

文章编号 1004-924X(2013)04-1047-07

联合梯度直方图和局部二值模式特征的人体检测

黄 炎, 范赐恩*, 朱秋平, 张 虎, 邓德祥
(武汉大学 电子信息学院, 湖北 武汉 430079)

摘要:针对采用单一梯度方向直方图(HOG)特征进行人体检测时易受竖直梯度分量干扰的缺点,提出了将分块局部二值模式(LBP)特征加入 HOG 特征的方法。首先,将检测窗口划分为大小为 16×16 的不重叠块,以块为单位统计 LBP 特征直方图,并通过大量实验获得了 LBP 算子的最佳参数;然后用优化过的插值方式计算 HOG 特征,将两者组成联合直方图。最后,用线性支持向量机(SVM)通过 Bootstrapping 的方式训练,得到判别模型。在 INRIA 人体库上的测试表明,检出率在误检率(FPPW)为 10^{-4} 时由原始的 89% 提高到 95%,单窗口检测速度由 0.625 ms 提高到 0.533 ms。本文将纹理特征加入原始描述轮廓的 HOG 特征中,排除了部分梯度干扰信息造成的误检,提高了检出率。

关键词:梯度方向直方图;分块局部二元模式;支持向量机;人体检测

中图分类号:TP391 **文献标识码:**A **doi:**10.3788/OPE.20132104.1047

HOG-LBP pedestrian detection

HUANG Yan, FAN Ci-en, ZHU Qiu-ping, Zhang Hu, DENG De-xiang

(School of Electronic Information, Wuhan University, Wuhan 430079, China)

* Corresponding author, E-mail: fce@whu.edu.cn

Abstract: This paper proposed a method to concatenate a cell-structured Local Binary Pattern(LBP) feature into Histogram of Gradients(HOG) to solve the problem that HOG was vulnerable to the interference of vertical background gradient information in pedestrian detection. Firstly, the detection window was divided into 16×16 non-overlapping blocks, then the LBP histogram of each block was calculated and his parameters were obtained by extensive experiments. Afterwards, the HOG was computed by the optimized interpolation method, and it was combined with LBP histogram to constitute a joint histogram. Finally, a discriminative model was trained by Bootstrapped linear Support Vector Machine(SVM). Based on the test of the INRIA pedestrian dataset, it is shown that the detection rate has been increased from 89% of the HOG feature to 95% when False Positive Per Window(FPPW) is 10^{-4} , and the detection speed has been raised from 0.625 to 0.533 ms per window. It is concluded that the proposed method in this paper eliminates the false detection caused by the interference of gradient information and improves the detection rate by describing both contour and texture information.

Key words: Histogram of Gradient(HOG); cell-structured Local Binary Pattern(LBP); Supporting Vector Machine(SVM); pedestrian detection

收稿日期:2012-12-11;修订日期:2012-12-31.

基金项目:国家自然科学基金面上项目(No. 61072135)

1 引言

人体检测是目标检测^[1]中的一个典型问题,主要应用在机器人、视频监控、汽车安全等领域。由于人体高矮、姿态多变及易受遮挡、环境光线等因素的影响,使得人体特征存在较大的类内杂散度^[2]。目前主要采用基于形状的特征来描述人体。

受到仅关注感兴趣点的稀疏描绘子尺度不变特征变换(SIFT)^[3]算子的影响,Dalal 等人提出了基于梯度方向直方图(HDG)的稠密描述子 HOG,在 MIT 人体库上得到了接近完美的结果,为此建立了更具挑战性的 INRIA 人体库^[4]。在此基础上产生了多种联合 HOG 的多特征描绘子,如 Walk^[5]等人将 128 个块的颜色信息加入到 HOG 算子中,充分利用了人体肤色的局部相似性,但颜色空间的转换计算量很大。针对 HOG 特征丢失了梯度矢量坐标位置的缺陷,文献^[6]提出的共生 HOG(CoHOG)考虑了像素间的 31 种偏移关系,缺点是其特征空间高达 34 704 维。Maji^[7]等在大小可变的非重叠块上计算各方向边缘的能量特征,这一简单的特征联合直方图交叉核支持向量机(SVM)时取得了优于 HOG 的效果,但其在线性 SVM 下性能下滑了约 47%。Felzenszwalb^[8]等人以金字塔结构上经主成份分析(PCA)降维后的 HOG 特征作模板,利用隐式 SVM 来求解人体部件的位置偏移这一参数,一定程度上解决了人体的形变和遮挡问题,但其需要在较高分辨率的图片上建立人体部件模型,且训练过程复杂。

文献^[9]采用了联合特征的方法,但其计算过程存在大量的卷积操作,作者并未提及其计算时间。文献^[10]采用了积分直方图^[11]来计算多种尺寸块的 HOG 特征,Adaboost 级联分类器的应用使检测速度提升了约 70 倍。此外,利用 GPU 加速可使积分直方图的计算、窗口检测过程并行化,如文献^[12],但这些方法都弱化了检测器性能。

本文从 HOG 特征存在的缺陷出发,加入反映图像纹理结构的局部二值模式(LBP)特征来补充单一的梯度特征,且对单调变化的灰度具有不变性、计算复杂度低。实验表明,本算法减小了类

似人体腿部——有着平行垂直分量物体的误检率,提高了人体轻微形变时的检出率。

2 HOG 特征及其缺陷

以 64×128 的窗口为单位提取 HOG 特征,将其划分成 8×8 的子块(cell),邻近的 4 个子块组成 1 个 32×32 的块(block),如图 1 所示,每个子块在 4 个不同块中各出现一次。

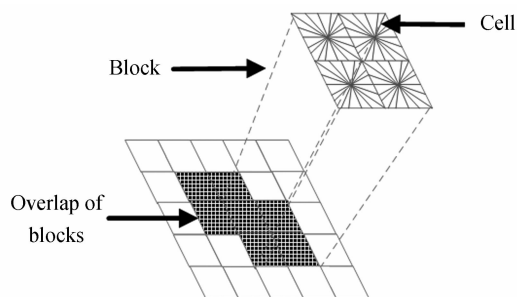


图 1 块划分

Fig. 1 Block division

再计算 RGB 3 个通道下的梯度,以最大值作为该点的梯度幅值,将无方向的梯度角度 $0 \sim 180^\circ$ 划分为 9 个箱子(bin),以线性插值的方式分配到邻近的箱子中,如图 2 所示。以子块为单位构建梯度方向直方图,窗口中有 $7 \times 15 \times 4$ 个子块,顺次连接成 3 780 维的特征向量^[13]。

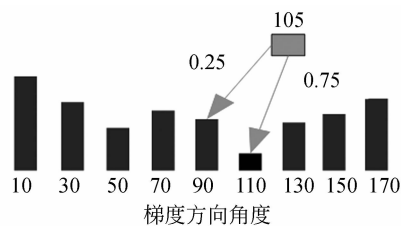


图 2 角度插值运算

Fig. 2 Angle interpolation

原始 HOG 算子的各像素点梯度幅值根据式(1)双线性插值到 4 个 cell 构建的直方图中,式中: $|\nabla f(x, y)|$ 表示梯度幅值, d_x 、 d_y 表示 cell 的尺寸, (x_i, y_i) 表示 cell 中心坐标, (x, y) 表示该点坐标,如图 3 所示^[14]。

$$|\nabla f(x, y)| \left(1 - \frac{x - x_i}{d_x}\right) \left(1 - \frac{y - y_i}{d_y}\right) \quad i=1, 2, 3, 4. \quad (1)$$

而本文采取如下插值法:将 4 个子块进一步

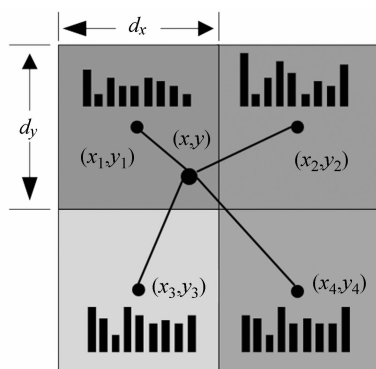


图 3 空间插值

Fig. 3 Spatial interpolation

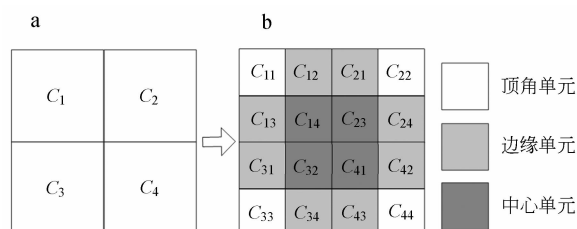


图 4 改进的空间插值

Fig. 4 Improved spatial interpolation

划分成 16 个单元,如图 4 所示, C_{11} 、 C_{22} 、 C_{33} 、 C_{44} 称为顶角单元,只对所在 cell 直方图有贡献;而 C_{14} 、 C_{23} 、 C_{32} 、 C_{41} 称为中心单元,对 4 个 cell 都要进行插值运算;其余的称为边缘单元,对与其紧邻的 2 个 cell 有贡献。 C_1 对应的直方图按式(2)计算,而原始计算方式如式(3),式中 hist 均表示直方图,可见加法次数减少为 9/16 倍。

$$hist(C_1) = \sum_{i=1}^4 hist(C_{1i}) + hist(C_{21}) + hist(C_{31}) + hist(C_{23}) + hist(C_{41}) + hist(C_{32}), \quad (2)$$

$$hist(C_1) = \sum_{i=1}^4 \sum_{j=1}^4 hist(C_{ij}). \quad (3)$$

HOG 特征最显著的优点是在坐标空间及角度方向都采用了插值运算,缓解了子块划分可能造成的不一致性。然而,HOG 特征对梯度敏感,如对人体衣服图案、周围背景的描述都是干扰信息,降低了人体特征描述的分辨力。其次,HOG 特征将存在大量细小曲线的人体轮廓用切向直线近似表达,描述欠准确,易对类似人体躯干,上方呈 Ω 形、下方有着强烈垂直分量的物体造成误检,如图 5 所示,树木、电线杆、建筑物、花瓣等是常见的误检。



图 5 HOG 算子误检

Fig. 5 False positive by HOG

本文采取特征融合的方法克服以上弱点,将近年来流行于人脸检测的 LBP 特征加入到人体检测中。

3 HOG-LBP 联合特征向量

LBP 算子最初用于纹理分类,是一种统计和结构相结合的方法。LBP 算子定义在圆圈上,如图 6 所示,表达式如式(4), x_i 表示圆周上点的灰度值, x_c 表示圆心点的灰度值, p 表示圆周上点的个数, r 表示圆的半径^[15]。

$$LBP_{p,r} = \sum_{i=0}^{p-1} f(x_i - x_c) \cdot 2^i \quad f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}. \quad (4)$$

该函数舍弃了灰度差值来获得对单调灰度变化的不变性。LBP 算子的值每位都为 0 或 1 时表示平坦区域或单个点;对于边缘,沿着梯度方向的值为 1,顺着边缘方向的值为 0,1\0 呈对称分布;而角落则呈不对称分布,如图 6 所示。这些对纹理结构的描述加强了人体的特征表达。

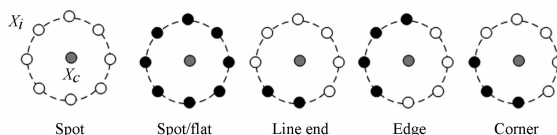


图 6 LBP 纹理结构表达

Fig. 6 LBP texture structure

当 $LBP_{8,1}$ 编码中 0\1 或 1\0 跳变次数不超过两次时,被称为一致(uniform)模式。一致模式的编码有 58 种,将所有非一致模式归为一种,经过一致性映射后,编码方式由 2^8 种缩短为 59 种。通过对人体图像的统计得出约 90% 的模式是一致的,非一致模式则代表了噪声或繁杂的纹理,而这些往往是干扰信息。利用一致性,既表达了主

要的纹理特性,又削弱了干扰区域的特征表达。

联合特征向量的计算过程如图 7 所示:分别计算 HOG 特征直方图和 LBP 特征直方图,连接起来形成联合直方图。

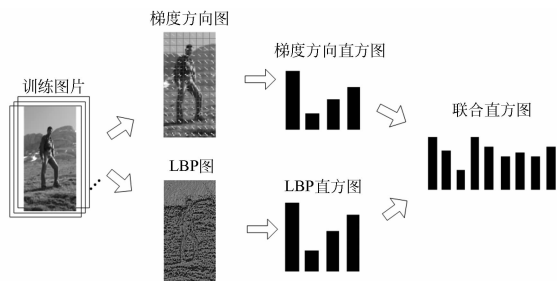


图 7 联合直方图

Fig. 7 Joint histogram

4 训练算法流程

样本集的特征计算完毕后进入到训练过程,由于本文采用的特征向量维数较高,线性 SVM 即可保证分类效果,本文采用专门针对线性 SVM 的 LIBLINEAR^[16]。采用如表 1 所示的训练方法发掘困难负样本以降低误检率。困难负样本的定义如式(5), x 表示样本特征, y 表示样本标记, N 表示样本集, β 表示判别模型, H 表示被错分或离分界面近的本样本。

$$H(\beta, N) = \{(x, y) \in N \mid y \cdot \beta(x) < 1\}. \quad (5)$$

表 1 Bootstrapped SVM 训练过程

Tab. 1 Bootstrapped SVM training

算法 1 Bootstrapping 的训练算法
 在初始样本集 N_0 上生成初始模型 β_0
 For 轮次 $i=0, \dots, m$
 在 N_i 上发掘困难负样本集 $H(\beta_i, N_i)$
 $N_{i+1} = N_i + H(\beta_i, N_i)$
 在新样本集 N_{i+1} 上更新模型为 β_{i+1}
 End For

5 实验过程及结果分析

样本图片来自 INRIA^[13]库,大小都为 64×128 ,共 2 416 张正样本,负样本为在不包含人体

的图片中随机裁剪出的窗口,共 12 180 张。HOG 特征的计算如前所述,计算 LBP 特征时,取 $p=8, r=1$,将图片划分成 16×16 的不重叠块,以块为单位计算直方图,每张图片上的 LBP 特征向量长度为 $8 \times 4 \times 59 = 1\ 888$ 维,联合 3 780 维的 HOG 特征,构成总长度为 5 668 维的特征向量。

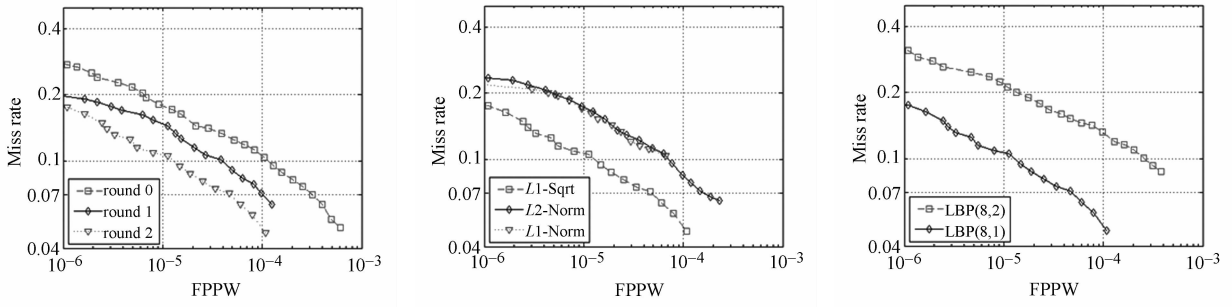
正测试库包含 1 126 张图片,负测试库包含 453 张图片,在负测试库上的检测窗口数为 362 400 个。采用对数坐标下的 DET (Detection Error Tradeoff) 曲线来评价检测器的性能,即在正测试样本上计算漏检率 (Miss Rate),用纵坐标表示;在负测试样本上计算误检率 (False Positive Per Window, FPPW),用横坐标表示。

图 8(a)为 Bootstrapping 次数 m 取值对检测结果的影响,线性核 SVM 的支持向量数目较少, $m > 2$ 后支持向量几乎无变化,性能趋于稳定。较不经过该过程的检测器在 FPPW 为 10^{-4} 时提高了 5%,这种重新抽样的方法有效地缓解了由于样本数目限制对分类器性能的影响。

HOG 算子采用和文献[4]一致的 L_2 范数截断归一化,而 LBP 算子的直方图不包含幅度信息,应与 HOG 分别进行归一化处理。如图 8(b),采用如式(6)的 L_1 -Sqrt 归一化比 L_2 、 L_1 范数归一化在 FPPW 为 10^{-4} 时提高了约 4%, 10^{-5} 时提高了约 7%。

$$L_1 - Sqrt: v = \sqrt{v / (\|v\|_1 + \epsilon)}. \quad (6)$$

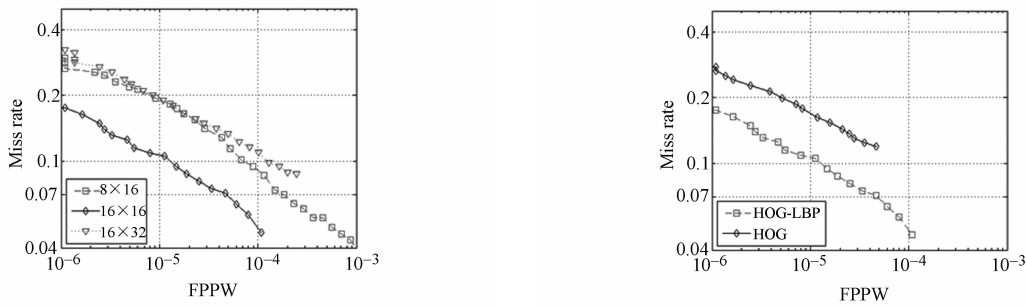
LBP 算子参数 p 过大会引起特征向量长度指数级增长,过小则采样点数不足,通常取为 8。另一个参数 r 的选取如图 8(c), $r=2$ 比 $r=1$ 时,检测效果显著下降,这是由于 $r=2$ 时,较多的点被归类为非一致模式,降低了特征的分辨力。如图 8(d)所示本文比较了 3 种尺寸相近的子块划分法, 8×16 的块划分使得特征向量长度为 $3\ 780 + 3\ 776$,高维数的特征向量使得检出率在 FPPW 为 10^{-3} 时检出率达 96%,但过分强调纹理特征会造成对枯树叶、花瓣等纹理丰富物体的误检;而 16×32 的块划分使得特征向量长度为 $3\ 780 + 944$,纹理特征向量长度显著小于梯度特征,且在大范围内统计直方图造成了纹理特征模糊,检测率与单一 HOG 算子几乎无区别; 16×16 的子块划分较为均衡。HOG-LBP 算子与 HOG 算子的对比如图 8(e)所示,在 FPPW 为 10^{-4} 时检出率达 95%,高于文献[4]中提到的 89%,可见具有互



(a) Bootstrapping 次数的影响
(a) Effect of Bootstrapping times

(b) 归一化方式的影响
(b) Effect of normalization

(c) LBP 参数 r 的选择
(c) Effect of r with LBP



(d) LBP 块尺寸的影响
(d) Effect of block size with LBP

(e) HOG 与 HOG-LBP 的比较
(e) Comparison between HOG and HOG-LBP

图 8 实验结果(DET 曲线)

Fig. 8 Experiment results(DET curve)



(a) 误检
(a) False positive

(b) 漏检
(b) False negative



(c) 对原始检测结果的纠正
(c) Correction of HOG

图 9 部分检测结果

Fig. 9 Some detection results

补特性的联合特征较显著地提升了单一特征检测器的性能。本文所采用分类器造成的部分误检如图 9(a)所示,漏检则是由于人体的形变过大或遮挡,如图 9(b)所示,图 9(c)展示了 HOG-LBP 算子检出而 HOG 未检出的图片。此外,LBP 算子的计算都是简单的算术运算,算法基于 OpenCV 2.3 修改,运行在 VS2008 环境下,P4 3.0 G,1 G 内存的电脑上。对所有窗口检测时间的统计平均表明,优化后的 HOG 联合 LBP 算子单窗口检测时间为 0.533 ms,小于原始 HOG 算子的 0.625 ms。

6 结 论

本文首先分析了 HOG 特征的梯度敏感性,

参考文献:

- [1] 李欣,赵亦工,陈冰,等. 基于模糊分类的弱小目标检测方法[J]. 光学 精密工程, 2009, 17(9): 2311-2320.
LI X, ZHAO Y G, CHENG B, *et al.*. Approach to dim and small target detection based on fuzzy classification [J]. *Opt. Precision Eng.*, 2009, 17(9): 2311-2320. (in Chinese)
- [2] 刘亚洲. 基于时空分析和多粒度特征表示的人体检测方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2009.
LIU Y ZH. *Spatial-temporal analysis and multi-granularity representation based human detection* [D]. Harbin: Harbin Institute of Technology, 2009. (in Chinese)
- [3] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [4] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*, 2005: 886-893.
- [5] WALK S, MAJER N, SCHINDLER K, *et al.*. New features and insights for pedestrian detection [C]. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*, 2010: 1030-1037.
- [6] WATANABE T, ITO S, YOKOI K. Co-occurrence histograms of oriented gradients for human detection [J]. *IPSN Transactions on Computer Vision and Applications*, 2010, 2: 39-47.
- [7] MAJI S, BERG A C, MALIK J. Classification using intersection kernel support vector machines is efficient [C]. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*, 2008: 1-8.
- [8] FELZENSZWALB P F, GIRSHICK R B, MCALLESTER D. Cascade object detection with deformable part models [C]. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*, 2010: 2241-2248.
- [9] WANG X Y, HAN T X, YAN SH CH. An HOG-LBP human detector with partial occlusion handling [C]. *Computer Vision, IEEE International Conference*, 2009: 32-39.
- [10] ZHU Q, YEH M C, CHENG K T, *et al.*. Fast human detection using a cascade of histograms of oriented gradients [C]. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*, 2006: 1491-1498.
- [11] PORIKLI F. Integral histogram: a fast way to extract histograms in Cartesian spaces [C]. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*, 2005: 829-836.
- [12] BILGIC B. *Fast human detection with cascaded ensembles* [D]. Massachusetts: Massachusetts Institute of Technology, 2010.
- [13] DALAL N. *Finding people in images and videos* [D]. France: the French National Institute for Research in Computer Science and Control, 2006.
- [14] PANG Y W, YUAN Y, LI X L, *et al.*. Efficient HOG human detection [J]. *Signal Processing*,

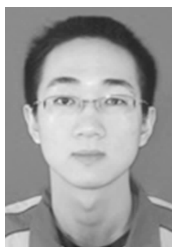
即易因与人体局部形状相似而造成误检,易因轻微形变产生漏检。然后将 LBP 特征加入到 HOG 特征中,从纹理结构的角度加强了特征的区分度。在 INRIA 库上的测试表明,在 FFPW 为 10^{-4} 时检出率达到了 95%。为了缓解新特征增加的运算量,还简化了 HOG 特征的插值方式,使得联合特征的计算时间短于原始特征。虽然在更短的计算时间里取得了更好的效果,但本文的特征维数较高,存在冗余,采用多核筛选机制 (multi-kernel learning) 的方法能更有效地联合多种特征。部件模型^[8]的方法表现出对人体形变更大的宽容度,将本文的特征加入到这一模型中也是作者以后研究的方向。

2011, 91: 773-781.

- [15] GUO ZH H, ZHANG L, ZHANG D. A Completed modeling of local binary pattern operator for texture classification [J]. *IEEE Transactions on Image Processing*, 2010, 19(6): 1657-1663.

- [16] FAN R E, CHANG K W, HSIEH, *et al.*. LIB-LINEAR: a library for large linear classification [J]. *Journal of Machine Learning Research*, 2008, 9: 1871-1874.

作者简介:



黄炎(1988-),男,湖北荆州人,博士研究生,2008年于武汉大学获得学士学位,主要从事机器视觉、模式识别等方面的研究。E-mail: up2mail@163.com



范赐恩(1975-),女,浙江慈溪人,博士,讲师,2012年于武汉大学获得博士学位,主要从事图像处理、机器视觉等方面的研究。E-mail: fce@whu.edu.cn



朱秋平(1986-),男,江苏东台人,博士研究生,2009年于武汉大学获得学士学位,主从事目标跟踪、模式识别等方面的研究。E-mail: 307469139@qq.com



张虎(1986-),男,天津人,博士研究生,2009年于武汉大学获得学士学位,主要从事图像处理、机器视觉等方面的研究。E-mail: zhanghu0722@126.com

导师简介:



邓德祥(1961-),男,湖北荆州人,教授,博士生导师,1982年于武汉测绘科技大学获得学士学位,1985年于武汉测绘科技大学获得硕士学位,主要从事空间图像处理、片上系统的研究。ddx@eis.whu.edu.cn

(版权所有 未经许可 不得转载)