

基于改进HOG特征的建筑物识别方法

杨松^{1,2,3}, 李盛阳^{1,2}, 邵雨阳^{1,2}, 郑贺^{1,2,3}

YANG Song^{1,2,3}, LI Shengyang^{1,2}, SHAO Yuyang^{1,2}, ZHENG He^{1,2,3}

1. 中国科学院 空间应用工程与技术中心, 北京 100094

2. 中国科学院 太空应用重点实验室, 北京 100094

3. 中国科学院大学, 北京 100049

1. Technology and Engineering Center for Space Utilization, Chinese Academy of Sciences, Beijing 100094, China

2. Key Laboratory of Space Utilization, Chinese Academy of Sciences, Beijing 100094, China

3. University of Chinese Academy of Sciences, Beijing 100049, China

YANG Song, LI Shengyang, SHAO Yuyang, et al. Building recognition method based on improved HOG feature. Computer Engineering and Applications, 2018, 54(7): 196-200.

Abstract: With the development and the extensive application of machine learning methods, the building recognition technology is developed rapidly. For the building recognition, the recognition accuracy is still the focus of attention. The traditional gradient method of Histogram of Oriented Gradients(HOG) can not effectively describe the boundary characteristics of the building and it directly affects recognition result. The HOG method based on the steerable filters is proposed, and Support Vector Machine(SVM) is used as the training method. The experimental result is analyzed according to the average accuracy rate, *TP*, *FP*, Recall, Precision and *F₁*. The results show that the proposed method has better performance than the steerable filtered-based building recognition method, and it is proven that the proposed method can effectively identify buildings.

Key words: building recognition; histogram of oriented gradients; feature extraction; steerable filters; support vector machine

摘要:随着机器学习方法的广泛应用,建筑物识别技术得到了快速的发展,识别的准确性一直是人们关注的重点。梯度方向直方图(HOG)特征提取方法中的梯度求解方式不能有效提取建筑物的边界特征,直接影响了识别的准确性,提出基于方向可控滤波器的HOG算法,利用支持向量机学习方法实现建筑物的识别。实验结果表明,该方法在平均准确率、*TP*、*FP*、召回率、精确率和 F_1 值等指标上优于基于方向可控滤波器的建筑物识别方法,证明了该方法可以有效识别建筑物。

关键词:建筑物识别;梯度方向直方图;特征提取;方向可控滤波器;支持向量机

文献标志码:A **中图分类号:**TP391 **doi:**10.3778/j.issn.1002-8331.1611-0332

1 引言

建筑物识别是计算机视觉、模式识别领域的研究热点之一,它能够使人们根据图像快速获取建筑物的位置、名称、描述等相关信息,在建筑物定位、建筑设计、建筑物标记等领域有着重要的应用价值,而如何有效地描述建筑物的特征是建筑物识别的关键问题。

李松霖等人^[1]提出了一种基于特征线匹配的城市建筑物识别系统,该系统可以满足移动条件下建筑物快速识别的需求,但是准确率不高。董肖等人^[2]提出了一种快速稳健的建筑物图像识别系统,该系统在建筑物识别上具有很高的准确性,但是图像识别的速度较慢。齐沁芳等人^[3]提出了一种应用于增强现实系统的建筑物识别

基金项目:中国科学院空间科学研究院培育项目(No.Y5031211WY)。

作者简介:杨松(1991—),男,硕士研究生,研究领域为图像处理、机器学习;李盛阳(1976—),男,博士,研究员,研究领域为图像处理与解译、机器学习, E-mail: shyli@csu.ac.cn;邵雨阳(1985—),男,助理研究员,研究领域为航天数据利用技术;郑贺(1988—),男,在读博士研究生,研究领域为图像处理、机器学习。

收稿日期:2016-11-18 **修回日期:**2017-02-06 **文章编号:**1002-8331(2018)07-0196-05

CNKI网络优先出版:2017-05-10, <http://kns.cnki.net/kcms/detail/11.2127.TP.20170510.1115.010.html>

算法,可通过手持设备实时获得建筑物的相关信息,图像识别速度较快,但准确率不高。Li等人^[4]提出基于方向可控滤波器的建筑物识别方法(Steerable Filter-based Building Recognition,SFBR),首次将方向可控滤波器应用在建筑物识别中,提取建筑物不同方向的边界信息作为特征。蔡兴泉等人^[5]使用SIFT算法进行图像识别,但SIFT计算量较大,响应时间较长。

梯度方向直方图(Histograms of Oriented Gradient,HOG)是近年来在目标识别研究中应用较为广泛的特征提取方法,HOG最早由Dalal等人^[6]提出,应用于行人检测中,并取得了较好的效果。HOG是一种基于图像梯度方向的特征描述子,能够捕捉局部的轮廓信息,较好地描述物体边缘形状,同时,对图像几何和光学形变保持很好的不变性,而且计算量较小^[7]。由于建筑物存在丰富的边缘信息,因此HOG适用于提取建筑物特征。徐云云^[8]将方向可控滤波器和HOG方法结合,并通过特征匹配的方式实现掌纹识别,该方法需要实验确定滤波器方向,而且识别效率较低。

本文深入研究了方向可控滤波器在建筑物边界提取的特点,提出了一种基于改进HOG的建筑物识别方法。该方法利用方向可控滤波器替换传统HOG方法中的滤波器模版,改进HOG的梯度求解方法,更好地提取水平方向和垂直方向的边缘信息,并结合支持向量机器学习方法对建筑物进行分类,通过实验验证了方法的有效性。

2 基于方向可控滤波器的HOG方法

2.1 方向可控滤波器

Freeman等人^[9]提出方向可控滤波器(Steerable Filters)的概念,并介绍了设计方法,可控滤波器在图像边缘检测、纹理分析等领域表现出较好的优势。方向可控滤波器是基于函数极坐标形式的傅里叶级数展开式,由若干个基滤波器函数的线性组合而成^[10]。

方向可控滤波器的一般形式为:

$$F^\theta = \sum_{i=1}^N k_i(\theta) F_i$$
 (1)

其中 N 为基滤波器的个数, F_i 为第 i 个基滤波器, $k_i(\theta)$

为仅与方向 θ 有关的滤波器系数, F^θ 为 θ 方向的滤波器。

由于高斯函数的各阶导函数都可以表示为一个圆对称的窗函数和一个多项式的乘积,常常采用基于高斯函数的方法来构造可控滤波器。本文采用基于高斯函数二阶导数所构成的可控滤波器,形式如下:

$$G_2^\theta = k_1(\theta)G_2^0(x,y) + k_2(\theta)G_2^{\pi/3}(x,y) + k_3(\theta)G_2^{2\pi/3}(x,y)$$
 (2)

其中 $G_2^0(x,y)$ 、 $G_2^{\pi/3}(x,y)$ 、 $G_2^{2\pi/3}(x,y)$ 分别为高斯函数在对应方向上的二阶导数,具体表达式如下所示:

$$\begin{cases} G_2^0(x,y) = 0.9213(2x^2 - 1)e^{-(x^2+y^2)} \\ G_2^{\pi/3}(x,y) = 1.8430xye^{-(x^2+y^2)} \\ G_2^{2\pi/3}(x,y) = 0.9213(2y^2 - 1)e^{-(x^2+y^2)} \end{cases}$$
 (3)

对应系数分别为:

$$\begin{cases} k_1(\theta) = \cos^2(\theta) \\ k_2(\theta) = -2\cos(\theta)\sin(\theta) \\ k_3(\theta) = \sin^2(\theta) \end{cases}$$
 (4)

2.2 改进的HOG特征提取方法

HOG方法中往往采用较为简单的梯度求解方法。如: $[-1 \ 0 \ 1]$, $[-1 \ 0 \ 1]^T$ 。传统的梯度求解方法不能有效地提取目标的边界特征,本文研究利用可控滤波器在边缘信息表达上的优势,提出改进的HOG算法,其特征提取流程如图1所示,具体步骤如下:

(1)将彩色图像转化为灰度图像。

(2)构造水平和竖直方向的方向可控滤波器(令滤波器的方向分别取 0 、 $\pi/2$),分别记为 F^0 、 $F^{\pi/2}$ 。计算灰度图像 (x,y) 处像素点在水平和竖直方向的梯度值 $G_x(x,y)$ 、 $G_y(x,y)$:

$$G_x(x,y) = F^0 * I$$
 (5)

$$G_y(x,y) = F^{\pi/2} * I$$
 (6)

式中 I 为灰度图像。

分别计算 (x,y) 处像素点的边界方向和幅值:

$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \approx |G_x(x,y)| + |G_y(x,y)|$$
 (7)

$$\theta(x,y) = \arctan\left(\frac{G_y(x,y)}{G_x(x,y)}\right)$$
 (8)

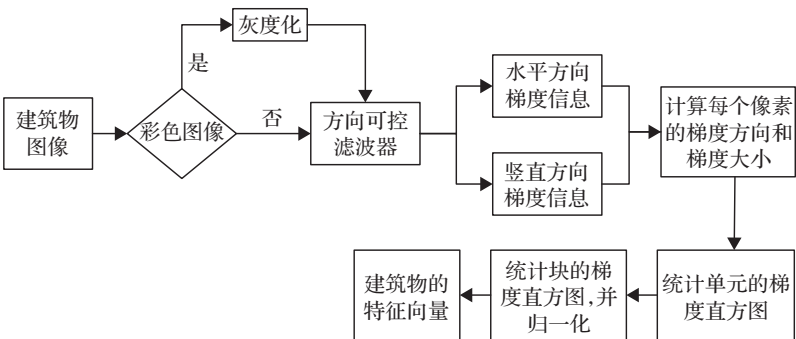


图1 改进HOG方法流程图

如果 $\theta(x, y) < 0$, 令 $\theta(x, y) = \theta(x, y) + \pi$, 使得 $\theta(x, y) \in [0, \pi)$ 。

(3) 将图像划分成大小相同的单元(Cell), 将相邻的单元组合成有重叠的块(Block), 有效利用重叠的边缘信息, 如图2所示。

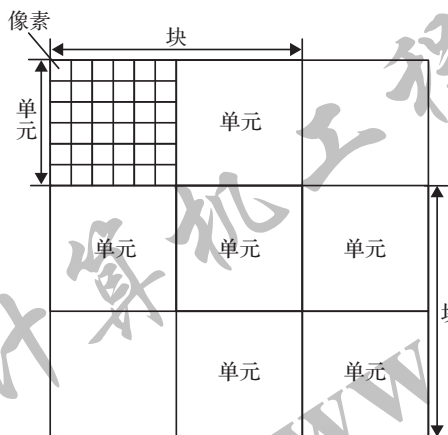


图2 像素、单元、块之间的关系

(4) 图像中每个像素点的梯度方向和梯度幅值各不相同, 按梯度方向划分为若干个均匀的区域(Bin), 将单元中的每个像素点的梯度幅值累加到对应 Bin 中, 生成单元的梯度方向直方图。然后统计整个块的直方图特征, 并对每个块内的梯度直方图进行 L_2 归一化, 即对于向量 v , 令:

$$v \leftarrow \frac{y}{\sqrt{\|v\|^2 + \epsilon^2}} \quad (9)$$

(5) 将所有块的 HOG 特征组合形成图像的 HOG 特征向量, 其维度 d 为:

$$d = \left(\frac{w-b}{s} + 1\right) \times \left(\frac{h-b}{s} + 1\right) \times \left(\frac{b}{c}\right)^2 \times p \quad (10)$$

其中 w 、 h 为图像的宽度和高度; b 、 c 分别为块、单元的大小; p 为单元内梯度方向的数目; s 为块移动的步长。

3 支持向量机

支持向量机(Support Vector Machine, SVM)是 Vapnik^[11]提出的一种监督式学习方法, 它建立在统计学习理论中 VC 维理论和结构风险最小原理的基础上, 通过适当地选择函数子集及该子集中的判别函数, 使得学习机器的实际风险达到最小^[12]。支持向量机具有结构简单、全局最优、泛化性能好、学习和预测效率高等优点^[13], 能够成功地处理分类问题, 因此被广泛应用于文本识别^[14]、图像识别^[15]等领域。SVM 的原理是寻找一个满足分类要求的最优分类超平面, 使得该超平面在保证分类精度的同时, 能够使分类间隔最大化。

以二分类问题为例, 给定训练样本集 (x_i, y_i) , $i = 1, 2, \dots, l$, $x \in R^n$, $y \in \{\pm 1\}$, 超平面记为 $(w \cdot x) + b = 0$, 为能够对所有样本正确分类并且具备分类间隔, 需满足如下约束条件:

$$y_i[(w \cdot x) + b] \geq 1, i = 1, 2, \dots, l \quad (11)$$

可计算出分类间隔为 $2/\|w\|$, 可将原问题转化为下述约束式:

$$\min \Phi(w) = \frac{1}{2} \|w\|^2 = \frac{1}{2} (w' \cdot w) \quad (12)$$

为了求解该最优化问题, 引入 Lagrange 函数:

$$L(w, b, a) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l a_i (y_i ((w \cdot x) + b_i) - 1) \quad (13)$$

其中 $a_i > 0$ 为 Lagrange 乘数。

将问题转化为相应的对偶问题, 即:

$$\max Q(a) = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j (x_i \cdot x_j) \quad (14)$$

$$\text{s.t. } \sum_{i=1}^l a_i y_i = 0 \quad a_i \geq 0 \quad (15)$$

解得最优解 $a^* = (a_1^*, a_2^*, \dots, a_l^*)^T$ 。从而计算出最优权值向量 w^* 和最优偏置 b^* :

$$w^* = \sum_{i=1}^l a_i^* y_i x_i \quad (16)$$

$$b^* = y_j - \sum_{i=1}^l a_i^* y_i (x_i \cdot x_j) \quad (17)$$

得到最优分类器超平面 $(w^* \cdot x) + b^* = 0$, 相应地, 最优分类函数为:

$$f(x) = \text{sgn}\{(w^* \cdot x) + b^*\}, x \in R^n \quad (18)$$

对于线性不可分情况, 对 x 做 Φ 变换映射到特征空间 H , 即

$$x \rightarrow \Phi(x) \quad (19)$$

则最优分类函数为:

$$f(x) = \text{sgn}\{(w^* \cdot \Phi(x)) + b^*\} \quad (20)$$

本文采用 libsvm 库^[16]进行 SVM 训练, 建筑物识别是多分类问题, libsvm 使用的是“一对一”方法, 即在任意两类样本之间建立一个 SVM, k 个类别的样本则建立 $k(k-1)/2$ 个 SVM。当对一个样本进行分类时, 得票最多的类别即为该样本的类别。

4 实验结果及分析

为了客观评价方法在建筑物识别中的有效性, 本文进行的建筑物识别实验采用与 SFBR 方法相同的数据集——SBID^[17], 并对结果进行分析。

4.1 数据集介绍

SBID 数据集包含 40 类建筑物, 共 4 178 张图像, 每张图像均放缩为 160×120 像素。这些图像拍摄时间不同, 拍摄角度不同, 拍摄远近不同, 如图 3 所示。

4.2 实验平台及参数设置

实验配置为处理器 Intel® Core™ CPU i7-3770 @ 3.40 GHz, 4 GB 内存, 操作系统为 32 位 Windows 7, 开发软件是 Visual Studio 2010, 采用 C/C++ 语言编程实现。



图3 SBID建筑物图像(从左到右,从上到下依次代表1~40类)

方向可控滤波器 F^0 、 $F^{\pi/2}$ 的大小均为 7×7 , w 、 h 为 120、160; c 、 b 、 p 、 s 分别取 20、80、9、40, 特征向量的维度 d 为 864。通过大量实验, 支持向量机选用 C_SVC 分类器和线性核函数时, 效果最好。

4.3 实验结果

本文在 SBID 数据集上进行 20 次独立的实验, 每次实验随机从每类建筑物中选取一半数量的图像 (共 2 098 幅图像) 进行特征提取并训练 SVM 分类器, 剩余的图像 (共 2 080 幅图像) 进行特征提取作为测试样本。

实验结果如图 4 所示, 由于每次实验选取的训练样本和测试样本不同, 同类图像在不同实验中的误分个数有所变化。各类图像在不同实验中的误分个数均在 15 以内, 其中有多处误分个数为 0, 如第 4 次实验中第 1、2 类建筑物全部分类正确; 第 18 次实验中第 16 类建筑物全部分类正确; 第 20 次实验中第 4、7、8、9 类建筑物全部分类正确。

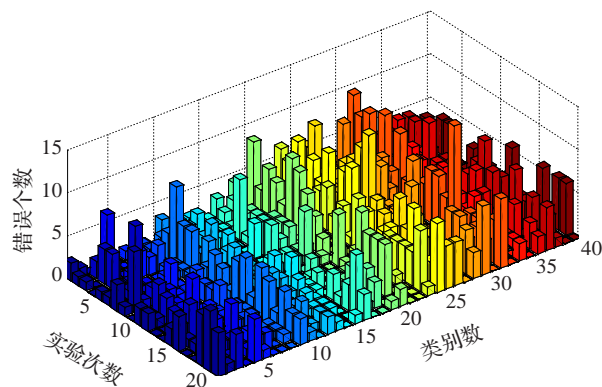


图4 实验结果

记 β 为识别准确率, 其计算方法为:

$$\beta = \frac{M}{N} \times 100\% \tag{21}$$

其中 N 为测试样本的个数, M 为识别正确的样本个数。

记 β_i 为第 i 类建筑物的识别准确率, 其计算方法为:

$$\beta_i = \frac{M_i}{N_i} \times 100\% \tag{22}$$

其中 N_i 为第 i 类建筑物测试样本的个数, M_i 为第 i 类建筑物识别正确的样本个数。

如图 5 反映了各类图像的平均准确率, 第 23 类建筑物的平均准确率最高, 接近 100%, 第 32 类建筑物的平均准确率最低, 介于 85%~90% 之间。除第 32 类建筑物以外, 其他建筑物的平均准确率均大于 90%, 其中有 30 类建筑物的平均准确率高 于 95%。

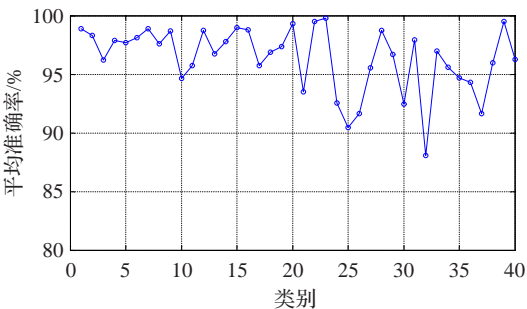


图5 每类图像的平均识别准确率

如图 6 反映了每次实验的识别准确率, 可以看出本文方法的准确率在 97% 附近波动, 计算得出 20 次实验的平均准确率为 96.56%。SFBR 方法的平均准确率为 94.66%, 与 SFBR 方法相比, 本文方法的平均准确率提高了 1.9%。

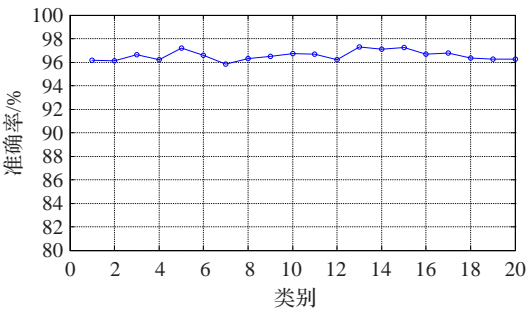


图6 每次实验的识别准确率

为了进一步验证方法的性能,实验采用召回率 *Recall*、精确率 *Precision*、 F_1 值等评价指标进行分析。

记 TP 为“真”正例的个数, FP 为“伪”正例的个数, TN 为“真”负例的个数, FN 为“伪”负例的个数。召回率、精确率、 F_1 值的计算公式分别如下:

$$Recall = \frac{TP}{TP + FN} \quad (23)$$

$$Precision = \frac{TP}{TP + FP} \quad (24)$$

$$F_1 = \frac{2TP}{2TP + FP + FN} \quad (25)$$

本文求得统计量在每类图像的分类性能,然后求均值作为最终的结果,结果如表1所示。从表1可以看出,本文方法在 TP 、 FP 、*Recall*、*Precision*、 F_1 值等指标上均优于 SFBR 方法。

表1 实验结果

方法	TP	FP	<i>Recall</i>	<i>Precision</i>	F_1
SFBR	49.150	2.850	0.956	0.942	0.947
本文方法	50.000	2.000	0.967	0.959	0.962

5 结束语

本文将方向可控滤波器应用到 HOG 算法中,结合支持向量机方法,提出一种建筑物识别方法。以 SBID 数据集作为实验数据,使用 SVM 对提取的建筑物特征进行训练与识别。实验结果表明,在建筑物识别中,本文方法得到了较高的准确率,并且识别效果优于 SFBR 方法。因此本文提出的建筑物识别方法可以提取建筑物的关键特征,增强建筑物的区分性,有效地对建筑物进行识别。

在后续的工作中,将进一步研究方向可控滤波器的特点,提取更有效的边界特征从而提高识别的准确率和鲁棒性。此外,研究提取颜色信息作为建筑物识别的辅助特征,进一步提高建筑物识别率。

参考文献:

[1] 李松霖,范海生,陈秀万.基于特征线匹配的城市建筑物识别方法研究[J].遥感技术与应用,2012,27(2):190-196.

- [2] 董肖.快速稳健的建筑物识别算法与系统[D].广州:华南理工大学,2012.
- [3] 齐沁芳.应用于增强现实系统的建筑物识别算法的研究与实现[D].北京:北京邮电大学,2013.
- [4] Li J, Allinson N. Building recognition using local oriented features[J]. IEEE Transactions on Industrial Informatics, 2013, 9(3): 1697-1704.
- [5] 蔡兴泉,柳静华.建筑物图像识别系统设计与实现[J].现代计算机:专业版,2015(14):18-20.
- [6] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//Proceedings of IEEE Conference on Computer Vision & Pattern Recognition, 2013: 886-893.
- [7] 王阳,穆国旺,睢佰龙.基于 HOG 特征和 SVM 的人脸表情识别[J].河北工业大学学报,2013(6):39-42.
- [8] 徐云云.面向智能手机的掌纹识别技术研究[D].合肥:合肥工业大学,2015.
- [9] Freeman W T, Adelson E H. The design and use of steerable filters[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1991, 13(9): 891-906.
- [10] 赵渊洁.结合可控滤波器与偏微分方程的图像去噪方法[D].天津:天津理工大学,2013.
- [11] Vapnik V N. An overview of statistical learning theory[J]. IEEE Transactions on Neural Networks, 1999, 10(10): 988-999.
- [12] 丁世飞,齐丙娟,谭红艳.支持向量机理论与算法研究综述[J].电子科技大学学报,2011,40(1):2-10.
- [13] 郭明玮,赵宇宙,项俊平,等.基于支持向量机的目标检测算法综述[J].控制与决策,2014(2):193-200.
- [14] 刘晓亮,丁世飞,朱红,等.SVM 用于文本分类的适用性[J].计算机工程与科学,2010,32(6):106-108.
- [15] 李雪花,许姜滢宇,于安军,等.基于 SVM 多类分类器的字符识别[J].信息技术,2016(1):20-22.
- [16] Chang C C, Lin C J. LIBSVM: A library for support vector machines[J]. ACM Transactions on Intelligent Systems & Technology, 2011, 2(3): 389-396.
- [17] Li J, Allinson N M. Subspace learning-based dimensionality reduction in building recognition[J]. Neurocomputing, 2009, 73(1/3): 324-330.