

智能视频监控中基于机器学习的自动人数统计*·论文·

贾慧星, 章毓晋

(清华大学 信息科学与技术国家实验室; 电子工程系, 北京 100084)

【摘要】 提出了智能视频监控中基于机器学习的自动人数统计系统。该系统通过机器学习的方法对视频序列中人的头肩部位进行准确地检测, 克服了传统检测方法如连通域分析和简单模板匹配的不足, 对光线变化和人群拥挤等问题具有较好的稳健性, 在具体场景的初步测试中取得了较满意的效果。

【关键词】 智能视频监控; 人数统计; 目标检测与跟踪; 机器学习

【中图分类号】 TP277

【文献标识码】 A

Automatic People Counting Based on Machine Learning in Intelligent Video Surveillance

JIA Hui-xing ZHANG Yu-jin

(Tsinghua National Laboratory for Information Science and Technology; Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

【Abstract】 An automatic people counting system based on machine learning in intelligent video surveillance is proposed in this paper. The system detects head-shoulder component based on machine learning, which overcomes the disadvantages of traditional detection methods such as connected component analysis or simple template matching. The system demonstrates its robustness to illumination change and crowd in practical applications.

【Key words】 intelligent video surveillance; people counting; object detection and tracking; machine learning

1 引言

现有的基于视频的人数统计系统所采用的算法可以分为3类: 基于底层特征的拟合方法, 基于特征点跟踪的方法和基于目标检测跟踪的方法。基于底层特征的方法^[1-3]忽略了个体的概念, 跳过了对单个目标的检测跟踪过程, 虽然简单易用, 但难以得到精确的人数信息, 一般用于粗略的人群密度估计。基于特征点跟踪的方法^[4-5]首先跟踪一些运动的特征点, 然后对特征点的轨迹进行聚类分析, 从而得到人数信息。该方法的优点是受摄像机视角影响较小, 缺点是特征点本身难以稳定地跟踪, 计数精度较差。基于目标检测跟踪的方法直接显式地对目标进行检测跟踪, 精度较高, 所以本文采用了基于目标检测跟踪的方法。

现有的基于目标检测跟踪的方法在通过背景建模得到前景区域之后, 一般采用连通域分析或者模板匹配的检测方法^[6-7]。近几年计算机视觉领域中目标检测的研究表明基于机器学习的目标检测方法^[8-9]的精度远强于连通域分析或者模板匹配的方法, 所以本文采用基于机器学习的目标检测方法, 并在此基础上设计了自动的人数统计系统, 具有较高的精度和较好的鲁棒性, 在商场客流统计等领域具有较大的应用价值。

2 系统概述

如图1所示, 本文的自动人数统计系统包括视频获取和视频分析两个主要的组成部分。由于监控系统中摄像机视角的多样性和严重的遮挡行为, 人全身难以稳定地检测跟踪, 而头肩部位的形状和姿态较为稳定并且遮挡较小, 较易检测跟踪, 所以笔者通过对头肩部进行检测跟踪, 从而达到计数的目的。

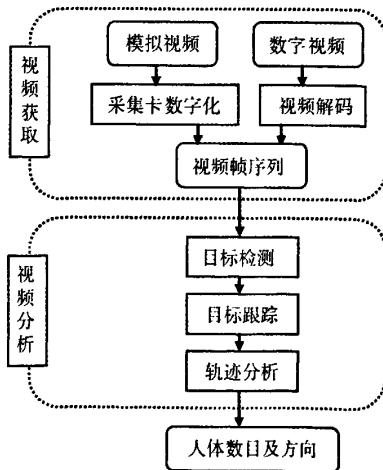


图1 系统结构

* 国家自然科学基金(60872084); 教育部高等学校博士学科点专项科研基金(20060003102)

3 目标检测

本文采用 Dalal^[8] 提出的基于梯度方向直方图特征 (Histograms of Oriented Gradients, HOG) 和支持向量机分类器 (Support Vector Machine, SVM) 的目标检测方法。HOG 特征将目标分块, 分别提取每块内梯度在各个方向上的分布情况, 得到直方图特征。通过合理的分块策略, HOG 特征可以有效地编码目标的局部形状信息, 同时保持对目标局部细微的平移和旋转运动的不变性。SVM 是 Vapnik 提出的基于结构风险最小化原理 (Structural Risk Minimization, SRM) 的机器学习理论, 这使得它比基于经验风险最小化 (Empirical Risk Minimization, ERM) 的人工神经网络方法具有更好的泛化能力。SVM 有线性和非线性两种形式, 由于线性 SVM 的运算速度较快, 所以笔者采用了线性 SVM。HOG 特征结合线性 SVM 在人体检测上达到了极好的效果^[9], 笔者将其直接应用到头肩部位的检测上。

检测器的工作流程如图 2 所示, 分为离线训练和在线检测两个部分。

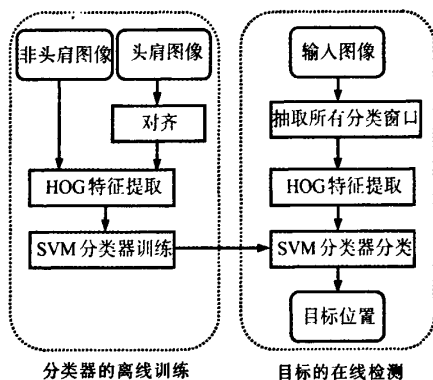


图2 基于 HOG 特征和 SVM 分类器的目标检测方法

训练集合的一些典型图像如图 3 所示, 正面样本共 2 954 个, 包括一系列摄像机俯视的头肩部。这些样本光线条件多变, 所处的背景也较为复杂, 同时为了突出目标的轮廓信息, 特意保留了头部周围的背景区域。反面样本从 3 012 个背景图片中获取。背景图片的选择对性能的影响比较大。该背景图片集合包括室外和室内场景, 同时为了防止检测器检测人体的其余部位, 还加入了一部分不包含头肩部的肢体图片作为背景。在训练过程中采用了“自举”方法以增强分类器的分类性能, 即首先训练一个分类器, 然后从背景图片中检测出虚警加入负面样本中, 重新训练得到最终分类器。

针对计数这种应用, 只需要检测运动的人体, 所以可以利用背景建模的方法加快检测器的搜索速度。利用背景建模得到前景图像之后, 首先计算每个待分类窗口



(a) 正面头肩图像



(b) 负面背景图像

图3 训练数据集合

内前景像素的比例, 如果该比值小于一定阈值, 则直接作为背景舍去, 否则再运行分类器看是否包含目标。任意矩形窗口内前景像素的比例利用前景图像的“积分图像”^[10]可以由 4 次查表运算得到, 大大加快了搜索速度。由于该方法主要依靠分类器进行判断, 所以对背景建模模块的要求较低, 允许前景中有较多的虚警, 然后依靠后续分类器去除这些虚警。该方法并不需要将前景像素分割成单独的运动块, 所以当人聚集在一起时依然可以正确地分出每个行人。检测器的运行示例如图 4 所示。

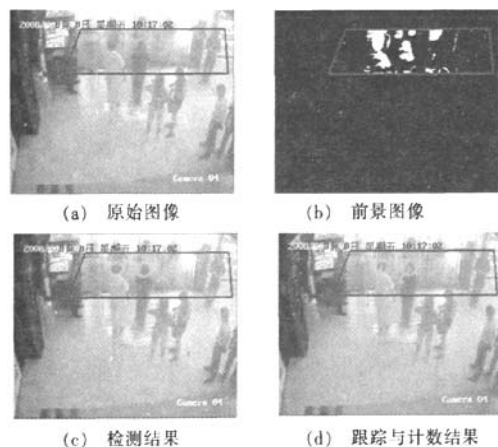


图4 计数过程中目标的检测跟踪

4 目标跟踪

如图 4c 所示, 目标检测得到的是目标在每一帧的空间位置, 要想达到计数的目的必须对目标进行跟踪得到目标的时空轨迹。如果检测器是完美的, 那么跟踪只需要对前后帧的检测结果进行简单的数据关联即可。但实际情况是检测器并不完美, 偶尔会出现虚警或者漏检, 此时必须利用跟踪器去除这些虚警, 预测漏检目标的位置。

如图 5 所示, 本文采用具有较好时间效率的最近邻匹配数据关联方法对目标进行跟踪。利用 Kalman 滤波器对目标轨迹进行建模, 其状态空间为 $s=(x_c, y_c, u, v)$, 其中 (x_c, y_c) 表示目标的中心, (u, v) 表示目标在两个方向上平移的速度。对于已经存在的轨迹, 首先通过最近邻匹配将该轨迹基于 Kalman 滤波器的预测位置与当前帧的

检测结果进行数据关联,如果有匹配的检测结果,则用匹配的检测结果作为 Kalman 滤波器的观测量更新该轨迹;否则直接用 Kalman 滤波器预测的位置作为目标在当前帧的位置。没有匹配轨迹的检测结果用来初始化新的目标轨迹,参与下一帧的最近邻匹配过程。每帧处理完毕后,对轨迹是否终止进行判断,如果目标运动出计数区域或者连续若干帧没有匹配的检测结果,则认为该轨迹终止,输出该轨迹,经简单的轨迹分析后用来更新计数器。

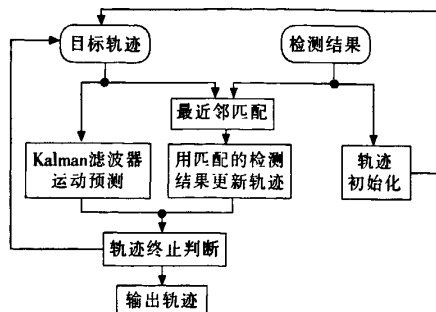


图5 目标跟踪

5 轨迹分析及目标计数

为了判断目标运动的方向、更新计数器并进一步降低虚警,需要对跟踪得到的轨迹进行分析。假设跟踪得到的目标轨迹用 $T=\{(x_i, y_i, w_i, h_i), (x_{i+1}, y_{i+1}, w_{i+1}, h_{i+1}), \dots, (x_{i_m}, y_{i_m}, w_{i_m}, h_{i_m})\}$ 表示,其中 (x, y) 表示目标在每帧的中心位置, (w, h) 表示目标在每帧的大小, i 表示目标出现时的帧编号, $i+n$ 表示目标消失时的帧编号。首先计算由点 (x_i, y_i) 到点 (x_{i_m}, y_{i_m}) 的矢量与用户定义的进出方向的夹角,从而判断出该目标是进还是出。其次计算点 (x_i, y_i) 到点 (x_{i_m}, y_{i_m}) 的距离 d , 如果该距离大于预先定义的阈值 θ , 则将对对应方向上的计数器加一, 否则作为虚警舍去。阈值 θ 可以作为一个用户定义的参数在实际应用时进行调节, 以在检测率和虚警率之间寻求较好的折中。

6 实验结果

笔者基于本文的算法开发了如图6所示的监控系统中的自动人数统计系统,在 Intel Core2 2.4 GHz CPU, 1 Gbyte RAM 的 PC 上可以实时完成 4 路视频的计数功能。图6显示的是北京某商场出入口的人流情况,4 段视频的测试结果如表1所示。在第3路和第4路视频中行人较为稀少的情况下,该系统的计数精度可以达到95%以上,在第1路和第2路行人较稠密的情况下,可以达到85%的计数精度,基本达到了实际应用的要求。

7 小结

采用了基于机器学习的目标检测方法对视频序列



图6 人数统计系统程序界面

表1 测试结果

	真实值	统计值	误差	错误率/%
序列1(进)	110	96	-14	-12.7
序列1(出)	70	80	+10	+14.3
序列2(进)	57	62	+5	+8.8
序列2(出)	55	61	+6	+10.9
序列3(进)	29	30	+1	+3.4
序列3(出)	39	40	+1	+2.6
序列4(进)	36	35	-1	-2.8
序列4(出)	16	16	0	0

中人的肩部部位进行准确的检测,结合目标跟踪和轨迹分析等模块,实现了视频监控系统中的自动人数统计。本文的基于机器学习的目标检测方法克服了传统的检测方法如连通域分析和简单模板匹配的不足,对背景中的光线变化和人群拥挤等问题具有较好的稳健性。本文的算法并不局限于人数统计,同样可以直接应用到车流统计上,在智能交通领域也有较大的应用价值。

参考文献:

- [1] PARAGIOS N, RAMESH V. A mrf-based approach for real-time subway monitoring [EB/OL]. [2008-12-20]. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?isNumber=21353&prod=CNF&arnumber=990644&arSt=+1034&ared=+1040&arNumber=990644.
- [2] KONG D, GRAY D, TAO H. A viewpoint invariant approach for crowd counting [EB/OL]. [2008-12-20]. <http://www.soe.ucsc.edu/~tao/papers/ICPR06Kong2.pdf>.
- [3] RABAUD V, BELONGIE S J. Counting crowded moving objects [EB/OL]. [2008-12-20]. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1640823.
- [4] ANTONINI G, THIRAN J P. Counting pedestrians in video sequences using trajectory clustering[J]. IEEE Trans. Circuits and Systems for Video Technology, 2006, 16(8):1008-1020.
- [5] KILAMBI P, RIBNICK E, JOSHI A J, et al. Estimating pedestrian count in groups [J]. Computer Vision and Image Understanding, 2008, 110(1):43-59.
- [6] DONG L, PARAMESWARAN V, RAMESH V, et al. Fast crowd segmentation using shape indexing[EB/OL]. [2008-12-20]. <http://ieeexplore.ieee.org/iel5/4408818/4408819/04409075.pdf?arnumber=>

4409075.

- [7] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features[EB/OL].[2008-12-20].http://www.cs.utexas.edu/~grauman/courses/spring2007/395T/papers/viola_cvpr2001.pdf.
- [8] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[EB/OL]. [2008-12-20]. <http://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf>.
- [9] VAPNIK V. The nature of statistical learning theory[M]. Berlin:

Springer-Verlag, 2000.

作者简介:

贾慧星(1981-), 博士生, 主要研究方向是智能视频监控、计算机视觉中目标的检测与跟踪;

章毓晋(1954-), 教授, 博士生导师, 主要科学研究领域为图像工程(图像处理、图像分析、图像理解及其技术应用)及相关学科。

责任编辑:任健男

收稿日期:2009-01-20

(上接第 77 页)

常有利。隔离变压器是一个非常成熟的产品, 结构简单, 价格低廉, 规格齐全, 可以满足各种功率等级的要求, 而且安全可靠, 使用上无风险。所以, 经常被应用在现代精密配电中心(PPC)上。实践证明, 隔离变压器的引入, 不仅能很好地解决零线压降随电流增大而提高, 从而导致负载端零地电位差升高的问题, 有效杜绝了由谐波带来的干扰, 而且还有效减弱了强雷电给工艺系统带来的电磁干扰。在实际应用时, 隔离变压器应尽量接在靠近负载端, 因为隔离后的电源零线要就近接入工艺地线, 这样做是为了确保负载零线与工艺地线电位的一致, 使零地并接后其电位差趋近于零, 上述干扰也就不复存在。

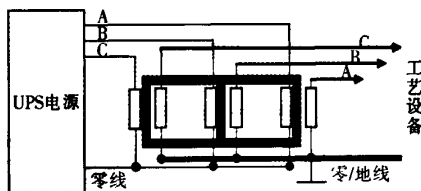


图2 从物理连接上将UPS与负载的零线分开

在日常的技术维护中, 因零地电位差过大而引起对广播电视工艺系统设备干扰的问题, 却普遍得不到人们的重视和研究, 把系统在运行时出现的一些古怪现象和故障表现, 简单归咎于设备本身, 使系统常常处在恶劣的电源环境和干扰之中运行, 而长期得不到改善。因此, 机房UPS三次谐波干扰与零电位扰动问题应引起充分重视。

3.4 措施四

合理的施工工艺是确保工程取得事半功倍成效之关键。设备的品质固然重要, 而其合理的布局和讲究的布线工艺对系统的质量影响尤其突出。从前面分析的情况看, 缩短零线长度, 增大零线截面积是一种减小零线电抗, 降低压降, 缩小零地电位差, 减低干扰的有效途径。从零线电抗计算公式 $Z_n = \rho \times L \div S$ 可以看出, 当导线长度 L 减小, 导线截面积 S 增大时, Z_n 随之减小, 零地电位差也随之降低。

当然, 合理地确定布线空间同样可以取得好的成效。根据有关技术资料说明, 对于双变换纯在线式UPS, 当逆变器工作时, UPS输入端零线电流理论上应该为零。

但由于配电机房电缆多数走的是大电流, 这些大电流设备包括大功率UPS、电梯变频器、空调压缩机等, 它们工作时多少会产生各种谐波, 若把这些电缆紧挨在一起走长线, 难免会产生各种谐波的耦合和调制, 并互为串扰, 产生新的谐波分量, 最终导致干扰的产生。所以, 有条件时应将UPS的火线和零线、地线分开走线, 两者的距离应该保证在20 cm以上, 其他动力电缆也应该远离UPS零线。如因场地原因不能拉开距离, 零线或地线可考虑使用屏蔽电缆, 也可达到同样效果。

4 工程案例

本案例是丽水水电中心新建工艺机房的高质量供电系统, 该系统采用模块式在线型UPS, 分3个区域布局设置。分别是: 1号区域为电视演播制作区, 额定容量120 kV·A, 主要负责大小演播厅6个, 新闻、栏目、广告等3个编辑机房及录音配音室的日常供电。在非常时期, 经卸载部分负荷, 可作为远程UPS向保障级别要求更高的广播电视播控系统, 提供更长时间的电力保障。2号区域为电视播控区, 额定容量40 kV·A, 系本地UPS, 专供具有4套电视节目播出容量的播总控系统的日常供电。3号区域为广播制播区, 额定容量20 kV·A, 系本地UPS, 负责最多可同时进行3套广播节目制播的电力保障。实践证明, 本设计(如图3所示)技术成熟, 安全有效。

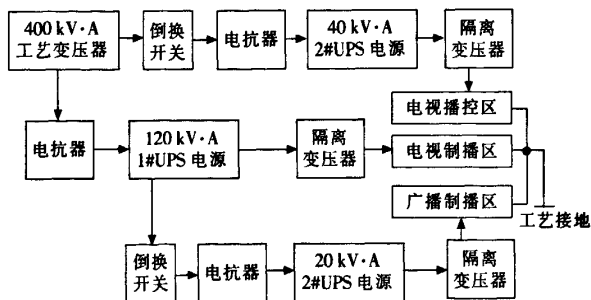


图3 机房供电系统结构

责任编辑:张家豪

收稿日期:2009-01-24

作者: 贾慧星, 章毓晋, JIA Hui-xing, ZHANG Yu-jin
作者单位: 清华大学, 信息科学与技术国家实验室, 电子工程系, 北京, 100084
刊名: 电视技术 **ISTIC PKU**
英文刊名: VIDEO ENGINEERING
年, 卷(期): 2009, 33(4)
被引用次数: 3次

参考文献(9条)

1. PARAGIOS N; RAMESH V [A mrf-based approach for real-time subway monitoring](#) 2008
2. KONG D; GRAY D; TAO H [A viewpoint invariant approach for crowd counting](#) 2008
3. RABAUD V; BELONGIE S J [Counting crowded moving objects](#) 2008
4. ANTONINI G; THIRAN J P [Counting pedestrians in video sequences using trajectory clustering](#) [外文期刊] 2006(08)
5. KILAMBI P; RIBNICK E; JOSHI A J [Estimating pedestrian count in groups](#) [外文期刊] 2008(01)
6. DONG L; PARAMESWARAN V; RAMESH V [Fast crowd segmentation using shape indexing](#) 2008
7. VIOLA P; JONES M [Rapid object detection using a boosted cascade of simple features](#) 2008
8. DALAL N; TRIGGS B [Histograms of oriented gradients for human detection](#) 2008
9. VAPNIK V [The nature of statistical learning theory](#) 2000

本文读者也读过(10条)

1. 樊秋月. FAN Qiu-yue [基于彩色视频图像处理的小区监视及人数统计](#) [期刊论文] - 通信技术 2010, 43(7)
2. 王强, 冯燕. Wang Qiang, Feng Yan [基于颜色和形状信息的快速人数统计方法](#) [期刊论文] - 计算机测量与控制 2010, 18(9)
3. 顾德军, 伍铁军. GU De-jun, WU Tie-jun [一种基于人头特征的人数统计方法研究](#) [期刊论文] - 机械制造与自动化 2010, 39(4)
4. 赵炯, 林旺城, 贾培源, 李中山 [公共娱乐场所人数统计系统研究](#) [期刊论文] - 自动化仪表 2009, 30(6)
5. 闫敬文, 樊秋月. YAN Jing-wen, FAN Qiu-yue [基于视频图像处理的人数统计方法](#) [期刊论文] - 汕头大学学报 (自然科学版) 2008, 23(2)
6. 邱磊, 唐桂军, 周书林 [自适应特征融合跟踪在进出人数统计中的应用](#) [期刊论文] - 科学技术与工程 2010, 10(4)
7. 衣淑凤 [基于小波变换和神经网络的人数统计方法研究](#) [学位论文] 2004
8. 张庚, 徐昊 [简单背景下基于灰度图象的人数统计](#) [期刊论文] - 中国新技术新产品 2009(1)
9. 刘毅安 [基于运动检测的人数统计系统](#) [学位论文] 2006
10. 苏航, 郑世宝, 杨华. SU Hang, ZHENG Shi-bao, YANG Hua [视频监控中人群流量和密度估计算法分析](#) [期刊论文] - 电视技术 2009, 33(11)

引证文献(3条)

1. 江帆, 王贵锦, 刘畅, 林行刚 [一种基于模型融合的行人跟踪算法](#) [期刊论文] - 电视技术 2010(3)
2. 袁理, 陈庆虎 [基于自适应3DLBP特征的人脸深度图像识别](#) [期刊论文] - 电视技术 2013(19)
3. 何华丽, 唐普英, 唐湘成, 杨建熙, 曾军, 黄自力, 李晶 [基于自适应多模板匹配的目标跟踪](#) [期刊论文] - 电视技术 2013(19)

引用本文格式：[贾慧星](#). [章毓晋](#). [JIA Hui-xing. ZHANG Yu-jin](#) [智能视频监控中基于机器学习的自动人数统计](#)[期刊论文]-[电视技术](#) 2009(4)