

# 成象光谱遥感图象的有限光谱 混合分析\*

刘政凯 岑曙炜

(中国科学技术大学电子工程与信息科学系 合肥 230027)

**摘 要** 本文提出了一种成象光谱数据处理的新方法——有限光谱混合分析方法 CSMA (Constrained Spectral Mixture Analysis), 该方法将图象中每一象素看作多种类别地物的混合, 利用成象光谱仪各波段反射值计算各个类别所占的比例, 这是一个求解混合光谱方程最优解的过程。本文提出的有限光谱混合分析方法是在约束条件下求混合光谱方程的最优解, 并提出了快速算法。该方法可以得到比传统遥感分类更有意义的结果, 使遥感技术更趋于定量化、实用化。

**关键词** 成象光谱, 光谱混合分析, 最优解。

## 1 引 言

成象光谱仪是近几年发展起来的一项新技术<sup>[1,2]</sup>, 它的特点是波段多, 光谱分辨率高, 它是遥感技术上的一次突破, 在找矿、水体、生态、大气等方面的遥感定量研究中表现出巨大的实用潜力。成象光谱仪通过探测地面景物的辐射量, 对同一地面景物可获得几十以至几百个波段的图象数据, 波段分辨率可达 10—20 nm。根据成象光谱仪获得的地面景物的近似连续的光谱曲线, 就可以象在实验室里根据物理光谱曲线进行物质化学成分分析一样, 对成象光谱图象中各象素所包含的地物类别进行分析, 显然这在遥感技术的定量化、实用化方面具有极其重要的意义。美国、加拿大、法国、德国等在 80 年代中期就开始研制成象光谱仪, 我国也于 1991 年研制出具有 71 个波段的航空成象光谱仪 MAIS。因此, 开展成象光谱数据分析的研究具有迫切的需要。

## 2 有限光谱混合分析方法

成象光谱仪得到的数据是每一象素所对应地球表面在各个波段的反射率, 由于不同的物体具有不同的反射特性曲线, 因此由成象光谱仪获得的数据即可对象素进行分类, 确定其所属类别。象素各波段的反射率可看成一多维矢量, 应用各种统计模式识别方法即可对这些矢量(即对象素)进行分类, 如最小距离法, 最大似然概率法等, 最近几年还出现了许多改进的适于遥感识别的分类方法, 如人工神经网络方法等等。

\* 本研究得到国家自然科学基金的支持, 特此致谢。

收稿日期: 1995 年 1 月 17 日; 收到修改稿日期: 1995 年 7 月 21 日

但是,成象光谱仪的每一个象素对应地面的大小至少在  $10\text{m} \times 10\text{m}$  的量级(如美国的 AVIRIS 每一象素对应  $20\text{m} \times 20\text{m}$ <sup>[11]</sup>),在这么大的尺寸里,大部分象素并不仅由同一类物体构成,往往是几类均占有一定的比例,传统的分类方法只能将每个象素判作确定一类,因此有必要采用一种新的处理方法,使其能准确计算每个象素中各类所占的比例。采用线性光谱混合分析的模型即能准确地反映各类别的合成情况,进而求解其值。其原理如下<sup>[3-5]</sup>:

以  $DN_b$  表示象素第  $b$  个波段的灰度值 ( $b=1,2,\dots,B$ );  $f_r$  表示该象素中第  $r$  类物体所占的比例 ( $r=1,2,\dots,e$ );  $R_{br}$  表示第  $r$  类物体第  $b$  个波段的反射值,是事先已知的,可以从光谱成象图中采样或者从光谱数据库中得到的。由于整个象素的反射值是各类别反射值依面积的加权平均,因此有

$$DN_b = \sum_{r=1}^e R_{br} f_r + E_b \quad b=1,2,\dots,B \quad (1)$$

此式称为混合光谱方程,其中,  $\sum_{r=1}^e f_r = 1$ ,考虑到地面起伏的影响,此条件可修正为  $\sum_{r=1}^e f_r \leq 1$ ,

$E_b$  为拟合的误差。令  $E = (E_1, E_2, \dots, E_B)^T$ , 当  $\|E\|$  (通常取  $\|E\|^2 = \sum_{b=1}^B E_b^2$ ) 达到最小值时,对应的  $F = (f_1, f_2, \dots, f_e)^T$  即准确地反映了各类别的面积比例。传统的遥感数据(如 Landsat 卫星获得的 TM 数据等)由于波段数较小,很难应用上述模型,而成象光谱仪的多波段数据为这种计算提供了可能。公式(1)中  $F$  的求解可由有关的线性方程组求解理论解决。将(1)式写成矩阵形式,即可知  $F$  是超定方程

$$DN = RF \quad (2)$$

的最优解,其中  $DN = (DN_1, DN_2, \dots, DN_B)^T$ ,  $R$  为由  $R_{br}$  构成的矩阵。此最优解为

$$F = (R^T R)^{-1} R^T DN \quad (3)$$

这里需要特别注意,由此求得的  $F$  并不一定满足  $\sum_{r=1}^e f_r = 1$  的条件,还可能出现  $f_r < 0$  及  $f_r > 1$  的情况,这在实际分析中是不能接受的。如果取

$$f'_r = \begin{cases} 0 & f_r \leq 0 \\ f_r & 0 < f_r < 1 \\ 1 & f_r \geq 1 \end{cases} \quad (4)$$

虽然  $F' = (f'_1, f'_2, \dots, f'_e)^T$  是一个可接受的解,但它却不是(1)式在约束条件

$$\begin{cases} 0 \leq f_r \leq 1 \\ \sum_{r=1}^e f_r \leq 1 \end{cases} \quad (5)$$

下的最优解。为此,我们提出一种有限光谱混合分析方法(CSMA),通过迭代法来求(1)式在此约束条件下的最优解。

先看无约束条件时  $F$  的迭代解法:由(1)式,得

$$E_b = DN_b - \sum_{r=1}^e R_{br} f_r, \quad b=1, 2, \dots, B \quad (6)$$

所以：

$$\frac{\partial E_b}{\partial f_r} = -R_{br} \quad (7)$$

$$\frac{\partial \|E\|^2}{\partial f_r} = \frac{\partial \sum_{b=1}^B E_b^2}{\partial f_r} = \sum_{b=1}^B 2E_b \frac{\partial E_b}{\partial f_r} = -\sum_{b=1}^B 2E_b R_{br} \quad (8)$$

由梯度法则，令

$$F^{(n+1)} = F^{(n)} - \delta \frac{\partial \|E\|^2}{\partial F} \quad (9)$$

其中  $\frac{\partial \|E\|^2}{\partial F} = \left( \frac{\partial \|E\|^2}{\partial f_1}, \frac{\partial \|E\|^2}{\partial f_2}, \dots, \frac{\partial \|E\|^2}{\partial f_e} \right)^T$ ， $\delta$  为迭代步长因子，通常取 0—1 之间的很小正数。

由此得到的最终值满足  $\|E\|^2$  最小，即与 (3) 式得到的结果相同。

在有约束条件时，令

$$\varepsilon = \|E\|^2 + A_1 g_1(F) + A_2 g_2(F) \quad (10)$$

其中  $g_1, g_2$  为约束条件的代价函数，

$$g_1(F) = \sum_{i=1, f_i < 0}^e f_i^2, \quad \frac{\partial g_1(F)}{\partial f_r} = \begin{cases} 2f_r, & f_r < 0 \\ 0, & f_r > 0 \end{cases} \quad (11)$$

$$g_2(F) = \begin{cases} 0, & \sum_{i=1}^e f_i \leq 1 \\ \left( \sum_{i=1}^e f_i - 1 \right)^2, & \sum_{i=1}^e f_i > 1 \end{cases}, \quad \frac{\partial g_2(F)}{\partial f_r} = \begin{cases} 0, & \sum_{i=1}^e f_i \leq 1 \\ 2 \left( \sum_{i=1}^e f_i - 1 \right), & \sum_{i=1}^e f_i > 1 \end{cases} \quad (12)$$

$A_1, A_2$  为常数，当  $A_1, A_2$  取很大的值时， $\varepsilon$  的最小值即为  $\|E\|^2$  在条件 (5) 下的最小值。因此，有约束条件时的迭代式为

$$f_r^{(n+1)} = f_r^{(n)} - \delta \frac{\partial \varepsilon}{\partial f_r} = f_r^{(n)} - \delta \left( \frac{\partial \|E\|^2}{\partial f_r} + A_1 \frac{\partial g_1(F)}{\partial f_r} + A_2 \frac{\partial g_2(F)}{\partial f_r} \right) \quad (13)$$

由上式即可求得满足约束条件的混合光谱分析的最优解。

在用 (13) 式进行迭代时，每次迭代均需对  $r (r=1, 2, \dots, e)$  计算  $\frac{\partial \|E\|^2}{\partial f_r} = -\sum_{b=1}^B 2E_b \cdot R_{br}$ ，在成象光谱数据中，由于波段数  $B$  较大 (几十以至几百)，以至 (13) 式的计算量非常大，因而有必要研究 (13) 式的快速算法，主要是降低  $\frac{\partial \|E\|^2}{\partial f_r}$  的计算量。本文提出的

CSMA 的快速算法原理如下：

由

$$\frac{\partial^2 \|E\|^2}{\partial f_i \partial f_j} = - \frac{\partial \sum_{b=1}^B 2 E_b R_{bi}}{\partial f_j} = 2 \sum_{b=1}^B R_{bi} R_{bj} \quad (14)$$

因此  $\frac{\partial^2 \|E\|^2}{\partial f_i \partial f_j}$  为一常数，记此常数为  $m_{ij}$ ，而  $\|E\|^2$  则是  $f_i$  的二次式。由于当  $F$  取(3)式的解时此二次式达到最小值，如记此解为  $F_m = (f_{m_1}, f_{m_2}, \dots, f_{m_e})^T$ ，则

$$\|E\|^2 = \sum_{i=1}^e \sum_{j=1}^e (f_i - f_{m_i})(f_j - f_{m_j}) + E_0 \quad (15)$$

$$\frac{\partial \|E\|^2}{\partial f_r} = \sum_{i=1}^e m_{ri} (f_r - f_{m_i}) \quad (16)$$

设  $M$  为由  $m_{ij}$  构成的矩阵，则上式可简记为：

$$\frac{\partial \|E\|^2}{\partial F} = M (F - F_m) \quad (17)$$

根据公式(8)和(15)的对比表明，公式(15)中的类别数  $e$  通常取 3 到 4，而公式(8)中的波段数  $B$  通常为几十到几百，因此公式(15)比公式(8)的计算量大大减少。波段数越多，计算量减少越多。

综合上述公式，在(5)式约束下  $F$  的求解过程归结为(部分参数取实验中的取值)：

- ① 由(3)式求  $F_m$ ， $F_m = (R^T R)^{-1} DN$ ，取  $F^{(0)} = F_m$ ， $n = 0$ ；
- ②  $n$  增 1，令  $\delta = \frac{6 \times 10^{-5}}{3n+1}$ ， $A_1 = 6n \times 10^{-5}$ ， $A_2 = 2n \times 10^{-5}$ ；
- ③ 计算  $\frac{\partial g_1(F)}{\partial f_r}$ ， $\frac{\partial g_2(F)}{\partial f_r}$ ， $\sum_{i=1}^e m_{ri} (f_r - f_{m_i})$ ， $r = 1, 2, \dots, e$ ；
- ④  $f_r^{(n+1)} = f_r^{(n)} - \delta \left( \sum_{i=1}^e m_{ri} (f_r - f_{m_i}) + A_1 \frac{\partial g_1(F)}{\partial f_r} + A_2 \frac{\partial g_2(F)}{\partial f_r} \right)$ ， $r = 1, 2, \dots, e$ ；
- ⑤ 若  $n$  未达到迭代次数，转向②，否则  $F = F^{(n+1)}$ 。

在实际应用中，由于此最优解与  $F_m$  相差总是很少，选择适当的  $\delta$ 、 $A_1$ 、 $A_2$ ，只需十几次迭代就能求得满足误差范围的最优解。

### 3 实验结果

取自上海技术物理研究所研制的成象光谱仪 MAIS 在宁夏黄河流域的航空测量数据<sup>[2]</sup>，数据共有可见光至近红外的 30 个波段，对应的波长为 470—1050nm，取样间隔为 20nm。图象的尺寸为 512×475 个像素，地面分辨率约 12m，像素的灰度值为从 0 到 255 的相对值。从图中选取特征较明显的 3 类做为参考值  $R$ ，其中第 1 类为水体，第 2 类为植被，第 3 类为土壤，其反射曲线如图 1 所示。

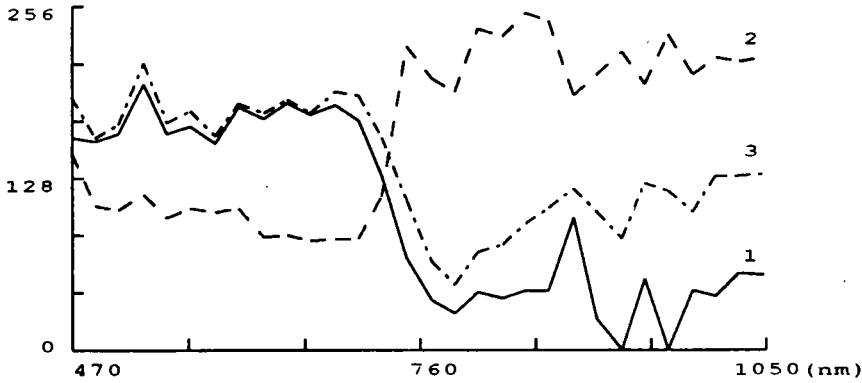


图 1 3 类地物的反射曲线  
Fig.1 Spectrum of 3 classes

图版 I 图 2 是原数据中第 24 波段的图象, 对应的波长为 930 nm; 图版 I 图 3 是第 4、12、21 波段的假彩色合成图 (R 对应 690 nm, G 对应 870 nm, B 对应 530 nm); 图版 I 图 4 是水体的分布图, 各象素的灰度大小即象素的水体含量; 图版 I 图 5 是依象素中最大类别进行显示的结果, 其中, 蓝色为水体, 红色为土壤, 绿色为植被。\$E\_b\$ 的相对平均值为 6.6%,  $\left| \sum_{r=1}^c f_r - 1 \right|$  的平均值为 11.9%, 据 (4) 式取 \$F'\$ 时 \$E\_b\$ 的相对平均值为 9.8%,  $\left| \sum_{r=1}^c f_r' - 1 \right|$  的平均值为 15.6%。实验在 VAX-8700 上进行, 有约束条件时所用的计算时间为 12 分钟, 如若波段数增加, 采用 CSMA 算法的优越性将会更明显。

## 4 结 论

本文提出的有限光谱混合分析法, 是在一定的约束条件下求解混合光谱方程, 它可分析出成像光谱图象中每个象素所含地物主要类别所占的比例, 使遥感数据的解译更趋于定量化、实用化, 这是一个值得重视的研究方向。在实验中, 对一幅 30 波段的遥感图象进行处理只需十几分钟, 基本上可以实用。

由于条件的限制, 缺乏足够的先验知识, 未能对成像光谱仪的图象作大气校正, 因而不能应用野外实测值作为各类的参考反射值, 而只能在图上取样。合理地选择取样类别的个数及样本的位置, 使这些样本能包括图象中的所有类别, 而各个样本又仅由同一类构成, 这对计算结果是非常重要的。因此在实际分析中, 最好对地面实况有一定的了解, 了解的越多, 选取样本也越容易。

同时实验中发现, 在选取了适当的样本后 (一般为 3、4 类), 再简单增加样本数并不能有效地减少  $\|E\|$  的平均值, 这说明上述几类已有效地代表了地物的构成情况, 同时也说明了算法的合理性。

## 参 考 文 献

- [ 1 ] Vane, G., Green, R. O., Chrien, T.G., Enmark, H.T., Hansen, E.G., and Porter, W.M., The Airborne Visible /Infrared Imaging Spectrometer (AVIRIS). *Remote Sensing of Environment*, 1993, **44** : 127 — 143.
- [ 2 ] 薛永祺. 机载扫描成象系统的技术发展. *红外与毫米波学报*, 1992, **11**: 168 — 191 .
- [ 3 ] Kruse, F.A., Lefkoff, A. B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., and Goetz, A.F.H., The Spectral Image Processing System (SIPS)-Interactive Visualization and Analysis of Imaging Spectrometer Data. *Remote Sensing of Environment* . 1993 , **44** : 145 — 163 .
- [ 4 ] Smith, M., Spectral Mixture Analysis-New Strategies for the Analysis of Multispectral Resolution Data. *ISPR*, 1992 Nov., 23 — 27 .
- [ 5 ] Roberts, D.A., Smith, M. O., Adams, J.B., Green Vegetation, Nonphotosynthetic Vegetation, and Soils in AVIRIS Data. *Remote Sensing of Environment*. 1993, **44** : 255 — 269 .

## 作 者 简 介

刘政凯 男, 1940年8月生, 1964年7月毕业于中国科技大学无线电遥控遥测专业, 现为教授, 博士生导师。主要从事遥感图象处理、人工神经网络及模式识别方面的研究, 已发表论文 60 余篇, 出版专著 6 本。

## Spectral Mixture Analysis of Imaging Spectrometer Data

Liu Zhengkai    Cen Shuwei

(*University of Science and Technology of China, Hefei 230027*)

**Abstract** Imaging Spectrometer is a new developed remote sensing instrument. A new processing method of its data-Constrained Spectral Mixture Analysis (CSMA) is presented in this paper. CSMA looks every pixel as mixture of some endmembers. The content of each endmember in a pixel can be calculated using Imaging Spectrometer data. This is a problem of getting optimal solution of the spectral mixture equation. CSMA gets the optimal solution of the equation with some conditions and a fast algorithm is proposed here. This result is more interesting than traditional classification because it is more quantitative.

**Key words** Multispectral mapping, Constrained spectral mixture analysis, Optimal solution