

基于光谱统计和混合像元分解的遥感影像分类

王淑香¹ 岳东杰²

河海大学测绘系, 江苏南京(210098)

E-mail: shuxiang1007@163.com

摘要: 针对各类目标地物的特性, 本文采用了分层分类方法依据光谱统计特征进行土地利用和植被覆盖分类研究。在此基础上, 对于难以区分的地物, 采用混合像元分解的方法提取各地类丰度, 按单个像元内各地类所占比例提取。结果表明, 与传统的最近邻法分类相比, 本文方法的总体精度和 KAPPA 系数都有了较大的提高。

关键词: 遥感影像 分类 光谱特征 混合像元分解

中图分类号: P23

1. 引言

分类问题是遥感图像信息提取中的最基本的问题之一。遥感图像分类过程实际上是通过遥感影像各类地物的光谱信息和空间信息等进行分析, 选择特征, 并用一定的手段将特征空间划分为互不重叠的子空间, 然后将影像中的各个像元划分到各个子空间去。将图像中的每一个像元点或区域归属到若干个类别中的一类, 分类的结果是将图像空间划分为若干个子区域, 每个子区域代表一种地物类型。分类后的图像广泛用于专题制图、专业信息提取、动态变化监测和遥感信息库的建立。

光谱特征是遥感影像的最基本特征, 传统的监督分类和非监督分类大都是利用光谱特征进行的分类, 在这方面, 已经有很多学者和研究者发表了大量的著作和研究成果。本文充分利用 TM 影像的多光谱波段特点, 采用针对地物特殊特征的分层分类进行分类实验。

2. 分层分类介绍

不同土地覆盖与土地利用类型赖以区分的特征集不同, 对所有类型的地物采用相同的特征集进行分类效果并不显著, 一方面不仅加大特征提取的工作量, 另一方面可能综合利用所有特征导致信息冗余。不同的特征和分类方面对不同的地物提取分类的精度不一样, 综合利用现有的各种成熟技术, 针对各类地物的特点, 采用适合的分层特征和方法分别进行处理, 当一种地物成功提取后, 通过图像处理方法, 将其从原始图像中去除, 以避免它对其它地物提取的影响。这种方法优点在于每层处理时, 目标明确, 只考虑与区分目标有关的最佳变量, 问题相对简单。

在区分事物时我们常常遵循“突出特征、先易后难”的原则, 对数据集进行几次部分分类, 直到全部数据集分类完毕, 对遥感图像分层分类同样适用此原则。对于大多数遥感图像来说, 还是有几种地物占主导地位, 即从光谱研究的角度看, 地面特征可以唯一表达。一般占主导地位的地物主要是水体、植被、岩石和土壤, 其他更多层次的细节可根据需要和影像信息的规律(实际可分性)不断补充^[1]。

2.1 光谱统计特征分析

1 光谱基本统计量

自然界中不同类型的地物具有各自不同的电磁波谱特性, 遥感数字图像中像元的不同数值(亮度值)反映了相应地物的波谱特性。因此, 通过计算机对图像像元的数值的统计、运算、对比和归纳, 对像元进行分类, 即可达到对地物的自动识别和提取。

单波段基本统计量包括灰度的均值、中值、众数、置于、反差、灰度直方图等，反映了与位置无关的不同灰度值出现的情况。多波段间的统计量有相关系数、协方差等。

对于图像分类问题，地物覆盖类型单一的地区，利用这些基本统计量就能够很好地将其提取出来，既方便又能达到应用精度要求。

表 1 样本光谱统计值

类型 \ 波段 统计值	统计 值	TM1	TM2	TM3	TM4	TM5	TM7
水体	范围	[135,152]	[58,69]	[60,72]	[46,72]	[23,70]	[8,39]
	均值	145	63	66	50	30	13
植被 (高密度)	范围	[134,148]	[57,64]	[57,70]	[65,108]	[70,111]	[23,66]
	均值	141	60	62	80	89	38
植被 (低密度)	范围	[133,149]	[58,64]	[60,72]	[63,77]	[69,125]	[30,78]
	均值	139	60	66	70	105	61
建筑区	范围	[137,150]	[58,66]	[61,72]	[54,64]	[61,88]	[38,57]
	均值	143	61	66	59	74	46
水泥地	范围	[142,158]	[63,72]	[70,84]	[70,88]	[77,163]	[41,124]
	均值	148	66	76	77	120	76

2 植被指数分析

常用的植被指数有归一化植被指数(NDVI),差值植被指数(DVI)、比值植被指数(RVI)和垂直植被指数(PVI)、绿度植被指数(GVI)等,这些指数法主要是通过将各波段反射率以不同形式进行组合来消除外在因素的影响,这些线性组合或波段比值的指数可满足特定的遥感应用^[2]。

NDVI 是最早、应用最广的植被指数,常用于指示植被的数量特征以及用于检测植被的季节变化和用于土地覆盖研究,也可用于植被区域和非植被区域的识别。MNDWI 是改进的归一化水体指数,适合于水体的识别提取。

$$NDVI = \frac{TM4 - TM3}{TM4 + TM3} \quad MNDWI = \frac{TM5 - TM2}{TM5 + TM2}$$

3 K-L 变换

K-L (Karhunen—Loeve) 变换也称为主成分或主分量分析,是一种统计特征基础上的多维正交线性变换。通过变换,将其分量相关的原随机向量转化成其分量不相关的新随机向量,把有用的信息集中到数目尽可能少的主分量中。

原始 TM 影像的各个波段之间存在很大的相关性,不同波段之间存在相当多的冗余信息。经过 K-L 变换,原空间中具有较大的相关性的各分量在新的空间中正交且相互独立,相关系数几乎为零,且 94% 以上的信息集中在前三个分量上,后三个分量存在相当多的噪声。因此,可以用主成分前三个分量代替原始 6 个波段,既充分地利用了原始影像信息,去除了噪声信息的干扰,又实现了信息的压缩。

2.2 线性混合像元分解

遥感像元对应的光谱曲线往往由几种不同物质的光谱曲线混合而成,这样的像元称为混合像元。对光谱分辨率较低的 TM 影像,混和像元是影响其分类精度的主要原因之一。尤其在城市地区,建筑物、草地、裸地相互混合的情况很多,光谱异质性复杂,仅仅利用像元光谱统计特征进行分类有很大的难度。

线性混合光谱模型是使用最多也最简单的一种混合光谱模型,它是建立在像元内相同地物都有相同的光谱特征以及光谱线性可加基础上的,构造简单、理论科学、物理含义明确,用于混合像元分解效果比较理想^[3]。

3 分层分类实验

3.1 分类过程

根据 TM 影像的特性和研究区的情况,在确定分层提取的顺序时,首先考虑光谱特征明显易于提取的类别。在本研究区内水体、林地相对容易区分,而城市建筑区、草地盖区和裸地,混合的情况比较复杂。

(1) 水体的提取

根据 MNDWI 指数分析, $MNDWI < 0$ 可以提取大部分的水体,但是有一部分建筑区混淆在内,而在 TM7 上,水体和建筑区的反射值范围分别为[8,39]和[38,57],具有很大的可分性。因此,通过反复实验,最终水体的提取模型为:

$$MNDWI < 0 \quad \& \quad TM7 < 30$$

(2) 植被提取

由于获得的研究区影像为夏季,高密度植被覆盖区(多为林地)和低植被覆盖区(草地、农田等)光谱差异较大。因此,将植被区分为低密度和高密度区分别提取,以确保能够得到较高的分类精度。理想情况下,植被的 NDVI 值应该大于 0,而实际情况往往不可能如此。在此,选择一定的阈值,确保大量的高密度植被能够提取处理,其提取模型为:

$$NDVI > 0.07$$

(3) 其它类提取

方案一:基于光谱统计特征提取

利用原始光谱统计特征,比较分析各样本统计值特性,不同地类使用不同的统计提取模型。由表 1 对光谱统计值的分析,水泥地在 TM3 波段像元反射值同其它类型反射值在不同的范围,但与建筑区有少量的混合;建筑区在 TM4 波段能够较好的提取。因此,我们建立这两种地物类型的提取模型:

水泥地: $70 < TM3 < 84$

建筑区: $54 < TM4 < 64$

对于低密度植被区用光谱统计特征很难提取,考虑到分层提取的特点,我们将其它四种地类提取后,剩余的不确定地类统一归为低密度植被类。这样的处理虽然避免了阈值提取时未分类像元的存在,但是影响了最终分类的精度。

方案二:针对方案一,使用光谱统计提取时的缺点,对于难易区分的三种地类,使用线

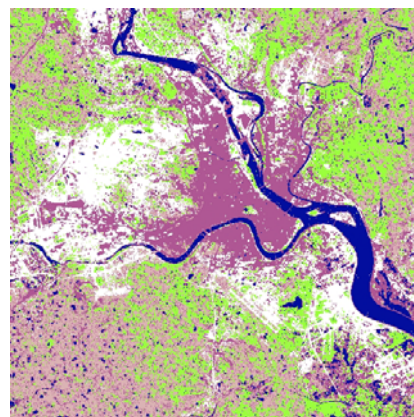
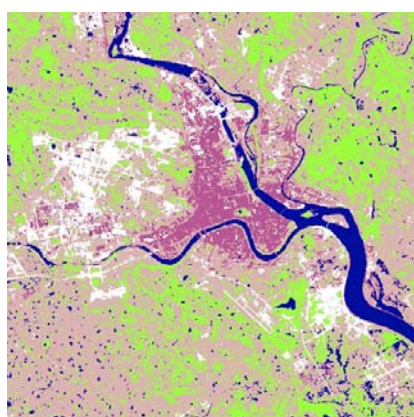
性混合光谱分解方法进行分类。分解波段时，采用经过 K-L 变换后的前三个波段，这样就保证了信息量的丰富，又降低了原始波段之间的强相关性。混合光谱分解后得到的各地类丰度，根据单个像元中各地类所占面积比的大小，确定其地类归属。在地类认定时，原则上在单个图斑内有不同土地利用性质时，依据最大面积所属的性质确定地类。因此，面积比最大的地类即为该像元的地类，具体提取模型如下。其中，f 表示丰度，s 水泥地，dz 表示低密度植被。

$$\text{水 泥 地: } f-s(i) > f-j(i) \quad \& \quad f-s(i) > f-dz(i)$$

$$\text{建 筑 区: } f-j(i) > f-s(i) \quad \& \quad f-j(i) > f-dz(i)$$

$$\text{低密度植被: } f-dz(i) > f-s(i) \quad \& \quad f-dz(i) > f-j(i)$$

上述两种分类方案的分类结果分别如图 1 和图 2 所示，



■ 水体 ■ 林地 ■ 草地 ■ 水泥地 ■ 建筑区

图 1 光谱统计特征分类结果

图 2 光谱特征和混合像元分解分类结果

3.2 分类结果与精度评价

为了对分类精度有一个纵向的比较，本文还使用了最小距离进行分类。其分类精度如表 2。从分类结果来看，本文采用的两种分类方法总体精度和 KAPPA 系数比最小距离分类有了提高，尤其是经过混合像元分解后的分类方法总体精度达到 90% 以上。

表 2 分类精度比较

分类方案	总体精度	KAPPA 系数
最小距离 KL123	84.99%	0.809
最小距离 (TM345)	83.36%	0.78
光谱统计量	88.56%	0.84
光谱+混合分解	91.34%	0.88

4 结论与展望

从实验结果来看，结合目标地物的分层分类是一种精度相对较高的分类方法。对于光谱特征难易区分的地物，进行混合像元分解增加了提取的精度。然而不足之处在于，仅仅利用

了地物的光谱特征,空间纹理特征和颜色特征,以及非遥感信息的 DEM 都可以用来辅助分类,提高分类精度。因此,结合多特征进行遥感影像的分类是我们下一步研究的重点。

参考文献

- [1] 周成虎.遥感影像地学理解与分析. 科学出版社,1999-12
- [2] 陈述彭,赵英时。遥感地学分析[M].北京.测绘出版社,1990.
- [3] 吕长春,王忠武,钱少猛.混合像元分解模型综述[J].遥感信息 2003 (3) :55—60.

based on spectral statistics feature and mixed-pixel decomposition of RS classification

Wang Shuxiang

Department of Survey and Mapping,Hohai University ,nanjing, (210098)

Abstract

In this paper, the experiment of a hierarchical classification has been made for all kinds of statistical characteristics of spectral features. And on this basis, to distinguish difficultly between the features , we used mixed-pixel decomposition to extract. The results show that the overall accuracy and kappa has been greatly improved compared with the minimum distanc classification.

Keywords: rs image , classification , spectrol feature , mixed-pixel decomposition