

高光谱图像模糊识别分类及其精度评价

闫永忠, 万余庆

(中国煤炭地质总局航测遥感局遥感应用研究院, 西安 710054)

摘要:高光谱遥感是在特定光谱域以高光谱分辨率同时获取连续波长的地物光谱图像,并能提供更多的精细光谱信息,有利于选择各种单一地物光谱差异明显的波段,使得高光谱遥感应用于更广的地物识别中。文中采用模糊识别的绝对值指数法对高光谱图像进行识别分类,并对分类结果进行像元级的评价。

关键词:高光谱遥感;模糊识别;绝对值指数;精度评价

中图分类号:X87

1 引言

高光谱遥感是在特定光谱域以高光谱分辨率同时获取连续的地物光谱图像,使得遥感着重于在光谱维上的应用。早期高光谱遥感主要应用于矿物识别,实现了单矿物识别和填图。由于高光谱遥感能提供更多的精细光谱信息,有利于选择各种单一地物光谱差异明显的波段,这将高光谱遥感应用于更广的地物识别中,如用作为西部陕西省靖边县,边邻毛乌素沙漠,地域广阔平坦,农作物品种单一,并且单种农作物种植面积较大,有利于地物的自动识别。

遥感影像分类大多数是根据地物单元的数理统计特征,将特征空间划分为互不重叠的子空间,然后将影像中的单元划归到各个子空间区。但是事实上在遥感影像单元中所包含的并不一定是单纯的一种地物信息,特别是空间分辨率低或者反映地表比较复杂的影像单元中,往往是多种混合地物类型的综合反映。因此用传统的统计进行影像分类或信息提取,所获得的精度与预期会相差很大,而且中间过程中丢失的信息成分较严重。而模糊分类为代表的“软”分类器的研究是遥感影像分类的一个重要趋势。

2 高光谱图像的模糊识别原理

利用高光谱遥感能提供更多的精细光谱信息,

对每个像元的光谱数据,对比光谱库中已知地物的光谱数据,根据样本间的统计相似度或者样本空间距离判别它属于何种地物。模糊识别的基本方法有最大隶属度原则和择近原则。文中采用的是模糊识别中的绝对值指数法,该方法属于模糊聚类分析距离法中一种分析方法,是直接利用距离或者类似距离建立相关程度对高光谱影像识别分类。

设 $U=\{x_1, \dots, x_n\} \subseteq R^m$ 其中 $x_i=(x_{i1}, \dots, x_{im}), i=1, \dots, n$, 可得到数据矩阵

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

建立 U 中的一个模糊相似关系 R , 表示要建立的模糊相似矩阵,利用绝对值指数距离法确定

$$r_{ij} = R(x_i, x_j) = \exp\left\{-\sum_{k=1}^m |x_{ik} - x_{jk}|\right\}$$

对遥感影像分类结果的精度评价,主要通过两种方法进行。第一种是基于像元级的评价,即在分类结果图上,以一定的间隔(地理位置)选取样本数据,通过实地情况的了解与比较,获得初步的分类结果。第二种评价方法,是基于特征级的精度评价,分别对各类别或者类别内部结构组成的影像特征与真实特征进行比较,通过特征的各属性域的差异,获得对精度的评价结果。大多数遥感影像分类方法是基于光谱信息的像元级运算过程,因此大多

收稿日期:2004-11-12; 修回日期:2005-04-24.

资助项目:本文为国家 863-308 主题“西部金睛行动”的部分研究成果。

作者简介:闫永忠(1972-),男,1997年毕业于长安大学地球科学与国土资源学院,现从事遥感信息研究工作。

数精度分析是通过第一种方法进行的。本次高光谱遥感影像分类结果精度的评价分析也采用第一种方法。像元级分类精度的计算是以误差矩阵(Error Matrix)为基础,误差矩阵的形式如表 1,假设其中有 X 类别,则综合分类精度为:

$$\text{分对率} = \sum_{i=1}^r x_{ii}/N \quad \text{其中: } r \text{ 为分类类别数, } N \text{ 为}$$

总样本数。

这种精度分析方法反映了各类别精度及总的分类效果。除此还采用 Kappa 系数对本次遥感影像分类结果进行精度估计。这种精度分析也是建立在误差矩阵的基础之上,对分类器的总体有效性能进行量化的评价。

$$\overline{K} = \frac{\theta_1 - \theta_2}{1 - \theta_2}$$

其中 $\theta_1 = \sum_{i=1}^r x_{ii}/N$ $\theta_2 = \sum_{i=1}^r x_{i+i}/N^2$ (表示行或列的求和, r 为分类类别数, N 为总样本数)。

3 实验内容

2001 年 7 月 12 日采用 OMIS1 高光谱成像光谱仪获取陕北靖边县城附近的航空高光谱遥感影像,同时采用 FieldSpec FR (美国分析光谱设备 ASD 公司生产)的便携式分光辐射光谱仪作飞行时的同步定标测试。此时植被正处于生长旺盛阶段,有利于不同地物分类识别。

由于飞行高度低,受大气影响严重,再加上航空遥感平台的位置和姿态变化十分复杂,所以高光谱图像上存在较为严重的几何畸变和辐射畸变,使其应用价值受到严重影响。要恢复图像的本来面貌,对图像进行几何校正和辐射校正是两项必不可少的工作。另外在图像处理前还必须对仪器和传输过程中混入图像的相干性噪音进行清除,使得后继的处理工作可以在具有较高信噪比的图像上进行。对高光谱图像预处理的一般性技术流程如图 1。

3.1 影像的模糊分类识别

首先在影像上选择训练样区(ROI),训练样区的选择必须是在影像范围内,选取有代表性的地物,也就是尽可能的使选择训练样区的像元纯度提

高;在实验中根据不同地物或不同作物长势选取好玉米、中玉米、差玉米、土豆地、葵花地、糜子地、苹果园、烟草地、裸地(1)、裸地(2)、裸地(3)、休耕地、土路及杨树阴影 14 种样本。

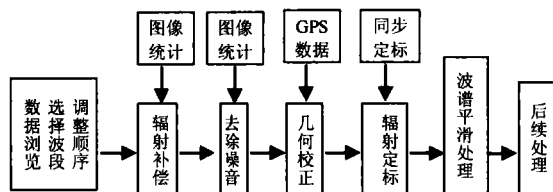


图 1 高光谱图像处理流程

Fig.1 The flow chart of hyperspectral image processing

然后对各种训练样区求取均值光谱曲线,建立光谱库;

最后利用训练样区的光谱库,根据样本间的统计相似度或者样本空间距离对影像进行模糊分类识别。由于遥感影像的复杂性和不精确性等特点,预先很难获得所有有代表性的样本的各类别的精确含量,分类结果也就不可能有很高的精度,因此可以对初步分类结果进行精度评价,根据误差矩阵的结果对训练样区进行重新定义,去除掉错选的部分,这样会得到更理想的纯像元,重新建立光谱库,再次对影像进行识别分类,结果可以得到较高的分类精度。模糊分类结果如图 2b。

3.2 分类识别结果及其精度评价

对分类识别结果图整体像元数统计获得:图像总像元数为 102 400,其中未分类占 4 394 (4.2910%)、好玉米地占 10 802 (10.5488%)、中玉米地占 8823 (8.6162%)、土豆地占 6 443 (6.2920%)、葵花地占 7 809 (7.6260%)、差玉米地占 5 360 (5.2344%)、糜子地占 10 489 (10.2432%)、土路占 4 269 (4.1689%)、裸地(1)占 8 395 (8.1982%)、休耕地占 1 3706 (13.3848%)、阴影占 1 909 (1.8643%)、裸地(2)占 5949 (5.8096%)、裸地(3)占 1 439 (1.4053%)、苹果园占 915 (0.8936%)、烟草地占 11 698 (11.4238%)。

在样本数据集的基础上对识别分类结果进行统计,得到测试样本的误差矩阵(Error Matrix)见表 1,分类结果精度评价见表 2,从识别结果图与原始影像对比来看,各类地物基本上被分出,总体分对率为 1528/1667=91.6617%,Kappa=0.9076,像元的纯度较高,分类识别效果较好,Kappa 值表示该分

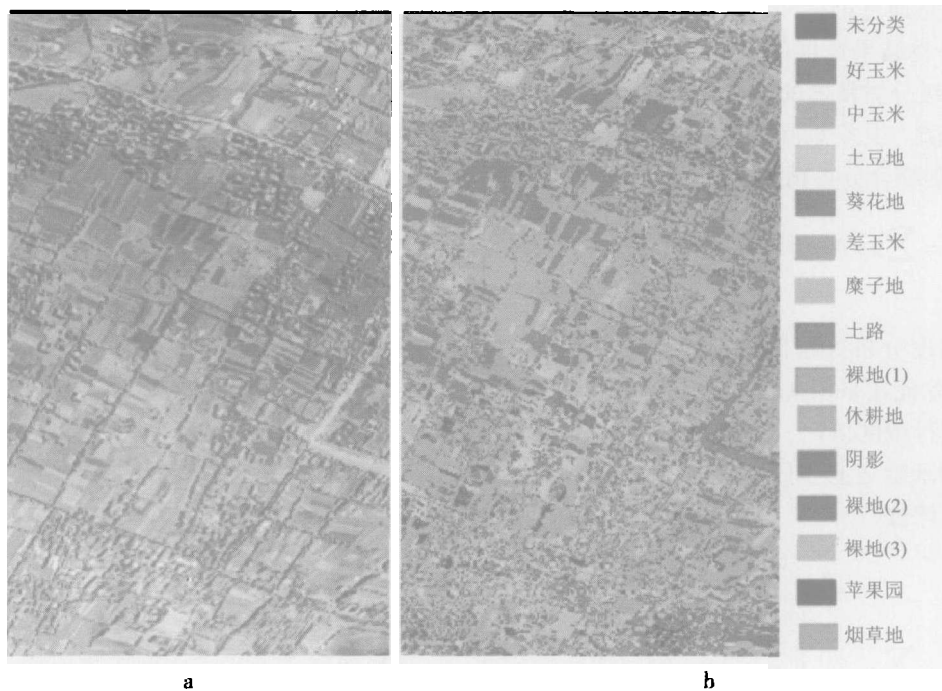


图 2 原始影像与分类结果对比图 (a 为原始影像图; b 为模糊识别分类图)

Fig.2 The original image (a) and the classification result (b)

表 1 模糊分类识别方法误差矩阵

Tab.1 The error matrix of fuzzy classification (单位: Pixels (1528/1667=91.6617%) Kappa = 0.9076)

样区 类别	土豆地	葵花地	糜子地	土路	休耕地	裸地 (1)	裸地 (2)	裸地 (3)	中玉米	好玉米	差玉米	烟草地	苹果园	阴影	合计
未分类	0	0	0	0	0	9	0	1	0	0	0	0	3	5	18
土豆地	94	0	0	0	0	0	1	0	0	0	0	0	0	0	95
葵花地	0	130	0	0	0	0	0	0	0	0	0	0	0	7	137
糜子地	6	0	14	0	0	0	0	0	0	0	0	0	0	0	20
土 路	0	0	0	215	0	0	0	0	0	0	0	0	0	0	215
休耕地	1	0	0	0	169	4	2	0	0	0	0	0	0	0	176
裸地(1)	0	0	0	0	1	223	4	0	0	0	0	0	0	0	228
裸地(2)	2	0	0	8	0	0	230	0	0	0	0	0	0	0	240
裸地(3)	0	0	0	0	0	0	0	74	0	0	0	0	1	0	75
中玉米	0	0	0	0	0	0	0	0	219	0	0	0	1	0	220
好玉米	0	0	0	0	0	0	0	0	3	88	0	0	0	0	91
差玉米	0	0	0	0	0	0	0	0	0	0	61	0	2	0	63
烟草地	0	2	0	0	0	0	0	0	0	0	0	32	0	0	34
苹果园	0	0	0	0	0	0	0	0	0	0	0	0	24	0	24
阴 影	0	0	0	0	0	0	0	0	0	0	0	0	0	31	31
合 计	103	132	14	223	170	236	237	75	222	88	61	32	31	43	166

类方法要优于随机赋予各点一类别的方法 90.76%,从误差矩阵结果数据可以看出,其中好玉米、差玉米、糜子地、烟草地分类结果达到 100%,这是由于这些地物面积相对较大,植被覆盖度相对大

些,基本不受背景土壤的影响,像元的纯度较高。其他地物识别结果虽有误分现象,但精度也相对较高,分对率分别为休耕地 169/170=99.41%、裸地(3) 74/75=98.67%、中玉米 219/222=98.65%、葵花地

表 2 模糊识别分类结果精度统计表
Tab.2 The precision table of fuzzy classification

类 别	Prod.Acc		User.Acc		Commission		Omission	
	像元	(%)	像元	(%)	像元	(%)	像元	(%)
土豆地	94/103	91.26	94/95	98.95	1/95	1.05	9/103	8.74
葵花地	130/132	98.48	130/137	94.89	7/137	5.11	2/132	1.52
糜子地	14/14	100.00	14/20	70.00	6/20	30.00	0/14	0.00
土路	215/223	96.41	215/215	100.00	0/215	0.00	8/223	3.59
裸地(1)	223/236	94.49	223/228	97.81	5/228	2.19	13/236	5.51
休耕地	169/170	99.41	169/176	96.02	7/176	3.98	1/170	0.59
阴影	31/43	72.09	31/31	100.00	0/31	0.00	12/43	27.91
裸地(2)	230/237	97.05	230/240	95.83	10/240	4.17	7/237	2.95
裸地(3)	74/75	98.67	74/75	98.67	1/75	1.33	1/75	1.33
苹果园	24/31	77.42	24/24	100.00	0/24	0.00	7/31	22.58
中玉米	219/222	98.65	219/220	99.55	1/220	0.45	3/222	1.35
好玉米	88/88	100.00	88/91	96.70	3/91	3.30	0/88	0.00
差玉米	61/61	100.00	61/63	96.82	2/63	3.17	0/61	0.00
综合类	32/32	100.00	32/34	94.12	2/34	5.88	0/32	0.00

注：假设 X 为分类选择某一类地物 ROI 的总像元数,Y 为分类选择的所有地物 ROI 内被分作某类地物的总像元数，

w 为分类选择某一类地物 ROI 内正确识别像元数。 $Prod.Acc=\frac{w}{X}$ $User.Acc=\frac{w}{Y}$ $Commission=\frac{Y-w}{Y}$ $Omission=\frac{X-w}{X}$

130/132=98.48%、裸地(2) 230/237=97.05%、土路 215/223=96.41%、裸地(1) 223/236=94.49%、土豆地 94/103=91.26%、苹果园 24/31=77.42%、阴影 31/43=72.09%。其中阴影和苹果园的精度较低,是因为阴影下的地物不同等因素所致,而苹果园是由于苹果树生长程度不同,覆盖度不高,由土壤背景因素的影响所致。

4 结论

作为国家 863 计划“西部金睛行动”项目的高光谱遥感应用研究,采用多种方法对黄土高原物种识别分类研究中,绝对指数距离法模糊识别分类获得的结果精度最高,分类效果最好,在该方法研究中总结出最为有效的分类识别流程是:选择训练样区(ROI),建立光谱库,然后对影像进行模糊识别分类。其中 ROI 区的选择影响着光谱库的建立,也直接影响分类结果的精度,所以根据分类统计结果重新精选训练样区(ROI),使选择的样区像元纯度达到理想化程度,重新建立光谱库进行识别分类,分类结果会更好。对于面积较大,地物单一,像元纯度

较高,利于选择理想化的纯像元的 ROI 区,分类的精度也就很高。同时该方法对土地利用和黄土高原生态环境监测有实用价值,在其他方面也有着广阔的应用前景。

参考文献

[1] 汪培庄,韩立岩. 应用模糊数学(修订版). 北京:科学出版社,1998.
[2] 张俊福,邓本让. 应用模糊数学. 北京:地质出版社,1988.
[3] 周成虎,骆剑承 等. 遥感影像地学理解与分析. 北京:科学出版社,2001.
[4] 陈述彭,童庆禧,郭华东主编. 遥感信息机理研究. 北京:科学出版社,1998,139~212.
[5] 方红亮,田庆久. 高光谱遥感在植被监测中的研究综述. 遥感技术与应用,1998.3(1):62~68.

Classification and Precision Evaluation About Fuzzy Diagnosis in Hyperspectral Images

YAN Yongzhong, WAN Yuqing

(*Aerophotogrammetry and Remote Sensing Center of China Coal, Xi'an 710054, China*)

Abstract: Hyperspectral remote sensing acquires sequential object spectral images with high spectral resolution in special spectral ranges, and provides more subtle spectral information. It is propitious to select clear bands for identifying single object spectrum, which made it largely used in object identification. The article used absolute exponent as one of the methods of fuzzy diagnosis to identify and classify hyperspectral images, then evaluated precision of classification result based on pixel level.

Key words: hyperspectral remote sensing; fuzzy diagnosis; absolute exponent; precision evaluation

上接 P19

The Design and Realization of a Content-based Remote Sensing Image Retrieval System

ZHU Guangliang¹, LU Lizhen²

(*1 Department of Geography, Zhejiang Education Institute, Hangzhou 310012, China;*

2 Zhejiang Provincial Key Lab of GIS, Zhejiang University, Hangzhou 310028, China)

Abstract: Retrieval and management of mega remote sensing image data need efficient approaches of content-based image retrieval. Now researches of content based remote sensing image retrieval focus on exploring the approaches for extracting low level vision features, such as color, texture, shape and so on. We are short of practical CBIR system. In this paper we present the architecture of RSIQuery system and discuss the data model and data organization. Finally we address some of the preliminary results on content-based remote sensing image retrieval and future work.

Key words: CBIR; Quin-tree decomposition; remote sensing image data model; data organization; RSIQuery