

基于人体检测理论的室内智能监控系统

陈良仁,潘泉松,侯北平

(浙江科技学院 自动化与电气工程学院,杭州 310023)

摘要: 智能监控已随着社会发展成为主流,但当前智能监控系统的人体检测模块尚不能满足实时检测要求。为解决这一问题,结合梯度方向直方图特征和 SVM(支持向量机)级联分类器,提出室内监控系统的人体检测策略,并基于以太网平台,实现远程实时室内监控。实验结果表明,该监控系统精度较高、速度较快,能在实时人体检测的基础上进行预警,达到了智能监控的目的。

关键词: 人体检测;梯度方向直方图;支持向量机;级联分类器

中图分类号: TP391.41; TP277.2

文献标志码: A

文章编号: 1671-8798(2012)04-0297-06

Indoor smart surveillance system based on human detection theory

CHEN Liang-ren, PAN Quan-song, HOU Bei-ping

(School of Automation and Electrical Engineering, Zhejiang University of
Science and Technology, Hangzhou 310023, China)

Abstract: In order to solve the problem that existing intelligent surveillance system cannot detect humans well in real time, we incorporate histograms of oriented gradients (HOG) features with SVM (support vector machine) cascade classifiers to speed up detection process. Besides, based on Ethernet platform, this system has the ability of remote indoor monitoring. The experiments show that this surveillance system can process frames at relatively high detection speed and detection precision, basically realizing real-time human detection function.

Key words: human detection; histograms of oriented gradients; SVM; cascade classifier

人体检测是继人脸检测后机器视觉的又一个重要研究领域,它正逐渐应用于各个行业,如人机接口、行车辅助、视频监控和机器人等。基于人体检测的监控系统有报警、跟踪功能,优于常规的监控系统,但由于人的姿态及环境的多样性,导致人体检测的实现存在一定难度。

收稿日期: 2012-04-16

基金项目: 浙江省公益性技术应用研究计划项目(2011C31014);浙江科技学院学科交叉预研项目(2010JC02Y)

作者简介: 陈良仁(1990—),男,浙江省瑞安人,2009级自动化专业本科生。

通讯作者: 侯北平,副教授,博士,主要从事图像处理、机器视觉方面研究。

近年来,出现了大量的人体检测算法。Mu 和 Wang 等人^[1-2]采用 LBP(local binary patterns)作为人体检测特征。近期人们提出的算法^[3-4]是合并多种人体特征,如色彩、位置、纹理、边缘等,这种算法在提高检测精度的同时,也增加了检测时间。Dalal 等人^[5]提出梯度方向直方图(histograms of oriented gradients, HOG)结合支持向量机(support vector machine, SVM)的方法,但该方法每秒钟只能处理 1 帧 320×240 像素的图像,无法满足实时检测要求。李同治等人^[6]在 Dalal 算法的基础上,通过一个由简到繁的级联线性 SVM 分类器,将级联拒绝机制与梯度方向直方图特征相结合,实现了快速、准确的人体检测算法。曾春等人^[7]提出基于感兴趣区梯度方向直方图的行人检测方法,有效地减少了特征向量维数,缩短了检测时间。笔者提出的方法对文献[5]进行了加速,利用 Adaboost 挑选出少数具有鉴别性的 block 特征,并结合积分梯度方向直方图和级联 SVM 分类器,开发了室内人体检测系统,经实际运行,取得了较好的效果。

1 梯度方向直方图

梯度方向直方图类似于 SIFT (scale invariant feature transform)特征^[8],具有旋转不变性、尺度不变性和光照强度不变性,其核心思想是物体的轮廓和纹理可以被梯度幅值和梯度方向较好地描述。在人体检测中,轮廓特征的提取尤其重要,笔者选择梯度方向直方图来描述人体的轮廓特征,算法实现步骤如图 1 所示。

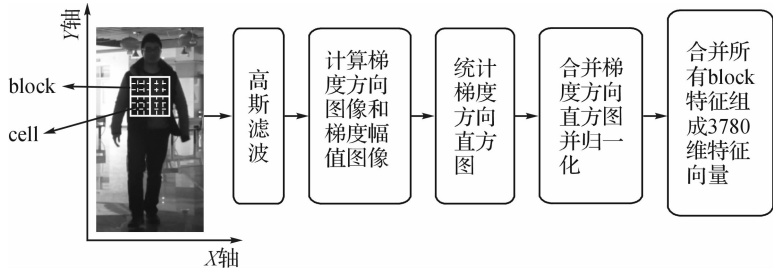


图 1 HOG 特征提取流程图

Fig. 1 Flow chart of HOG features extraction

第一步 为消除光照强度对人体检测的影响,使算法具有更好的鲁棒性,通过式(1)和式(2)对图像进行高斯滤波。

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2} \tag{1}$$

$$S(x,y) = G(x,y,\sigma) \cdot I(x,y) \tag{2}$$

其中, $G(x,y,\sigma)$ 为高斯函数, $I(x,y)$ 为输入图像, $S(x,y)$ 为平滑后的图像。

第二步 采用 $[-1,0,1]$ 梯度模板计算梯度幅值图像 $M(x,y)$ 和梯度方向图像 $\theta(x,y)$,分别用式(3)和式(4)表示:

$$M(x,y) = \sqrt{(S(x+1,y) - S(x-1,y))^2 + (S(x,y+1) - S(x,y-1))^2} \tag{3}$$

$$\theta(x,y) = \arctan\{[S(x,y+1) - S(x,y-1)]/[S(x+1,y) - S(x-1,y)]\} \tag{4}$$

第三步 将 64×128 像素的检测窗口图像划分为 8×8 像素的细胞单元(cell),在每块 cell 上构建九通道梯度方向直方图,梯度方向取值为 $[0^\circ, 180^\circ]$,通过权值投票,对梯度方向直方图进行统计,每一票的权值为像素的幅值。

第四步 将 4 个 cell 组合成空间上连通的大小为 16×16 像素的区间块(block),在 block 上对 4 个梯度方向直方图进行合并,得到一组 36 维的 block 特征,采用二范数对 block 特征归一化,其算法表示形式:

$$L_2 - \text{normal}, V \leftarrow v / \sqrt{\|v\|_2^2 + \epsilon^2}.$$

第五步 沿图像的 X 轴和 Y 轴方向,以 8 像素为间隔获得 105 个 block,合并检测窗口内所有 block 特征,最后得到 3 780 维特征向量。

2 快速人体检测研究

实验采用 INRIA(French national institute for research in computer science and control)提供的图片作为训练样本,包括 2 416 张正样本及 1 218 张负样本,其中正样本为 64×128 像素的人体图片,负样本为不含人体的背景图片,样本图片见图 2。Dalal^[5]提出的算法平均检测速度为 1 帧/s,而 Zhu^[9]提出的算法平均检测速度为 12 帧/s,使用多尺度 block,并引入积分梯度方向直方图,在统计 cell 上的梯度方向直方图时,对内存的操作只需 4×9 次,相比原来遍历读取梯度信息,该方法使特征提取速度有了较大提升。



图 2 INRIA 中的样本图片

Fig. 2 Some sample images from INRIA

2.1 多尺度 block

Dalal^[5]算法中 block 尺度固定不变 (16×16 像素),包含的人体特征少。Zhu^[9]根据人脸检测中 Haar 特征的多样性,提出多尺度 block 特征,包含的人体轮廓信息更丰富。因此在实验中构造了 5 031 个 block,尺度从 12×12 像素到 64×128 像素,宽与高之比为 $1:1,1:22,2:1$,提取间隔分别为 4、6、8 像素。

2.2 级联分类器

当人体尺寸逐渐增大以至超过检测窗口大小时,常常会导致人体检测失败,因此需按一定比例对图像进行缩放,然后在缩放后的每张图像上通过平移检测窗口来判断是否含有人体。以 320×240 像素的图像为例,图像的缩放比例取 1.2,并用 64×128 像素的检测窗口按照 8 像素步长来搜索人体,共需判断

1 900 多个窗口,而且大部分窗口都是背景图像。为了达到实时检测要求,笔者提出了如图 3 所示的 SVM 级联分类器。它由多个 SVM 分类器串联而成:前面几级分类器结构比较简单,使用的特征数较少,能够快速滤除那些与人体特征差异比较大的负样本;后面几级的分类器使用更多的特征,具有更复杂的结构,可以分类与人体相似的负样本。当检测窗口被前面一级的分类器判断为人体时才被送入后面的分类器继续处理,反之则被认为是背景图片,能通过所有分类器的检测窗口为含有人体的图像。

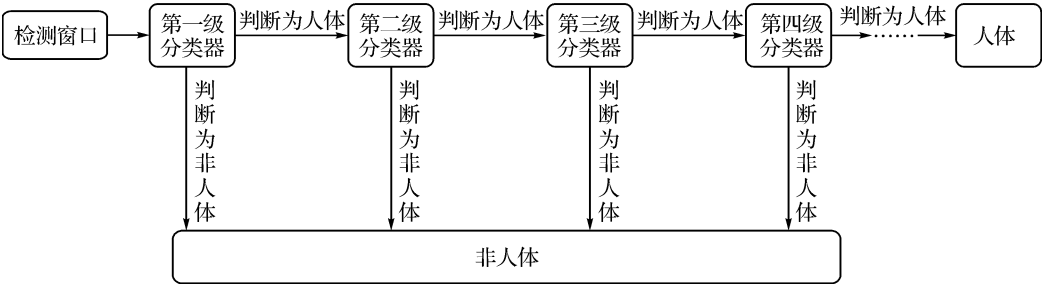


图 3 级联分类器人体检测的机制

Fig. 3 Mechanism of cascade classifier on detecting humans

计算所有线性 SVM 分类器要花费大量时间,因此在每层级联分类器训练过程中从 5 031 个 block 随机选取 250 个 block,计算其对应的线性 SVM 分类器,并用 Adaboost 筛选出少数具有鉴别性的 block 特征,算法实现如图 4 所示。

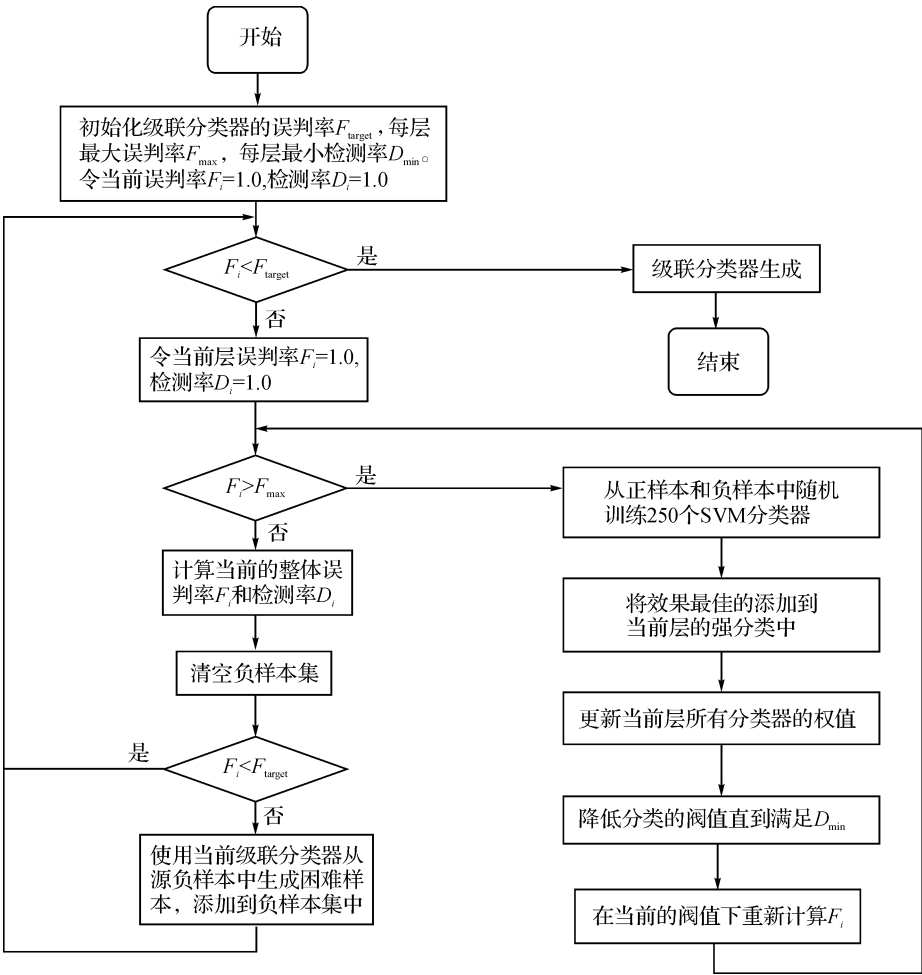
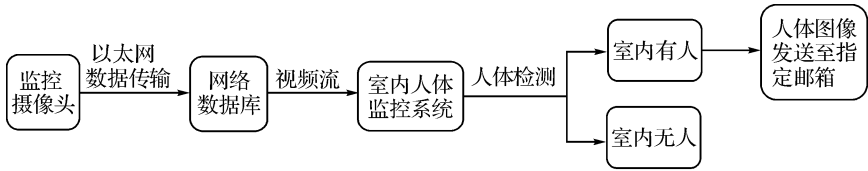


图 4 级联分类器训练流程图

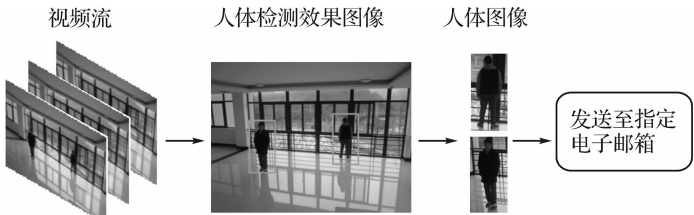
Fig. 4 Flow chart of cascade classifier training

3 基于人体检测室内监控系统的网络平台实现

为解决远程智能监控问题,如图 5(a)所示,实验开发了实时网络传输平台。由于网络数据库能够存储庞大的图像数量,而且可以查看历史记录,因此先通过以太网把监控摄像头获取的图像传输到网络数据库,由智能监控软件获取视频图像,并经人体检测软件模块进行人体检测,判断室内有人则将视频图像中的人体区域发送到指定邮箱,通过移动设备和电脑查看邮箱即可获得室内信息,从而起到远程预警功能。



(a) 远程控制系统网络平台框架图



(b) 人体图像获取

图 5 远程监控系统网络平台

Fig. 5 Network platform of remote surveillance system

如图 5(b)所示,在测试的实时视频中截取检测到的人体图像,并实时进行发送。以本次实验为例,视频图像分辨率为 320×240 像素,截取到的两幅人体图像大小分别为 43×81 像素和 42×80 像素,发送时将其压缩为 JPG 格式图像,实际大小分别为 2.51 kB 和 2.17 kB,测试使用的电子邮箱为 126 邮箱,其容量为 3 GB,可容纳多幅人体图像的连续发送,用户可以根据监控实际情况定期清理邮箱数据;经实验测试发送一幅图像平均时间小于 1 s,能满足异常报警与取证的基本要求。

基于人体检测的室内监控系统与传统的监控系统相比,能对视频流进行智能分析,只有判断出有人时才进行报警,从而降低了工作人员的劳动强度,提高监控智能化。本系统在实验室等实际场景进行了多次测试,取得了良好的人体检测效果,为后继的智能监控研究奠定了基础。

4 实验结果分析

实验平台为研华工业控制计算机,CPU 为酷睿双核 2.80 GHz,内存为 2 G,硬盘为 320 G,开发环境为 Microsoft Visual Studio 2008,使用的数据库为 SQL Server。在多个场合进行了测试,检测效果如图 6 所示。图 6 中,(a)和(b)为实验楼人体检测结果,(c)为图书馆人体检测结果,(d)为学生宿舍人体检测结果,(e)和(f)为学校餐厅人体检测结果。

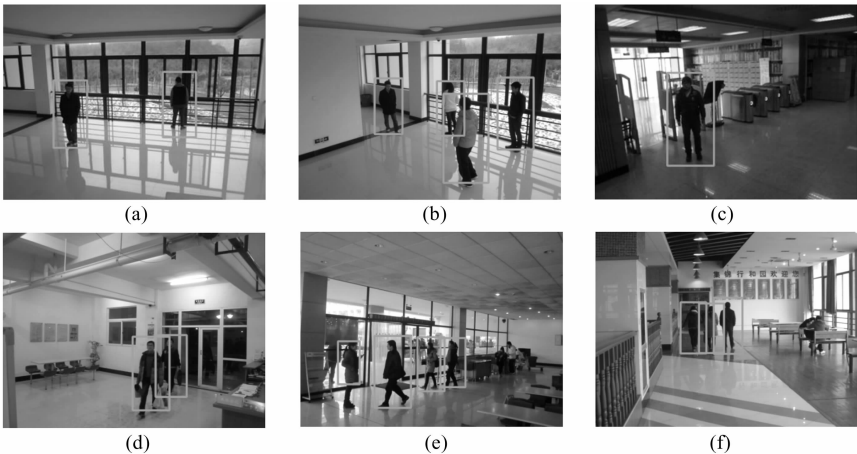


图 6 人体检测结果

Fig. 6 Human detection results

由表 1 可见,增大图像缩放比例和检测窗口步长可以提高检测速度。在多种环境进行人体检测,发现图像缩放比例为 1.5,检测窗口步长为 8 像素时效果最佳,在保证检测率较高的同时,检测速度平均为 90 ms/张;在系统整体测试中,每张人体图像发送邮箱时间小于 1 s,基本达到实时监控要求,为后继的智能监控系统设计、产品开发奠定了良好的基础。

表 1 检测窗口的步长和图像的缩放比例对人体检测速度的影响

图像缩 放比例	检测速度/(ms·张 ⁻¹)		
	A	B	C
1.1	700	260	200
1.3	460	130	90
1.5	360	90	70
1.7	240	80	60
1.9	230	70	50

注:A、B、C 分别代表检测窗口步长为 4、8、12 像素。

5 结 语

通过将梯度方向直方图特征和 SVM 级联分类器相结合,使基于人体检测的室内监控系统在保证较高检测率的同时检测速度提高到 12 帧/s,满足实时检测的基本要求,在此网络远程监控系统中取得良好的预警效果,具有较好的应用前景。

参考文献:

[1] Mu Y, Yan S, Liu Y, et al. Discriminative local binary patterns for human detection in personal album[C]// Conference on Computer Vision and Pattern Recognition, Alaska: IEEE,2008.

[2] Wang X, Han T X, Yan S. An HOG-LBP human detector with partial occlusion handling[C]// International Conference on Computer Vision (ICCV), kyoto: IEEE,2009:32-39.

[3] Dollar P, Tu Z, Perona P, et al. Integral channel features[C]// British Machine Vision Conference, London: BMVC,2009.

[4] Schwartz W R, Kembhavi A, Harwood D, et al. Davis. Human detection using partial least squares analysis[C]// International Conference on Computer Vision (ICCV), kyoto: IEEE,2009:24-31.

[5] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]// Conference on Computer Vision and Pattern Recognition (CVPR), San Diego: IEEE,2005:886-893.

[6] 李同治,丁晓青,王生进. 利用级联 SVM 的人体检测方法[J]. 中国图像图形学报,2008,13(3):566-570.

[7] 曾春,李晓华,周激流. 基于感兴趣区梯度方向直方图的行人检测[J]. 计算机工程,2009,35(24):182-184.

[8] Lowe D. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004,60(2):91-110.

[9] Zhu Q. Fast human detection using a cascade of histograms of oriented gradients[C]// Computer Society Conference on Computer Vision and Pattern Recognition, New York: CVPR,2006:1491-1498.