

一种遥感图象的混合多尺度 Hurst 参数分类模型

李 艳 彭嘉雄

(华中理工大学图象识别与人工智能研究所图象信息处理与
智能控制教育部重点实验室, 武汉 430074)

摘 要 针对遥感图象分类较困难的问题, 提出了一个遥感图象的分类模型——扩展的自相似模型(ESS), 该模型是一种广义的分形布朗模型(fBm), 它的多尺度 Hurst 参数与粗糙度之间是对应的, 同时不必像分形维数那样要求粗糙度的尺度不变性, 因而比 fBm 更接近于实际情况. 另外, 由于它的参数可以作为很好的分类特征, 而且特征维数低、计算快, 其方向性 Hurst 参数还描述了纹理在 4 个方向上的粗糙度, 因此可将它们与灰度的均值和标准差一起作为一组特征, 来构造一个混合多尺度 Hurst 参数分类模型, 将其用于卫星遥感图象分类, 获得了较高的分类正确率.

关键字 纹理分类 Hurst 参数 自相似模型 分形布朗运动

中图法分类号: TP751 文献标识码: A 文章编号: 1006-8961(2001)06-0537-05

A Mixed Multiscale Hurst Parameter Classification Model of Remote Sensing Image

LI Yan, PENG Jia-xiong

(Institute of Pattern Recognition & Artificial Intelligence, Huazhong University of Science & Technology,
State Education Commission Laboratory of Image Information and Intelligence Control, Wuhan 430074)

Abstract A classification model for remote sensing imaging is presented. The Extended self-similar model(ESS) is a general fractional Brownian motion(fBm) model. Its multiscale Hurst parameters have relations with roughness. At the meantime, it doesn't require the roughness to be scale-invariant as fractal dimensions do. The ESS is closer to the realities than fBm. Moreover, ESS gives multiscale parameters to provide more accurate interpretation of textures while the fBH gives only one. The multiscale Hurst parameters can discriminate a large number of natural textures and are suitable to be the features for texture classification. These features' dimension is lower compared with many other texture features, so that the computation intensity is less. Directed Hurst parameters describe the roughness at four orientations and multiscales. In this study they are mixed with the mean and standard deviation of gray level to be the feature vector. A new classification model of mixed multiscale Hurst parameters is constructed based on Bayes theorem. In this model we suppose that the conditional possibility distribution function of each feature is Gaussian, and the features are independent with each other. The a priori possibilities are decided by the highest rate of correct classification of the training set. For remote sensing texture classification, the performance of the new model is compared to other features, such as co-occurrence matrix features and Kaplan's features, etc. These classification algorithms are all based on Bayes theorem and the assumption that the a priori possibilities of all the classes are equal. Our experiments show that higher rate of correct classification to SPOT image is obtained by this new model.

Keywords Texture classification, Hurst parameter, Self-similar model, Fractional Brownian motion

0 引 言

遥感数据的分类问题是遥感应用的关键问题, 而

且遥感图象分类是随着模式识别理论的发展而发展的, 利用计算机通过对遥感图象中各类地物的光谱信息和空间信息进行分析, 选取地物的模式特征, 用一定的手段和规则将特征空间划分为互不重叠的子空

间,然后再将遥感图象中的各个象元划归到各个子空间去.由于遥感影象纹理反映着地物的空间分布状况,代表着物体表面的特征,是人们进行遥感图象目视判读的重要标志之一^[1],故纹理信息对实现影象的自动解译和提高解译的可靠性都起着一定的作用.

空间数据的复杂性和不确定性,导致了数据集之间互相重叠,从而给正确分类带来了很大的困难,是影响分类的一个瓶径.许多纹理分析的方法,如基于空间灰度共现矩阵的分类方法^[2],虽然能很好地解决一些典型纹理的分类问题,但用于遥感图象的分类仍然不能有令人满意的结果,其原因是遥感图象的纹理与一般纹理册上所收集的纹理有着很大的区别.遥感图象纹理类型非常丰富,除一些特殊景物外,在高分辨率的卫片上很难找到两块同样或非常相近的纹理,即使分类很细,纹理类内的散布程度仍将是很大的,其训练结果本身就不太理想,所以要进行正确的分类,其首要问题是选择合适的特征来描述纹理,即所选特征应该能够使不同纹理之间有尽可能大的区别,在这种情况下,高斯马尔可夫随机场模型^[3]、分形特征^[4]、基于小波变换的特征等相继被提出.一般来讲,特征维数越高,分类的能力越强,但过高的维数会使分类能力变差,因为所需的样本数将达不到要求,而且计算代价也太大.扩展的自相似模型(ESS)是广义的分形布朗模型(fBm),它的多尺度 Hurst 参数与粗糙度之间是对应的,同时不必像分形维数那样要求粗糙度的尺度不变性,比 fBm 更接近于实际情况,但它的参数可以作为很好的分类特征,而且特征维数低,计算快.

1 多尺度 Hurst 特征

扩展的自相似(ESS)模型是对分形模型的一种扩展.为了描述自然纹理中没有表现出准周期性结构的那些纹理,Mandelbrot 将自相似的分形布朗运动模型(fBm)一般化,并以一个 Hurst 参数表示^[5],它控制着该过程在所有尺度上的可见粗糙度.虽然一个自然纹理可在很大范围的尺度内呈现出相似的粗糙度,但由于在实际应用中,假设粗糙度在任何尺度上都是恒常的并不合适,因此人们提出了各种拓展 fBm 的方法,其中 ESS 便是对 fBm 的一种拓展^[6].因为由 ESS 引入的多尺度 Hurst 参数提供了表达不同尺度上纹理粗糙度的一种特征,所以本研究利用多尺度 Hurst 参数与传统的统计特征相结合的方法来进行遥感图象分类.

自相似模型可以描述为一个 n 维 fBm 过程 $B(t)$,它是一个零均值的高斯随机场,其在原点处为零,该过程是统计自相似的,即

$$\text{VAR}[B(t+\tau)-B(t)]=\sigma^2\|\tau\|_n^{2H},\forall t,\tau\in R^n \quad (1)$$

式中, $\text{VAR}(\cdot)$ 表示一个过程的方差, $H(0<H<1)$ 是 Hurst 参数. fBm 是分形的,且维数 D 满足 $D=n+1-H$,但自然界的纹理不是严格的自相似,且 f_d 的幂率关系不可能在所有尺度上保持不变,所以需要更加一般的随机分形模型.扩展的自相似过程 ESS 为

$$\text{VAR}[B(t+\tau)-B(t)]=\sigma^2f(\|\tau\|_A),\forall t,\tau\in R^n \quad (2)$$

其中, $\|\tau\|_A=\sqrt{\tau^T A \tau}$, $f(\cdot)$ 是规范化的结构函数, $f(1)=1$.这样,数字图象就可看作是满足式(2)的 2 维过程.在二进尺度下,对一幅 $N\times N$ 大小的,只包含单一纹理的图象,考虑在 4 个方向上(图 1)的多尺度 Hurst 参数.

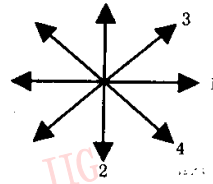


图 1 多尺度 Hurst 参数计算时的 4 个方向

其值估计为

$$\tilde{H}_s^\theta = \frac{1}{2} \log_2 \left(\frac{f_s^\theta}{f_s^0} \right), \text{for } \theta = \{1, 2, 3, 4\} \quad (3)$$

其中,

$$f_s^\theta = \frac{1}{N(N-2^s)} \sum_m \sum_n |\Delta_s^\theta(m, n)|^2, \text{for } \theta = \{1, 2, 3, 4\}$$

$$\Delta_s^1(m, n) = B(m+2^s, n) - B(m, n)$$

$$\Delta_s^2(m, n) = B(m, n+2^s) - B(m, n)$$

$$\Delta_s^3(m, n) = B(m-2^s, n+2^s) - B(m, n)$$

$$\Delta_s^4(m, n) = B(m+2^s, n+2^s) - B(m, n)$$

由于多尺度 Hurst 参数有很好的性质,即它是偏差和对比度不变的,因此可以用于区分类似云状的没有周期性的纹理与谐和的纹理.方向性 Hurst 参数 \tilde{H}_s^θ 描述了纹理在 4 个方向上的相似性,而且提供了较好的一组区分纹理的特征.因为遥感图象的纹理方向性差别很大,所以在本研究中使用 4 个方向的参数,以提高分类正确率.由于 Hurst 参数的计算是基于各尺度的灰度差,因而仍然有其局限性.因为某些海洋类影象纹理的 H 参数与山地类影象纹理的 H 参数十分接近,以致有时产生错误分类,但在某个波段的卫星图象上这两种类别的灰度却存在明

显差异,所以这里还选择了灰度的均值和标准差来作为分类特征.

2 混合多尺度 Hurst 参数分类模型

分类实验所用的遥感图象是马赛地区一块有海洋、城市、山地、植被等地类的大尺寸 SPOT 图象,分辨率为 30m. 根据经验将图象划分为 16×16 的子图比较合适. 在这样的子图上,纹理类型是多样的,海洋类影象包含不同形态的纹理,如平缓的或变化很大的;城市等也一样. 如果只将之划分成海洋、城市、山地、植被 4 个概念类进行训练和分类的话,则类内散布程度太大,而且类间重叠的可能将增加,产生错误分类的可能也加大,故采用多层次分类机制比较合适,如采用二层分类机制,其第 1 层为 4 个概念类别,第 2 层为第 1 层类别的子类,即每个概念类都包含多种类别纹理. 这样其分类最终结果是将样本空间划分成 4 个概念类子空间.

构造一个高斯分布的分类模型,设 Ω 为样本的特征空间, $X \in \Omega$. 对每个子图,多尺度 Hurst 特征的尺度 s 取为 0, 1, 2, 3, 共得到 12 个 Hurst 参数,加上灰度均值和标准差,则 X 维数是 14, 这远小于通常的几十甚至上百维,因而减少了计算强度.

假定特征是独立的,且其分布为正态函数,则贝叶斯距离提供了最大相似度分类,其相似度函数为

$$P(X|\omega_i) = r_i \prod_{d=1}^{N_f} \frac{1}{\sqrt{2\pi}\sigma_{d,i}} \exp\left(-\frac{(x_d - \mu_{d,i})^2}{2\sigma_{d,i}^2}\right)$$

$$0 < r_i < 1, i = 1, 2, \dots, C; \sum_{i=1}^C r_i = 1$$

式中, $\mu_{d,i}, \sigma_{d,i}$ 代表第 i 个子类第 d 个特征的均值和标准差, ω_i 代表第 i 个子类,则样本与第 i 子类的 Bayes 距离为

$$g_i(X) = \sum_{d=1}^{N_f} \left[2\ln\sigma_{d,i} + \left(\frac{x_d - \mu_{d,i}}{\sigma_{d,i}} \right)^2 \right] - 2\ln r_i$$

使这个距离最小的类别就是该样本所属的子类别. 根据该子类所属的上层概念类将之归入上层类中. r_i 是一组变量,通过对其迭代,可使训练样本的平均正确分类率达到最大. 据此可以看出, r_i 相当于各类的先验概率. 这里,假设特征间是相互独立的,这样不仅可使计算简单,而且可避免协方差矩阵可能奇异的弊病,这与假定特征不独立所得到的分类正确率是相当的.

3 实验结果及分析

在样本空间中,选择一半的数据作目视判读,并将之分成层次结构的各类别样本,而将另一半数据用做测试. 根据实验用图所包括的地貌类型,将之分成海洋、城郊、山地、植被 4 个概念类,其中,海洋类影象纹理形态有明显区别,即靠近海岸线区域的纹理十分平滑,而深海区域纹理则比较粗糙,所以海洋概念类包括 4 个子类纹理;城郊类影象纹理一致性也较差,房屋密集区、居住区及郊区影象纹理各成一类,因此这一类纹理包括 3 个子类纹理;山地类影象纹理包括两个子类纹理;植被影象纹理一致性较强,没有子类. 训练样本数,海洋类 4 个子类依次为 212、34、62、36;城市类 3 个子类依次为 220、58、167;山地类两个子类依次为 29、44. 植被类为 86.

图 2 中的纹理图块是训练样本集中的一部分样本,每个图块对应一个子类.

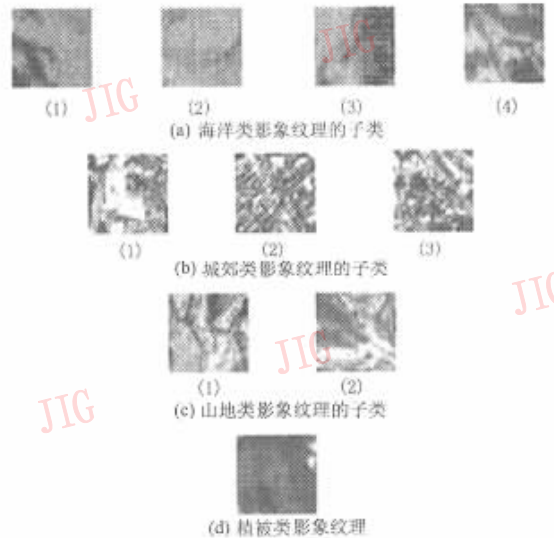


图 2 卫星遥感影象训练样本集例

通过对从训练样本集中提取的特征均值的分析可以发现,海洋类影象纹理的 H 参数一般较高,这说明其纹理粗糙度较低,这也符合 H 值与分形维数 D 之间的关系. 若某些尺度的 H 参数出现负值,则说明这些纹理的相关函数衰减得太慢,如周期性纹理的情况^[7].

为了测试分类的正确率,本文在另一半测试用的图象上按类别随机取出若干子图块. 因为这些子图块不包含训练数据在内,所以测试结果是有说服力的. 表 1 列出了这些图块的测试正确率和错误率,每一类

都取两个测试图,表中显示出每个测试图被正确分类的比例,以及被错误地归入其他各类中的比例。

表 1 测试样本用本文方法分类的正确率和错误率列表
(单位:%)

	海洋	城郊	山地	植被
海洋图 1	84.38	0	10.94	4.69
海洋图 2	92.86	0	3.57	3.57
城郊图 1	0	62.50	36.11	1.39
城郊图 2	0	96.43	3.57	0
山地图 1	8.33	14.58	70.83	6.25
山地图 2	5.56	16.67	72.22	5.56
植被图 1	0	0	20	80
植被图 2	12	8	20	60

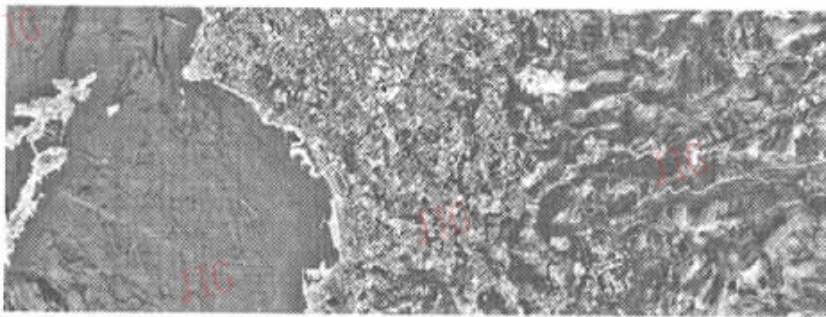
从表 1 可见,本文方法对海洋类影象纹理的正确分类率是最高的,而且其错误分类均集中在山地和植被类别上,即将海洋类影象错分为山地、植被类,因为海洋类影象纹理的第 3 个子类与植被类影象纹理比较相似(图 2),海洋类影象纹理的第 4 个子类与山地的第 2 子类影象纹理也比较相似。另外由于城郊和山地影象较难区别,二者易产生混淆,故分类正确率较低。从表 1 还可以看出,海洋类影象纹理与城郊类影象纹理错误分类率最低,说明这两类影象纹理与其他类影象纹理的区别最大。为了与其他方法进行比较,选择了 K-L 变换的特征值, x 和 y 方向共现矩阵的惯性、聚类投影(Cluster

shade)、聚类显度(Cluster Prominence)、局部一致性、能量、熵特征, Kaplan 特征集^[7]等特征测试,这些方法所使用的分类器均为只考虑了条件概率的贝叶斯距离分类器,其正确分类率如表 2 所示。

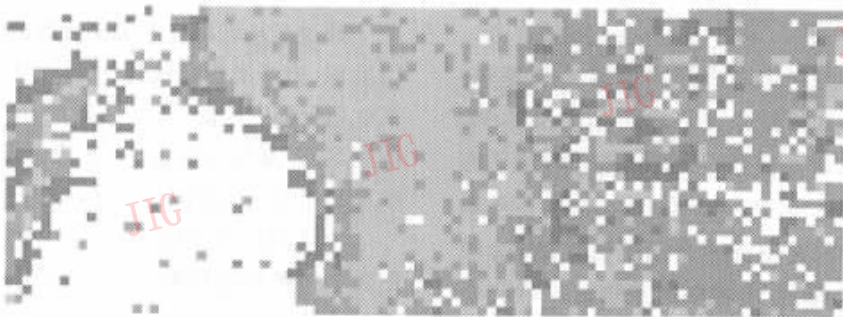
表 2 与利用其他一些特征和分类方法进行分类的正确率比较
(单位:%)

	K-L 变换特征	共现矩阵特征	Kaplan 特征	本文方法
海洋图 1	17.19	82.56	76.56	84.38
海洋图 2	91.07	75.58	91.07	92.86
城郊图 1	56.94	33.72	41.67	62.50
城郊图 2	91.07	58.14	87.50	96.43
山地图 1	83.67	27.19	89.58	70.83
山地图 2	72.22	30.23	94.44	72.22
植被图 1	10.00	68.70	70.00	80.00
植被图 2	0.00	60.47	48.00	60.00
平均正确率	52.77	54.65	74.85	77.40

比较而言,本文分类模型的综合分类能力优于上述方法,并且共现矩阵特征的分类能力明显不如本文方法,其原因在于,它虽提供了图象的二阶统计特征,但共现矩阵的定义却限制了它只是在一定方向和距离上的灰度共现,因而它不具备多尺度的能力。另外,由于自然纹理在不同尺度下,呈现出的特点有所不同,因此用多尺度的方法进行纹理分析,比在一个尺度下的效果要好。图 3 示出了用本文方法对包含 4 类纹理的某地区卫星遥感图象进行分类的结果。



(a) 原始 SPOT 图象



(b) 分类结果

图 3 本方法对一部分图象的分类结果

4 小 结

数据的复杂性和不确定性,导致了遥感图象数据集之间的互相重叠,从而影响了正确分类,而基于扩展的自相似模型的多尺度 Hurst 参数则具有能够区分大部分自然纹理(包括具有一定周期性的和没有周期性的纹理)的能力. 本文在 Hurst 参数及经典的统计特征的基础上,提出了一个混合的多尺度 Hurst 参数分类模型,并由其对一幅大尺寸卫星遥感图象的分类实验结果可见,本文方法对实际遥感图象上地物影象的分类能力是比较高的. 当然也应看到,许多特征选择方法虽各有其优点,但简单地将它们混合起来应用,反而可能会导致分类正确率的下降. 小波变换也是一种很好的多尺度分析方法,也可进一步将这一方法引入到遥感图象分类的研究中.

参 考 文 献

- 1 黄桂兰. 航空像片影象纹理分类研究[博士论文]. 武汉测绘科技大学. 1996,7.
- 2 Conner R W. Segmentation of a high resolution scene using texture operators, CVGIP, 1984, 25(3): 273~310.
- 3 Descombe X, Sigelle M, Preteux F. Estimating Gaussian markov random field parameters in a nonstationary framework: Application to remote sensing imaging. IEEE Trans. on Image Processing, 1999, 8(4): 490~503.

- 4 Solka J L, Marchette D J, Wallet B C *et al.* Identification of man-made regions in unmanned aerial vehicle imagery and videos. IEEE Trans. on PAMI, 1998, 20(8): 852~857.
- 5 Mandelbrot B B, Ness J W V. Fractional Brownian motions, fractional noises and applications. SIAM Rev., 1968, 10(10): 422~437.
- 6 Kaplan L M, Kuo C C J. Texture roughness analysis and synthesis via extended self-similar (ESS) model. IEEE Trans. on PAMI, 1995, 17: 1043~1056.
- 7 Kaplan L M. Extended fractal analysis for texture classification and segmentation. IEEE Trans. On Image Processing. 1999, 8(11): 1572~1585.

李 艳 1968 年生, 讲师, 1990 年获武汉水利电力大学工学学士学位, 1995 年获华中理工大学工学硕士学位, 现为华中理工大学图象识别与人工智能研究所在职博士生. 研究方向为机器视觉、模式识别等.

彭嘉雄 1934 年生, 教授, 享受国家特殊津贴的有突出贡献专家, 华中理工大学模式识别与智能系统博士点首位博士生导师. 主要从事模式识别、人工智能等领域的研究. 发表论文 200 余篇.