

基于空间角度理论的卫星光学遥感数据认知与挖掘*

马建文 马超飞

(中国科学院遥感应用研究所遥感信息科学开放实验室, 北京 100101)

E-mail: Jianwen@irsa. irsa. ac. cn

摘要 随着科学技术的发展,新的探测器不断地升空,卫星光学遥感数据经历了长周期的积累和更新。如何有效地开发与利用巨量遥感数据,从中挖掘出隐藏的信息和信息知识化是当前面临的严峻挑战。首先分析了地面物质和结构光谱与卫星遥感信息之间的关系,建立了空间角度模型;然后以多维遥感信息的象元矢量和波段趋势面分析为基础,分析了当前用来处理光学遥感数据的商业化“最优化”算法在处理遥感数据时经常出现的“欠优”结果的原因;最后,通过对美国陆地卫星 TM 数据的处理与挖掘说明基于空间角度算法在处理多波段遥感数据时的数学能力。

关键词 光学遥感数据 物质与结构光谱 数据挖掘 信息知识化

0 引言

数据是事物原始信息的载体。数据挖掘与知识发现 MD & KDD (Data Mining and Knowledge Discovery) 是从巨量数据中分离出有效信息,并形成能被人理解知识模式的多阶段高级处理过程。数据挖掘是一项非常复杂的高新技术和系统工程,需要有体现科学规律的概念模型设计^[1]。近年来,MD & KDD 在电子商务、地理信息等方面发展很快。1995年在加拿大召开了第一次 KDD-95 国际会议,至今已经召开了 4 次国际和多次地区会议。1999年5月在我国召开了亚太地区 MD & KDD 国际会议。这些国际会议有力地促进了国际交流与合作,目前 MD & KDD 无论在理论方面还是在实践方面均取得了较大的发展,展示了巨大的发展前景和应用潜力。尽管如此,当前人们利用的遥感信息仅占全部获取信息的 5%,而深层次的信息开发则更少^[2]。因此,巨量遥感数据信息挖掘与知识化是对 MD & KDD 提出的新需求。

卫星遥感数据是通过传感器对地球表面扫描,以数字方式记录的结果。在每个象点的位置,信号被重采样和量化。遥感数字图象是由离散象元集合组成,每个象元有 2 个基本属性:位置和灰度。遥感数据的两个空间维和一个光谱维构成了反映地表能量变化

的三维数字阵列,多波段遥感数据构成一个多个灰度维向量空间。多维向量空间更能直观表达不同物质与表面结构形成的电磁场,因此已成为揭示地球系统各圈层之间的物质循环、能量转换、演化规律的基本工具。由于不同地物的反射、辐射强度不同,相对较弱的信息有时是解释宏观规律的标志信息,是重要的信息。在利用“商业化”遥感数据处理软件挖掘这些弱信息时存在一定的困难。因为① 较弱信息在较强的反射、辐射场中被强信息抑制,这种效应被称之为临地效应或单元效应^[3-7];② 地物光谱在可见光和短波红外波段的变化是连续的,遥感多个波段的光谱域信息更能表达连续变化的地物光谱特征,但多波段数据的色彩表达只能通过三个波段的合成完成,限制了多维遥感数据场中的有效信息^[8-10]。尽管如此,遥感界专家还是从不同的方面探索表达地表能量场与遥感数据场之间的固定模式,以便能快捷、真实地反演遥感图数据中的特征信息,如植被指数 NDVI,二向性反射模型 BRDF^[11,12],临地效应模型和尺度效应模型等的研究。

商业化图象处理软件是建立在统计意义上的“最优化”基础上的,经典统计学在数据采集方式、数据类型、属性特点等方面均与遥感数据有较大的差别(见表 1),因此软件算法在深入挖掘遥感数据中弱化信息时难以出现“最佳”效果^[13]。本文介绍了我们在研制开发“适用性”算法方面的初步研究结果。

* 本文研究受中国科学院“九五”重大和特别支持项目 (No. KZ95T-03-04-05) 资助

收稿日期:1999-10-27

表 1 统计学数据特点与遥感数据特点比较

类 型	采样方式	数据类型	属性特点	边界条件	维数	其 它
统计样本	直接	离散	确定	清晰	低	无
遥感数据	高空感知	离散	不确定	过度	高	临地效应,尺度效应,相同和相异效应,大气,地形等因素

1 多维向量空间角度关系

1.1 柯西-许瓦兹不等式

利用柯西-许瓦兹不等式说明两个三维向量间的关系^[14,15]。下面用到它时,简称为(C.-S.)不等式。现先定义向量的长度,设 x 是内积空间 V 的任一向量,则非负实数 $\sqrt{(x,x)}$ 称为向量 x 的长度,并记为 $|x|$,即定义向量 x 的长度为:

$$|x| = \sqrt{(x,x)} \quad (1)$$

由于向量的内积 (x,y) 在内积空间中是个实数,因此利用长度概念,(C.-S.)不等式又可以表示为:

$$|(x,y)| \leq |x| \times |y| \quad (2)$$

当 x,y 都不是零向量时,由此不等式可得:

$$\frac{|(x,y)|}{|x| \times |y|} \leq 1 \quad (3)$$

即:

$$-1 \leq \frac{(x,y)}{|x| \times |y|} \leq 1 \quad (4)$$

因此,可以用等式:

$$\cos\varphi = \frac{(x,y)}{|x| \times |y|} \quad (5)$$

来定义两个非零向量 x,y 的夹角 φ ,且限制 φ 的取值范围为 $0 \leq \varphi \leq \pi$ 。当 $(x,y)=0$ 时,则称 x,y 是正交的,且记为 $x \perp y$ 。

1.2 洛伦兹协变定理

多维数据向量角度空间是由标量、矢量、张量定义的。改变一个角度会不会使原来坐标系定义的物理量发生异变?这个问题需要引入洛伦兹协变定理^[16]。(1)坐标中矢量各分量的变换规律同坐标的变换规律一定是密切相关的;(2)同一点上两个矢量的分量的具体数值可以各不相同,但都遵循相同的规律进行变换;(3)在坐标变换过程中标量、矢量和张量等物理量具有协变的性质。下面是二维坐标角度变换的几何图形与变换矩阵(见图 1)。

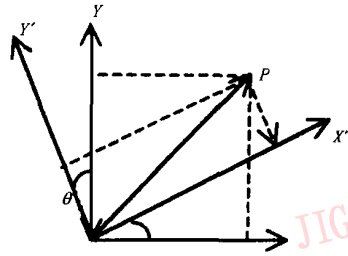


图 1 坐标变换几何图形

坐标 XY 转过 θ 角变成新坐标 $X'Y'$,任一点 P 的坐标变换关系是:

$$\begin{cases} X' = \cos\theta \times X + \sin\theta \times Y \\ Y' = -\sin\theta \times X + \cos\theta \times Y \end{cases} \quad (6)$$

(6)式中包括 4 个变换系数,排列成矩阵形式为:

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad (7)$$

(7)式矩阵称为坐标变换矩阵。

2 多维遥感数据场角度分解模型

一个波段遥感数据记录的象元灰度值可以看成象元的一维向量。多波段遥感数据的象元向量维数是多维数的,象元向量的表示为:

$$DN_{ij}^{[k]} = [DN_{ij1} DN_{ij2} \cdots DN_{ijk}]^T = \begin{bmatrix} DN_{ij1} \\ DN_{ij2} \\ \vdots \\ DN_{ijk} \end{bmatrix} \quad (8)$$

(8)式中 i,j 为象元的位置下标, k 代表维数,也就是波段数。同样对于整个波段的数据也可以当做一个向量,表示为:

$$DN_k^{[p]} = [DN_{k1} DN_{k2} \cdots DN_{kp}]^T = \begin{bmatrix} DN_{k1} \\ DN_{k2} \\ \vdots \\ DN_{kp} \end{bmatrix} \quad (9)$$

(9)式中 p 为整幅图象的象元数,如果图象的大小为 $M \times N$,则 $p=M \times N,k$ 为波段数。

上述两种表示方法的基向量是不同的,当整幅

图象表示为一个向量时,维数可以表示为图象的象元数;对于象元的 k 维向量,一组基向量为一个单位矩阵,维数为 k 。以3个波段为例,象元向量在三维坐标空间的位置^[17]如图2所示。

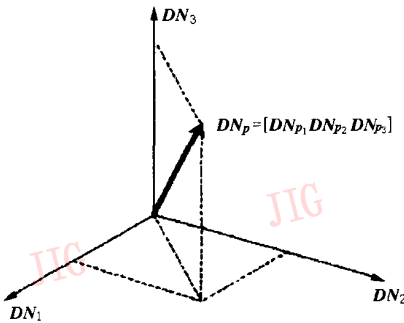


图2 遥感数据象元向量在三维坐标空间的位置

建立了象元和波段向量关系的矩阵和几何表示,就可以利用向量空间变换的数学工具完成遥感数据的特征选择、处理结果的物理解释和指导开发多波段遥感数据处理的新方法。在线性空间中,可以进行向量的加和数乘运算,讨论向量的相关性、向量组的秩和线性空间的基。但是,如果要对向量进行度量、定量分析时,需要在内积空间中进行,内积空间中可以定义向量的长度、距离、正交性,并且可以构造空间的标准正交基。遥感数据场的角度分析是基于内积空间象元向量和波段向量的分析。

3 图象分析与处理

从挖掘多光谱遥感数据中有效弱信息的目的出发,探索了基于遥感光谱多维角度变换的遥感数据分析方法。

3.1 多因子逐步正交变换

多因子逐步正交变换是矩阵对角化分解,对多波段遥感数据方差协方差矩阵 A 进行多步旋转,每一步旋转特征轴表示的角度不同,以达到分离目标信息的效果。

$$C = p_1 p_2 p_3 \cdots p_n$$

C 是正交矩阵, $C^T = C^{-1}$ 。上式表示 C 可分解为一系列简单正交矩阵的乘积,其中:

$$p_i = U_{pq} \tag{10}$$

(10)式矩阵表示为:

$$\begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & 1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cos\theta & \cdots & \sin\theta & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & 1 & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & -\sin\theta & \cdots & \cos\theta & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 1 & 0 & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{bmatrix} \tag{11}$$

(11)式矩阵中对角线上省略元素为1,非对角线上省略元素为0。得到:

$$\begin{aligned} b_{pp} &= a_{pp}\cos^2\theta - 2a_{pq}\sin\theta\cos\theta + a_{qq}\theta\sin^2\theta \\ b_{qq} &= a_{pp}\sin^2\theta + 2a_{pq}\sin\theta\cos\theta + a_{qq}\cos^2\theta \\ b_{pq} &= b_{qp} = (a_{pp} - a_{qq})\sin\theta\cos\theta + \\ &\quad a_{pq}(\cos^2\theta - \sin^2\theta) \\ b_{pj} &= a_{pj}\cos\theta - a_{qj}\sin\theta \quad j \neq q \\ b_{qj} &= a_{pj}\sin\theta + a_{qj}\cos\theta \quad j \neq p \\ b_{ij} &= a_{ij} \quad i, j \neq p, q \end{aligned} \tag{12}$$

为了使(12)式非对角元素为0,即:

$$\begin{aligned} u_{pp} &= u_{qq} = \cos\theta \\ u_{pq} &= -u_{qp} = \sin\theta \\ u_{ii} &= 1 \\ u_{ij} &= 0 \quad i \neq j \\ B &= u^T A u \end{aligned} \tag{13}$$

(13)式 U_{pq} 矩阵的作用就是将主对角线上的元素不断增加,非对角线上的元素的值不断降低,每一次旋转,就得到 p 和 q 两个因子方差最大的正交轴,也就是使得 p 和 q 两个新向量之间互相独立,协方差为零。通过两两建立相互正交轴,反复迭代,最终能够获得全部的方差最大的正交轴,当方差协方差矩阵旋转成对角矩阵时,得到的特征值和特征向量为成分分析的结果。通过对旋转过程中变量载荷因子的分析,得到不同旋转位置在原始数据的载荷分配,从而达到突出和提取有用目标信息的目的。

3.1.1 多波段遥感数据角度变换统计数据分析

以华北地台北缘峪耳崖金矿带为例说明多因子逐步正交变换处理图象的效果:图象大小为236km²,TM数据有7个波段,其中波段6属热红外波段,不参与旋转分析。首先对原始数据进行统计分析,计算6个波段的方差协方差矩阵。统计结果如表2所示。

表 2 6 个波段的均值和协方差矩阵

均值:					
80.768070	37.249001	40.608448	69.195240	97.684002	47.537426
协方差矩阵:					
53.967155	33.535265	69.660571	-15.800603	117.254421	103.295697
33.535265	23.264519	45.931809	0.750338	84.171626	69.506811
69.660571	45.931809	99.826045	-22.889488	171.121876	151.388773
-15.800603	0.7050338	-22.889488	140.418399	43.231309	-25.636412

根据原理中的公式(14)计算旋转矩阵。

$$b_{pq} = b_{qp} = 0$$

$$(a_{pp} - a_{qq})\sin\theta\cos\theta + a_{pq}(\cos^2\theta - \sin^2\theta) = 0$$

$$\text{ctg}(2\theta) = \frac{a_{pp} - a_{qq}}{2a_{pq}} \quad (14)$$

在旋转矩阵的作用下,每次旋转产生一个载荷因子矩阵,即因子轴在原始数据上的权重。旋转过程中,坐标轴两两旋转, $U_{12}, U_{13}, U_{14}, U_{15}, U_{16}, U_{23}, U_{24}, U_{25}, U_{26}, \dots, U_{56}$ 为一个周期,反复迭代,最后得到方差协方差矩阵的对角线元素为特征值,非对角线上元素为 0 的矩阵,旋转因子的乘积为主成分的载荷矩阵,迭代数据如下:

(1) 旋转因子的乘积为主成分的载荷矩阵,数据迭代 3 周期 45 次之后,得到的旋转特征值和特征向量。

原始的方差协方差矩阵通过迭代旋转得到的结果矩阵具有下列特点:对角线上的元素为特征值,非对角线上的元素趋于 0。如表 3、表 4 所示。

表 3 对角线上的元素为特征值的特征向量

2.48054	-0.05615	-0.09361	-0.04251	0.00056	-0.00124
-0.05615	0.65072	-0.03876	0.00268	-0.12310	0.00010
-0.09361	-0.03876	165.98595	0.00641	0.00001	0.00008
-0.04251	0.00268	0.00641	18.97274	0.00003	0.00000
0.00056	-0.12310	0.00001	0.00003	775.23418	-0.00329
-0.00124	0.00009	0.00008	0.00000	0.07493	7.03870

表 4 非对角线上的元素趋于 0 的特征向量

2.48054	-0.05615	-0.09361	-0.04251	0.00056	-0.00124
-0.05615	0.65072	-0.03876	0.00268	-0.12310	0.00010
-0.09361	-0.03876	165.98595	0.00641	0.00001	0.00008
-0.04251	0.00268	0.00641	18.97274	0.00003	0.00000
0.00056	-0.12310	0.00001	0.00003	775.23418	-0.00329
-0.00124	0.00009	0.00008	0.00000	0.07493	7.03870

(2) 旋转矩阵的乘积为特征向量与主成分分析

结果,特征值比较表 5 与表 6。

表 5 旋转矩阵的乘积为特征向量

0.62667	-0.26724	0.13091	0.60439	0.23414	-0.31406
-0.03573	0.91651	0.01341	0.36302	0.16146	-0.02660
-0.73550	-0.28107	0.18056	0.48219	0.33796	0.02671
0.01180	-0.09252	-0.91278	0.28556	0.00712	0.27665
-0.04715	0.02779	-0.28101	-0.41289	0.69985	-0.50771
0.25037	-0.01641	0.19478	-0.13567	0.56130	0.75210

表 6 主成分分析变换特征值和特征向量

特征值:					
18.966984	0.683514	2.4816441	65.986014	775.234213	7.010476
特征向量:					
0.61662	-0.29996	-0.61098	-0.13063	0.23415	-0.29090
0.35896	0.91499	0.01864	-0.01323	0.16132	-0.08596
0.47192	-0.25826	0.74928	-0.18106	0.33801	0.04631
0.28453	-0.07294	-0.00455	0.91276	0.00716	0.28376
-0.41080	-0.00696	0.04012	0.28101	0.69980	-0.51078
-0.13602	0.02728	-0.25163	-0.19464	0.56136	0.75128

3.1.2 因子载荷空间特征变化趋势面

通过以上分析可以看出:不同角度的因子轴,在原始数据上的载荷不同。从而整个旋转过程构成了一个向量载荷空间。为了直观地显示多因子逐步变换对区分岩性信息的效果,在因子轴当中,选择主要包含岩性信息的 TM5、7 波段(5,6 因子轴)的因子作为主因子轴做载荷空间曲面(如图 3,图 4)。图 3 选择 TM3-X, TM5-Y, TM7-Z 三维坐标,曲面为第 5 个因子轴在三个波段上的载荷分布。其中最高值分布区为主成分载荷,在该区域相关性最小,这时在 TM5 上的载荷为中等 $Y=0.70$ 。岩性信息主要分布在 TM5 贡献较大的区域, $Y=0.85$ 。图 4 选择 TM1-X, TM7-Y, TM5-Z 三维坐标,曲面为第 6 个因子轴在三个波段上的载荷分布。其中最低值分布区为主成分载荷,在该区域相关性最小。岩性信息主要分布在 TM7 上的载荷为 $Y=0.81$, TM5 贡献较大的区域, $Y=-0.38$ 。

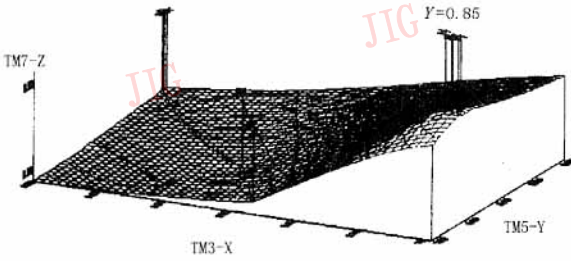


图 3 因子轴 5 在原始波段 TM5(Y)、3(X)、7(Z)上的载荷空间曲面图

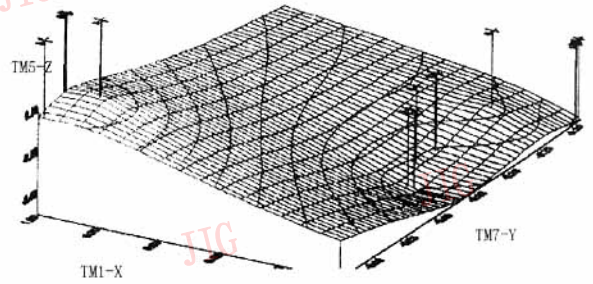


图 4 因子轴 7 在原始波段 TM7(Y)、1(X)、5(Z)上的载荷空间曲面图

根据上述分析选择载荷因子 526 合成因子图象,将含矿超基性岩有效地区分出来,为以后的地质详细勘测提供了决策依据。

4 多维向量空间角度投影

在多维向量空间内,任何一组相互独立的向量,都可以利用正交投影方法找到一组该向量的正交基,构造算法如下:

设:

$$\{\eta_1, \eta_2, \dots, \eta_i\}$$

是一组相互独立的向量,利用下列方式构造该向量组的正交基。

$$\begin{aligned} \beta_1 &= \frac{\eta_1}{\|\eta_1\|_2} \\ \beta_i &= \frac{\eta_i - \eta_i - \hat{\eta}_i}{\|\eta_i - \hat{\eta}_i\|_2} \\ \hat{\eta}_i &= \sum_{k=1}^{i-1} \langle \eta_i, \beta_k \rangle \beta_k \end{aligned} \quad (15)$$

多波段遥感数据是遥感传感器所记录的地表反射的地物目标的电磁能量,它是有限的,即平方可积的离散函数。我们将整个波段 $M \times N$ 大小的图象当

$$\begin{aligned} E\{\hat{TM}_k(TM_k - \hat{TM}_k)\} &= E\left\{ \left[\sum_{j=1}^n E(TM_k)\eta_j \right] \left[TM_k - \sum_{j=1}^n E(TM_k\eta_j)\eta_j \right] \right\} \\ &= E\left\{ TM_k \sum_{j=1}^n E(TM_k\eta_j)\eta_j \right\} - E\left\{ \sum_{i=1}^n E(TM_k\eta_i)\eta_i \sum_{j=1}^n E(TM_k\eta_j)\eta_j \right\} \\ &= \sum_{j=1}^n [E(TM_k\eta_j)\eta_j]^2 - \sum_{i=1}^n \sum_{j=1}^n E(TM_k\eta_i)E(TM_k\eta_j)E(\eta_i\eta_j) \end{aligned} \quad (19)$$

根据标准正交基的性质 $E(\eta_i\eta_j)=0$,得到上式为 0,所以 $\beta_i \perp \beta_j$ 。 β_1 为 TM 波段 1 归一化的结果, β_2 为 TM 波段 2 和波段 1 的矢量差,依次类推。 β_i 为 TM_i 到 $\beta_1, \beta_2, \dots, \beta_{i-1}$ 所组成的正交空间的最小均方距离的方向。通过实践证明,该方法在分离新疆康西瓦铅-锌-金-银矿化弱信息和区分林地和耕地信息方

成一个 $M \times N$ 维的向量,所有波段的遥感数据组成一组。利用正交投影方法构造多波段遥感数据的正交向量。这种正交分解方法从几何上讲,每一个正交量是去掉相关分量后的剩余分量。而 η_i 在 $\beta_1, \beta_2, \dots, \beta_{i-1}$ 上的投影分量是在 $\beta_1, \beta_2, \dots, \beta_{i-1}$ 组成的向量空间中对 η_i 的最小均方估计。原始 TM 影像 TM1,2,3,4,5,7 被看作向量空间线性随机变量序列。将遥感多波段数据 TM 数据波段 1,2,3,4,5,7 正交分解为 $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_7$ 的算法步骤如下:

$$\beta_1 = \frac{TM_1}{\|TM_1\|} \quad (16)$$

$$\beta_k = \frac{TM_k - \hat{TM}_k}{\|TM_k - \hat{TM}_k\|} \quad k = 2, 3, \dots \quad (17)$$

(17)式中 \hat{TM}_k 是 TM_k 在 $\beta_1, \beta_2, \dots, \beta_{k-1}$ 组成的线性子空间上的投影。

$$\hat{TM}_k = \sum_{i=1}^{k-1} \langle TM_k, \beta_i \rangle \beta_i \quad (18)$$

分解后, β_i, β_j 相互垂直,为了证明这一点,只要证明和 $TM_k - \hat{TM}_k$ 内积的数学期望为 0 即可。设 $\eta_1, \eta_2, \dots, \eta_n$, 为一组标准正交基,则:

面体现了较强的能力。

5 结束语

从巨量遥感数据中挖掘有用信息是遥感技术应用长期以来所追求的目标。应用 MD & KDD 的原

则和思想,深入分析多波段遥感数据空间向量的角度特征和数学模型,可以帮助我们进一步理解数据特点,根据数据特点选取有效的数学处理和信息挖掘工具。可以从多波段遥感数据中挖掘出更多信息,从而促进遥感信息的深层次开发与利用。

参 考 文 献

- 1 陈述彭. 数字地球战略及其制高点. 遥感学报, 1999, (3).
- 2 童庆禧. 遥感信息机理研究. 前言. 北京: 科学出版社, 1998.
- 3 田国良. 遥感辐射特性研究的现状与展望. 见: 遥感科学新进展, 北京: 科学出版社, 1995.
- 4 Cracknell A P. Synergy in remote sensing-What's in a pixel? Int J Remote Sensing, 1998, 19(11): 2025~2047.
- 5 马建文. 基于临地效应、尺度效应的多波段遥感数据深层信息分离——概念、方法和应用. 中国科学院遥感信息科学开放研究室 1998 年报, 1999.
- 6 Charles Ichoku, Arnon Karnieli. A review of mixture modeling techniques for Sub-pixel land cover estimation. Remote Sensing Reviews, 1996, 13: 161~186.
- 7 Eklund P W, Kirkby S D, Salim A. Data mining and salinity analysis. Int J Geographical Information Science, 1998, 12(3): 247~268.
- 8 马建文. 基于临地效应的多光谱数据深层弱信息分离方法与应用. 见: 第十届全国遥感学术会议论文集, 1999.
- 9 马建文. 利用 TM 数据快速提取含金蚀变带方法研究. 遥感学报, 1997, 1(3): 208~213.
- 10 Ma Jianwen, Slaney V R, Harris J. Use of landsat TM data for the mapping of limonitic and altered rocks in the sulphurets area, central British columbia. In: Proceedings of 14th Canadian Symposium on Remote Sensing, Calgary, Alberta, 1991, 419~422.
- 11 Li Shuqing. Electronmagnetic scattering from two adjacent cylinders, IEEE Transactions on Geosciences and Remote Sensing, 1998, 36(6): 1981~1985.
- 12 Justics C O. The moderate resolution imaging spectro-radiometer(MODIS): Land remote sensing for global change research, IEEE Transactions on Geosciences and Remote Sensing, 1998, 36(4): 1228~11249.
- 13 马建文. 从空间数据到科学知识的转变. 科学周刊, 1999, (3).
- 14 罗家洪. 矩阵分析引论. 广州: 华南理工大学出版社, 1997.
- 15 陈大新. 矩阵理论. 上海: 上海交通大学出版社, 1997.
- 16 郑庆璋. 相对论与时空. 太原: 山西科学技术出版社, 1998.
- 17 张贤达. 信号处理中的线性代数. 北京: 科学出版社, 1997.

Understanding and Mining of Optical Satellite Data Based on the Space Angel Theory

Ma Jianwen, Ma Chaofei

(Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing 100101)

Abstract With the development of science and technology, new remote sensing devices are lunched on schedule. Optical satellite data have been undergone a long time renewal and accumulation. How to reveal hidden information from huge amount of remote sensing data and let it knowledgeable is new challenge for us. In this article, first we analysis the relation between material, structure spectral and remote sensing information, formulate space angel model. Then based on pixel vector and band vector tendency face of multi-band remote sensing information, we analysis the results of “not proper” in the cases of mining some information from remote sensing data when using the software algorithms which are selected according to “the bast optimum” of the statistics. Finally, we illustrate the mathematical principles and effectiveness of algorithms that we used in mining multi-band TM data.

Keywords Optical satellite data, Material, Structure spectral, Data mining, Information knowledgeable