANC). So many iterations which may result in big computational complexity are needed in the traditional Fully Constrained LS (FCLS) methods to obtain the optimal solution. In this paper, an analysis of impacts on abundance estimation of hyperspectral image in different simplex shapes was implemented and a fully constrained linear unmixing method based on simplex regularization was proposed which could get optimal solution under limited iteration when the hyperspectral image was spanned into a regular simplex. The proposed method was carried out by three steps. Firstly, the simplex of hyperspectral image was regularized by the known endmembers' whitening matrix. Secondly, the analytical solution of abundance coefficients was obtained under Abundance Sum-to-one Constraint (ASC). Then for every pixel, the FCLS solution was achieved by eliminating the endmembers with negative abundance coefficients and solving the ASC equation iteratively. Experiments on simulated and real hyperspectral images indicate that the proposed method can obtain consistent results with

收稿日期: 2015-08-18, 修回日期: 2016-01-07

Received date: 2015-08-18, revised date: 2016-01-07

基金项目: 中国地质调查局地质调查项目(1212011120226); 国家 863 计划课题(2012AA12A308); 中国科学院科技服务网络计划项目(KFJ-EW-STS-046)

Foundation items: Supported by the Geological Survey Program of China Geological Survey (1212011120226); the National High-tech R&D Program of China (2012AA12A308) and the Science and Technology Services Network Program of Chinese Academy of Sciences (KFJ-EW-STS-046)

作者简介 (Biography): 许宁(1982-), 男, 四川邛崃人, 博士生, 研究方向为光学遥感影像配准、融合以及高光谱图像处理.

* 通讯作者 (Corresponding author): E-mail: x_ning@aliyun.com

traditional FCLS method and decrease the computational burden efficiently.

Key words: Hyperspectral Imagery, spectral unmixing, endmember whitening, simplex regularization, FCLS

PACS: 07.05. Pj, 95.75. Mn

引言

由于遥感器件空间分辨率的限制,图像单个像元中往往包含不同的地表覆盖类型,另外在辐射能量大气传输过程中地物光谱会发生多次散射和混合效应,从而导致遥感图像中普遍存在的混合像元现象^[1-2]. 针对这类混合像元,目前主要采用光谱解混的方法来解决这种图像的软分类问题,以求得遥感数据各个像元中纯净地物类型以及其所占的比例系数(丰度)^[2-3]. 高光谱遥感数据由于光谱分辨率高,非常利于混合像元的光谱解混处理^[1]. 根据对高光谱数据成像过程中与不同地物的辐射传输过程进行建模的差异,可以将光谱解混方法分为线性和非线性光谱解混两类^[2],线性光谱解混由于模型简单、易于理解和分析,在光谱解混的研究和应用中占据着主导地位. 线性光谱解混方法可划分为代数统计、单形体几何以及稀疏解混三大类^[2],典型的代数解算类方法有最小二乘法^[4]、非负矩阵分解^[5]、独立成分分析^[6]等,其采用代数求解方式获得丰度系数;稀疏解混思想来源于稀疏表示、基追踪以及压缩感知理论^[2],其采用过完备的光谱数据库进行迭代求解,运算量较大. 而基于单形体几何的方法源于 Boardman^[7]引入的凸面几何学理论,这类方法直观、易于理解、计算量较小,在高光谱遥感界得到了广泛关注,出现了基于单形体体积比^[8]、质心坐标法^[9]、距离几何法^[10-11]等算法.

从高光谱图像中是否存在纯净像元(端元)的角度,基于凸面几何的高光谱数据解混方法可归结为两类,一类假设高光谱数据中存在端元,则端元必位于所有像元张成的凸面单形体的顶点,这类算法包括基于 PPI^[7]、N-Findr^[12]、SGA^[13]等端元提取的解混方法;一类假设数据中不存在纯净端元,则出现了以最小体积变换(MVT)^[14]为基础的端元生成及光谱解混算法. 在基于凸面几何的高光谱数据解混研究中,普遍认为,由于数据获取过程中仪器误差、光谱变异以及噪声等因素的存在,使得一些像元在高维特征空间中分布在端元张成的单形体之外,成为外点像元^[15](outliers,后简称外点). 另外,这些端元张成的单形体也可能为非正单形体. 在任意单形

体及外点存在的情况下,传统的非负约束(ANC)解混需要对端元进行迭代的丰度求解,计算量大;而在正单形体情况下,可首先求得像元的和为一约束(ASC)解,对于存在丰度负值的像元,在端元矩阵中直接剔除掉丰度负值对应的端元,以此可避免 ANC 的迭代求解. 基于此考虑,提出一种快速的基于单形体正化的高光谱数据线性光谱解混方法,提升高光谱数据丰度估计的处理效率.

1 基于单形体正化的高光谱数据线性解混方法

1.1 外点位置和单形体形态对解混的影响

在采用线性混合模型对高光谱数据进行光谱解混时,考虑到丰度系数的物理意义,通常在模型中(公式 1)增加 ASC 和 ANC 约束以提升解混精度^[2,4].

$$\mathbf{y}_i = \mathbf{D}\mathbf{x}_i + \boldsymbol{\varepsilon} \text{ s. t. } \mathbf{x}_i \geq 0, \mathbf{L}^T \mathbf{x}_i = 1, i \in [1, 2, \dots, N] \quad (1)$$

其中, $\mathbf{y}_i \in \mathbf{R}^m$ 表示像元向量, m 为波段数, N 为像元个数, $\mathbf{D} \in \mathbf{R}^{m \times p}$ 表示端元矩阵, p 为端元数量, $\mathbf{x}_i \in \mathbf{R}^p$ 表示丰度系数向量, $\boldsymbol{\varepsilon}_i \in \mathbf{R}^m$ 为误差, \mathbf{L} 为元素全为 1 的列向量的转置.

在实际应用中,由于高光谱数据场景不同,像元光谱向量所在地物的特性差异,可能会导致所有像元向量在特征空间形成的单形体(凸多面体)成不规则形状(非正单形体),在二维空间中,可将此种情况描述如下图 1 所示.

从图 1 中可以看出,在二维空间中,三个端元形成一个三角形(此时三角形为一个单形体),端元即为三角形顶点,大多数的像元分布在三角形内部,称为内点像元,它们的丰度值均满足 ASC 和 ANC;而外点则分布在三角形的外部. 当利用 ASC 对所有像元(包括外点像元)进行线性解混时,图 1 左图所示外点像元-1 的丰度解算结果对于端元 1 和端元 2 为负值(例如基于单形体质心坐标^[9]等算法(见公式-2 左侧带符号公式,在二维空间为面积比,即图 1 中虚线表示的内点或外点像元与任意两个端元的面积与三个端元形成的面积之比)). 而对于 ANC 约束求解,最简单的处理方法是将该两个丰度负值端元

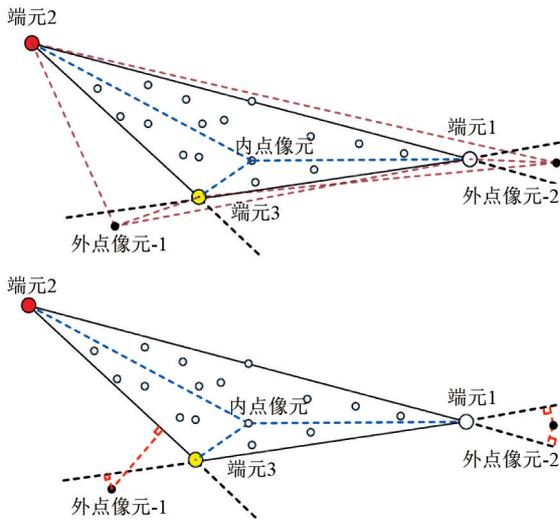


图1 二维空间中非正单形体外点解混示意图

Fig. 1 Unmixing for outlier of irregular simplex in 2-dimensional space

抛弃(对应的丰度值赋0值),再以剩下端元解求丰度系数,直到满足非负约束条件^[16];另一类方法如FCLS^[4]则采用迭代实现ANC丰度系数求解.对于前者方法,从图1右图可以看出,在非正单形体钝角处的外点-1,当将其向单形体投影时,该外点的垂足落在端元2-端元3边上和端元1-端元3延长线上,此时可以保证该外点对于端元1为负值,而对于端元2,其垂足在端元2-端元3边上,其丰度值应为正值,在此情况下抛弃端元2会造成端元误删以及丰度赋0值错误;而对于后者FCLS方法,采用ANC求解时迭代运算量大.因此,在非正单形体(二维空间三角形有钝角存在)情况下,常规方法基于两个物理约束解求丰度系数时,要么会造成端元的误删,要么需要多轮的迭代,最终影响了对部分外点的丰度系数的正确求解和效率.

$$x_i = \frac{V(\mathbf{E}_i)}{V(\mathbf{E})} = \frac{\det \begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 \\ \mathbf{e}_1 & \cdots & \mathbf{p}_i & \cdots & \mathbf{e}_p \end{pmatrix}}{\det \begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 \\ \mathbf{e}_1 & \cdots & \mathbf{e}_i & \cdots & \mathbf{e}_p \end{pmatrix}};$$

$$V(\mathbf{E}) = \frac{1}{(p-1)! \left| \det \begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 \\ \mathbf{e}_1 & \cdots & \mathbf{e}_i & \cdots & \mathbf{e}_p \end{pmatrix} \right|}, \quad (2)$$

上式中, $\det(\cdot)$ 表示行列式, $\mathbf{E} = \{\mathbf{e}_1 \cdots \mathbf{e}_i \cdots \mathbf{e}_p\}$, $i \in [1, \cdots, p]$ 表示高光谱数据中的 p 个端元向量组成的矩阵, $\mathbf{e}_i \in \mathbf{R}^m$ 为某个端元向量, $\mathbf{E}_i = \{\mathbf{e}_1 \cdots \mathbf{p}_i \cdots \mathbf{e}_p\}$ 表示将端元 \mathbf{e}_i 替换为某个像元 \mathbf{p}_i 后形成的矩阵, $V(\mathbf{E})$ 表示高维空间中单形体的体

积公式.式(2)上方公式可进行丰度值正负值的确定,因此该公式为一种带符号的丰度求解公式.

而正单形体(类似二维空间中三角形锐角处外点,图1右图外点像元-2)则不会出现ANC约束线性解混时,外点求解丰度系数的端元误删情况,其主要思想在二维空间的表示如图2.可看出,在正单形体情况下,当外点分布在与图1相似位置时,其通过体积公式(2)解混得到端元3丰度为正值且大于1,而端元1和端元2的丰度值均为负值,进一步将外点向单形体作投影,其投影点肯定落在单形体面^[17](facet,二维空间中,分别为线段端元1-端元3,线段端元2-端元3)的外部.从这类外点到单形体的距离来看,其到端元的距离即为最短距离,满足误差最小,因此,不会出现向单形体投影的投影点落在两个端元之间线段上的情况,即不会出现非正单形体中的图示外点在端元2、端元3丰度系数为正的情况.当然,二维空间的分析可以推广到高维空间.

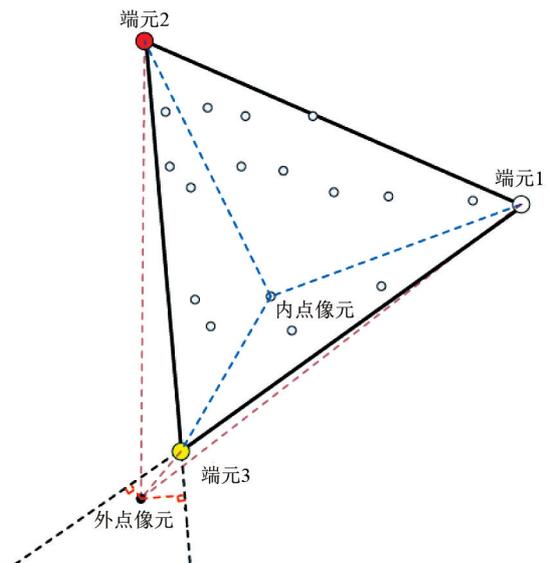


图2 二维空间中正单形体外点解混示意图

Fig. 2 Unmixing for outlier of regular simplex in 2-dimensional space

1.2 基于单形体正化的光谱解混方法

通过前文对非正单形体(有钝角)外点分布对解混的影响分析,引入端元白化^[18]和单形体正化处理方法,得到高光谱数据在特征空间的正单形体分布,以此消除外点ASC解混得到负丰度值对应的端元误删,并减少FCLS算法在ANC求解时的多次迭代,提升解混处理的效率.

假设通过N-findr^[12],SGA^[13]等端元提取算法获得了 p 个端元向量组成的端元矩阵 \mathbf{E} ,利用公式

(3)对 \mathbf{E} 以及高光谱数据像元矩阵 \mathbf{R} 进行中心化处理:

$$\bar{\mathbf{E}} = \mathbf{E} - \frac{1}{p} \sum_{i=1}^p \mathbf{e}_i; \bar{\mathbf{R}} = \mathbf{R} - \frac{1}{p} \sum_{i=1}^p \mathbf{e}_i, \quad (3)$$

再利用下式计算端元白化矩阵^[18]:

$$\bar{\mathbf{E}} = \mathbf{U}\mathbf{D}\mathbf{U}^T; \mathbf{F} = \mathbf{U}\mathbf{D}^{-1/2}, \quad (4)$$

公式(4)中, \mathbf{U}^T 为矩阵 \mathbf{U} 的转置, \mathbf{F} 即为求得的白化矩阵. 再利用公式(5), 得到白化后的端元矩阵 $\tilde{\mathbf{E}}$ 和单形体正化后的高光谱数据 $\tilde{\mathbf{R}}$:

$$\tilde{\mathbf{E}} = \mathbf{F}^T \bar{\mathbf{E}}; \tilde{\mathbf{R}} = \mathbf{F}^T \bar{\mathbf{R}} \quad (5)$$

通过高光谱数据的正化处理, 可以将不规则形状的数据单形体转换为正单形体, 再利用传统的 ASC 约束最小二乘算法^[4] 实现丰度系数的求解(公式6), 而 ANC 约束, 则通过对端元的剔除来实现, 减少不必要的迭代处理.

$$\mathbf{x}_{asc} = \mathbf{P}\mathbf{x}_i + (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{L}(\mathbf{L}^T(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{L})^{-1}, \quad (6)$$

其中, $\mathbf{P} = \mathbf{I} - (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{L}(\mathbf{L}^T(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{L})^{-1}\mathbf{L}^T$, $\mathbf{x}_i = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y}$ 表示伪逆, $\mathbf{L} = (1, \dots, 1)^T$ 为全元素为 1 的列向量, \mathbf{L} 的元素个数为 p , 即为端元的个数. 通过以上分析和流程设计, 提出一种基于单形体正化的高光谱数据全约束线性解混方法(Simplex Regularization-based Linear Spectral Unmixing, SROLSU), 其算法如表 1 所示.

表 1 基于单形体正化的线性解混算法

Table 1 Simplex regularization-based linear spectral unmixing algorithm

输入: (1)高光谱数据 \mathbf{R} , (2)端元数量 p
输出: 解混丰度系数矩阵 $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N\}$
初始化: MNF 降维处理, 选择前 $p-1$ 组分, 利用 N-Findr 端元提取算法得到 p 个端元向量组成的端元矩阵 \mathbf{E}
1. 利用公式(3)(4)计算得到端元白化矩阵 \mathbf{F}
2. 利用公式(5)计算白化端元矩阵 $\tilde{\mathbf{E}}$ 和单形体正化数据矩阵 $\tilde{\mathbf{R}}$ 单形体正化前 SROLSU 采用下面循环求解丰度系数
3. For $i = 1, 2, \dots, L$ % L 为所有像元数量
4. For $j = 1, 2, \dots, p$ % p 为端元数量
5. 利用公式(6)计算每个像元中各个端元所占的丰度向量 \mathbf{x}_{asc} , % 满足和为 1 约束
6. While ($x_i < 0$) % x_i 为第 i 个端元对应的丰度
7. 删除掉第 i 个端元, 剩下的端元再利用公式(6)求解丰度系数
8. $x_i = 0$
9. $p = p - 1$
End % of while
End % of For
End % of For

2 实验结果与分析

分别采用模拟数据和美国 Cuprite 地区 AVIRIS

数据进行实验分析和验证. 在方法对比方面, 主要采用传统 FCLS 算法^[4] (其中模拟数据采用了 Matlab 的 lsqnonneg 函数) 和 SROLSU 进行对比和分析. 实验采用计算机硬件 CPU AMD A6-3420 1.50 GHz, RAM 4 GB, 软件采用 Matlab 2012a. 对于模拟数据, 由于已知丰度真值, 采用信号重构误差 (Signal Reconstruction Error, SRE) 和平均绝对误差 (Mean Absolute Error, MAE) 进行量化比对 (公式 7), 而对于真实数据, 从视觉上进行定性比较和分析.

$$SRE(\text{dB}) = 20 \lg\left(\frac{\|\mathbf{X}\|_F}{\|\mathbf{X} - \hat{\mathbf{X}}\|_F}\right);$$

$$MAE = \frac{1}{N} \sum_{i=1}^N \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_1 \quad (7)$$

上式中, \mathbf{X} 和 $\hat{\mathbf{X}}$ 分别表示丰度矩阵真值和估计值, $\|\cdot\|_F$ 表示矩阵的 F -范数, $\|\cdot\|_1$ 表示第 i 个像元的丰度向量, 表示向量 1-范数. 从公式(7)可以看出, SRE 越大或 MAE 越小, 说明算法丰度求解越准确.

2.1 模拟数据实验

为便于直观的显示和分析, 模拟数据采用在二维空间设定三个端元和一组丰度系数的方式生成, 即通过已知的端元和丰度系数按照线性混合模型(1)合成模拟数据. 假设三个端元在二维空间中为非正单形体(存在钝角)结构, 设定其坐标为端元-1 (10, 40), 端元-2 (50, 20), 端元-3 (100, 30), 生成服从狄利克雷分布 (Dirichlet distribution, 满足 ASC 和 ANC 条件) 的 1 000 组丰度系数, 这 1 000 组丰度系数与端元生成的模拟数据像元点分布在单形体(三角形)内部; 为检验算法的有效性, 随机生成两组外点丰度系数数据, 它们主要分布在三角形的钝角位置, 即端元-2 附近(图 1 中钝角处), 由于锐角处外点类似正单形体条件, 不会出现端元误删情况, 本次实验没有设计相应的外点. 在模拟外点的丰度系数值时, 为实现最终的外点像元分布在钝角处, 需要保证针对钝角处端元丰度系数大于 1, 而另外两个端元的丰度系数有正有负.

实验共生成含 10 个外点和 20 个外点的模拟数据各一组, 其中模拟的 10 个外点满足 ASC 约束, 而另一组 20 个外点的数据为随机生成(不满足 ASC 约束, 为比较接近真实数据, 丰度值介于 -1 ~ 2 之间). 为更真实地反映算法的处理结果, 对模拟数据添加了不同信噪比 ($SNR = \|\mathbf{D}\mathbf{x}\|^2 / \|\boldsymbol{\varepsilon}\|_2^2$) 的高斯白噪声, 生成的第一组模拟数据及其白化处理结果在二维空间的分布如下图 3 所示 ($SNR = 40$), 其中端元-1 为最左侧小圆圈, 端元-2 为中部小圆圈, 是钝

角处端元,端元-3 为右侧小圆圈.可看出,通过单形体正化处理,实现了不规则单形体向正单形体的变换.在模拟数据实验中,由于端元已知,实验中不再采用端元提取算法进行端元提取.

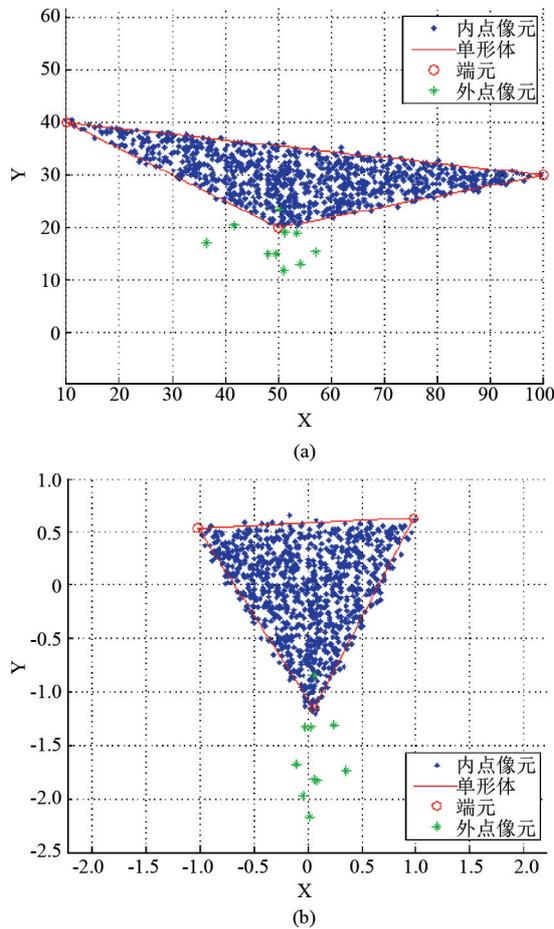


图3 二维空间中第一组模拟数据正化处理(左:原始数据,右:正化后数据, $SNR = 40$ dB)

Fig. 3 The regularization of simulated data for data of group one in 2-dimensional space (Left: the original data. Right: the regularized data, $SNR = 40$ dB)

由于在二维空间中,只能采用二维坐标表示模拟数据的像元,即端元矩阵 $D \in R^{2 \times 3}$,这样在进行解混时出现方程少,变量多的情况,无法求解唯一的丰度系

数.因此,在线性混合模型求解方程中添加了 ASC 约束,这样得到的模拟数据丰度求解方程如下(对于真实数据实验,由于波段数大于端元数,采用经典的 FC SL 算法^[4]求解):

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & e_3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad (8)$$

其中, x, y 表示二维空间中像元坐标, $e_i \in R^2$ 表示第 i 个端元, a_i 表示第 i 个端元对应的丰度系数.

实验分别采用 Matlab 提供的非负约束最小二乘 (lsqnonneg) 函数和 SRLSU 算法进行丰度系数的求解,获取的丰度解与模拟丰度真值进行了量化比较(公式 7).由于不同方法对于像元内点,其结果一致,因此只对外点的丰度解进行精度评价.两组模拟数据的 SRE 和 MAE 评价结果如表 2 所示.可以看出,对于模拟数据,数据正化前采用端元剔除的全约束解混 (b-Proposed) 可以得到更好的结果,这主要是由于这些外点中,多数外点对于锐角处端元的模拟丰度系数为负值,而钝角处为大于 1 的正值,采用 lsqnonneg (b-FCLS) 非负约束求解时,会迭代将负值系数转为非负,导致钝角处端元产生更大的丰度解误差;而 SRLSU 方法剔除了负值丰度对应的端元,在解算大于 1 端元时,拟合值接近于 1,因而误差更小.而对于单形体正化后的数据,两种方法 (a-FCLS 和 a-Proposed) 得到的丰度估计值完全一致,并且精度均得到了提升.对于第一组数据,当信噪比较低 ($SNR = 20$ dB) 时,解混结果出现了异常情况,通过同一组测试数据进行多组测试(噪声发生变化)发现,误差指标 SRE/MAE 出现了不稳定,变化较大,这主要是由于添加的噪声较大,而外点数量较少,造成了噪声对数据的影响增大,在表 2 中展示了第一组数据在 $SNR = 20$ dB 时某一次测试的结果.

第一组数据实验中,还分别计算了 SNR 为 40 dB 时 10 个外点解算得到的丰度解相对于 3 个端元的丰度真值的误差(图 4-a, b, c),即 $\Delta a_{ij} = \hat{a}_{ij} - a_{ij}$,

表 2 两组模拟数据单形体正化前后的解混结果评价

Table 2 Comparison of evaluation index values for unmixing of different simulated data

数据集		原始数据 (SRE/MAE)		单形体正化数据 (SRE/MAE)	
		lsqnonneg (b-FCLS)	SRLSU (b-Proposed)	lsqnonneg (a-FCLS)	SRLSU (a-Proposed)
一组	$SNR = 40$ dB	0.9331/1.387	9.950/0.510	15.175/0.264	15.175/0.264
	$SNR = 30$ dB	0.758/1.443	9.810/0.559	14.53/0.356	14.53/0.356
	$SNR = 20$ dB	2.293/1.264	2.483/1.222	7.717/0.767	7.717/0.767
二组	$SNR = 40$ dB	8.870/0.514	8.912/0.506	9.256/0.454	9.256/0.454
	$SNR = 30$ dB	6.480/0.699	6.488/0.697	6.535/0.682	6.535/0.682
	$SNR = 20$ dB	2.628/1.168	2.628/1.168	2.785/1.140	2.785/1.140

其中 \hat{a}_{ij} 表示估计值, a_{ij} 为丰度真值, $i \in \langle 1, 2, 3 \rangle$ 表示第 i 个端元的系数, $j \in \langle 1, 2, \dots, N \rangle$ 表示外点个数. 另外, 也得到三个端元误差的绝对值之和 $\sum_{i=1}^3 \alpha_{ij}$ (图 4-d). 从图 4 中可看出, 对于端元-1 (图 4-a), 正化前作者采用的端元剔除方法 (b-Proposed) 得到的外点丰度估计值误差相较 lsqnonneg 方法 (全约束最小二乘, 用 b-FCLS 表示) 基本一致, 只在第 9 个外点有所改善, 单形体正化后两种方法 (a-Proposed、a-FCLS) 误差完全相同, 相较正化前 lsqnonneg 方法只在第 9 外点有提升, 但第 10 外点时误差增加. 对于端元-2 (图 4-b), 正化前作者所采用方法相较 lsqnonneg 误差较小, 在进行正化处理后两种方法结果一致, 且误差降至最小. 从端元-3 误差曲线 (图 4-c) 也可看出, 在非正单形体情况下, SRSU 方法相较传统的 FCLS 方法得到了更小的误差, 而单形体正化后的数据采用两种解混方法得到的结果一致. 图 4-d 为三个端元误差的绝对值和, 也可看出, b-FCLS 方法误差最大, b-Proposed 次之, 正化后两种方法误差最小且完全相同, 这也证实了正单形体情况下, 处于图(2)位置的外点可直接删除, 不会对解混结果产生影响. 在模拟数据实验中, 由于数据量较小, 不再进行运算效率的讨论, 而在真实数据实验中进行处理效率的比较.

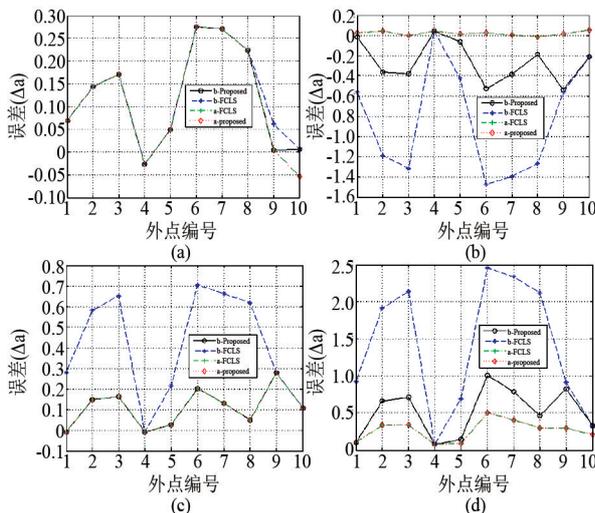


图 4 第一组模拟数据外点不同方法得到的估计值与真值的差异比较 (SNR = 40 dB)

Fig. 4 The error between true and estimated values by different methods for outliers in simulated data of group one (SNR = 40 dB)

2.2 AVIRIS 数据实验

真实数据实验采用美国 Cuprite 地区的 AVIRIS 机载高光谱数据¹, 大小为 250 (行) × 190 (列) × 188

(波段), 于 1997 年完成航飞成像, 已经在高光谱数据应用中进行了大量的实验验证.

实验采用 N-Findr^[12] 端元提取算法得到 10 个端元, 再分别基于 FCLS^[4] 和 SRSU 方法对单形体正化前后的数据进行处理, 得到 10 个端元的丰度系数图像. 限于篇幅, 挑选了 6 种获取的比较典型的矿物端元丰度图像 (如图 5 所示) 进行比较和分析. 从图中可看出, 不同算法得到的结果在区域形态上基本一致, 但也存在差异; 端元的丰度值范围在单形体正化前后发生了变化, 单形体正化前丰度系数的范围为 0 ~ 1, 而正化后丰度最大值均小于 1, 该现象说明: 1) 在真实的高光谱数据中, 并非所有的像元都处于利用 N-Findr 获得的端元形成的单形体之内 (单形体内部的像元白化前后丰度值一致); 2) 单形体正化处理会改变丰度估计值. 另外, 我们对单形体正化后的两种方法的丰度结果进行了差值比较, 结果均在 10 ~ 15 量级, 可认为二者结果相同, 这也验证了模拟数据在进行正化后 10 ~ 15 丰度解相同的结论. 从图 5 中第二端元下部以及第五个端元丰度图可看出, 两种方法的结果存在一定的差异, 这主要是由于两种方法在进行非负解混时策略不同造成, FCLS 需要对所有像元进行迭代解算, 而 SRSU 方法在进行数据的单形体正化前, 采用了直接剔除负值端元的策略, 对于高光谱数据单形体外点, 会导致一定的误差.

从关键的运行效率角度, 对不同方法的处理时间 (Running Time, RT) 进行记录, 其结果如下图 6 所示. 从图中可以看出, 对于 AVIRIS 实验数据, 采用 SRSU 方法可以较大地提升处理效率, 相较传统的 FCLS 算法速度提升达到了 7 倍以上, 其中, 与解混所需的时间相比, 端元白化及正单形体正化处理在整个处理过程中只占据了不到 2% 的时间, 相较 FCLS, 白化时间所占比重更小. 从方法的逻辑上分析, 效率提升主要是因为对于单形体正化后的数据, SRSU 方法直接剔除了丰度负值端元以获得非负约束解, 这样避免传统的 FCLS 算法中 ANC 约束求解的多次迭代, 从而提升了算法的效率.

3 结语

以高光谱数据中端元张成的单形体形态并非均为正单形体为出发点, 考虑外点在单形体外部的空间分布特性, 提出一种基于单形体正化的高光谱数据全约束线性解混方法. 该方法引入了端元白化和单形体正化步骤, 将高光谱数据在高维空间的单形

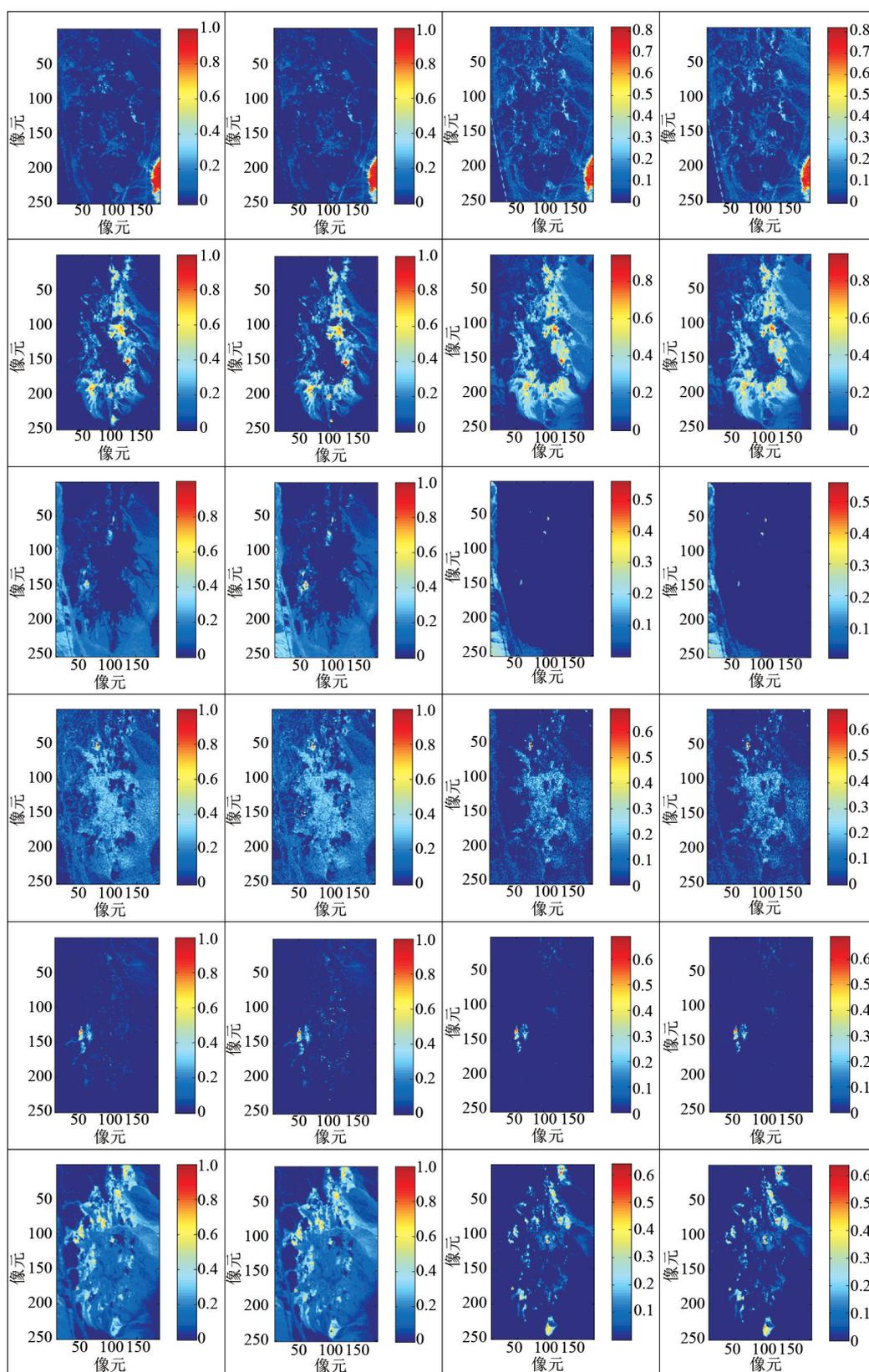


图5 N-Findr获得的六个端元对应的丰度估计图像(a)正化前-FCLS,(b)正化前-SRSLSU,(c)正化后-FCLS,(d)正化后-SRSLSU

Fig. 5 The abundance maps of six endmembers extracted by N-Findr algorithm(a) results of FCLS method before simplex regularization, (b) results of SRSLSU method before simplex regularization, (c) results of FCLS method after simplex regularization, (d) results of SRSLSU method after simplex regularization

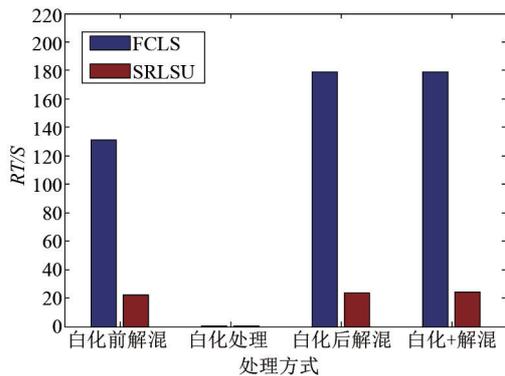


图 6 不同算法及环节的处理时间比较

Fig. 6 Comparison of computing time for different methods and processes

体不规则分布转换为在正单形体形态. 通过正单形体条件下高光谱数据外点空间分布的特性分析, 认为对于 ASC 约束解混得到的负值丰度估计端元, 可直接将该端元剔除, 而不会影响最终的解混结果, 以此替代 FCLS 中的 ANC 迭代求解, 从而提升全约束光谱解混处理效率. 模拟数据和真实高光谱数据的实验结果表明, 单形体正化处理会改变端元张成的单形体外点的丰度估计结果, 而内点保持不变; 模拟数据基于单形体正化的方法可以获得更好的丰度估计结果, 而在真实数据情况下, 丰度估计结果与单形体正化前的丰度估计空间分布基本一致; 通过真实数据处理效率的比较, SRLSU 方法相较传统的 FCLS 方法大大减小了高光谱数据进行线性光谱解混的计算开销, 提升高光谱数据的解混处理效率.

References

- [1] TONG Qing-xi, ZHANG bin, ZHENG Lan-fen. *Hyperspectral Remote Sensing* [M]. Beijing: Higher Education Press (童庆禧, 张兵, 郑兰芬. 高光谱遥感-原理、技术与应用. 北京: 高等教育出版社), 2006.
- [2] BIOCAS-DIAS J. M., PLAZA A., DOBIGEON N. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches [J]. *IEEE Journal of selected topics in applied earth observations and remote sensing*, 2012, **5**(2): 354–379.
- [3] KESHAVA N, MUSTARD J F. Spectral unmixing [J]. *IEEE Signal Process. Mag.*, 2002, **19**(1): 44–57.
- [4] HEINZ D C, CHANG C I. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery [J]. *IEEE Trans. Geosci. Remote Sensing.*, 2001, **39**(3): 529–545.
- [5] TANG Yi, WAN Jian-Wei, XU Ke, et al. Hyperspectral unmixing based on material spatial distribution characteristic [J], *J. Infrared Millim. Waves* (汤毅, 万建伟, 许可, 等. 地物空间分布特性的高光谱遥感图像解混算法, *红外与毫米波学报*), 2014, **33**(5): 560–571.
- [6] BAYLISS J, GUALTIERI J A, CROMP R. Analysing hyperspectral data with independent component analysis [C], in Proc. SPIE, 1997, **3240**: 133–143.
- [7] BOARDMAN J W. Automating spectral unmixing of AVIRIS data using convex geometry concepts [C]. In 4th Annual JPL Airborne Geo Science Workshop. Pasadena: JPL Pub., 1993.
- [8] GENG Xiu-Rui, ZHANG Bin, ZHANG Xia, et al. An unmixing method of hyperspectral imagery based on convex volume in high dimensional space [J]. *Progress in nature sciences* (耿修瑞, 张兵, 张霞, 等. 一种基于高维空间凸面单形体体积的高光谱图像解混算法. *自然科学进展*), 2004, **14**(7): 810–814.
- [9] HONEINE P, RICHARD C. Geometric unmixing of large hyperspectral images: A barycentric coordinate approach [J]. *IEEE Trans. Geosci. Remote Sens.*, 2012, **50**(6): 2185–2195.
- [10] WANG Li-Guo, ZHANG Jing, LIU Dan-Feng, et al. Distance measurement based methods from endmember selection to spectral unmixing [J]. *J. Infrared Millim. Waves* (王立国, 张晶, 刘丹凤, 等. 从端元选择到光谱解混的距离测算方法. *红外与毫米波学报*), 2010, **29**(6): 471–475.
- [11] PU Han-Ye, XIA Wei, WANG Bin, et al. A Fully Constrained Linear Spectral Unmixing Algorithm Based on Distance Geometry [J]. *IEEE Trans. Geosci. Remote Sens.*, 2014, **52**(2): 1157–1176.
- [12] WINTER M E. Fast autonomous spectral endmember determination in hyperspectral data [C]. In: Paper Presented at the Proceedings of SPIE. 1999.
- [13] CHANG C I, WU Chao-Cheng, LIU Wei-Ming, et al. A new growing method for simplex-based endmember extraction algorithm [J]. *IEEE Trans. Geosci. Remote Sens.*, 2006, **44**(10): 2804–2819.
- [14] CRAIG M. Minimum-volume transforms for remotely sensed data [J]. *IEEE Trans. Geosci. Remote Sens.* 1994, **32**(3): 542–552.
- [15] AUDREOU C, KARATHANASSI V. Estimation of the number of endmembers using robust outlier detection method [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.* 2014, **7**(1): 247–256.
- [16] YANG Hua-Dong, AN Ju-Bai, ZHU Chuang. Subspace-projection-based geometric unmixing for material quantification in hyperspectral imagery [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.* 2014, **7**(6): 1966–1975.
- [17] CHAN T H, MA W K, CHI C Y, et al. A convex analysis framework for blind separation of non-negative sources [J]. *IEEE Trans. Signal Process.* 2008, **56**: 5120–5134.
- [18] GENG Xiu-Rui, JI Lu-Yan, ZHAO Yong-Chao, et al. A new endmember generation algorithm based on a geometric optimization model for hyperspectral images [J]. *IEEE Geosciences and Remote Sensing Letters.* 2013, **10**(4): 811–815.