

陆地卫星数据的 LBV 变换

曾志远

(中国科学院南京土壤研究所)

LBV TRANSFORMATION OF LANDSAT DATA

Zeng Zhiyuan

(Institute of Soil Science, Academia Sinica, Nanjing)

自从陆地卫星发射以来,已经提出了不少关于陆地卫星数据信息提取的方法^[3-6],其中最重要的为 Karhunen-Loève 变换和 Kauth-Thomas 变换。这些变换基本上都是先用已有的数学统计方法(主成分分析等)对图象数据进行运算而不管其结果如何,仅仅在通过运算得到新变量之后,将新变量与实际地物对照,才能大致地确定这些新变量的意义。

本研究提出的新方法,是先分析实际地物的具有物理意义的遥感数据(辐射值),进而总结出地物的三个主要遥感特性(辐射特性),然后再考察已有的数学方法,看是否有一种数学方法能提取出地物的这三个主要遥感特性,从而决定用该种数学方法来提取。这样设计出来的信息提取方法在不少方面优于已有的信息提取方法。

一、地物遥感特性分析

迄今为止,通用的陆地卫星数据主要是 CCT 磁带数码。此外,有时也使用图象扫描或测量密度值。但无论前者或后者,都只是一些相对数,而不是有明确意义的物理量。新方法设计的第一个想法就是要使用具有明确物理意义的遥感数据来进行分析。因此,我们先将非物理量——磁带数码或图片密度值转换为物理量——辐射值。

将磁带数码值转换为辐射值的计算公式,依赖于数据的卫星号、接收站、接收时段和数据的分辨率(0—255, 0—127 或 0—63)。其通式为:

$$R_i = \frac{D_i}{D_{i\max}} (R_{i\max} - R_{i\min}) + R_{i\min} \quad (1)$$

式中 D_i 为波段 i ($i = 4, 5, 6, 7$) 的磁带数码值; $D_{i\max}$ 为波段 i 数据分辨率的最高数码值,即 255, 127 或 63; $R_{i\max}$ 为波段 i 的最高辐射值; $R_{i\min}$ 为波段 i 的最低辐射值; R_i 即为要求的波段 i 的辐射值。

对于图片(透明正负卫片或洗印卫片)扫描或测量密度值 D_i , 可按作者所建议的方法计算辐射值^[1]。

我们现在来分析地物的辐射随波段变化的规律——地物的波段辐射曲线。

设有一地物,它在陆地卫星 4 个波段中的辐射值分别为 R_4 、 R_5 、 R_6 和 R_7 。以各波段的某一波长值 λ_4 、 λ_5 、 λ_6 和 λ_7 (通常是波段中心的波长值,也可为该波段的其它波长值,例如 0.55, 0.65, 0.75 和 0.90 微米)为横坐标值,以各波段相应的辐射值 R_4 、 R_5 、 R_6 和 R_7 (毫瓦/厘米²·立体角)为纵坐标值,可以点绘出该地物在 $\lambda-R$ 平面中的一条曲线(实为折线)。我们可以称它为波段辐射曲线。

分析世界上一些代表性地物的波段辐射曲线,我们可以得到如下的印象:

第一个印象是,各种地物的波段辐射曲线在 $\lambda-R$ 平面中有些处于较高的位置,即它们的总的辐射水平较高;另一些则处于较低的位置,即它们的总的辐射水平较低。这就是说,地物的总的辐射水平的高低是地物的第一个重要遥感特性(辐射特性)。

我们得到的第二个印象是,各种地物的波段辐射曲线在 $\lambda-R$ 平面中有一些向左倾斜,即它们的红外辐射相对较强,可见辐射相对较弱;另外的一些则向右倾斜,即它们的红外辐射相对较弱,可见辐射相对较强。这就是说,红外辐射与可见辐射的强弱对比关系,或它们的辐射平衡关系是地物的第二个重要遥感特性。

我们得到的第三个印象是,各种地物的波段辐射曲线从一个波段到另一个波段变化的方向和强度不一样。有些地物的波段辐射曲线,从第 4 波段到第 5 波段是上升的,从第 5 波段到第 6 波段是下降的,从第 6 波段到第 7 波段又转为上升;有些地物的辐射曲线,从第 4 波段到第 5 波段是下降的,从第 5 波段到第 6 波段是上升的,从第 6 波段到第 7 波段也是上升的等等。而且,各曲线的上升下降的速度(曲线的陡度)也不相同。由此看来,辐射随波段变化的方向和程度,是地物的第三个重要的遥感特性。

正是地物的这三个遥感特性在遥感数据或遥感图象上有敏感的反应,它们也正是我们分辨地物的主要依据。

二、信息提取原理与计算方法

我们已经看到, R_4 、 R_5 、 R_6 和 R_7 的物理含义都是辐射值,只是分属于不同的波段而已。使用回归分析的方法,可以将 R_4 、 R_5 、 R_6 和 R_7 都分解为具有不同物理意义的三部分 l_i 、 b_i 和 v_i ($i = 4, 5, 6, 7$); 然后,将各波段意义相同或相近的那一部分分别组合起来,构成几个新的具有清楚物理含义的新变量。

首先,经研究发现,各波段辐射值被这样分解而成的三部分: l_4 、 l_5 、 l_6 、 l_7 ; b_4 、 b_5 、 b_6 、 b_7 ; v_4 、 v_5 、 v_6 、 v_7 具有如下的性质:

(1) l_i 的性质

$$\text{数学性质 I: } l_4 = l_5 = l_6 = l_7 \quad (2)$$

我们用 L_0 来表示它们,即

$$L_0 = L_4 = l_5 = l_6 = l_7 \quad (3)$$

此 L_0 的物理意义是它能反映一个地物的总的辐射水平的高低。即前述地物的第一个重要遥感特性。它是地物最一般的重要遥感特性。

(2) v_i 的性质

$$\text{数学性质 II: } v_4 + v_6 = -(v_5 + v_7) \quad (4)$$

数学性质 II₂: v_4 和 v_6 的数值同符号; v_5 和 v_7 的数值同符号; v_5 和 v_7 的数值的符号与 v_4 和 v_6 的数值的符号相反。

数学性质 II₃: v_4 、 v_5 、 v_6 和 v_7 的数值(绝对值)的大小成固定的比例,即

$$v_4:v_5:v_6:v_7 = a:b:c:d \quad (5)$$

式中 a 、 b 、 c 、 d 是一些固定的常数。它们与地物无关,不随地物各波段辐射值的大小而改变;而是只与作波段辐射曲线时选定的坐标系的横坐标的值有关。在 λ_4 、 λ_5 、 λ_6 、 λ_7 分别选定为 0.55, 0.65, 0.75, 0.90 的情况下,

$$a:b:c:d = 0.15:0.42:0.35:0.08 \quad (6)$$

数学性质 II₃ 的推论 I: 在 λ_4 、 λ_5 、 λ_6 、 λ_7 选定的情况下, a 、 b 、 c 、 d 的绝对值的大小顺序固定不变。例如在取 $\lambda_4 = 0.55$, $\lambda_5 = 0.65$, $\lambda_6 = 0.75$, $\lambda_7 > 0.85$ 的情况下, a 、 b 、 c 、 d 的绝对值的大小顺序为:

$$|b| > |c| > |a| > |d| \quad (7)$$

数学性质 II₃ 的推论 II: 若 v_5 为 0, 则 v_7 必为 0, 且 v_4 和 v_6 亦必为 0; 若 v_6 为 0, 则 v_4 必为 0。

v_4 、 v_5 、 v_6 和 v_7 具有上述这些数学性质对信息提取是极端重要的。这意味着,我们只需知道其中任何一个的数值便可确切地知道其它三个的数值。也就是说,其中任何一个数值都包括了 4 个数值所包含的信息的全部。因此,虽然从原有不同波段的 4 个辐射值分解出来的 v_4 、 v_5 、 v_6 、 v_7 也有 4 个值,但原则上说我们可以取其中任何一个作为新变量来代表它们 4 个。这大大有利于降低被处理的新变量的个数或数据的维数。

考虑到整个数学性质 II 及其推论,我们可以定义一个特殊的和,即

$$V_0 = -v_4 + v_5 - v_6 + v_7 \quad (8)$$

来作为新变量。

此 V_0 的物理意义是它能反映一个地物在整个卫星接收波段 (MSS 波段) 范围内从一个波段到另一个波段辐射变化的方向和强度,即前述地物的第三个遥感特性。而这个特性又反映地面植被与裸地的对比关系,或者说植被的复盖度和生活强度的总情况。若 V_0 为负,则地面植物密度较大,植物生长旺盛; V_0 的负值的绝对值愈大,则植被的总状况愈好。若 V_0 为正,则地面较裸露;正值愈大,地面裸露程度愈大。因此, V_0 是地面植物生产量的良好指标。 V_0 还是水体混浊度的良好指标: V_0 值愈大,水愈清,反之则愈浑。此外 V_0 还有其它的作用。

(3) b_i 的性质

数学性质 III₁: b_i 是另外两个量 B_0 和 C_0 的函数。即

$$\left. \begin{aligned} b_4 &= (B_0 - 0.55)^2 C_0 \\ b_5 &= (B_0 - 0.65)^2 C_0 \\ b_6 &= (B_0 - 0.75)^2 C_0 \\ b_7 &= (B_0 - 0.90)^2 C_0 \end{aligned} \right\} \quad (9)$$

数学性质 III₂: b_i 的值主要取决于 B_0 , 且 C_0 可以表示为 B_0 的近似函数。

数学性质 III₁ 和 III₂ 的推论: b_i 的信息主要包含在数 B_0 中,而 C_0 是可以忽略的。

由上述情形来看,可以取 B_0 为另一个新变量。

此 B_0 的物理意义是它能反映一个地物的可见光辐射与红外光辐射的对比关系,或者说可见光与红外光的平衡关系,即前述地物的第二个重要遥感特性。这一平衡值 B_0 在其高值区反映地面水分的多少。 B_0 值愈大,水分状况愈好。当地面完全为水占据时,具有最高的一些 B_0 值。因此 B_0 是地面水体存在与水分含量的一个良好指标。在 B_0 中的低值区, B_0 亦可以反映地面植物密度和生活强度的情况,即 B_0 值愈小,植物状况愈好。但这一关系不是十分稳定。因此, B_0 仅可作为地面植物状况的一个辅助指标。 B_0 还有其它的功能。

下面来介绍 L_0 、 B_0 、 V_0 等的计算方法。

对任何一地物或卫星图象上的任一点,以其 λ_4 和 R_4 , λ_5 和 R_5 , λ_6 和 R_6 , λ_7 和 R_7 分别组成数据对,以 λ_4 、 λ_5 、 λ_6 、 λ_7 为自变量,以 R_4 、 R_5 、 R_6 、 R_7 为倚变量,作出一元二次多项式回归方程

$$\hat{R}_i = a + b\lambda_i + c\lambda_i^2 \quad (10)$$

式中 \hat{R}_i ($i = 4, 5, 6, 7$) 是实际 R_i 值的回归估计值。

此回归曲线的最低点的纵坐标值即第一个新变量 L_0 (毫瓦/厘米²·立体角); 其最低点的横坐标值即第二个新变量 B_0 (微米); 其回归残差(即 v_4 、 v_5 、 v_6 和 v_7) 的特殊定义和 [式 (8)] 即为第三个新变量 V_0 (毫瓦/厘米²·立体角)。通过推导简化并改变形式, 可以得如下的几个计算 L_0 、 B_0 、 V_0 的简单方程式:

$$\left. \begin{aligned} V_0 &= -0.457604R_4 + 1.28129R_5 - 1.06774R_6 + 0.195271R_7 \\ C_0 &= 19.3411R_4 - 14.1550R_5 - 21.5375R_6 + 13.0811R_7 \\ B_0 &= (30.6010R_4 - 19.6827R_5 - 31.9311R_6 + 16.8103R_7)/(2C_0) \\ L_0 &= 11.9112R_4 - 6.35144R_5 - 11.2071R_6 + 5.3179R_7 - C_0B_0^2 \end{aligned} \right\} \quad (11)$$

式中 C_0 是第四个新变量。但 C_0 的信息基本上已包括在 L_0 、 B_0 、 V_0 之中, 主要是包括在 B_0 之中。因此, 实际图象处理中仅使用 L_0 、 B_0 、 V_0 就够了。这里 C_0 仅作为计算 B_0 和 L_0 的一个桥梁。一般也不必输出。

方程组 (11) 的系数是固定的, 不依卫星号、接收站、接收时段、象幅或数据分辨率而变更, 更不依赖于特定的地域条件。所以使用起来是很方便的。

考虑到世界范围内各种地物的 L_0 、 B_0 和 V_0 的数值分布性质, 本方法设计了二次方程拉伸—压缩模式, 使 L_0 的数值在中低值部分得到拉伸, 在高值部分得到压缩; 设计了三次方程拉伸—压缩模式, 使 B_0 和 V_0 , 特别是 B_0 的数值在中值部分得到拉伸, 在低值和高值部分都得到压缩。同时, 为了图象处理系统计算机处理的方便, 模式设计中还考虑到要使信息提取最后输出的值 L 、 B 、 V 都是 0—255 的整数, 综合拉伸—压缩模式为:

$$\left. \begin{aligned} L &= 245L_0 - 56L_0^2 \\ B &= 486 - 2950B_0 + 5428B_0^2 - 2714B_0^3 \\ V &= 128 - 320V_0 + 300V_0^3 \end{aligned} \right\} \quad (12)$$

式中, L 、 B 、 V 分别为 L_0 、 B_0 、 V_0 拉伸—压缩后取值范围为 0—255 的数值。需要注意, V 公式的设计使得当 V_0 具有最小值(绝对值最大的负值)时 V 具有最大值, 当 V_0 具有最大值(最大的正值)时 V 具有最小值。这是为了合成“真”彩色卫星图象的需要。

最后, 由于现在通用的磁带数据仍然是以数码的形式存在, 而使用本方法进行信息提

取又必须先获得经转换而成的辐射值,所以在我们的实际图象处理运算中,我们并不使用包含有辐射值 R_i 的方程组 (11), 而是使用直接从现有磁带上读取数码值的综合公式。

综合公式的推导是将方程 (1) 代入方程组 (11), 经过整理简化得到如下新的方程组:

$$\left. \begin{aligned} V_0 &= v_{00} + v_{04}D_4 + v_{05}D_5 + v_{06}D_6 + v_{07}D_7 \\ C_0 &= c_{00} + c_{04}D_4 + c_{05}D_5 + c_{06}D_6 + c_{07}D_7 \\ B_0 &= (b_{00} + b_{04}D_4 + b_{05}D_5 + b_{06}D_6 + b_{07}D_7)/(2C_0) \\ L_0 &= l_{00} + l_{04}D_4 + l_{05}D_5 + l_{06}D_6 + l_{07}D_7 - C_0B_0^2 \end{aligned} \right\} \quad (13)$$

式中, $D_i (i = 4, 5, 6, 7)$ 是波段 i 的磁带数码值, $v_{0i}, c_{0i}, b_{0i}, l_{0i} (i = 0, 4, 5, 6, 7)$ 都是常数。将各陆地卫星、各接收站和各接收时段的 $R_{i\max}, R_{i\min}$ 以及各种数据分辨率的 $D_{i\max}$ 等分别代入方程 (1), 然后将方程 (1) 代入方程组 (11), 即可算得这些常数的值。

依据方程组 (13) 的各套系数, 已分别编成不同的子程序。使用者可根据自己的具体磁带数据的陆地卫星号、接收站、接近时段和数据分辨率, 直接调用自己所需要的子程序, 非常方便。

对于卫片扫描密度数据, 使用者须输入该卫片的各灰标等级的密度值, 以便计算机根据设计程序自动计算出密度—灰标—辐射值转换方程, 然后算出辐射值, 再用方程组 (11) 进行信息提取。

三、新变量与新方法的意义

新变量 L, B, V 的意义前面已有叙述。在这里, 分别根据其意义, 我们将 L 命名为地物的总辐射水平值 (general radiance level); 将 B 命名为地物的可见—红外辐射平衡值 [visibleinfrared (radiation) balance]; 将 V 命名为地物波段辐射变化方向和强度(速度)值 [band (radiance) variation direction and speed] 或波段辐射变化向量值 [band (radiance) variation vector]。

本方法提取出来的新变量, 与陆地卫星原有 4 个波段的数据相比, 有如下优点: (1) 数据维数降低, 一般只需使用 3 个变量。由于在许多图象处理运算中, 维数降低一维, 运算量就成倍地减少, 故此法将大大降低图象处理的成本。(2) 物理含义明确清晰。(3) 分辨地物的能力强。(4) 提供了世界范围内不同地区、不同时间、不同卫星、不同象幅、不同数据分辨率的图象及其处理结果定量比较的可能性。

与 Karhunen-Loève 变换相比, 这种新的信息提取方法有下述优点: (1) 计算非常简单, 这将进一步大大降低图象处理的成本。(2) 转换方程的系数是已知的, 不随地域条件和所用象幅或其子区而改变, 不需要每次图象处理时都重新计算转换方程的系数。(3) 新变量的物理含义明确。(4) 分辨地物的能力强。(5) 所得图象及其处理结果有世界范围内的数值可比性。

与 Kauth-Thomas 变换相比, 这种新信息提取方法的优点是: (1) 转换方程的系数不依赖于特定地域条件和原始试验点的选择, 且其基本方程的系数不因卫星不同而变化。(2) 具有广大范围直至世界范围内的适用性。(3) 新变量物理含义明确。(4) 分辨地物

的能力强。(5) 所得新图象及其处理结果具有世界范围的数值可比性。

本方法提取出来的第三个新变量 V , 其物理含义独特, 分辨地物能力很强。无论是 Karhunen-Loève 变换, 还是 Kauth-Thomas 变换, 或是其它变换, 都没有产生出这样的类似变量来。本方法提取出来的第二个新变量 B , 在对于水分和水体以及某些其它地物的分辨方面, 也比较独特和有效, 为其它方法所不及。本方法提取出来的第一个新变量 L , 比较接近于前两种变换提取出来的第一个变量[常被叫做亮度 (brightness)], 但所含信息更加明确单纯。

新变量 L 、 B 、 V 用于彩色合成时, 比原来的 4 个波段数据的任何三个波段的组合都好。因为这三个新变量包含的信息比原波段任何三个的组合包含的信息都多。它们正好能充分利用颜色的三原性(红、绿、兰)。不仅如此, 由于 L.B.V. 的特定性质, 它们还能合成出一种近于自然色彩的“真”彩色合成图象来。

新变量用于密度分割, 可得到有明确特征地物分级的结果。如植物密度等级、水分含量等级等。

此外, 目前困扰着许多遥感科学家的小于象元或不可分辨的云, 同时使用 L 和 V 可以识别。

作者已经将本方法应用到世界范围内的资料, 取得了良好结果。部分结果见文献 [2]。其它将另文发表。

参 考 文 献

- [1] 曾志远, 1981: 卫星图象密度—辐射量算与土壤资源探测。自然资源, 第 4 期, 52—56 页。
- [2] 曾志远, 1988: 新的卫星数据信息提取方法在世界范围内的应用。第 5 届全国遥感技术交流会论文集(广州, 深圳), 154—159 页。
- [3] Gray, T. I. and McGray, D. G. 1981: The Environmental Vegetation Index, a Tool Potentially Useful for Arid Land Management. Agristars Report, No. EW-NI-04076, JSC-17132.
- [4] Kauth, R. J. and Thomas, G. S. 1976: The Tasseled cap —A Graphic Description of the Spectral-temporal Development of Agricultural Crops AS Seen by Landsat. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette Indiana.
- [5] Santisteban, A. and Munoz, L. 1978: Principal Components of a Multispectral Image: Application to a Geological Problem. IBM J. Res., Develop., Vol. 22, No. 5, September.
- [6] Schowengerdt, R. A., 1983: Techniques for Image Processing and Classification in Remote Sensing. Academic press. INC. pp. 167—170.

新书出版消息

《砂姜黑土综合治理研究》, 张俊民主编, 1988 年, 安徽科技出版社, 332 页, 5.65 元。由南京土壤研究所地理室、编译室和安徽省土肥所销售。