

文章编号:1004-4213(2010)01-0178-5

# 基于多特征融合的城市遥感图像自动解译方法\*

孙显<sup>1,2</sup>, 王宏琦<sup>1</sup>, 张道兵<sup>1</sup>, 胡岩峰<sup>1</sup>, 巩大亮<sup>3</sup>

(1 中国科学院空间信息处理与应用系统技术重点实验室 北京 100190)

(2 中国科学院电子学研究所, 北京 100190)

(3 北京市遥感信息研究所, 北京 100085)

**摘要:**为更全面有效地解译城市遥感图像,提出了一种新的基于多特征融合的自动解译方法.该方法定义对象网络来表达图像结构并获取更为准确的处理单元.在此基础上,综合分析颜色、纹理、形状和位置等众多特征,通过自适应的概率学习训练最优分类器并标记目标类别.方法中还结合上下文信息进行空间平滑,大大消除了噪音、遮挡等影响,矢量标绘后得到最终解译结果.实验表明,该方法准确率高、鲁棒性好,适用于多种遥感图像城市场景的自动解译.

**关键词:**图像处理;自动解译;特征融合;基于对象;城市遥感图像

**中图分类号:** TP391.4

**文献标识码:** A

**doi:** 10.3788/gzxb20103901.0178

## 0 引言

近年来,利用解译技术分析遥感图像中的城市场景日益成为遥感应应用研究的热点课题,这大大加快了城市地理信息系统的建设.在高分辨率遥感图像中,城市由一系列复杂目标组合而成,主要包含建筑物、道路和植被等地物目标<sup>[1]</sup>.其中,建筑物和道路目标的数量较多,建筑物一般为规则外形、排列较整齐;而道路纵横交错,将城区划分为若干小街区.树冠等植被位于建筑物周围或道路两侧,成块状面积分布.

现有的遥感图像解译方法大致分为两类.一类通过在图像中寻找符合目标特性的形状元素,如直线、圆弧、多边形等,然后分析元素间的关系来标绘目标的几何外形<sup>[2-3]</sup>.该类方法直观形象,但对每类目标都需要一个对应的模板,考虑地物目标种类的多样性,需要的计算开销较大,且在检测复杂目标时的效率较低.另一类则通过分析和使用地物目标的若干区域统计特征,例如颜色均值、灰度方差等,找出目标存在的候选区域并进行后处理<sup>[4-6]</sup>.该类方法在图像分类及部分地物目标的识别中已经取得了一定的效果,但由于特征种类较为单一且需要人工交互,而且往往忽略了地物间的空间关系,在实际应用中存在局限性.

为此,本文提出了一种新的城市遥感图像自动解译方法.与传统的方式不同,该方法参照基于对象的思想<sup>[7]</sup>,定义对象网络来获取图像结构.这种新的表达方式将图像分割和识别相结合,不但可以得到真实客观的对象实体,还可以完整描述图像内容和空间信息,为目标解译提供了准确的依据.方法中统计了图像对象的颜色、纹理、形状、位置等多类特征,通过自适应的概率学习来融合处理,所需的人工干预少,分析全面,能够快速定位地物目标区域并识别目标类别.此外,本文还采用了结合上下文信息的空间平滑,不但有效消除了遥感图像中绝大部分的背景噪音,而且可以避免受图像来源、光照不同的限制以及目标遮挡的影响.矢量标绘后可以得到城市场景的解译结果.实验表明,本文提出的这种方法准确率高、鲁棒性好,适用于多种遥感图像城市场景的自动解译.

## 1 对象网络的构建

遥感图像中地物目标的内部一般特征分布较为均匀,而边缘部分由于存在干扰噪音,特征差异明显.如果采用传统的滑动窗等方式处理,必然会出现 MAUP 问题<sup>[8]</sup>,导致处理准确度下降.为此,本文首先对遥感图像进行多尺度分割构建对象网络,为后续处理奠定良好的基础.

### 1.1 多尺度分割

本文提出了一种 Pyramid-cut 算法来进行多尺度分割.该算法在 Normalized-cut 算法<sup>[9]</sup>的基础上添加多尺度序列,将图像表达为金字塔结构.金字塔的顶层为图像整体,底层为所有像素,中间  $L$  个层

\* 国家自然科学基金(40871209),国家高技术研究发展计划(2006AA12Z149)和中国科学院电子学研究所青年创新基金资助

Tel: 010-58887208-8933 Email: sunxian0918@hotmail.com

收稿日期: 2008-07-14

修回日期: 2008-11-16

次以尺度因子  $\alpha$  逐层递减, 每一个尺度都代表对图像的一个分割, 每个下级尺度都表示对上级尺度的细化, 所有的尺度合并起来, 既可以得到准确的区域描述, 又能够全面反映区域间的空间关系。

考虑大尺度下生成的对象可表达区域的总体信息, 相比小尺度下生成的对象更具指导意义。在分割过程中, 将 Normalized-cut 算法与尺度因子相关联, 自上而下逐层实现图像的全局最优划分

$$K_l = \alpha \cdot K_{l-1}, \quad l=2,3,\dots,L \quad (1)$$

式中,  $K_l$  表示第  $l$  层的分割对象数,  $K_1$  定义为

$$K_1 = \lceil \sqrt{\max(M, N)} \rceil \cdot \lambda \quad (2)$$

$M$  和  $N$  分别为待处理图像的长、宽像素数,  $\lceil \cdot \rceil$  为上取整高斯函数,  $\lambda$  为分割系数。

Pyramid-cut 算法从第一层起按照  $K_1$  完成初始分割, 图层中的每个对象随后继续分割成  $\alpha$  块小对象, 由这些小对象组成下一图层, 依次递归处理。这样一方面能够将目标对象从复杂背景中较好地划分出来; 另一方面严格保证了小尺度对象都位于大尺度对象范围内, 避免出现边界重叠或交叉等错误情况。

## 1.2 对象空间关系

为构建对象网络, 还需要对各个层次对象间的空间关系进行分析。本文将空间关系概括为同层对象间的拓扑关系和相邻层对象间的继承关系两类。依据对象间相对位置的变化, 将拓扑关系细分为相离、相邻、包含三种; 依据父子对象间的对应数量和尺度因子大小, 将继承关系分为一对二(父对象具备两个子对象)和一对多(父对象具备两个以上子对象)两种, 如图 1, 以此在对象间建立关联, 形成一个网络结构。

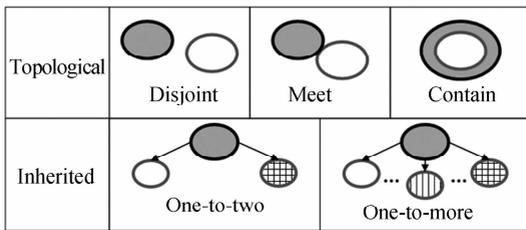


图 1 对象空间关系

Fig. 1 Examples of spatial relations in objects network

## 2 多特征融合的概率学习

方法将遥感图像中的目标对象分为建筑物、道路、树冠三类, 剩余地物都归为背景类, 并从不同类别中抽取各个尺度的样本对象, 人工标记后组成训练集。考虑遥感图像中场景复杂、目标多样, 仅使用单一种类的特征进行处理难度较大, 结合实际应用将多类特征融合在一起, 采用概率学习的方式得到

分类器, 用于评估各类目标在图像中存在的可能性。

### 2.1 特征描述

地物目标特征集由颜色、纹理、形状、位置等四大类组成, 每类特征的计算方式为:

#### 1) 颜色特征

颜色直方图是图像色彩特征的一个很好表征。本文利用了遥感图像的灰度、彩色的 RGB 和 HSV 共 7 个通道的直方图来统计图像的颜色特征。

选取样本对象  $B_i$ , 统计其灰色直方图  $G_i$ , 并在训练集上度量该对象的灰色直方图相似差异比  $D_i$

$$D_i = \frac{\sum_{j=1}^{T^+} d(G_i, G_j^+) / T^+}{\sum_{j=1}^{T^-} d(G_i, G_j^-) / T^-} \quad (3)$$

式中,  $T^+$  和  $T^-$  分别代表同类和非同类样本对象的数量, 采用  $\chi^2$  距离<sup>[10]</sup> 统计对象间的直方图相似度  $d(G_i, G_j)$

$$d(G_i, G_j) = \frac{\sum_{r=0}^{R-1} (G_i(r) - q(r))^2}{q(r)} \quad (4)$$

$$q(r) = \frac{G_i(r) + G_j(r)}{2} \quad r=0, \dots, R-1 \quad (5)$$

式中,  $R$  为图像灰度级, 取  $R=256$ 。依此可以计算得到各个通道的直方图相似差异比。

#### 2) 纹理特征

纹理可由一些基本单元的规则模式, 如滤波器组等得到。本文使用的滤波器组有 Gabor 滤波器组、LM 滤波器组、S 滤波器组和 MR 滤波器组。按照尺度和参量的不同, 每个滤波器组由 32 种具备不同表达形式的滤波器组合而成, 总计 128 个滤波器。处理过程中, 由于对象实体的外形并不规则, 先获取对象的外接矩形并得到滤波器的响应, 然后统计对象内部的实际响应均值来近似拟合计算结果。

#### 3) 形状特征

颜色和纹理特征都是基于区域的描述, 而基于边界的形状特征可以较好描述目标结构, 提高识别效果。

本文通过计算图像  $I$  中所有像素的方向能量  $OE$ <sup>[11]</sup> 对形状特征建模。方向能量定义为

$$OE = (I \cdot f_1)^2 + (I \cdot f_2)^2 \quad (6)$$

式中  $f_1$  为高斯差分滤波器,  $f_2$  为  $f_1$  的变形

$$\begin{cases} f_1(x, y) = \frac{d^2}{dy^2} \left( \frac{1}{Z} \exp\left(-\frac{y^2}{\sigma^2}\right) \exp\left(-\frac{x^2}{l^2 \sigma^2}\right) \right) \\ f_2(x, y) = \text{Hilbert}(f_1(x, y)) \end{cases} \quad (7)$$

式中,  $\sigma$  为尺度,  $l$  为方向比,  $Z$  为归一化常量。

对每个对象而言, 取其内部  $OE$  值最大的 10 个像素来代表该对象的形状特征。

#### 4) 位置特征

位置特征用来描述对象在图像整体中所占的大概比重和相对位置. 本文利用中心距, 即对象中心坐标与图像中心坐标的距离, 和面积比, 即对象面积占整幅图像场景的比例这两个特征来衡量.

## 2.2 概率学习

学习的目的是寻找一种适当的特征融合方式, 训练出高效准确的分类器, 实现目标与非目标的分类.

Boosting 作为一种强大的集成学习算法, 已经在文本分析、数据挖掘等领域得到了广泛的应用. 但是一般的 Boosting 算法主要用于区分两类目标, 在处理多分类问题时, 计算效率偏低.

本文利用 Joint Boosting 算法<sup>[12]</sup>来训练多特征分类器. 该算法在特征处理上做了改进, 达到了特征共享, 对多目标分类有着较好的效果. 令样本对象的类别列表为  $S(n) = \{\text{建筑物、道路、树冠、背景}\}$ , 分别对应类别标记  $c = \{1, 2, 3, 4\}$ . 将样本对象的特征向量和类别标记输入分类器进行训练. 每个特征视为一个弱分类器, 每一轮迭代选取一个最优弱分类器

$$h_t(v, c) = \begin{cases} a\delta(v_m^f > \theta) + b & c \in S(n) \\ k^c & c \notin S(n) \end{cases} \quad (8)$$

式中,  $t$  表示迭代轮次,  $(a, b, \theta, k)$  为分类器训练参数,  $v_m^f$  表示第  $f$  维特征在第  $m$  个训练样本上的响应. 经过  $T$  轮迭代训练后, 可以得到强分类器

$$H(v, c) = \sum_{t=1}^T h_t(v, c) \quad (9)$$

标记过程中, 利用强分类器  $H$  计算待检测对象  $B_i$  属于  $c$  类目标的标记概率<sup>[13]</sup>

$$p(B_i, c) = \frac{1}{1 + \exp[-H(v, c)]} \quad (10)$$

概率值越大, 说明该对象属于  $c$  类目标的可能性越大.

## 3 图像解译

完成概率学习和标记后, 利用对象网络中的上下文关系进行空间平滑, 矢量标绘后得到最终结果.

### 3.1 空间平滑

本文 1.2 节中已经分析了对象网络中的拓扑和继承关系, 定量计算后对结果进行空间平滑.

同一图层中, 对象间的相互影响与距离成反比, 相邻对象的信息对类别判断起到一定的辅助作用. 例如, 建筑物内部的遮挡、树冠中的阴影, 可以通过考察相邻对象得到平滑纠正. 拓扑平滑的计算公式为

$$p'(B_{i,l}, c) = p(B_{i,l}, c) \sum_u^U p(B_{u,l}, c) \cdot \frac{\gamma_{u,l}}{\gamma_{i,l}} \quad (11)$$

式中,  $B_{i,l}$  表示第  $l$  层的  $i$  个对象,  $B_{i,u}$  表示第  $u$  个与  $B_{i,l}$  相邻的对象,  $U$  为相邻对象总数,  $P(B_{i,l}, c)$  和  $P'(B_{i,l}, c)$  分别为  $B_{i,l}$  拓扑平滑前和平滑后的标记概率,  $\gamma_{i,l}$  为  $B_{i,l}$  边界的长度,  $\gamma_{u,l}$  为  $B_{i,l}$  和  $B_{i,u}$  共有边界的长度.

同样的, 父子对象间也存在严格的继承关系, 某些大尺寸对象的误差可以通过其子对象的统计信息得到改进. 继承平滑的计算公式为

$$p'(B_{i,l}, c) = p(B_{i,l}, c) \sum_v^V p(B_{v,l}, c) \cdot \frac{\zeta_{v,l}}{\zeta_{i,l}} \quad (12)$$

式中,  $B_{i,v}$  表示  $B_{i,l}$  的第  $v$  个子对象,  $V$  为  $B_{i,l}$  子对象的总数,  $\zeta_{i,l}$  和  $\zeta_{v,l}$  分别为  $B_{i,l}$  和  $B_{i,v}$  的面积.

按不同类别标记, 自下而上逐层交替对所有对象进行拓扑平滑和继承平滑后, 可依据最大概率  $c^* = \operatorname{argmax}_c p'(B_{i,l}, c)$ , 来确定对象的目标类别. 然后, 投影所有尺度图层到同一图层, 得到标记结果.

### 3.2 目标标绘

对标记结果使用掩膜法, 提取目标对象的二值边缘, 形态学处理去除毛刺. 结合目标特性, 简要标绘目标要素的位置和轮廓界线, 作为输出结果. 建筑物目标大多可以用规则多边形表示, 利用 SUNSAN 算子<sup>[14]</sup>检测关键角点并连接, 获取矢量多边形. 道路一般为连续的具有一定宽度的狭长区域, 通过 Hough 变换搜索平行线, 进行边缘链码跟踪和折线段逼近后可以得到用折线表示的道路目标. 树冠多数为类圆形结构, 角点检测后以不同直径的圆形来近似拟合.

## 4 实验结果及分析

为验证本文提出的基于多特征融合的城市遥感图像自动解译方法的有效性, 本节中利用该方法对图像数据进行了验证. 实验数据集由从因特网上搜集到的 200 幅包含各类城市场景的高分辨率遥感图像组成, 这些图像的分辨率和目标尺度范围基本近似. 实验中取其中 50 幅用于训练, 剩余 150 幅用于测试和评估.

图 2 给出了自动解译的基本流程. 对待处理图像先做多尺度分割, 令尺度个数  $L=4$ , 尺度因子  $\alpha=2$ , 分割系数  $\lambda=100$ . 分割效果如图 2(b), 多个尺度结合能够实现目标的正确划分. 从每个尺度抽取四类样本对象各 1 000 个, 总计 16 000 个构成样本库. 训练得到融合多类特征的分类器, 标记各个尺度图层, 如图 2(c). 图 2(d)为空间平滑后的结果, 有效消除了错分情况. 矢量标绘结果如图 2(e)和(f).

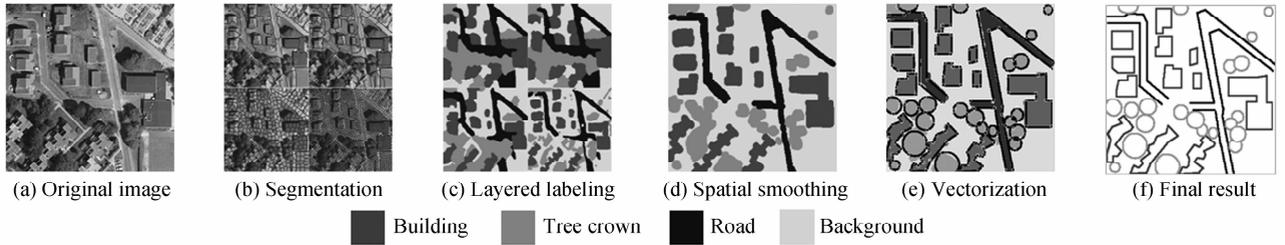


图2 本文提出方法的解译流程  
Fig. 2 Interpretation procedure of our approach

本文采用 ROC 曲线<sup>[15]</sup>和像素分类正确率来定量评估算法性能. ROC 曲线是显示分类模型真正率 (TPR)和假正率 (FPR)之间折中的一种图形化方法. 像素分类正确率  $Recall=1-FPR$ .

处理方法融合了颜色、纹理、形状和位置四大类特征. 表 1 统计了各类特征被忽略后对正确率的影响, 下降数值越大, 说明该类特征的重要程度越大. 结果表明, 这四类特征在处理过程都起到了一定的作用. 相比而言, 纹理、颜色和形状特征所起的作用较大, 这也基本符合遥感地物目标的特性.

表 1 各类特征被忽略后对标记正确率的影响

Table 1 The drop in accuracy caused by removing each type of feature

Feature types	Color	Texture	Shape	Position
Drop proportion/(%)	16	27	13	7

本文提出的方法以多尺度分割后得到的对象为处理单元, 分割效果对解译准确度有一定的影响. 图 3

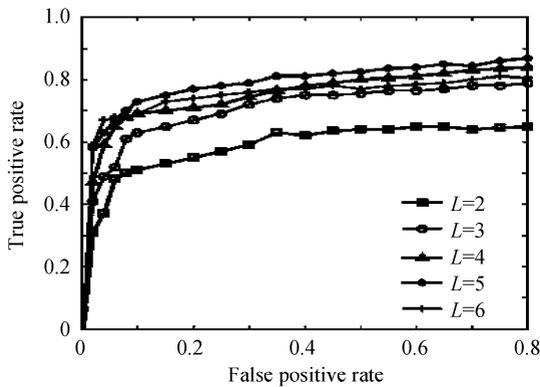


图3 不同尺度个数对解译结果的影响  
Fig. 3 Performance of interpretation with different scale numbers

给出了方法在不同尺度个数  $L$  下的 ROC 曲线. 可以发现, 尺度个数的增加一定程度上会提高处理准确度, 这是由于特征描述准确度和空间信息得到增加; 然而, 尺度个数过多也会导致对象网络过于复杂, 增加了计算负担, 部分冗余信息甚至会导致处理效率的下降.

图 4 比较了本文提出方法和一般滑动窗方法的处理结果. 两种方法采用相同的训练图像和特征信息. 结果表明, 本文提出的方法性能更优.

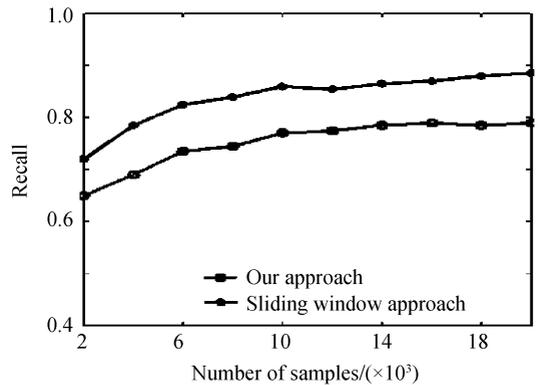
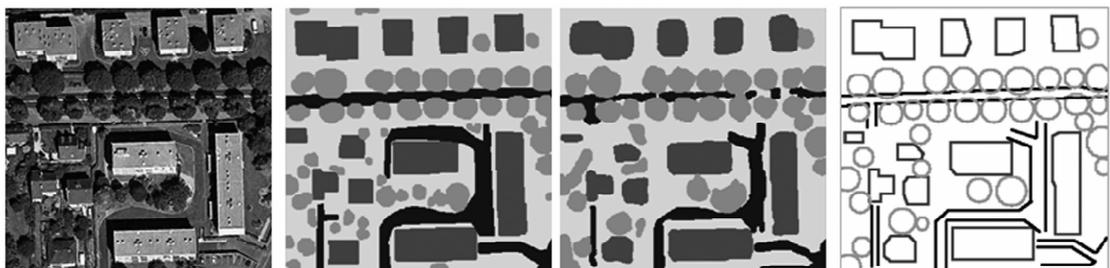


图4 本文提出的方法和一般滑动窗方法的比较  
Fig. 4 Comparison between our approach and general sliding window approach

图 5 给出了其余部分测试图像及其解译结果. 本文提出的方法不仅能准确处理目标清晰、分布规则的场景, 对于背景纷乱、结构复杂的场景也有较好的效果. 一些建筑物和树冠之间存在遮挡, 以及道路和建筑物之间边界模糊, 通过空间平滑和矢量标绘可以得到校正. 此外, 由于图像间的目标差异较大, 少量目标可供采样的样本对象不够全面, 导致性能指标出现波动, 如果适当增加样本图像的数量, 这种



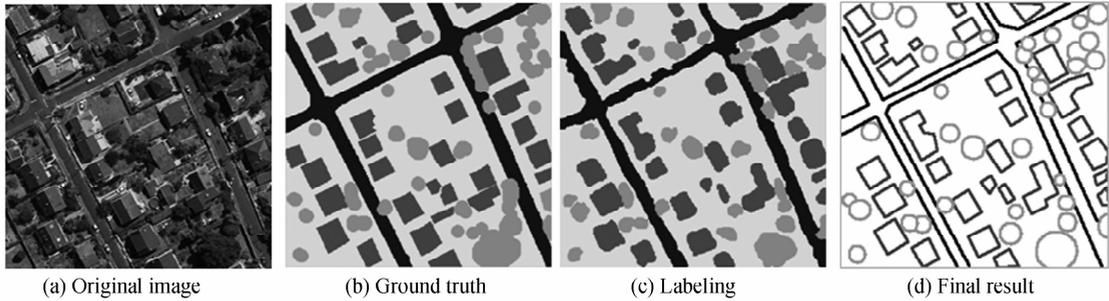


图5 部分测试图像及解译结果

Fig. 5 Some representative interpretation results

情况可以得到改善. 总体而言, 本文提出的这种基于多特征融合的城市遥感图像自动解译方法是有效实用的.

## 5 结论

本文提出了一种基于多特征融合的城市遥感图像自动解译的方法. 方法充分利用遥感图像中包含的丰富信息, 综合分析了颜色、纹理、形状和位置等众多特征, 通过概率学习训练最优分类器并实现多类目标的标记. 方法中还采用了多尺度的对象网络来定义图像结构, 以符合客观的对象为处理单元, 获取准确的特征信息. 此外, 还结合对象间的上下文关系, 有效消除了遥感图像中的背景噪音, 并避免了来源、光照及目标遮挡等影响.

实验结果表明, 本文提出的这种方法能有效解译绝大多数高分辨率下的城市遥感图像, 在地理信息系统和数字城市系统建设中具有一定的实用价值. 同时, 如何挖掘更具针对性的目标特征和图像信息, 进一步提高解译效率, 是今后值得继续研究的课题.

### 参考文献

- [1] BARALDI A, PARMIGGIANI F. Urban area classification by multispectral spot images[J]. *IEEE Trans On Geosci Remote Sensing*, 1990, **28**(4): 674-680.
- [2] ZHANG Dao-bing, LIU Bo, WANG Hong-qi. Parallel active contour model for main road segments extraction from high-resolution remote sensing images[J]. *Acta Photonica Sinica*, 2007, **36**(7): 1345-1350.  
张道兵, 刘波, 王宏琦. 基于平行活动味道模型的高分辨率遥感影像城区主干道路段提取[J]. *光子学报*, 2007, **36**(7): 1345-1350.
- [3] ZHANG Dao-bing, ZHANG Hui, ZHANG Zheng, et al. A robust main direction extraction method for road segment[J]. *Acta Photonica Sinica*, 2007, **36**(6): 326-330.  
张道兵, 张慧, 张正, 等. 一种稳健的道路主方向提取算法[J]. *光子学报*, 2007, **36**(6): 326-330.
- [4] ZHOU Bing, WANG Yong-zhong, SUN Li-hui, et al. Study on local entropy used in small target detection[J]. *Acta Photonica Sinica*, 2007, **36**(6): 326-1003.
- [5] ZHANG Zhi-an, FENG Hong-wei. A new retrieval algorithm based on the texture-spatial feature [J]. *Acta Photonica Sinica*, 2008, **37**(2): 400-404.  
张志安, 冯宏伟. 一种新的基于纹理和空间分布特征的图像检索[J]. *光子学报*, 2008, **37**(2): 400-404.
- [6] SU Juan, LIU Dai-zhi. An object-level change detection algorithm for remote sensing images [J]. *Acta Photonica Sinica*, 2007, **36**(9): 1764-1768.  
苏娟, 刘代志. 一种目标级的遥感图像变化检测算法[J]. *光子学报*, 2007, **36**(9): 1764-1768.
- [7] DEFINIENS IMAGE GMBH. Ecognition user guide [R]. Germany, 2002. <http://www.definiensimaging.com>.
- [8] MARCEAU J. The scale issue in the social and natural sciences [J]. *Canadian Journal of Remote Sensing*, 1999, **25**(4): 347-356.
- [9] SHI Jian-bo, MALIK J. Normalized cuts and image segmentation [J]. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2000, **22**(8): 888-905.
- [10] BAO Qian, GUO Ping. Comparative studies on similarity measures for remote sensing image retrieval based on histogram[J]. *Journal of Remote Sensing*, 2006, **10**(6): 893-900.  
包倩, 郭平. 基于直方图的遥感图像相似性检索方法比较[J]. *遥感学报*, 2006, **10**(6): 893-900.
- [11] MALIK J, BELONGIE S, LEUNG T, et al. Contour and texture analysis for image segmentation [J]. *International Journal of Computer Vision*, 2001, **43**(1): 7-27.
- [12] TORRALBA A, MURPHY P, FREEMAN T. Sharing visual features for multi-class and multi-view object detection[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2007, **19**(5): 854-869.
- [13] SHOTTON J, BLAKE A, CIPOLLA R. Multi-scale categorical object recognition using contour fragments [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2007, **30**(7): 1270-1281.
- [14] SMITH S M, BRADY J M. Susan: A new approach to low level image processing [C]. *International Journal of Computer Vision*, 1997, **23**(1): 45-78.
- [15] FAWCETT T. An introduction to ROC analysis [J]. *Pattern Recognition Letter*, 2006, **27**(8): 861-874.

## An Automatic Interpretation Approach for Urban Remote Sensing Image Based on Multiple Features Integration

SUN Xian<sup>1,2</sup>, WANG Hong-qi<sup>1</sup>, ZHANG Dao-bing<sup>1</sup>, HU Yan-feng<sup>1</sup>, GONG Da-liang<sup>3</sup>

(1 *Key Laboratory of Spatial Information Processing and Application System Technology, Chinese Academy of Sciences, Beijing 100190, China*)

(2 *Institute of Electronic, Chinese Academy of Sciences, Beijing 100190, China*)

(3 *Beijing Remote Sensing Institute, Beijing 100190, China*)

**Abstract:** A new automatic approach based on multiple features integration is proposed to interpret urban remote sensing images more effectively and comprehensively. The approach builds a hierarchical objects network to organize image structure and gets precise processing units. Then the probabilistic learning integrating multiple features including colour, texture, shape and position are performed to train a best classifier, and label all of the objects according to their classification values. The approach also applies spatial smoothing which incorporates contextual information to eliminate the adverse effects caused by background disturbance, occlusion and so on. After vectorization procedure, final results are given. Experiments demonstrate that proposed approach achieve high exactness and robustness in interpreting manifold urban remote sensing images.

**Key words:** Image process; Automatic interpretation; Features integration; Object based; Urban remote sensing image



**SUN Xian** was born in 1981. He is a Ph. D. degree candidate at Institute of Electronic, CAS. His research interests focus on remote sensing image processing and machine learning.



**WANG Hong-qi** was born in 1964. He received the Ph. D. degree from Institute of Electronic, Chinese Academy of Sciences in 1994. Now he is a professor, and his research interests focus on signal and information processing and remote sensing processing.