

辅助数据在遥感影像计算机分类中的应用

杨凯 陈军

(武汉测绘科技大学)

1986年1月17日收到

遥感影像计算机自动分类技术的基本原理,是根据地物类别在遥感影像上反映出来的光谱、纹理和时相等特征而对地物进行分类的,并在影像上确定每一地物类别的平面分布。以光谱特征为基础的模式识别方法,是当前遥感应用中最常见的计算机自动分类方法,它是根据不同地物类别的不同反射光谱特性、以地物影像光谱特征的相似性作为判别标准来分类的。多年来的研究和实践表明,这种分类方法有一定的局限性。因此,除了纯光谱分类算法自身的改进和在分类中逐步考虑地物影像的纹理特征之外,人们还致力于研究各种辅助数据在遥感影像计算机分类中的应用,以不断提高分类精度和分类过程中的人工智能水平^[1-2]。

辅助数据的来源较为广泛,除地形图、像片判读草图、土壤分布图、植被图、地质图等各种图件资料外,还包括有关的地面实测数据和统计资料等。人们在根据需要选择好有关辅助信息资料之后,往往要对之进行数字化处理,使之与数字遥感影像几何配准。在某些场合下,还需要根据这些原始的辅助数据,生成新的辅助数据。

辅助数据在遥感影像分类处理中的作用表现为多种形式,包括对遥感影像进行分类前的预处理、直接参加分类判别、对分类结果进行后处理以及判定和评定类别属性等。本文总结了笔者近年来的研究实践^[3-5],着重介绍辅助数据在遥感影像计算机分类中应用的三种方法。

一、配合有太阳入射角数据的分类法

遥感影像任一像元的太阳入射角,定义为相应地物表面法线方向与太阳入射方向之间的夹角。从遥感影像地物辐射亮度的一般表达式^[6]

$$L_p = \frac{1}{\pi} \rho(\lambda)(E_{s0} \cos \theta + E_{\lambda d}) + L_u \quad (1)$$

可知,每一像元辐射亮度 L_p 不仅取决于地物的光谱反射特性 $\rho(\lambda)$, 而且取决于摄影时刻的大气条件 $E_{\lambda d}$ 和 L_u 以及以太阳入射角 θ 表征的地表起伏情况。对单时域遥感影像来说,大气的影响虽使得影像反差降低,但当地形影响可以忽略不计时,具有相同光谱反射特性的地物仍将具有相同的辐射亮度值。然而,在地形影响不可忽略的情况下,一方面光谱反射特性不同的地物可能具有相同的辐射亮度,另一方面光谱反射特性相同的地物可能具有不同的辐射亮度。图1和图2分别示出了湖北省九宫山地区1978年10月16日成象的陆地卫星MSS6影像和与之几何配准的太阳入射角影像(图版IV)。这两个影

像之间存在着惊人的相似之处,直观地证实了山区地形对遥感影像有很大的影响。因此,单纯利用地形起伏较大地区的原始 MSS 影像分类,必然导致一些分类错误。表 1 给出了对九官山地区原始 MSS 影像进行非监督分类后,采用“系统-随机抽样”方法^[7]进行精度评定后得出的混淆矩阵。从中不难看出,纯 MSS 分类结果中混类现象严重,主要原因之一是山区地物辐射亮度受地形影响很大。为了改善分类精度,必须消除地形的影响。把太阳入射角数据作为新加入的特征子空间,与原始 MSS 影像一起参加分类,这样就考虑到地形的影响了。大致有以下几种情况:

表 1 纯光谱分类混淆矩阵 I

Table 1 Confusion Matrix of Pure Multispectral Classification

实地类别 分类结果	针 杉	针 楠	楠	阔	疏 灌	非 林
A	20.68	3.13	7.62	10.34	23.53	0
B	17.24	12.5	10.47	3.45	0	0
C	13.81	50	56.19	41.38	0	0
D	17.24	28.12	18.1	3.45	11.76	0
E	17.24	6.25	7.62	41.38	11.76	1.76
F	0	0	0	0	47.05	17.64
G	13.79	0	0	0	5.9	80.6
Σ	100	100	100	100	100	100

(1) 如果两像元在灰度空间中十分接近,而它们的太阳入射角数值相差较大时,则不能认为这两个像元对应的地物具有相同的光谱反射特性,因而不宜分至同一集群类别。引入了太阳入射角特征量后,就有可能把这两个像元分至不同的集群类别;

表 2 配合有太阳入射角数据的多光谱数据分类混淆矩阵

Table 2 Confusion Matrix of Multispectral Classification with solar incident Angle Data

实地类别 分类结果	针 杉	针 楠 (针阔)	楠	阔	疏 灌	非 林
针 杉	76.47	15.38	6.54	4.54	7.69	0
针楠(针阔)	11.76	55.38	9.34	6.82	0	0
楠	5.89	23.08	79.43	29.54	11.54	0
阔	2.94	3.08	4.69	65.85	15.38	0
疏 灌	0	3.08	0	2.29	61.54	11.76
非 林	2.94	0	0	0	3.85	88.24
Σ	100	100	100	100	100	100

(2) 如果这两个像元的太阳入射角数值和光谱数值都很接近,则有理由认为它们属于具有相同或相近光谱反射特性的地物,可以分为同一类别;

(3) 如果这两个像元的太阳入射角数值及光谱数值都相差较大, 则将它们分至两个不同的集群类别, 至于这两个类别是否属于同一种地物类别, 尚需进一步分析判断。

按此方法对原始陆地卫星 MSS 四个波段数据和相应的太阳入射角数据进行非监督分类, 并将那些太阳入射角数据统计均方根差较大的类别, 进一步细分为若干个子类别, 以有助于使每一分类类别所含的像元都具有比较接近甚至相同的光谱数值和光照条件。此外, 根据山南基本上没有楠竹而山北一侧成片以及楠竹一般只生长在海拔 800 米以下地区的特点, 利用边界信息和数字地面模型, 对分类结果进行了后处理。表 2 给出了利用太阳入射角数据后的分类混淆矩阵。从中可以看出, 主对角线上的元素, 均大于同一列的其它元素, 这表明分类精度大有改善。

二、辐射校正分类法

本方法将改变每一像元的辐射亮度值, 亦即在分类前按下式^[4]对原始影像进行辐射校正, 以改正地形的影响:

$$\left. \begin{aligned} L &= \frac{(L_p - L_u)(\cos\theta_0 + H_p E_{\lambda d}/E_{u0})}{\cos\theta_p + H_p E_{\lambda d}/E_{u0}} + L_u & \theta < 90^\circ \\ L &= \frac{(L_p - L_u)(\cos\theta_0 + H_p E_{\lambda d}/E_{u0})}{H_p E_{\lambda d}/E_{u0}} + L_u & \theta \geq 90^\circ \end{aligned} \right\} \quad (2)$$

式中: L_p 为某一像元的辐射亮度, 一般可由影像灰度值换算而得^[8];

L 为校正后的像元辐射亮度值;

S_p 和 θ_p 是该像元相应的地面坡度角和太阳入射角, 可由 DEM 计算得出;

$$H_p = 1 - \frac{S_p}{\pi};$$

S_0 和 θ_0 是校正参考平面的坡度角和太阳入射角, 可预先选定;

$E_{\lambda d}/E_{u0}$ 和 L_u 一般未知, 笔者曾提出过利用陆地卫星 MSS 数据和地面坡度数据估求这两个参量的方法^[4]。

按(2)式对影像进行辐射校正, 实质上是把所有像元的辐射亮度值校正到统一的参考平面上去, 以消除地形起伏对影像的影响。图 3 和图 4 分别示出了校正后的 MSS5 和 MSS6 影像(图版 IV)。对校正后的影像进行非监督集群分类处理后可知, 地形对集群迭代结果的影响亦大有消除。分类混淆矩阵(表 3)表明, 辐射校正分类法的分类结果比起纯 MSS 分类大有改善。

值得指出的是, 从表 2 和表 3 可以看出, 配合有太阳入射角数据的分类结果, 比辐射校正法的分类结果要好一些。其原因之一, 是在根据 MSS 数据、地面太阳入射角数据和坡度角数据估求 $E_{\lambda d}/E_{u0}$ 和 L_u 这两个参数的过程中, 造成误差的因素较多; 此外, 在推导公式(2)的过程中所做的关于地物具有朗伯反射特性的假设不一定合理, 而影响了辐射校正影像的质量。这尚有待于进一步研究解决。

表 3 辐射校正分类混淆矩阵

Table 3 Confusion matrix of Classification of radiometric-corrected images

实地类别 分类结果	针 杉	针 楠 (针楠)	楠	阔	疏 灌	非 林
针 杉	40	11.11	2.83	8.57	4.76	0
针楠(针阔)	3.33	41.66	10.37	0	0	0
楠	36.67	38.89	74.52	5.71	4.76	0
阔	0	2.78	1.91	65.72	4.76	0
疏 灌	16.67	2.78	10.37	20	66.67	5.56
非 林	3.33	2.78	0	0	19.05	94.44
Σ	100	100	100	100	100	100

三、按高程分层分类法

由于自然界中某些不同地物的光谱辐射亮度值,在没有地形影响的情况下也是相同的或相近的,因此,即使在分类中正确地改正了地形影响,也仍然难以根据地物的光谱特性把它们区分开来。当某些地物类别在生长环境上具有按高程不同作垂直分布的特点时,高程信息将成为混类中再分类的附加信息。前面已提到笔者曾利用植被垂直分布信息对分类结果进行后处理,以区分一些光谱反射特性相同或相近的地物的情况。按高程分层分类法,是根据 DEM 数据和代表不同地面类别的样区数据,统计各主要地物类别的垂直分布特点,继而把研究区域划分成若干高程带,分别对每一高程带的遥感影像进行分类处理,然后把各高程带的分类影像叠加,形成整个研究区域的分类结果。其原理是在似然分类中考虑第 h 高程带中 i 类地物出现的条件概率 $P(w_i/r_h)$ 。在无高程信息引入的情况下,似然分类的概率判决函数由下式表出^[7]:

$$d_i(X_k) = L_n |\Sigma_i| + (x_k - M_i)^T (\Sigma_i)^{-1} (x_k - M_i) - 2 \ln p(w_i) \quad (3)$$

其中: $d_i(X_k)$ 是像元 K 在光谱特征空间中相距第 i 类中心的加权距离;

x_k 是像元 K 的波谱亮度向量;

Σ_i 是第 i 类的协方差矩阵;

$P(w_i)$ 是第 i 类地物出现的先验概率,一般用 i 类地物占整幅图像的面积率来估计。

显然,当图像中不同区域内各类地物出现的概率 $p'(w_i)$ 与 $p(w_i)$ 相差较大时,将会导致局部分类的不准确。引入 $p(w_i/r_n)$ 后,由于 $p(w_i/r_n)$ 具有局部统计意义,它比 $p(w_i)$ 更符合地物分布的实际,因而按高程分层分类优于整体分类。我们在利用陆地卫星 MSS 数据进行湖北省咸宁县土地利用状况测定时^[9],根据该县 DEM 数据和六个样区的地面数据,把全县划分为五个高程带,以每个高程带中某地物类别出现的先验概率近似地代替 $p(w_i/r_n)$,对五个高程带的特征影像(由两个时期的生物量指标影像组成)分别进行分类处理,并对样区范围内的整体分类和按高程分层分类的结果与样区实地类别数

表 4 引进高程信息前后分类试验结果的混淆矩阵 P
Table 4 Confusion Matrixes of Multispectral Classification before and after Introducing of DHM data

项目	C_j		混 合 类 别										S_j	
	w_i	P_{ij}	耕	林	水	草 (林)	林 (草)	耕 (林)	耕 + (林)	林 (耕)	耕 (水)	水 (耕)		(耕) + (草, 林)
BIO ₈₉ 图 象 整 体 分 类	耕		72.8	11.8	16.3			60.6	45.7			43.2		
	林		12.2	72.1	5.8		24.1	42.7				6.8		
	水		5.9	0.3	76.9		3.6	0.7				49.1		
	草		3.1	10.0	0		3.4	6.1				0		
	园		2.6	2.3	1.0		6.1	3.7				0.5		
	城		3.4	0	0		2.1	0.4				0.5		
	裸		0	3.5	0		0.1	0.7				0		
		S_i	3822	4642	104		6403	3440			220			
BIO ₈₉ 高 程 带 分 层 分 类	耕		80.5	8.6	9.9		58.6	50.1		26.0	62.3	29.0	36.2	61.3
	林		7.9	78.6	0	34.7	26.8	41.9		64.7	6.2	0	27.6	10.9
	水		4.3	0.2	90.1	0	2.8	0.7		0.5	23.7	66.4	0.8	11.6
	草		2.8	7.0	0	62.7	3.2	1.9		2.9	3.4	0.9	29.1	10.3
	园		2.7	2.1	0	0	6.4	3.3		2.8	0.6	0	5.5	2.8
	城		1.7	0	0	0	2.1	0.8		0.2	3.9	3.7	0	0
	裸		0	3.5	0	2.7	0.1	1.4		2.9	0	0	0.8	0
		S_i	4975	2394	91	75	6739	855	3772	355	107	254	604	

据进行比较。从分类混淆矩阵(表 4)可知,引入高程信息后分类精度有所提高,其中单一类别的平均分类精度提高了 7.5%。

四、其它辅助数据的应用

在遥感应用研究中,通常采用地面实况资料,辅助确定各分类类别所对应的实地地物类别。

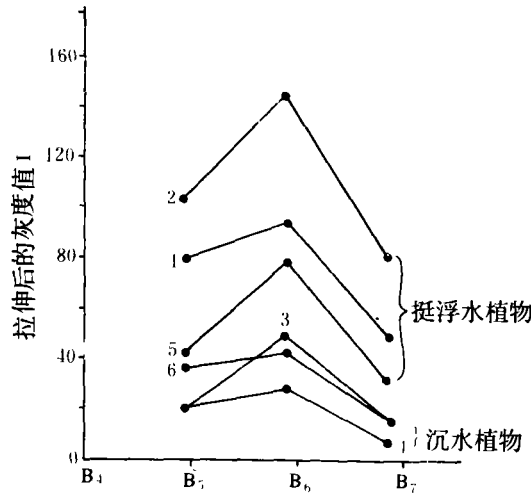


图 5 洪湖水生植物集群类别的波谱响应曲线

Fig. 5 Spectrum response curves of different aquatic plant classes in Honghu lake

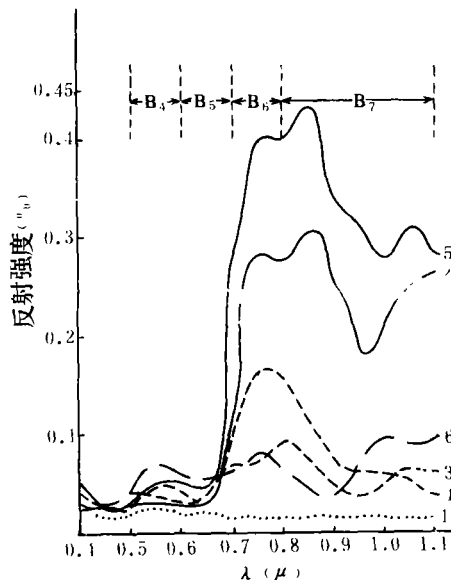


图 6 洪湖地区典型水生植物的光谱测定曲线

Fig. 6 Measured and Plotted Spectrum Characteristic Curves of Typical Aquatic Plants in Honghu Lake

1—水, 2—藕, 3—金鱼藻, 4—黄丝草, 5—菱, 6—藻

在洪湖水生植物的遥感调查中^[3], 我们通过对计算机非监督分类类别的波谱曲线和野外光谱测定曲线的对比分析, 参考植被叶绿素含量测定数据, 确定出每一集群类别所对应的水生植物类别(图5和6)。

对于地形起伏较大地区的分类结果, 可以通过分析每一集群类别的太阳入射角数值的波谱曲线, 结合实况资料, 并参考以下原则确定每一集群类别对应的地物类别属性^[4]:

(1) 对具有相同太阳入射角数值的类别, 光谱数值较高者对应于光谱反射能力较强的地物类别(图7);

(2) 往往某一地物类别对应着若干个集群中心, 即具有若干条标有不同太阳入射角数值的波谱曲线, 对于这些波谱曲线, 一般地讲, 太阳入射角愈小, 波谱曲线愈高(图8);

(3) 当两个集群中心的太阳入射角相差很大, 但光谱数值比较接近甚至完全相同时, 太阳入射角较大的那一集群中心, 对应着光谱反射能力较强的地物类别(图9)。

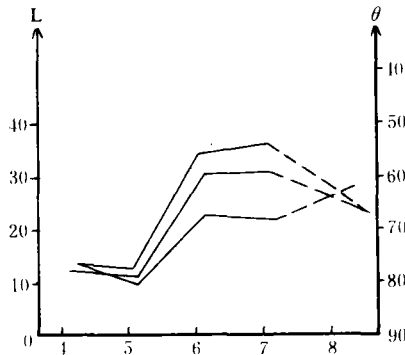


图7 太阳入射角为 60° 左右时一些树类的波谱曲线

Fig. 7 Spectrum Response Curves of Some Tree Species with Solar Incident Angles about 60°

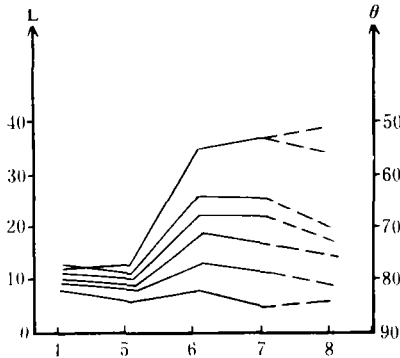


图8 楠阔类波谱曲线

Fig. 8 Spectrum Response Curves of Nan Bamboo and Broadleaf Trees

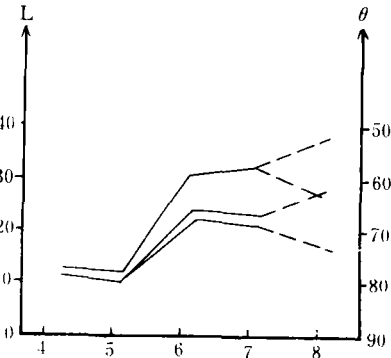


图9 光谱数值相等(或相近)但太阳入射角不同的一些树类的光谱曲线

Fig. 9 Spectrum Response Curves of Some Tree Species with Similar Spectrum values but Different Solar Incident Angles

此外,在数字遥感影像中加入研究区域的行政区划界线等边界信息,可以缩短影像分类处理的时间^[4]。

五、结 束 语

试验和研究结果表明,充分利用各种辅助数据可以不同程度地提高遥感影像的分类精度和效果。因此,应该进一步研究各种辅助数据在遥感影像分类处理中的可能应用及方法,分析研究每一方法的适用性和局限性,以不断提高遥感数字影像处理的人工智能水平。

参 考 文 献

- [1] B. C. Forster, Combining Ancillary and Spectral Data For Urban Applications, International Archives of Photogrammetry and Remote Sensing, Vol. XXV. Part A7, 1984, 6.
- [2] Friedrich Quiel, Trends and developments in the Classification of Multispectral data, International Archives of Photogrammetry and Remote Sensing, Vol. XXV, part A7, 1984, 6.
- [3] 杨凯、林开愚、张荣幸、谢其明, 遥感数字图像处理技术在洪湖水生植物分布调查中的应用, 植物学报, 25(5), 1983.
- [4] 陈军, DTM 在遥感影像分类中的应用, 武汉测绘学院学报, (1), 1984.
- [5] 杨凯、林开愚、陈军等, 遥感多时相特征图像按高程分层分类法——咸宁县土地利用面积的遥感测定, 测绘学报, 14(3), 1985.
- [6] Phillip. N. Slater, Remote Sensing-Optical and Optical Systems, Addison-Wesley Publishing Company, 1980.
- [7] Townshend. J. R. G, Terrain Analysis and Remote Sensing, London-Allen and Unwin, 1981.
- [8] Charles. J. Robinove, Computation with Physical Values from Landsat Digital Data, Photogrammetry Eng. S, 1982.

The Use of Auxiliary Information in the Classification of the Remote Sensed Data

Yang Kai Chen Jun

(Wuhan Technical University of Surveying and Mapping)

Abstract

In this paper, the authors have summarized their methods and experiments on the use of the auxiliary data in the classification and post-processing for the remotely sensed image data. The results show that a better classification accuracy can be obtained and the application effects are rather satisfactory.



图1 湖北省九官山陆地卫星影像
(1978. 10. 16. MSS6)

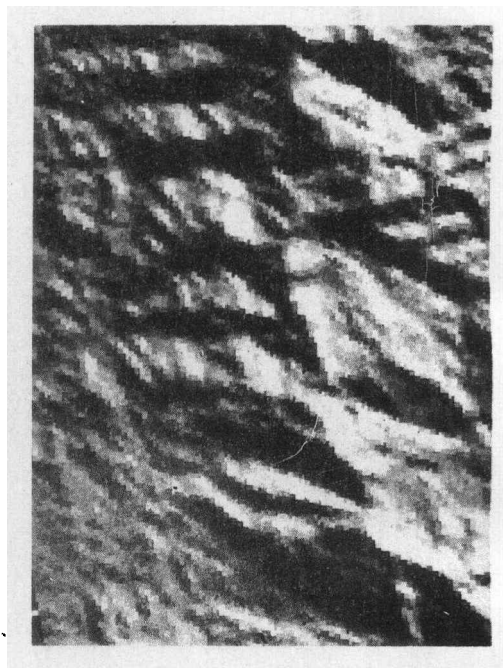


图2 与图1相应的几何配准
太阳入射角影像

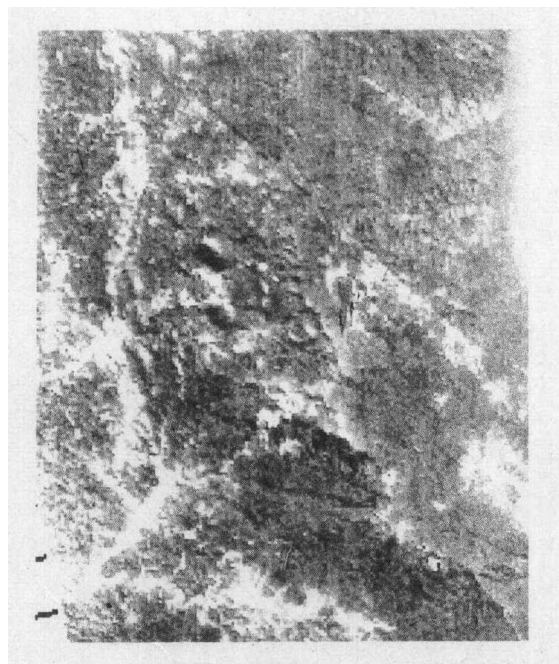


图3 校正后的MSS5影像



图4 校正后的MSS6影像