基于贝叶斯网的航拍图像建筑目标提取算法*

李玲玲[1]，金泰松[2a]，李翠华[2a]，李　渊[2b]

(1. 郑州航空工业管理学院 计算机科学与应用系，郑州 450015；2. 厦门大学 a. 信息科学与技术学院；b. 建筑与土木工程学院，福建 厦门 361005)

摘　要：提出一种基于贝叶斯网的建筑目标提取算法。该算法通过多场景航拍图像进行训练后建立纹元字典，将实际图像中的纹元映射到纹元字典获得图像的场景类信息；然后使用朴素贝叶斯网建模建筑目标与场景类空间上下文的关系约束，将建筑目标提取转换为求解贝叶斯网类别节点的后验概率问题。与同类方法的对比实验表明，提出的算法能有效地提取航拍图像中的建筑目标。

关键词：图像分析；目标识别；贝叶斯分类器；航拍图像；纹元

中图分类号：TP391.41　　**文献标志码**：A　　**文章编号**：1001-3695(2012)04-1550-03

doi:10.3969/j.issn.1001-3695.2012.04.097

Automatic building extraction from aerial images based on Bayesian nets

LI Ling-ling[1], JIN Tai-song[2a], LI Cui-hua[2a], LI Yuan[2b]

(1. *Dept. of Computer Science & Application, Zhengzhou Institute of Aeronautical Industry Management, Zhengzhou* 450015, *China*; 2. *a. School of Information Science & Technology, b. School of Architecture & Civil Engineering, Xiamen University, Xiamen Fujian* 361005, *China*)

Abstract: This paper proposed an approach to building extraction from aerial image using Bayesian nets. Firstly, it represented texton dictionary by training aerial image sets. Secondly, acquired classification of a new image by mapping the textons to texton dictionary and obtained the scene categories in the whole image. Finally, it selected naive Bayesian nets to represent spatial constraints between objects and scene classes, and computed the posterior probability of category node to building extraction. The test on the dataset shows that the proposed approach yields substantial improvement over others on building extraction from the aerial image.

Key words: image analysis; target recognition; Bayesian classifier; aerial image; textons

0　引言

针对航空图像理解与建筑目标提取问题，国内外学者发表了很多相关论文。Porway 等人[1]提出一种基于场景上下文的航空图像层次理解模型。Karantzalos 等人[2,3]将统计学习理论引入到建筑目标提取过程，取得了不错的效果。中国科学院的孙显等人[4]通过基于对象的 Boosting 方法自动提取高分辨率遥感图像中的人造目标。清华大学的唐亮等人[5]充分结合物方空间和图像空间的各种有用信息，逐步推理得出一类人造目标的位置、高度和屋顶轮廓等信息，实现建筑目标的三维重构。

以往目标提取方法大多直接针对图像中的特定目标进行处理，往往忽视目标类别周围的环境信息，甚至将环境信息当做干扰因素予以剔除。人类视觉实验[6]表明："目标类别周围的环境信息对人眼感知起重要作用。"为了建模目标类别与环境之间的关系依赖，本文以航拍图像为例，引入贝叶斯网络建模建筑目标与环境的关系约束，将建筑目标提取转换为求解贝叶斯网中类别节点的最大后验概率问题。

1　基于贝叶斯网的建筑目标提取方法

本文中的航拍图像建筑目标提取过程包括特征提取、类别划分和贝叶斯网目标验证三个阶段。

对一幅输入图像，首先提取纹元特征，然后基于纹元特征的弱分类器对图像中的场景类别进行划分，最后对可能的建筑目标使用贝叶斯网进行验证，输出图像中的建筑目标。其整个过程如图 1 所示。

2　航拍图像中场景类别的划分

本文采用文献[7]的做法对航拍图像中的场景类别进行初步划分。该方法使用纹元特征，建立纹元字典，用于图像场景的分类。纹元字典是纹元与图像场景类之间多对一的映射。

收稿日期：2011-08-12；**修回日期**：2011-09-27　　**基金项目**：国家自然科学基金资助项目(41171341)；国家教育部新世纪优秀人才支持计划资助项目(2009)；河南省科技创新杰出青年计划(114100510006)；航空科学基金光电控制技术国防科技重点实验室资助项目(20095155008)；福建省自然科学基金资助项目(2010J05099,2011J01365)；河南省科技攻关资助项目(112102310082)；河南省高等学校青年骨干教师资助计划、郑州市科技创新培育计划(10PTGG342-1)

作者简介：李玲玲(1973-)，女，副教授，博士(后)，主要研究方向为计算机视觉与图像处理；金泰松(1978-)，男，讲师，博士，主要研究方向为图像处理(jintaisong@xmu.edu.cn)；李翠华(1960-)，男，教授，博导，主要研究方向为计算机视觉；李渊(1979-)，男，副教授，主要研究方向为测绘学.

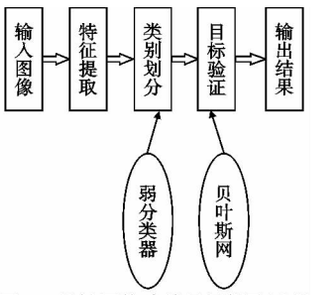

图1　航拍图像建筑目标提取过程

航拍图像场景一般包括森林、公路、河流、湖泊、建筑、山脉、桥梁、机场等类别,本文为每类场景图构建一个特征纹元集(textons),不同类别的纹元集形成一个数据字典。首先场景图与滤波器组作卷积运算,获得滤波响应图,该响应图通过 K 均值聚类方法产生纹元,并将该纹元标定为对应的场景形成映射关系,形成数据字典。整个过程如图 2 所示。

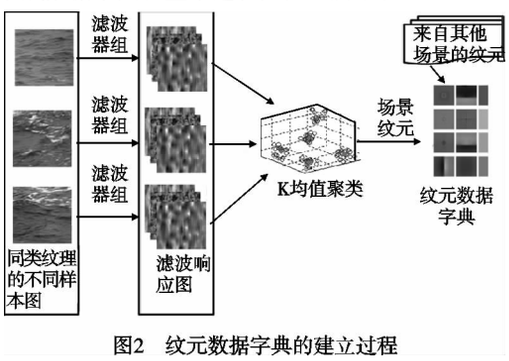

图2　纹元数据字典的建立过程

对一幅测试图像,采用相同方法获得图像的纹元模型,然后依靠最近邻分类器与 χ^2 统计方法衡量纹元字典与测试图像纹元模型间的距离,根据距离大小对测试图像中的场景类别进行划分。图 3 是对一幅测试图像的处理过程,其中不同颜色代表图像中的场景类。

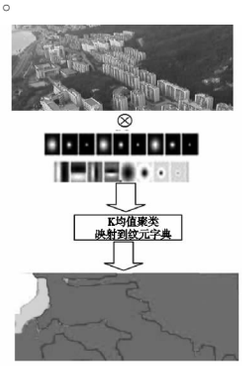

图3　测试图像类别划分的过程

3　基于贝叶斯网的建筑目标提取

3.1　朴素贝叶斯网的建立

航拍图像场景复杂,场景类别的初步划分往往不准确,建筑目标更是如此。因此验证弱分类器分类得到的场景类是十分必要的。

贝叶斯网是处理不确定性信息的重要工具,提供了一种表示因果信息的方法,已成功地用于医疗诊断、统计决策、专家系统等领域。本文注意到,建筑目标与周围环境空间上下文的关系约束可以作为建筑目标的证据,将建筑目标与场景之间的依赖关系通过贝叶斯网进行建模,通过计算它们之间的概率依赖大小,实现航拍图像中建筑目标的有效提取和验证。

常见的关系类型通常包括包含、相似、相邻等,其中相邻关系最常见。汉词中的“依山傍海”“高耸入云”等就是对此类关系的形象描述。本文选择相邻关系类型,作为建筑目标与场景类之间的因果关系类型,将建筑目标与环境的关系约束的量度表示为它与相邻场景类之间的依存度。显然,对类别提取与标注而言,相邻场景类的出现对建筑目标有启示作用。

使用贝叶斯网最先需要确定的是贝叶斯网的网络结构。本文采用朴素贝叶斯网络(naive Bayesian wetwork)作为描述目标与场景关系约束的工具,其中类别节点是建筑目标类,它是所有属性节点的父节点,并假定属性节点之间互相独立。除了类别节点和属性节点之间的有向边外,属性节点间再无连接,网络结构如图 4 所示。

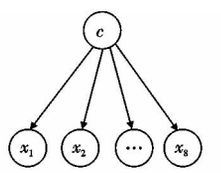

图4　朴素贝叶斯网分类器

城市场景航拍图像中常见的场景类有森林、公路、河流、湖泊、建筑、山脉、桥梁、机场等,由于本文提取图像中的建筑目标,因此将建筑目标定义为贝叶斯网的类别节点,其他场景类别定义为属性节点。每个节点的值表示为二值随机变量,取值为“真”表示对应场景类出现在图像中,取值为“假”表示对应场景类未出现在图像中。为了验证分类器获得建筑目标的真伪,本文使用贝叶斯网络将建筑目标提取转换为求解贝叶斯网类别节点的后验概率问题。

朴素贝叶斯网中属性节点相互独立,根据贝叶斯定理,目标类 c 的后验概率为

$$p(c|x_1,x_2,\cdots,x_8)=p(c)\prod_{i=1}^{8}\frac{p(x_i|c)}{p(x_i)}$$

$$p(x_i|c)$$

$$p(x_i=\text{"true"}|c=\text{"false"}) \tag{1}$$

其中:$p(c)$、$p(x_i)$分别表示目标类或其他场景类的先验概率;$p(c|x_1,x_2,\cdots,x_8)$表示给定场景类 $x_1,x_2,\cdots,x_8$ 的条件下目标类的后验概率。该概率值取决于 $p(c)$、$p(x_i)$和 $p(x_i|c)$。

本文选取相同数量的正负样本图像,其中正样本图像中包含建筑目标,负样本图像中不包含建筑目标。正负样本数量相同,先验概率 $p(c)=50\%$。$p(x_i)$是场景类 x_i 的先验概率,假设正样本图像中场景类 x_i 的个数为 n,正样本总数为 m,则概率 $p(x_i=\text{"true"})=(n/m)$。依此类推计算其他场景类的先验概率。$p(c)$、$p(x_i)$确定以后,条件概率 $p(x_i|c)$决定整个后验概率的取值。假设正样本图像中与建筑目标相邻场景类 x_i 的个数为 p,正样本总数为 m,则概率 $p(x_i=\text{"true"}|c=\text{"true"})=(p/m)$;负样本图像中与建筑目标相邻场景类 x_i 的个数为 q,负样本总数为 n,则概率 $p(x_i=\text{"true"}|c=\text{"false"})=(q/n)$。其他后验概率通过互补法则获得。

3.2　基于贝叶斯网的目标验证

给定一幅测试图像,使用文献[7]中的弱分类器进行分类,对图像场景类别进行初步划分;然后朴素贝叶斯网作为工具验证获取场景类中的建筑目标类。

航拍图像场景复杂,弱分类器划分的场景类可能是错误的。有两种可能情况,一种是将建筑目标分类为其他场景类;另一种是将其他场景类分类为建筑目标。为了准确地提取图

像中的建筑目标,本文对类别划分中的建筑目标类别进行如下处理:

假定该场景类为朴素贝叶斯网的建筑目标类节点,其他场景类是属性节点,并选取类别划分中与该类别的相邻场景类,将其对应的属性节点变量取值为真;其他场景类对应属性节点变量取值为假;最后判断目标类别的后验概率是否大于非目标类别。如果大于,这个目标属于建筑目标类;否则,属于非建筑目标类。

4 实验结果与分析

通过厦门大学视频与图像处理实验室购买的长 15 m,载重 50 kg 的飞艇设备获取航拍图像。实验中,航拍图像的传输、GPS 数据通过无线链路双向传输,获得了大量航拍图像作为训练样本集和测试样本集,其中 2 000 幅图像作为训练样本,1 000 幅图像作为测试样本。图 5 是本文实验中所使用的飞艇,图 6 是飞艇拍摄的一幅航拍图像。

图5 本文采用的飞艇工具 图6 飞艇拍摄的一幅航拍图像

首先,定义算法的性能评价标准。对航拍图像来说,复杂背景使得现有算法难以完全准确提取建筑目标,因此,本文以算法准确提取目标像素的比例为标准。定义目标准确提取的标准为算法准确提取的建筑目标像素达到 90% 以上。

然后,使用本文算法与同类算法分别进行了建筑目标提取实验。文献[8]提出一种基于局部特征子和概率统计方法的建筑目标提取算法。文献[9]通过目标类别的局部结构进行学习对测试图像中的目标类别进行提取。本文使用建筑图像样本库进行训练,提取图像中的建筑目标。

图 7 是建筑目标提取算法的性能曲线。其中本文算法的分类阈值是纹元模型的距离;文献[8]的分类阈值是密度函数;文献[9]的分类阈值是显著图(activation map)的阈值。从图 7 可知:在相同虚警率的条件下,本文算法的识别率高于其他算法,进而表明,本文使用贝叶斯网建模目标与环境关系约束的算法提取建筑目标的能力,高于文献[8,9]的方法。

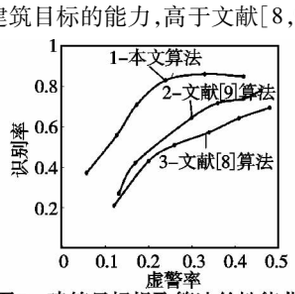

图7 建筑目标提取算法的性能曲线

对实验结果分析后还发现:文献[8]的方法难以有效提取航拍图像的局部特征子,影响了该算法提取建筑目标的性能;文献[9]没有建模环境上下文,导致对以复杂场景为特点的航拍图像提取建筑目标的效果不理想。

为了更好地理解分类器与贝叶斯网在建筑目标提取的作用,对本文算法删除贝叶斯网验证的步骤,并与原方法进行比较。实验结果如图 8 所示。其中完整模式是本文的完整算法,删除验证步骤是从原算法中删除贝叶斯网验证过程。从图 8 可知,删除贝叶斯网验证的步骤后,原算法的建筑目标提取能力有了明显降低,这证明了贝叶斯网验证步骤在整个算法中不可或缺。

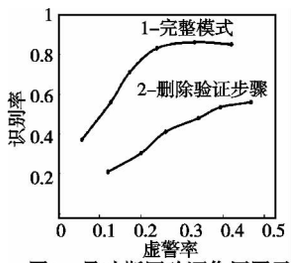

图8 贝叶斯网验证作用图示

本文使用主频为 2 GHz CPU,内存 1 GB 的计算机进行了实验。训练一幅图像平均耗时 2 s,主要花费在与滤波器进行的卷积运算上;待训练完毕后,处理单幅测试图像平均耗时为 0.5 s。

5 结束语

航拍图像场景复杂,现有方法难以有效提取建筑目标,本文分析航拍图像中建筑目标与相邻场景类之间的依赖关系,使用贝叶斯网工具对环境与建筑目标的关系约束进行建模,从而有效地提取和验证弱分类器得到的建筑目标。在航拍图像集上进行的实验表明:与同类算法相比,贝叶斯网能有效去除错误提取的建筑目标,提高了算法的识别率。进一步工作中,笔者将探讨其他贝叶斯网类型应用到建筑目标提取的可能性,进而更好地表示目标与环境之间的关系约束。

参考文献:

[1] PORWAY J, WANG K, ZHU S C. Hierarchical and contextual model for aerial image understanding[J]. International Journal of Computer Vision, 2010, 88(2): 254-283.

[2] KARANTZALOS K, PARAGIOS N. Recognition-driven two-dimensional competing priors toward automatic and accurate building detection[J]. IEEE Trans on Geoscience and Remote Sensing, 2009, 47(1): 133-144.

[3] KATARTZIS A, SAHLI H. A stochastic framework for the identification of building rooftops using a single remote sensing image[J]. IEEE Trans on Geoscience and Remote Sensing, 2008, 46(1): 259-271.

[4] 孙显, 王宏琦, 张正. 基于对象的 Boosting 方法自动提取高分辨率遥感图像中建筑物目标[J]. 电子与信息学报, 2009, 31(1): 177-181.

[5] 唐亮, 谢维信, 黄建军, 等. 从航空影像中自动提取高层建筑物[J]. 计算机学报, 2005, 28(7): 1199-1204.

[6] LI Fei-fei, IYER A, KOCH C, *et al*. What do we perceive in a glance of a real-world scene[J]. Journal of Vision, 2007, 7(1): 1-29.

[7] VARMA M, ZISSERMAN A. A statistical approach to texture classification from single images[J]. International Journal of Computer Vision, 2005, 62(1-2): 61-81.

[8] SIRMACEK B, ÜNSALAN C. Urban area and building detection using SIFT keypoints and graph theory [J]. IEEE Trans on Geoscience and Remote Sensing, 2009, 47(4): 1156-1167.

[9] AGARWAL S, AWAN A, ROTH D. Learning to detect objects in images via a sparse, part-based representation[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2004, 26(11): 1475-1490.