

一种利用纹理特征和朴素贝叶斯分类器 检测近景影像植被的方法

詹总谦^{1,2} 来冰华¹ 万杰³ 李楼¹

(1 武汉大学测绘学院,武汉市珞喻路 129 号,430079)

(2 国土环境与灾害监测国家测绘地理信息局重点实验室,徐州市三环南路,221116)

(3 中国科学院对地观测与数字地球科学中心,北京市海淀区邓庄南路 9 号,北京 100094)

摘要:针对植被对近景摄影测量自动滑坡监测影响严重的问题,提出了一种基于纹理特征和朴素贝叶斯分类器的滑坡植被区域检测算法。对算法的有效性、影像对比度拉伸以及样本训练通用性对检测植被区域的效果等若干问题进行了实验探讨。通过与基于视觉认知特征的检测方法比较,验证了本算法的实用性和有效性。实验结果表明,本文算法能够很好地检测出近景影像中的植被区域,结果比较满意。

关键词:纹理特征;植被检测;贝叶斯分类器;图像标记;凸包计算

中图分类号:P237.3

利用近景影像进行滑坡变形监测的应用中,影响该技术监测精度、可靠性和自动化程度的一个关键因素是滑坡表面覆盖的植被。植被在不同时相的影像上会有生长高度的变化,而且有时会随风摇摆偏离其中心一定距离,从而影响自动化匹配精度。对于完全被植被覆盖的滑坡区域,几乎很难实现自动化匹配,而对于部分或者零星植被覆盖的区域,影像上的植被及其阴影也严重影响了自动化匹配的精度和可靠性。因此,在近景影像的研究应用中,有必要预先剔除影像中植被对应的伪信息。

在传统的近景摄影测量处理中,往往通过人工交互方式检测植被区域,或直接删除误匹配滑坡体 DSM,或使用特殊标志点进行动态监测。由于工作量大、效率低等,不能快速全面、有效地获取较为准确的滑坡变化趋势,因而极大地制约了近景摄影测量在滑坡监测中的广泛应用。

目前,国内外已有许多关于植被检测的研究,但基本上都是针对卫星或遥感影像的^[1-5],近景影像植被检测方面的文献比较少见。近景影像中植被的颜色几乎不受空间分辨率影响,在影像上通常表现得很明显,故可以利用植被的颜色(即基于

视觉认知特征)来进行植被的自动提取^[1]。但是,利用颜色分量来进行植被提取也有一些缺陷,如枯萎植被的颜色不是绿色^[6-7]。本文主要针对近景影像中部分或零星植被覆盖的区域,利用纹理特征和朴素贝叶斯分类器进行植被的自动化检测。

1 植被检测算法

1.1 灰度-梯度共生矩阵和纹理统计量

纹理是图像的一个重要特性,反映了物体表面颜色和灰度的某种变化。灰度-梯度共生矩阵的定义及其构建方法与灰度共生矩阵类似,都是描述图像纹理特征的重要方法。基于灰度-梯度共生矩阵的常用纹理特征统计量有很多^[7]。本文先通过对植被区域和非植被区域的各种纹理特征进行统计,从中选取 8 个具有明显区别的纹理特征统计量用于后续的分类:大梯度优势、梯度分布的不均匀性、相关、灰度平均、梯度平均、梯度标准差、惯性矩、逆差矩。

1.2 朴素贝叶斯分类器

朴素贝叶斯分类器^[8]是一个简单、有效的分

收稿日期:2013-01-02。

项目来源:国家自然科学基金资助项目(41101418,41071292);国土环境与灾害监测国家测绘地理信息局重点实验室资助项目(LED2009B03);中央高校基本科研业务费专项资金资助项目(3101047)。

类器,可以预测一个给定样本属于某一类别的概率。朴素贝叶斯分类器假设特征与给定类的影响独立于其他特征,即特征独立性假设。

设有变量集 $U = \{A_1, A_2, \dots, A_n, C\}$, 其中, A_1, A_2, \dots, A_n 是实例的属性变量, C 是取 m 个值的类变量。假设所有的属性都条件独立于类变量 C , 即每个属性变量都以类变量作为唯一的父节点, 就得到朴素的贝叶斯分类器。使用朴素贝叶斯分类器进行分类的工作过程描述如下。

将每个没有类标号的数据样本用 n 维特征向量 $X = \{X_1, X_2, \dots, X_n\}$ 表示, 分别描述 X 在 n 个属性 $\{A_1, A_2, \dots, A_n\}$ 上的属性值。设有 m 个类 $\{C_1, C_2, \dots, C_m\}$, 则当且仅当 $P(C_i | X) > P(C_j | X)$ ($1 \leq j \leq m, j \neq i$) 时, 最大化 $P(C_i | X)$, 朴素贝叶斯分类将未知样本 X 分配给类 C_i , 其中 $P(C_i | X)$ 最大的类 C_i 为最大后验假定。根据贝叶斯定理:

$$P(C_i | X) = \frac{P(X | C_i)P(C_i)}{P(X)} \quad (1)$$

其中, $P(X)$ 对于所有类为常数。因类的先验概率未知, 故令 $P(C_1), P(C_2), \dots, P(C_m)$ 为等概率。先验概率可以通过 $P(C_i) = N_{C_i} / N$ 计算, N 为已知分类的样本总数, N_{C_i} 为已知分类的样本中类别为 C_i 的样本数。为简化计算, 可以假定属性值间相互独立, 则:

$$P(X | C_i) = \prod_{k=1}^n P(X_k | C_i) \quad (2)$$

概率 $P(X_k | C_i)$ 可以从训练样本估值得到:

$$P(X_k | C_i) = \frac{N_{C_{i,k}}}{N_{C_i}} \quad (3)$$

其中, $N_{C_{i,k}}$ 是在属性 A_k 上具有值 X_k 的类 C_i 的训练样本数。

1.3 后处理

利用纹理特征和朴素贝叶斯分类器检测出来的植被区域内部可能存在一些“空洞”, 或者区域过于琐碎。因此, 实际检测算法应该包括各种后处理工作, 如数学形态学分析、小面积植被区域滤除和区域合并等。其中, 区域合并是采用凸包算法^[9-10]将相隔比较近且面积比较小的植被区域合并起来, 从而降低植被区域的复杂度。

令 S 代表该植被区域的面积, D 表示两个植被区域重心的距离 (S 和 D 的单位为像素), 其合并的原则如下: ① 若面积 $S < 10, D = 50$ 的区域内无其他植被区域, 则删除该区域; ② 若面积 $S < 10, D = 50$ 的区域内有其他植被区域, 则进行区域合并; ③ 若面积 $S > 10, D = 50$ 的区域内有其他植被区域, 则进行区域合并; ④ 若面积 $S > 10,$

$D = 50$ 的区域内无其他植被区域, 则该区域独立, 不进行处理。

1.4 算法流程

1) 预处理。某些图像的反差较小, 需要对影像进行反差拉伸处理, 以达到更满意的检测效果。

2) 计算灰度-梯度共生矩阵。首先对原始影像进行图像灰度平滑、灰度归一化变换(降低比特数, 减少计算量)、梯度矩阵计算, 对梯度矩阵进行归一化处理, 进而计算出灰度-梯度共生矩阵和灰度-梯度的归一化共生矩阵^[11]。

3) 计算样本点纹理统计量, 并选取分类器训练文件。选取一定数量的样本点(包括植被和若干非植被目标), 计算其对应的 8 个统计量。作为贝叶斯分类器的训练文件, 并贴上特殊标签。

4) 利用贝叶斯分类器进行植被检测。遍历影像的每个像元, 计算其纹理统计量, 并用贝叶斯判别函数判定其属于标签中的哪一类, 进而将影像二值化(如植被为 255, 非植被为 0)。

5) 后处理。对初始分类结果进行后处理, 包括腐蚀膨胀处理、小面积植被区域滤除、凸包计算以及植被区域编码输出等。

2 实验分析

实验原型系统使用 VC++ 和 OpenCV 进行开发。硬件系统的配置为宏基 Aspire 4710G, 主频为 2.00 GHz, 内存为 1 GB。实验影像大小为 $4\ 368 \times 2\ 912$ 。

2.1 算法有效性

利用纹理特征和朴素贝叶斯分类器进行原始近景影像植被检测, 图 1(a) 是原始图像, 图 1(b) 是检测结果, 图 1(c) 是经过区域合并处理后的结果。如图 1 所示, 本文算法能很好地将大多数植被区域检测出来。



图 1 植被检测效果

Fig. 1 Vegetation Detection Results

2.2 图像反差对检测结果的影响

有些原始影像的反差较小或者裸露地与植被纹理差异较小, 需要预先对图像进行对比度拉伸, 然后再利用本文算法进行植被检测。图 2 是原始

图像及其植被检测的结果。从图 2 可以看出,检测效果很差,而且存在很多误差检测。图 3 是经过对比度拉伸之后的图像及其检测结果。可以看出,对影像进行对比度拉伸后,植被和非植被的纹理和轮廓已经变得十分清晰,检测效果非常明显。

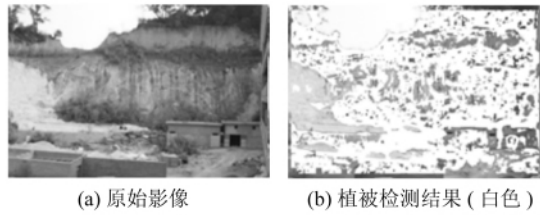


图 2 对比度拉伸前植被检测效果图

Fig. 2 Vegetation Detection Result Before Image Contrast Stretching

2.3 特征统计量的适用性

为了减少样本选择的工作量,同一场景只在

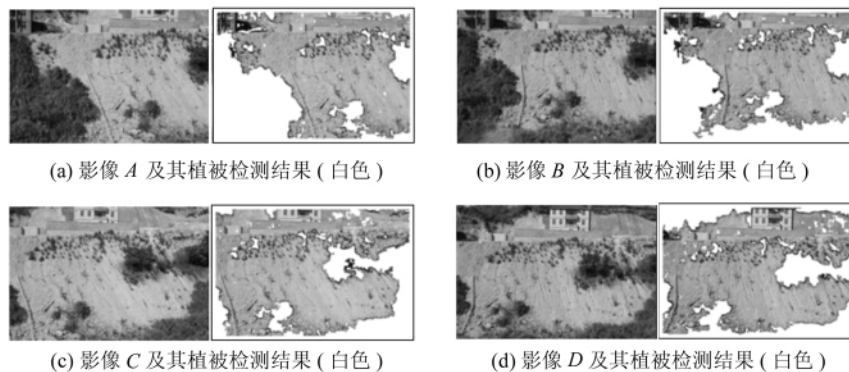


图 4 同一场景影像植被检测效果图

Fig. 4 Vegetation Detection Results at the Same Scene

2.4 基于视觉认知特征方法对比

按照文献[1]和[6]介绍的原理,这里将利用颜色分量检测植被的结果与本文算法进行对比。图 5(a)是原始图像,图 5(b)是本文算法检测的结果,图 5(c)是按文献[1]和[6]所提算法实现的结果。通过对比可以发现,两种方法基本上都能检测出影像中的大多数植被区域。其中,后者有一个明显的缺陷,即不能检测非绿色(枯草)植被。而利用纹理的方法则表现较好,只是在小面积植被区域存在漏检现象。

2.5 算法效率

通过实验发现,算法的主要运行时间消耗在贝叶斯分类上(使用 OpenCV 提供的算法)。若采用逐像素计算纹理特征值并用贝叶斯分类器预测该像素的类属,算法效率极低。因此,在实际处理中我们采用金字塔技术缩小原始影像,然后再进行植被检测。对于大小为 $4\ 368 \times 2\ 912$ 的原

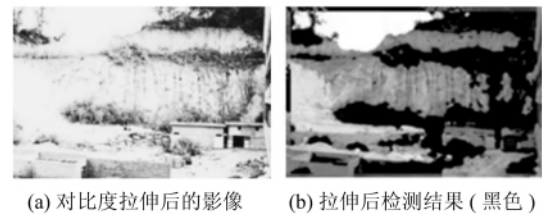


图 3 对比度拉伸后植被检测效果图

Fig. 3 Vegetation Detection Result After Image Contrast Stretching

一幅影像上选取样本。实际检测时,将该幅图像的特征统计量应用到不同位置或不同角度拍摄的近景影像中。图 4 是使用实验一(图 1)的特征统计量对其余图像进行检测的结果。从结果可以看出,同一场景中的纹理特征统计量具有较好的适用性,检测效果较好。

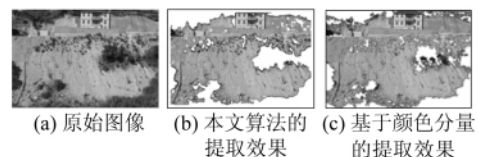


图 5 与基于视觉认知特征方法的对比结果

Fig. 5 Comparison Result of This Method with Visual Cognition Features Method

始影像,若金字塔模板窗口大小采用 5×5 ,则处理前后的耗时分别为 40 s 和 1.5 s。

3 结 语

针对零星或部分植被覆盖的滑坡近景影像,本文利用纹理特征和朴素贝叶斯分类器进行植被检测,重点分析了算法流程及其有效性和适用性等。实验表明,本文算法能够较好地检测出近景影像中的植被,极大地促进了近景摄影测量在滑

坡自动化监测中的应用。但是,算法的漏检和错检分析、算法复杂度以及如何对检测精度进行定量分析等,需要进一步探讨和研究。

参 考 文 献

- [1] 柳稼航. 高分辨率光学遥感影像视觉认知与自动识别技术研究[D]. 上海:上海交通大学,2010
- [2] 陈忠. 高分辨率遥感图像分类技术研究[D]. 北京:中国科学院研究生院,2006
- [3] 张倩,黄昕,张良培. 多尺度同质区域提取的高分辨率遥感影像分类研究[J]. 武汉大学学报·信息科学版,2011,36(1):117-121
- [4] Zhao Yindi, Zhang Liangpei, Li Pingxiang. Texture Feature Fusion for High Resolution Satellite Image Classification[C]. 2005 International Conference on IEEE, New Trends, 2005
- [5] 闫利,聂倩,胡文元,等. 基于对象级的 ADS40 遥感影像分类研究[J]. 武汉大学学报·信息科学版, 2009,34(2):183-186
- [6] 李楼. 近景影像植被检测算法研究[D]. 武汉:武汉大学,2010
- [7] 万杰. 近景影像植被自动检测算法研究[D]. 武汉:武汉大学,2011
- [8] 周龙. 基于朴素贝叶斯的分类方法研究[D]. 合肥:安徽大学,2006
- [9] 王杰臣,陈焱明. 一种栅格辅助的平面点集最小凸包生成算法[J]. 武汉大学学报·信息科学版,2010, 35(4):403-406
- [10] 刘广忠,黄琳娜. 基于二叉树的散乱点集快速凸包算法[J]. 测绘科学,2008, 33(4): 87-89
- [11] 洪继光. 灰度-梯度共生矩阵纹理分析方法[J]. 自动化学报,1984,10(1):22-25

第一作者简介:詹总谦,副教授,研究方向为摄影测量与计算机视觉。

E-mail:zqzhan@sgg.whu.edu.cn

Vegetation Detection of Close-Range Images Based on Texture Features and Naive Bayes Classifier

ZHAN Zongqian^{1,2} LAI Binghua¹ WAN Jie³ LI Lou¹

(1 School of Geodesy and Geomatics, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

(2 Key Laboratory for Land Environment and Disaster Monitoring of NASMG, South Sanhuan Road, Xuzhou 221116, China)

(3 Center for Earth Observation and Digital Earth of CAS, 9 South Dengzhuang Road, Beijing 100094, China)

Abstract: Considering of the problems of serious effects caused by vegetation on automatic landslide monitoring in close-range photogrammetry, we present a method of detecting vegetation regions in landslide images based on texture features and naive Bayes classifier. Some meaningful discussions and analysis have been done mainly for the effectiveness of this algorithm, image contrast stretching and the generality problem of samples training. Comparing with another detection method based on visual cognition features, we prove the availability and the validity of this method. The experimental results show that the vegetation detection method can almost detect the vegetation regions from close-range images and the result is satisfying.

Key words: texture features; vegetation detection; naive Bayes classifier; image labelling; convex hull calculation

About the first author: ZHAN Zongqian, associate professor, majors in photogrammetry and computer vision.

E-mail: zqzhan@sgg.whu.edu.cn