

改进的线性混合模型用于高光谱分离实验模拟*

薛彬^{1,2} 赵葆常¹ 杨建峰¹ 吕群波^{1,2}

(1 中科院西安光学精密机械研究所空间室, 西安 710068)

(2 中科院研究生院, 北京 100039)

摘要 从考虑成份光谱的噪声出发,提出了改进的线性混合模型来补偿成份光谱噪声的影响,并在假定成份光谱和混合光谱噪声源于同一噪声源的前提下,给出了改进线性模型的数学求解过程及数值计算思路,且通过合成的含噪声的混合光谱对两种方法做了对比验证,实验证明:改进的线性混合模型有效地克服传统线性模型的不足,大大减弱了成份光谱的噪声对计算精度的影响。

关键词 谱分离技术;混合光谱;成份光谱;改进线性模型

中图分类号 TN751.1 **文献标识码** A

0 引言

高光谱图像处理很大的一个应用是探测图像覆盖区域内的物质成份。在遥感数据处理中,一个最通用的用来辨识像素是否包含使用者所关心的物种的方法是图像分类,目前国际上有许多分类方法被提出,诸如光谱匹配技术^[1]、光谱角度制图法^[1]、主成份分析法^[1]、人工神经网络法^[1,7]等。然而这些光谱分类方法均受到由于地元分辨率低而引起的混合象素问题的限制(如,美国2001年发射的强力小卫星(MightSatII)上搭载的一台干涉成像光谱仪,其地元分辨率是 $30 \times 30 \text{ m}^2$,而我们感兴趣的目标小于该面积)。由此为了进一步挖掘高光谱图像的信息,发展了一门新的技术——“光谱分离技术”,来预估每个像素的成份组成。

目前,国际上已经提出很多方法模型,其中最为成功的是线性混合模型^[1,5],它假设物质成份在空间布局上是离散片混合(如同围棋棋盘)^[2],忽略在不同类型物质之间的多重散射的量,即每个像元点的光谱幅度仅是这个像元中存在的各物质成份的光谱信号以相对贡献量的线性叠加。

然而对于细小微粒的混合(如盐和胡椒粉混合)^[2],这个假设条件显然是不满足的,故线性关系也就不成立。但是当我们分析目的只是辨识成份的空间分布和一定意义上的物理含量时,可以通过多次散射反照率(Multiple Scatter Albedo, MSA)转化成单次散射反照率(Single Scatter Albedo SSA, Hapke^[2,5,6])的方法,使其“线性化”,这样,线性模型也是可以使用的。

1 传统线性混合模型(Traditional Linear Mixture Model, TLMM)^[3,5]

图1为其原理图,即对于不含噪声的混合光谱

(Pure Mixing Spectrum, PMS)我们假定其严格遵从对成份光谱(Pure EndMember Spectrum, PEMS)^[1]的线性关系。又由于噪声无规律,不在线性关系,对于加入噪声后的混合光谱(Observed Mixing Spectrum, OMS)在反演时,我们可在假定噪声量最小的情况下可以得到各成份光谱的含量。这种模型通常假设:①成份光谱和混合光谱都是在相同的人射光强度下获得;②预知成份光谱要“纯”。

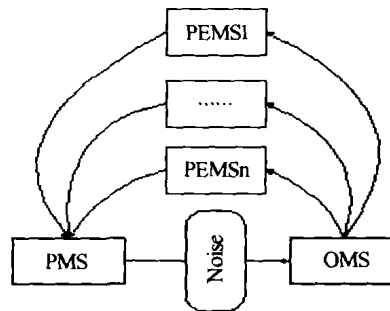


图1 TLMM原理图
Fig.1 Traditional linear mixture model

根据实际情况,混合像元中所含物质成份的含量估计值之和应为1,且不可以小于0,故丰度值 x_{iS} 须加归一化及非负的限制条件。如此,我们就可以用一个约束线性最小二乘问题(Constrained Least Squares, CLS)来描述,即

$$\begin{aligned} \min & \|r\|^2 \\ \text{s. t. } & Ax_{iS} = b + r \\ & \sum x_{iS} = 1, x_{iS} > 0 \end{aligned} \quad (1)$$

式中 b 是一个 $m \times 1$ 的矩阵,表示观察像素的混合光谱; A 是 $m \times n$ 的矩阵,其每一列(A_i)表示成份光谱,是混合像元中某一特定单一物质的光谱,通常 $m \geq n$; x_{iS} 是 $n \times 1$ 的矩阵,表示像元中各物质成份的丰度含量; r 为 $m \times 1$ 的矩阵,表示观察误差。

这儿 $\|r\|^2$ 表示向量 r 的内积,即向量 r 中每个元素的平方和。当矩阵 A 满秩时

$$x_{iS} = (A'A)^{-1}A'b \quad (2)$$

通过上式简单的运算,便可得到观察光谱 b 中成份光谱 A_i 的含量,然而,在实际运用中,会遇到如下困难:

*国家自然科学基金资助项目(No. 40301031)
Tel: 029-88484669 Email: lutai_xb@sohu.com
收稿日期: 2003-06-19

由于测量条件的差异,很难统一,即第一个假设往往不成立;

由于测量误差和模型误差的存在,第二个假设也往往不成立.

所以我们的 A 通常是有很大误差的,这样对于由纯光谱组成的混合光谱用含噪声的成份光谱去分离,误差必定较大,故我们采取了一种新的方法.

2 改进线性模型 (Modified Linear Mixture Model, MLMM) 及其解

2.1 成份光谱的选取

由于上面讲的测量条件及反演算法的不同,导致了光谱库中的成份光谱不纯,这会使模型误差较大,由此我们可以采用了一些方法,如 CCSM、二值编码匹配、SAM、SCM 等方法^[1,7] 在与混合光谱共处的同一幅光谱图(数据立方体)中,来找与光谱数据库中标准物质光谱相匹配的光谱数据,以其组成成份光谱,这样便可保证成份光谱和混合光谱在同一入射光强度下获得.

2.2 改进线性模型 (MLMM)

图 2 为其原理图,由于我们得到的成份光谱含噪声 (Observed EndMember Spectrum, OEMS),所以我们在传统的线性模型上做了一点改动,在成份光谱中加入一个噪声量,即将含噪声的混合光谱,用含噪声的成份光谱去分离.

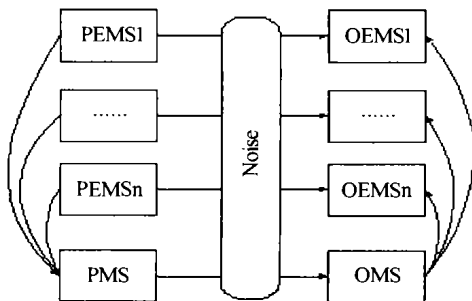


图 2 MLMM 原理图

Fig. 2 Modified linear mixture model

数学模型我们也可以在传统线性模型的基础上仅作一点修正,在纯光谱矩阵 A 上加一个噪声量 E ,其他不变. 这时,将误差定义为 $[E|r]$,于是模型变为

$$\begin{aligned} & \min \|E|r\|^2 \\ & \text{s. t. } (A_0 + E)x_{\text{TLS}} = b + r \\ & \quad \sum x_{\text{TLS}} = 1, x_{\text{TLS}} > 0 \end{aligned} \quad (3)$$

这可以看作一个约束总体最小二乘 (Constrained Total Least Squares, CTLS) 问题^[3,4]. 由于问题的针对性较强,我们可以作如下简化,以方便求解:

①由于我们认为 A_0 和 b 具有线性关系(并非严格的)——该算法的出发点,故可将 $[A_0|b]$ 化为最简非相关矩阵,得

$$[A|O_{m \times 1}] = [\beta_1, \beta_2, \dots, \beta_n] = [a_1 - b_0, a_2 - b_0,$$

$$\dots, a_n - b_0|O_{m \times 1}] \quad (4)$$

②由于 A_0 和 b 的测量是在同种状态下——通过我们上面提到的成份光谱的选取方法可知,即噪声来自同一噪声源,这样我们可以认为其每一光谱附加的噪声是相关的,由此可以得出 $[E|r]$ 的秩为 1,如果定义一个噪声源单位向量 u ,则

$$[E|r] = u[\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n, r] \quad (5)$$

在这两种情况下的,可大大简化了 CTLS 问题的求解. 根据式(5)得

$$\| [E|r] \|^2_r = \|r\|^2 + \sum_{i=1}^n \|\varepsilon_i\|^2 \quad (6)$$

同样,对给定的一个 u (单位向量),则

$$\|d_i\| = |(\beta_i)^2 u| \quad (7)$$

其中 $\beta_i = a_i - b$

$$\|r\| = \frac{\sum_{i=1}^n \|d_i\|}{(n+1)} \quad (8)$$

$$\|\varepsilon_i\| = \|d_i\| - \|r\| \quad (9)$$

将式(8)、(9)代入(6),并整理得

$$Error^2 = \left(\sum_{i=1}^n \frac{\|d_i\|}{(n+1)} \right)^2 +$$

$$\sum_{i=1}^n \left\| \|d_i\| - \frac{\sum_{i=1}^n \|d_i\|}{(n+1)} \right\|^2$$

简化得

$$Error^2 = [\|d_1\| \dots \|d_n\|] \cdot M \cdot [\|d_1\| \dots \|d_n\|]^t \quad (10)$$

式中 $M = I(n) - \frac{1}{n+1} \times \text{Ones}(n)$, $I(n)$ 为 $n \times n$ 单位矩阵, $\text{Ones}(n)$ 为 $n \times n$ 的所有元素均为 1 的矩阵.

根据式(7)中 $\|d_i\|$ 的定义

$$\|d_i\| = (\beta_i)^t \cdot u \Rightarrow Error^2 = u A M A^t u^t \quad (11)$$

其中 u 满足 $u^t u = 1, G u \geq 0$.

取 $u = \frac{Ax}{\|Ax\|}$ 带入上式得

$$Error^2 = \frac{x^t A^t A M A^t A x}{x^t A^t A x} \quad (12)$$

为了进一步简化计算机的运算量,我们还可以作进一步的简化.

①因为 $A^t A$ 是正定的对称矩阵^[3],所以我们将它作 Cholesky 分解,令 $G^t G = A^t A$,其中 G 为上三角矩阵;

②取 $Q = G M G^t$,令 $y = Gx$ 则问题可简化为,找一个 y 使 $y^t Q y$ 最小,受限于 $y^t y = 1$,且 $G^{-1} y \geq 0$. 这样, $y^t Q y$ 的值便是 Q 的奇异值,其最小值是 Q 的最小奇异值,其对应的特征向量便是所求 y ;

③最后对求的 $G^{-1} y$ 归一化即得要求的解 x .

3 模拟实验结果比较

3.1 数据获取

我们从美国 the Department of Geological Sciences of Brown University 的 Relab 数据库中取了

三种物质的光谱(PO-EAC-051 (OLV003 图 3 中实线), SC-EAC-063 (MET101 图 3 中长中断线), CR-

EAC-013(CHR103 图 3 中短中断线)).

表 1 采用的三种物质的光谱信息

样品名称	收集日期	光谱仪类型	起始波长/nm	终止波长/nm	波段数	分辨率/nm
PO-EAC-051	31-Dec-00	BD-VNIR	300	2600	461	5.0
SC-EAC-063	31-Dec-01	BD-VNIR	300	2600	461	5.0
CR-EAC-013	31-Dec-00	BD-VNIR	300	2600	461	5.0

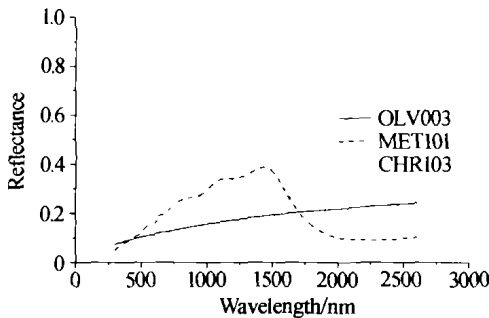


图 3 作为混合像元的三种成份的光谱
Fig. 3 Three PEMS of mixing pixel

3.2 模拟实验过程

由于试验条件限制,我们采用了数值模拟方法来验证,过程如下:

- 1) PEMS:由从 relab 数据库中取得的三种物质光谱来模拟;
- 2) OMS:由三种 PEMS 乘以混合因子(混合比例),加上一定白噪声来模拟;
- 3) OEMS:由 PEMS,加白噪声来模拟;
- 4) 由于在实际情况中,PEMS 是得不到的,我们从高光谱图像(数据立方体)中得到的成份光谱是 OEMS,我们以 OEMS 为成份光谱,分别采用 TLMM 法及 MLMM 法去分离 OMS.

3.3 分解结果比较及分析

取 $X = 0.15 \ 0.45 \ 0.40$ (信噪比为 10,噪声为白噪声)即表 2 中第三项的模拟实验过程见图 4.

表 2 解混前后的成份含量比较

原混合比 X	信噪比 (S/N)	TLMM 求解		MLMM 求解	
		混合比 x	RMS	混合比 x	RMS
0.15, 0.45, 0.40	50	0.1564 0.4495 0.3941	0.0087	0.1558 0.4496 0.3947	0.0079
	20	0.1602 0.4443 0.3955	0.0125	0.1568 0.4444 0.3988	0.0089
	10	0.1859 0.4491 0.3651	0.0501	0.1541 0.4485 0.3975	0.0050
0.15, 0.15, 0.70	50	0.1503 0.1501 0.6995	0.0006	0.1499 0.1501 0.7000	0.0001
	20	0.1537 0.1492 0.6970	0.0048	0.1504 0.1491 0.7005	0.0011
	10	0.1679 0.1504 0.6817	0.0256	0.1549 0.1497 0.6953	0.0068
0.75, 0.15, 0.10	50	0.7461 0.1499 0.1040	0.0056	0.7482 0.1499 0.1019	0.0026
	20	0.7346 0.1513 0.1140	0.0209	0.7462 0.1511 0.1027	0.0048
	10	0.6971 0.1462 0.1567	0.0776	0.7378 0.1452 0.1170	0.0215

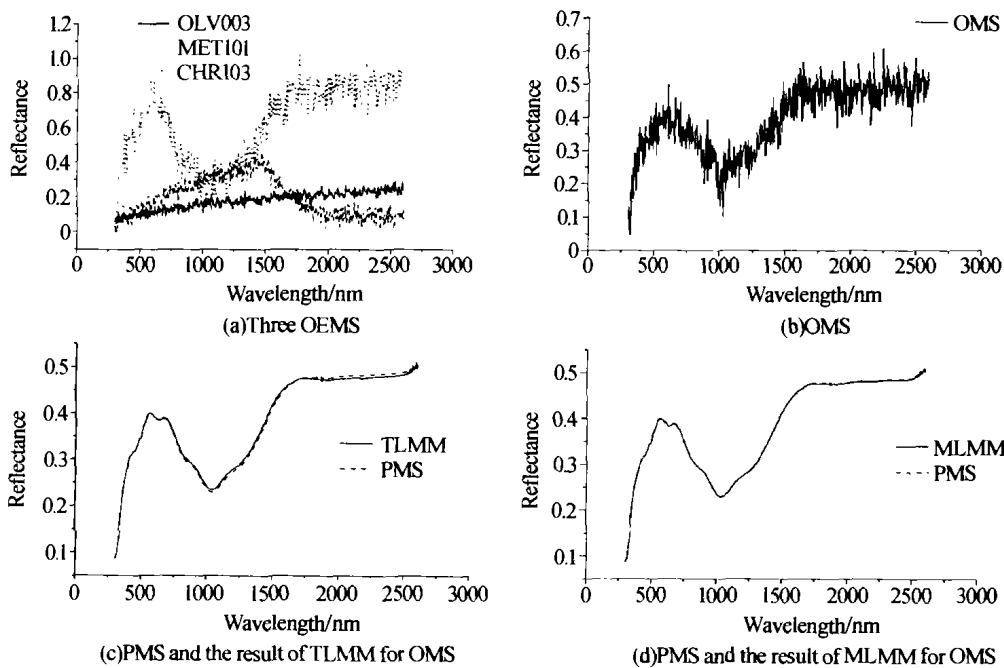


图 4 试验过程

Fig. 4 Experimental process

对表 2 的数据分析如下:

①同一信噪比:由于加入白噪声的随机性,不能

强认为其完全相关,故该改进算法还存在一定的 RMS,但我们可以看出,改进的线性模型 RMS 值要

比传统的线性模型效果好,尤其在信噪比较低的情况下,精度可以提高一倍多。

②不同信噪比:传统线性混合方法受信噪比影响较大,相比之下,而改进的线性混合法受到的影响较小。

4 结论

通过对线性模型及其改进模型的讨论,可以看出线性模型在光谱分解中效果是很好的,但由于成份光谱的缺乏或成份光谱的不纯,往往引起很大的误差,由此我们在传统线性模型的基础上提出了改进的线性模型,并在假定成份光谱和混合光谱噪声是源于同一噪声源的情况下,给出了其解决思路,同时我们利用模拟的混合光谱作了仿真实验比较,最终结果表明:改进的线性模型是更有效的。

参考文献

- 1 浦瑞良,宫鹏. 高光谱遥感及其应用. 北京:高等教育出版社,2000. 47 ~ 78
Pu R L, Gong P. Hyperspectral Remote Sensing and its Applications. Beijing: Higher Education Press, 2000. 47 ~ 78
- 2 Mustard J F, Li Lin, He Guoqi. The Importance of Nonlinear Mixing Modeling for Analysis of Lunar Multispectral Data. *Lunar and Planetary Science*. 1997. XXVIII: 995 ~ 996
- 3 Hu Y H, Lee H B, Scarpace F L. Optimal linear spectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 1999, 37(1): 639 ~ 644
- 4 Abatzoglou T J, Mendel J M. Constrained total least squares. *IEEE International Conf on Acoustics, Speech & Signal Processing*, Dallas, 1987. 1485 ~ 1488
- 5 甘甫平. 高光谱遥感信息提取与地质应用前景——以青藏高原为试验区. *国土资源遥感*, 2000, 45(3): 38 ~ 44
Gan F P. *Remote Sensing for Land & Resources*, 2000, 45(3): 38 ~ 44
- 6 Herzog S G, Mustard J F. Reflectance Spectra of Five-Component Mineral Mixtures: Implications for Mixture Modeling. *Lunar and Planetary Science*. 1996. XXVII: 535 ~ 538
- 7 Tchistiakov V, Ruckebusch C. Neural network modelling for very small spectral data sets: reduction of the spectra and hierarchical approach. *Chemometrics and Intelligent Laboratory Systems*, 2000, 54(1): 93 ~ 106

Modified Linear Mixture Model Applied to the Hyperspectral Unmixing Simulation

Xue Bin^{1,2}, Zhao Baochang¹, Yang Jianfeng¹, Lü Qunbo^{1,2}

¹ Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710068

² Graduate School of Chinese Academy of Sciences, Beijing 100039

Received date: 2003-06-19

Abstract Modified Linear Mixture Model (MLMM) is developed to compensate for the endmembers spectrum's error. Under the assumption that both endmembers spectra and mixing spectrum root from the same noise origin, the mathematical resolution and numerical computing are discussed, and this method is applied to the synthesized mixing spectrum. It is demonstrated that the MLMM eliminates the decline of operational precision caused by the noise of endmember spectrum.

Keyword Spectral unmixing technology; Mixing spectrum; Endmember spectrum; Modified linear mixture model



Xue Bin was born on Aug 1, 1979, in Shandong Province. He received his B. S. degree from Xidian University in 2001. Since then he come to work toward the MS degree in Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences. His main interest focuses on optical engineer.