

# 基于 Adaboost 和码本模型的手扶电梯出入口视频监控方法

杜启亮, 黎浩正\*, 田联房

(华南理工大学 自动化科学与工程学院, 广州 510640)

(\* 通信作者电子邮箱 466739850@qq.com)

**摘要:**针对传统视频监控方法无法对密集前景目标进行准确分割的问题,提出一种基于 Adaboost 和码本模型的多目标视频监控方法。首先,通过训练得到 Adaboost 人头分类器,利用码本算法为垂直拍摄的手扶电梯出入口图像建立背景模型,提取前景图像对其进行人头检测和跟踪;之后,剔除行人目标得到物件目标,对物件目标进行跟踪;最后,根据行人和物件的运动特征进行监控。对 12 段出入口视频序列的实验结果表明,监控方法能够准确稳定地跟踪行人和物件,完成逆行检测、客流统计、行人拥堵和物件滞留等监控任务,处理速度达到 36 帧/秒,目标跟踪准确率达到 94% 以上,行为监控准确率达到 95.8%,满足智能视频监控系统鲁棒性、实时性和准确性的要求。

**关键词:**Adaboost; 背景建模; 视频监控; 人头检测; 多目标跟踪

**中图分类号:** TP391.4      **文献标志码:** A

## Video monitoring method of escalator entrance area based on Adaboost and codebook model

DU Qiliang, LI Haozheng\*, TIAN Lianfang

(College of Automation Science and Engineering, South China University of Technology, Guangzhou Guangdong 510640, China)

**Abstract:** Aiming at the problem that the traditional video monitoring method can not divide the dense foreground objects accurately, a multi-target video monitoring method based on Adaboost and codebook model was proposed. Firstly, the Adaboost human head classifier was obtained by training, and the background model was established for the vertical elevator image of the elevator by the codebook algorithm. The foreground image was extracted and heads were detected and tracked. After that, the pedestrian target was removed to get the object target, and the object target was tracked. Finally, the movement of pedestrians and objects was monitored. The experimental results on 12 entrance area videos show that the method can track pedestrians and objects accurately and stably. It can accomplish the monitoring tasks of retrograde detection, passenger statistics, pedestrian congestion and object retention. With the processing speed of 36 frames per second, the tracking-accuracy rate is above 94% and the monitoring-accuracy rate is 95.8%. The proposed algorithm meets the intelligent video monitoring system robustness, real-time and accuracy requirements.

**Key words:** Adaboost; background modeling; video monitoring; head detection; multi-target tracking

## 0 引言

智能视频监控技术由于其准确直观和成本低廉的特点,被广泛应用于公共管理与决策领域,如商场出入口客流统计、公交车乘客人数统计和电梯口拥堵检测等。光照变化和目标遮挡是视频监控中造成误差的主要原因,为减少这两种因素的不良影响,一般采取摄像头垂直向下拍摄的方法。视频监控的核心任务是运动目标检测与跟踪,主要有两类方法:基于模板匹配的方法和基于统计分类的方法。

基于模板匹配的方法为目标建立特征模板,如人头的类圆模板和人体外轮廓的矩形模板等,通过拟合图像中与模板相匹配的前景目标进行检测与跟踪。该类方法计算量少,在检测单个或多个分离目标时效果较好,在工程上应用较广。潘浩等<sup>[1]</sup>用 Sobel 算子提取垂直拍摄的人体轮廓后进行曲线生长,能够适应公交车门处不同角度人体的检测需要,但是计算量大,处理速度只能达到 6 帧/秒。顾德军等<sup>[2]</sup>首先采用高

斯模型进行前景提取,之后根据轮廓特性和颜色特性为俯视的人头建立模板,最后对前景图像进行人头模板匹配,但是高斯模型无法提取静止的前景目标,因此无法进行滞留检测。张姗姗等<sup>[3]</sup>提出了一种基于深度图像的人头检测算法,利用立体匹配的方法有效解决了目标拥挤的问题,但是无法检测被遮挡的目标。上述算法的性能在很大程度上取决于前景提取算法的效果,因此存在对光线变化敏感和易受阴影干扰的问题。

基于统计分类的方法首先获取大量正负样本,之后提取样本的目标特征并训练分类器或建立神经网络进行目标检测,准确率高且鲁棒性强。文嘉俊等<sup>[4]</sup>和 Li 等<sup>[5]</sup>分别采用 Haar 特征和 LBP 特征训练得到 Adaboost 人头分类器,在垂直拍摄图像中检测人头并进行客流监控,解决了多人过线时传统方法难以准确完成的行人分割和计数,但是存在误检如背景中的椅子。朴春赫等<sup>[6]</sup>结合传统 ViBe 前景检测算法,利用 HOG 特征训练得到 SVM 分类器进行行人检测,有效消除目

**收稿日期:**2017-03-23; **修回日期:**2017-05-17。      **基金项目:**广州市产学研项目(201604010114);广东省前沿与关键技术创新专项资金资助项目(2016B090912001);广州市科信局国际合作项目(2012J5100001)。

**作者简介:**杜启亮(1980—),男,广东佛山人,副研究员,博士,主要研究方向:机器人、机器视觉;黎浩正(1994—),男,广东番禺人,硕士研究生,主要研究方向:计算机视觉、机器学习;田联房(1969—),男,山东济宁人,教授,博士,主要研究方向:模式识别、人工智能。

标残影但是无法检测水平拍摄图像中被遮挡的目标。Zhao 等<sup>[7]</sup>利用立体视觉和人工神经网络建立了准确率很高的行人检测系统但是处理速度只能达到 10 帧/秒。

本文在码本模型前景提取算法的基础上,提出基于 Adaboost 和码本模型的多目标跟踪算法。码本模型是高效的背景建模方法,但是无法对多目标前景进行分割,为此本文引入 Adaboost 人头分类器,利用检测目标之间的独立性完成目标分割,由此提高多目标跟踪的准确性。本文的视频监控方法主要应用于商场、地铁出入口和手扶电梯出入口等人流密集的场所,完成目标检测与跟踪、客流统计、行人拥堵和物件滞留等视频监控任务。

## 1 基于 Adaboost 和码本模型的前景提取

本文采取图 1 所示的摄像头垂直向下拍摄的方法以减少光照变化和目标遮挡的不良影响,首先利用码本算法提取包含行人目标和物件目标的前景图像,训练人头分类器对行人目标进行检测与跟踪,之后剔除行人目标对物件目标进行跟踪,最后根据行人和物件的运动特征进行目标行为监控,图 2 为算法流程。

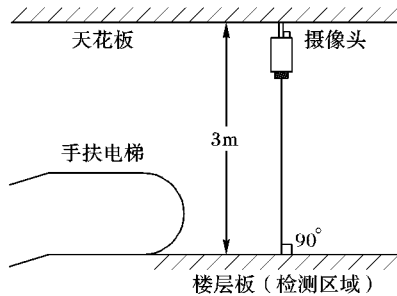


图1 摄像头安装位置示意图

Fig. 1 Location of camera installation

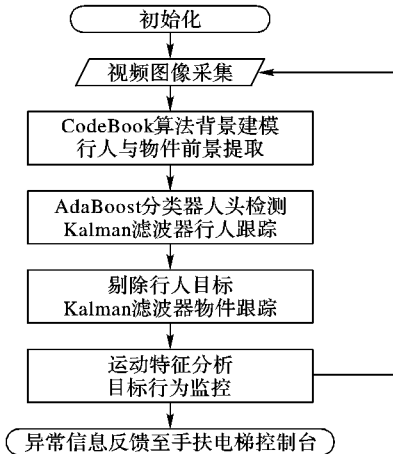


图2 视频监控算法流程

Fig. 2 Flow chart of video monitoring algorithm

### 1.1 基于码本模型的前景提取

视频监控的核心是背景建模和前景提取,在手扶电梯出入口等客流密集、背景复杂和光照变化频繁的应用场合,需要同时考虑算法对复杂背景的自适应能力以及算法复杂度对效率的影响。码本模型(Codebook Model)是一种高效的背景建模方法,基本思想是为图像所有像素建立一个时间序列,根据序列观察值对像素点进行分类。码本模型不需要背景像素的先验知识,只根据像素点时间特征提取前景图像,因此计算量

少且适用于复杂背景,步骤<sup>[8-9]</sup>如下:

1) 为图像每个像素建立一个码本(CodeBook, CB),每个码本由多个码字(CodeWord, CW)组成。CW为六元组,包含背景更新的学习上下界  $IH$  和  $IL$ ,当前像素的上下界  $I_{\max}$  和  $I_{\min}$ ,上一次的更新时间  $t_{\text{last}}$  和记录像素多久未被访问的陈旧时间  $t_{\text{st}}$ ,根据图像更新每个 CB 的状态;

2) 选择一帧或者多帧建立背景的 CB 模型,背景建立帧数一般为视频采集帧率的 1 ~ 2 倍;

3) 设前景图像的判定阈值上下界为  $IM_{\max}$  和  $IM_{\min}$ ,对图像中某一像素  $I(x, y)$ ,历遍背景 CB 模型中的每个 CW,如果存在一个 CW 的当前像素上下界  $I_{\max}$  和  $I_{\min}$  使得  $I_{\min} - IM_{\min} < I(x, y) < I_{\max} + IM_{\max}$ ,则  $I(x, y)$  为背景像素,否则为前景像素;

4) 隔一定帧数更新背景 CB 模型且进行时间滤波,设更新阈值  $T_{\text{st}}$  为更新次数的一半,历遍像素 CB 中的每个 CW,若陈旧时间  $t_{\text{st}}$  大于更新阈值,则移除该 CW。

采用带连通域分析的边缘生长算法对前景图像进行轮廓提取,加入轮廓周长等先验知识对轮廓图像进行二次校正,得到包含行人和物件两种目标的前景图像,图 3 为前景提取结果。

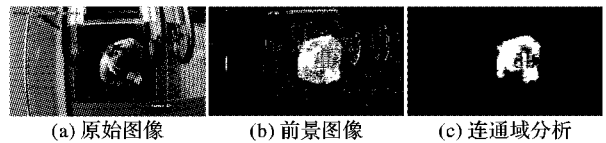


图3 前景提取结果

Fig. 3 Results of foreground extract

### 1.2 基于 Adaboost 的人头检测

基于轮廓特征的传统目标检测方法无法对密集前景目标进行准确分割,在垂直拍摄的图像中,人头有显著的椭圆形状特征,且不随前景目标密集程度发生变化,因此使用 HOG 特征的 Adaboost 分类器进行人头检测。HOG 通过计算和统计图像局部区域的梯度方向直方图来构成特征<sup>[10]</sup>,对目标图像进行 Gamma 压缩后计算梯度向量并建立方向直方图,本文的 HOG 描述子胞元(Cell)为  $8 \times 8$  的矩形像素块,梯度方向数为 9,块(Block)为  $2 \times 2$  的 Cell 串联块,也即  $16 \times 16$  的矩形像素块。

Adaboost 是一种结构简单的级联分类器,首先对同一个训练集训练不同的弱分类器,之后把弱分类器级联为强分类器,步骤<sup>[11]</sup>如下:

1) 设  $N$  为样本总数,将训练集所有样本的权重初始化为  $1/N$ ;

2) 以误差最小原则训练弱分类器,如果某一个训练样本已被正确分类,则在下一轮训练过程中降低该样本的权重,否则提高权重;

3) 弱分类器级联为强分类器,后者为前者的线性组合,误差率越小的弱分类器在强分类器中的权重越大。

本文选取正样本为 3850 张垂直拍摄的人头图片,大小为  $24 \times 24$ ,负样本为 8000 张不同场景下的其他图片,包括背景图片和非人头部位图片,大小范围为  $45 \times 45$  到  $105 \times 105$ ,图 4 为部分训练集图片。

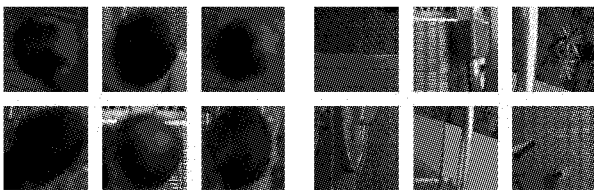
设置分类器训练的迭代层数为 20,每层的最小命中率为 0.999,最大误检率为 0.5,最大迭代次数为 100,表 1 为本文

Adaboost 分类器的训练结果,图 5 为人头目标检测结果。

表 1 Adaboost 分类器训练结果

Tab. 1 Training result of Adaboost classifier

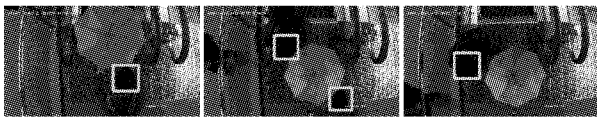
层数	弱分类器个数	权重	层数	弱分类器个数	权重
0	6	-1.949	10	100	-1.852
1	8	-1.680	11	100	-1.804
2	12	-1.799	12	100	-1.724
3	13	-1.722	13	100	-1.727
4	20	-1.690	14	100	-1.748
5	37	-1.981	15	100	-1.754
6	48	-1.760	16	100	-1.611
7	79	-1.792	17	100	-1.748
8	91	-1.549	18	100	-1.616
9	100	-1.760	19	100	-1.648



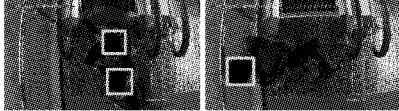
(a) 部分正样本 (b) 部分负样本

图 4 部分训练集图片

Fig. 4 Part of training set



(a) 第65920帧 (b) 第65981帧 (c) 第66014帧



(d) 第74072帧 (e) 第74342帧

图 5 目标检测结果

Fig. 5 Results of object detection

## 2 多目标跟踪

多目标跟踪的本质是目标特征的跟踪,视频监控任务中造成跟踪误差的主要原因有:1) 检测环境光照强度不足、不均或变化频繁,环境噪声造成分类器误检;2) 训练的分类器性能有限,出现漏检或误检;3) 训练集规模过小,不足以训练出高精度的分类器。本文针对前两点建立了一种运动目标的跟踪强度特征模型,能够自适应地增添新出现的目标,丢弃应该停止跟踪的过期目标,采用最小距离的目标匹配方法对多目标进行卡尔曼跟踪,图 6 为多目标跟踪算法流程。

### 2.1 目标的跟踪强度特征模型

根据分类器的检测结果,观测序列的运动目标数量和跟踪序列的运动目标数量存在三种情况:1) 观测数  $M$  小于跟踪数  $N$ , 此时可能是运动目标离开检测区域或分类器漏检 (False Rejection), 为保证跟踪序列的连续性, 应当增补  $N - M$  组观测值并降低未观测到的运动目标的跟踪强度;2) 观测数  $M$  大于跟踪数  $N$ , 此时可能是出现新的运动目标或分类器误检 (False Acceptance), 应当增补  $M - N$  组跟踪值并提高已检测到的运动目标的跟踪强度;3) 观测数  $M$  等于跟踪数  $N$ , 此时可能是分类器无差检测或漏检误检恰好使二者匹配, 应当

对观测序列与跟踪序列作最小距离目标匹配。本文将 1) 和 3) 合并为一种情况, 设  $cf(k)$  和  $cf(k-1)$  分别为某运动目标在  $k$  时刻和  $k-1$  时刻的跟踪强度,  $cfi(k)$  和  $efd(k)$  分别为目标被连续观测到和连续未被观测到的帧数,  $cfi(k) \cdot efd(k) = 0$ , 则该目标的跟踪强度特征按下式更新:

$$cf(k) = \begin{cases} cf(k-1) + \log_{1.5}(cfi(k)), & efd = 0 \\ cf(k-1) - 1.5^{efd(k)}, & cfi = 0 \end{cases} \quad (1)$$

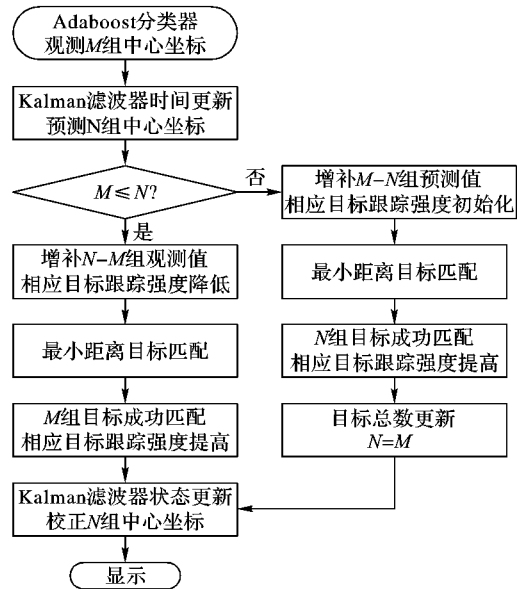


图 6 多目标跟踪算法流程

Fig. 6 Flow of multi-target tracking algorithm

### 2.2 最小距离目标匹配

观测序列与跟踪序列利用欧氏距离 (Euclidean Metric) 进行最小距离匹配, 当观测数  $M$  小于或等于跟踪数  $N$  时, 算法对  $M$  组已观测目标作最小距离匹配, 并在原地增补  $N - M$  组观测值即假设未观测到的运动目标静止不动; 当观测数  $M$  大于跟踪数  $N$  时, 算法对  $N$  组已观测目标作最小距离匹配, 并以未匹配的观测值为初始值新建  $M - N$  组跟踪序列。设观测序列为  $m$ , 跟踪序列为  $n$ , 图 7 为最小距离目标匹配的示意图。

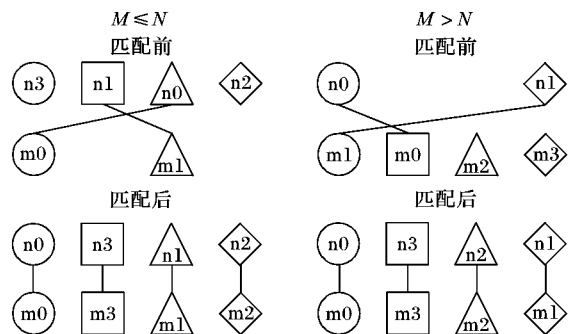


图 7 最小距离目标匹配示意图

Fig. 7 Schematic of minimum distance target matching

### 2.3 卡尔曼滤波器

卡尔曼滤波器 (Kalman Filter) 是一种拥有状态估计功能的滤波器, 通过迭代核心方程, 状态变量最终将收敛至一个最优自回归解, 即状态变量的最优估计, 利用卡尔曼滤波器预测下一时刻的系统状态, 从而完成多目标跟踪任务, 其时间与状态共五个核心更新方程<sup>[12]</sup>如下:

$$\begin{cases} \mathbf{xg}(k) = \mathbf{Ax}(k-1) \\ \mathbf{P}(k)' = \mathbf{AP}(k-1)\mathbf{A}^T + \mathbf{Q} \\ \mathbf{K}(k) = \mathbf{P}(k)' \mathbf{H}^T (\mathbf{HP}(k)' \mathbf{H}^T + \mathbf{R})^{-1} \\ \mathbf{x}(k) = \mathbf{xg}(k) + \mathbf{K}(k) (\mathbf{y}(k) - \mathbf{Hxg}(k)) \\ \mathbf{P}(k) = (\mathbf{I} - \mathbf{K}(k) \mathbf{H}) \mathbf{P}(k)' \end{cases} \quad (2)$$

对中心坐标  $\mathbf{P}_p = (x_p, y_p)$ ,  $\mathbf{x}(k) = [x_p \ y_p \ \Delta x_p \ \Delta y_p]^T$  为状态变量,  $\mathbf{y}(k) = [x_p \ y_p]^T$  为观测变量,  $\mathbf{A}$  为状态估计矩阵,  $\mathbf{H}$  为观测矩阵,  $\mathbf{q}$  和  $\mathbf{r}$  分别为估计噪声和观测噪声, 符合高斯分布,  $\mathbf{Q}$  和  $\mathbf{R}$  分别为其协方差矩阵,  $\mathbf{P}$  为误差协方差矩阵。本文建立行人和物件目标在手扶电梯出入口等环境下的线性运动模型并取值如下:

$$\begin{cases} \mathbf{x}(k) = \mathbf{Ax}(k-1) + \mathbf{q} \\ \mathbf{y}(k) = \mathbf{Hx}(k) + \mathbf{r} \end{cases} \quad (3)$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (5)$$

$$\mathbf{Q} = 10^{-5} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

$$\mathbf{R} = 10^{-1} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7)$$

以观测数  $M$  小于跟踪数  $N$  为例, 采用最小欧氏距离的目标匹配方法对两个行人目标进行卡尔曼跟踪, 算法对  $M$  组观测目标作最小距离匹配, 并在原地增补  $N - M$  组观测值即假设未观测到的运动目标静止不动, 同时降低未观测到的运动目标的跟踪强度, 图 8 为卡尔曼跟踪结果及其轨迹, 矩形框为观测结果, 圆形框为跟踪结果, 表 2 为跟踪过程中观测序列、跟踪序列、欧氏距离和跟踪强度的变化。

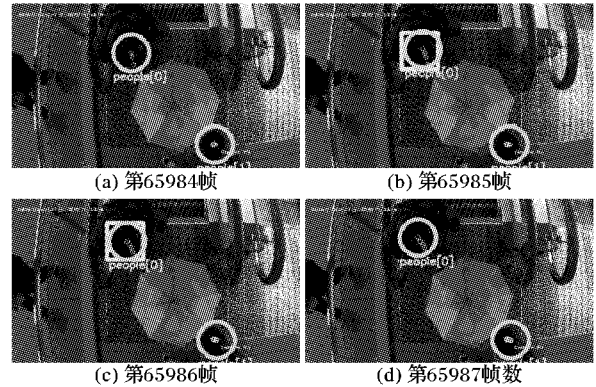


图 8 卡尔曼跟踪结果

Fig. 8 Results of Kalman tracking

表 2 跟踪过程中卡尔曼滤波器的参数变化

Tab. 2 Changes of Kalman filter parameters in tracking process

帧序号	匹配?	目标 0				目标 1			
		观测值	跟踪值	欧氏距离	跟踪强度	观测值	跟踪值	欧氏距离	跟踪强度
65984	匹配前	—	(198, 82)	0	11.55	—	(338, 232)	0	523.81
	匹配后	(198, 82)	(196, 76)			(338, 232)	(339, 233)		
65985	匹配前	(191, 79)	(196, 76)	5.83	13.26	—	(339, 233)	0	520.44
	匹配后	(191, 79)	(191, 71)			(339, 233)	(340, 233)		
65986	匹配前	(189, 71)	(191, 71)	2	15.67	—	(340, 233)	0	515.38
	匹配后	(189, 71)	(188, 65)			(340, 233)	(341, 233)		
65987	匹配前	—	(188, 65)	0	14.47	—	(341, 233)	0	507.78
	匹配后	(188, 65)	(185, 59)			(341, 233)	(342, 234)		

### 3 基于运动特征的视频监控

根据行人和物件的运动特征进行监控, 主要是由二者的中心坐标计算运动速度和运动方向, 判断运动目标行为从而达到视频监控的目的。本文对手扶电梯出入口上的运动目标进行下列 4 种行为监控: 逆行检测、客流统计、行人拥堵和物件滞留。设采样帧数为  $n$ , 运动目标在  $k$  时刻的跟踪强度为  $ef(k)$ , 中心坐标  $\mathbf{P}_p(k) = (x_p(k), y_p(k))$ , 则运动速度和运动方向如下式所示:

$$v(k) = \frac{1}{n} \sqrt{(x_p(k) - x_p(k-n))^2 + (y_p(k) - y_p(k-n))^2} \quad (8)$$

$$\theta(k) = \arctan\left(\frac{y_p(k) - y_p(k-n)}{x_p(k) - x_p(k-n)}\right) \quad (9)$$

#### 3.1 逆行检测

在手扶电梯出入口等存在正常前进方向的场所中逆行可能发生安全事故, 因此需要对该行为进行监控。以手扶电梯

前进方向为极轴, 取逆时针方向为正方向建立极坐标系, 则目标逆行的运动特征为:

$$-\pi/2 < \theta(k) < \pi/2 \quad (10)$$

#### 3.2 客流统计

在手扶电梯出入口监控区域设置客流计数线, 按手扶电梯前进方向将出入口分割为内部区域  $Area_{in}$  与外部区域  $Area_{out}$ , 从  $k-n$  时刻某目标出现在监控区域并被跟踪开始进行统计, 经过  $n$  帧后在  $k$  时刻目标通过客流计数线, 此时算法判定目标进入监控区域 (客流数加 1) 或离开监控区域 (客流数减 1), 当前客流统计过程结束, 如图 9 所示。

设  $k$  时刻该目标在内部区域存在的累计帧数为  $in(k)$ , 在外部区域存在的累计帧数为  $out(k)$ , 客流方向的起点阈值为  $T_{begin}$ , 终点阈值为  $T_{end}$ , 且  $T_{begin} > T_{end}$ , 定义符号“ $\in$ ”表示坐标包含于某区域内, 则目标进入监控区域的运动特征为:

$$\begin{cases} in(k) > T_{end} \\ out(k) > T_{begin} \\ P_p(k) \in Area_{in} \end{cases} \quad (11)$$

目标离开监控区域的运动特征为:

$$\begin{cases} in(k) > T_{begin} \\ out(k) > T_{end} \\ P_p(k) \in Area_{out} \end{cases} \quad (12)$$

本文客流统计的起点阈值  $T_{begin} = 5$ , 终点阈值  $T_{end} = 2$ , 运动目标通过客流计数线后, 其累计帧数  $in(k)$  和  $out(k)$  均要清零, 这样能够有效防止目标中心位置在计数线附近抖动所造成的客流误检。表 3 为某运动目标通过客流计数线进入或离开监控区域前后的运动特征变化, 其中符号“#”表示目标进入监控区域的时刻, 符号“\*”表示目标离开监控区域的时刻。

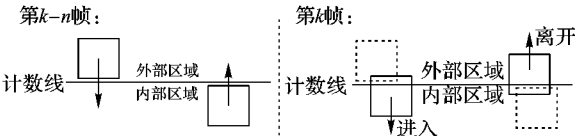


图 9 Process of passenger flow statistics

表 3 目标进入或离开监控区域时的运动特征变化  
Tab. 3 Changes of movement information while target is entering or leaving the area

帧 序号	进入		离开		帧 序号	进入		离开	
	$in(k)$	$out(k)$	$in(k)$	$out(k)$		$in(k)$	$out(k)$	$in(k)$	$out(k)$
1	0	0	0	0	11	1	0	7	1
2	0	0	0	0	12*	2	0	7	2
3	0	1	0	0	13	3	0	0	1
4	0	2	1	0	14	4	0	0	2
5	0	3	2	0	15	5	0	0	3
6	0	4	3	0	16	6	0	0	4
7	0	5	4	0	17	7	0	0	5
8	0	6	5	0	18	8	0	0	6
9	1	6	6	0	19	9	0	0	7
10#	2	6	7	0	20	10	0	0	8

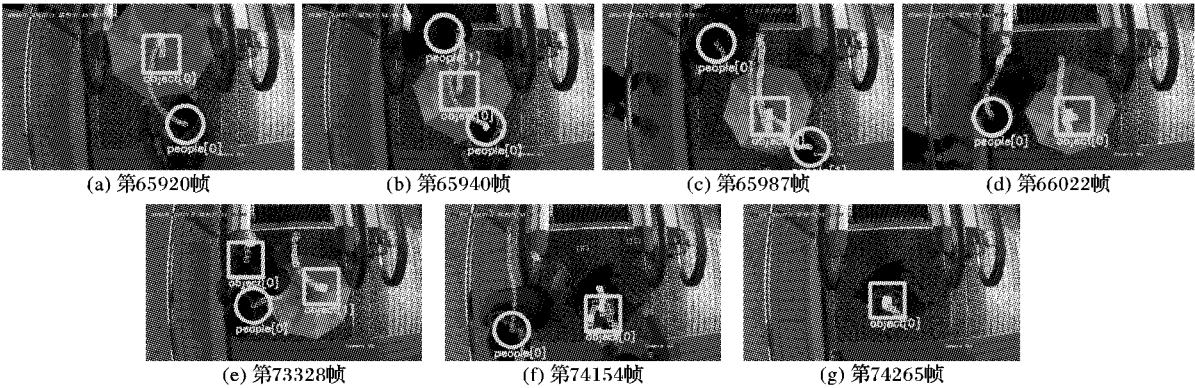


图 10 目标跟踪结果

Fig. 10 Results of object tracking

目标跟踪的结果和性能指标表明, 算法能够在行人目标和物件目标同时存在的情况下分别对其进行连续稳定的跟踪, 行人跟踪的总调和均值为 96.1%, 物件跟踪的总调和均值为 94.9%, 算法在光照充足且稳定的环境下性能达到最佳。

对行人跟踪性能指标进行分析, Adaboost 人头分类器能够有效抵抗检测环境光照强度及其稳定性的影响。由于光照不足或不稳所造成的召回率降幅约为 3%, 说明不良的光照

3.3 行人拥堵与物件滞留

若手扶电梯出入口区域存在行人拥堵或大物件滞留, 容易发生安全事故, 因此需要对两类行为进行监控, 由于出入口区域面积较小, 因此某目标在该区域内静止或反复作小幅移动均认为是拥堵或滞留, 设  $T_{cf}$  为拥堵滞留的跟踪强度阈值, 实验中  $T_{cf} = 2000$ , 则行人拥堵或物件滞留的运动特征如下:

$$cf(k) > T_{cf} \quad (13)$$

4 实验结果与分析

为分析基于 Adaboost 和码本模型的视频监控方法的效果, 对 12 段手扶电梯出入口视频序列进行实验, 视频包含不同光照环境下物件滞留、行人正常通过、逆行、滞留和拥堵等多种事件, 算法在 i5-6200U 2.40 GHz CPU、8 GB RAM、Windows 10 操作系统的计算机上用 C++ 编程实现。视频图像大小为  $480 \times 272$ , 处理速度达到 36 帧/秒。

4.1 目标跟踪实验结果

从某目标进入监控区域开始跟踪, 至其离开并结束跟踪, 设目标存在于监控区域期间经历的总帧数为  $n$ , 若算法能够连续稳定地跟踪目标位置且累计跟踪帧数达到  $0.95n$  以上, 则认为算法成功跟踪该目标。设整个视频监控过程中实际总目标数为  $TG(\text{Target})$ , 则按上述方法定义成功跟踪数为  $TP(\text{True Positive})$ , 若算法将行人肩部或背景中的阴影等误检为人头则计入误检数  $FP(\text{False Positive})$ , 若无法检测出某目标则计入漏检数  $FN(\text{False Negative})$ 。本文采用文献[13]的性能指标分析算法对行人目标及物件目标的跟踪效果, 分别是精确率  $PR(\text{Precision})$ 、召回率  $RE(\text{Recall})$  和调和均值  $F_1(F_1 \text{ Score})$ , 其中  $PR = TP / (TP + FP)$ ,  $RE = TP / (TP + FN)$ ,  $F_1 = 2TP / (2TP + FP + FN)$ 。图 10 为目标跟踪结果, 圆形框为行人跟踪结果及其轨迹, 矩形框为物件跟踪结果及其轨迹, 表 4 为目标跟踪性能指标。

环境造成了更多漏检, 这与 HOG 特征基于方向梯度的特性相符合。但即使是恶劣的光照环境下算法的调和均值也能保持在 94% 以上, 说明算法对环境光照的变化是鲁棒的。在此基础上, 客流的拥挤情况是造成算法性能下降的主要原因, 相同光照条件下, 算法对稀疏客流的跟踪性能比对拥挤客流的跟踪性能提高 3.5% 以上。在稀疏情况下, 俯视人体的肩部 HOG 特征与人头相似, 可能造成误检; 而在拥挤情况下, 人体边缘的粘连、衣物颜色过深和行人相距过近造成的阴影等因

素使人头的 HOG 特征不再明显,造成了更多漏检,图 11 为上述几种误检或漏检情况的例子。

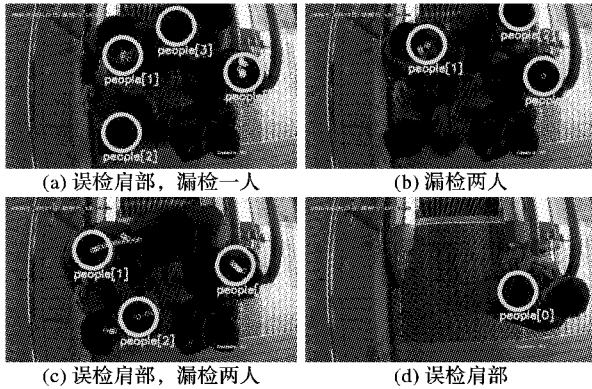


图 11 几种误检或漏检情况

Fig. 11 Several cases of false acceptance or false rejection

表 4 目标跟踪性能指标

Tab. 4 Performance of object tracking

序号	行人目标								物件目标							
	视频	TG	TP	FP	FN	PR/%	RE/%	F <sub>1</sub> /%	视频	TG	TP	FP	FN	PR/%	RE/%	F <sub>1</sub> /%
1	光照充足、稳定,客流拥挤	340	313	8	19	97.5	94.3	95.9	光照充足、稳定	88	86	2	0	97.7	100	98.9
2	光照充足、稳定,客流稀疏	327	316	3	8	99.1	97.5	98.3	光照充足、稳定	30	28	2	0	93.3	100	96.6
3	光照强烈、稳定,客流拥挤	30	29	1	0	96.7	100	98.3	光照强烈、稳定	109	101	7	1	93.5	99.0	96.2
4	光照强烈、稳定,客流稀疏	41	40	1	0	97.6	100	98.8	光照强烈、稳定	114	109	4	1	96.5	99.1	97.8
5	光照不足、稳定,客流稀疏	48	44	1	3	97.8	93.6	95.7	光照不足、稳定	99	86	10	3	89.6	96.6	93.0
6	光照不足、稳定,客流拥挤	299	269	11	19	96.1	93.4	94.7	光照不足、稳定	101	89	10	2	89.9	97.8	93.7
7	光照充足、不稳,客流拥挤	62	59	0	3	100	95.2	97.5	光照充足、不稳	42	38	4	0	90.5	100	95.0
8	光照充足、不稳,客流稀疏	73	68	1	4	98.6	94.4	96.5	光照充足、不稳	25	23	2	0	92.0	100	95.8
9	光照强烈、不稳,客流拥挤	322	303	7	12	97.7	96.2	97.0	光照强烈、不稳	39	33	6	0	84.6	100	91.7
10	光照强烈、不稳,客流稀疏	325	301	8	16	97.4	95.0	96.2	光照强烈、不稳	120	105	14	1	88.2	99.1	93.3
11	光照不足、不稳,客流稀疏	44	40	0	4	100	90.9	95.2	光照不足、不稳	52	44	8	0	84.6	100	91.7
12	光照不足、不稳,客流拥挤	313	273	14	26	95.1	91.3	93.2	光照不足、不稳	40	33	5	2	86.8	94.3	90.4
总计		2224	2055	55	114	97.4	94.7	96.1		859	775	74	10	91.3	98.7	94.9

#### 4.2 目标行为监控实验结果

本文完成了视频监控中逆行检测、客流统计、行人拥堵和物件滞留等监控任务,仍然采用文献[13]中的 PR、RE 和 F<sub>1</sub> 指标分析算法对目标行为的监控效果,图 12 为目标行为检测结果,表 5 为目标行为监控的性能指标。

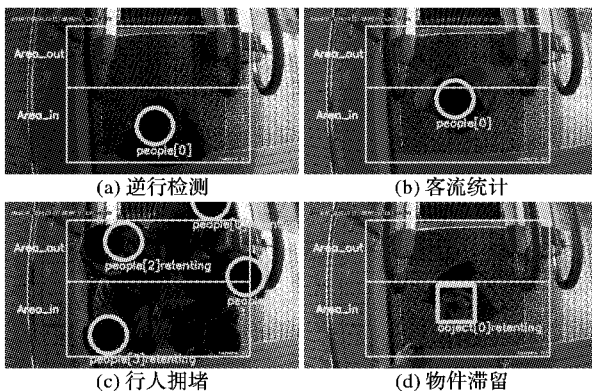


图 12 目标行为检测结果

Fig. 12 Target behavior analysis

分析目标行为检测的结果和性能指标,其中逆行检测和行人拥堵的效果仅决定于行人跟踪的性能,因此二者的性能指标与行人跟踪的相关指标相近。对客流统计而言,若无法

对物件跟踪性能指标进行分析,码本模型在手扶电梯出入口等光照变化缓慢均匀的环境下有较好的背景建模效果,但是作为一种传统的背景建模方法,码本模型无法分割拥挤情况下的粘连目标。对视频监控这一任务而言,物件跟踪一般为大件物品滞留判定提供有用信息,因此实际工程中视频监控对物件跟踪的性能要求较低。虽然在恶劣的光照环境下调和均值的降幅达到了 9% 以上,但引入某些先验知识及矫正条件后,算法的物件跟踪总计调和均值仍能够达到 94% 以上,足以完成滞留判定任务。此外,算法的精确率较低而召回率较高,说明码书模型一类的传统背景建模方法主要缺点在于容易造成误检。由目标跟踪的结果及性能指标来看,结合 Adaboost 和码本模型的多目标跟踪算法能够准确地建立行人及物件目标的位置序列,这为目标行为监控奠定了良好的基础。

保证行人跟踪的性能,则目标在通过客流计数线时可能多次触发即发生误检,表现在客流统计的精确率只有 91.3%,与其他目标行为的指标相比数值较低,因此该行为的调和均值也是最低的。物件滞留主要利用码本模型的相关背景建模与前景提取算法,在物件目标跟踪部分的分析中已经提到,在良好的光照环境下,较低的跟踪性能便足以完成该监控任务,因此物件滞留的调和均值能够达到最高的 97.2%。总体而言,算法能够在各种环境下准确地完成基本的视频监控任务,对各行为的检测调和均值均达到 93.9% 以上,平均调和均值为 95.8%,算法有较强的鲁棒性。

表 5 目标行为监控结果

Tab. 5 Performance of target behavior monitoring

目标行为	TG	TP	FP	FN	PR/%	RE/%	F <sub>1</sub> /%
逆行检测	587	534	22	31	96.0	94.5	95.3
客流统计	2014	1784	171	59	91.3	96.8	93.9
行人拥堵	1607	1511	44	52	97.2	96.7	96.9
物件滞留	832	787	37	8	95.5	99.0	97.2
平均					95.0	96.7	95.8

#### 5 结语

为了解决传统视频监控方法在密集目标场景下失效的问

题,满足智能监控系统的实时性需求,本文提出了一种基于 Adaboost 和码本模型的多目标视频监控方法,通过训练得到 Adaboost 人头分类器,利用码本算法为垂直拍摄的手扶电梯出入口图像建立背景模型,提取前景图像对其进行人头检测和跟踪,之后剔除行人目标得到物件目标,对物件目标进行跟踪,最后根据行人和物件的运动特征进行监控。实验表明,本文的监控算法能够准确稳定地跟踪行人和物件两类目标,通过分析目标的运动特征能够准确完成逆行检测、客流统计、行人拥堵和物件滞留等监控任务,处理速度达到 36 帧/秒,目标跟踪准确率达到 94% 以上,行为监控准确率达到 95.8%,满足智能视频监控系统鲁棒性、实时性和准确性的要求。但算法仍然存在不足,当检测环境过暗或光照产生猛烈突变时,算法容易监控失败。今后将会在多目标检测与跟踪方面改善算法的性能,增强算法在不同应用场合下的鲁棒性,使其更好地应用在各种智能视频监控系统当中。

#### 参考文献 (References)

- [1] 潘浩,高枝宝,何小海,等. 基于计算机视觉的公交系统人流量检测算法[J]. 计算机工程, 2007, 33(11): 216-218. (PAN H, GAO Z B, HE X H, et al. Pedestrian flow detection algorithm in public transport system based on computer vision[J]. Computer Engineering, 2007, 33(11): 216-218.)
- [2] 顾德军,伍铁军. 一种基于人头特征的人数统计方法研究[J]. 机械制造与自动化, 2010, 39(4): 134-138. (GU D J, WU T J. Pedestrian count method based on head feature[J]. Machine Building & Automation, 2010, 39(4): 134-138.)
- [3] 张姗姗,景文博,刘学,等. 一种基于深度信息的人头检测方法[J]. 长春理工大学学报(自然科学版), 2016, 39(2): 107-111. (ZHANG S S, JING W B, LIU X, et al. A head detection method based on depth information[J]. Journal of Changchun University of Science and Technology (Natural Science), 2016, 39(2): 107-111.)
- [4] 文嘉俊,徐勇,战荫伟. 基于 AdaBoost 和帧间特征的人数统计[J]. 中国图象图形学报, 2011, 16(9): 1729-1735. (WEN J J, XU Y, ZHAN Y W. People counting based on AdaBoost and inter-frame features[J]. Journal of Image and Graphics, 2011, 16(9): 1729-1735.)
- [5] LI B, ZHANG J, ZHANG Z, et al. A people counting method based on head detection and tracking [C]// Proceedings of the 2014 International Conference on Smart Computing. Washington, DC: IEEE Computer Society, 2014: 136-141.
- [6] 朴春赫,潘怡霖,赵海,等. 基于改进 ViBe 的多行人检测方法[J]. 东北大学学报(自然科学版), 2016, 37(4): 481-485. (PAK C H, PAN Y L, ZHAO H, et al. Multi-pedestrian detection approach based on improved ViBe algorithm [J]. Journal of Northeastern University (Natural Science), 2016, 37(4): 481-485.)
- [7] ZHAO L, THORPE C E. Stereo- and neural network-based pedestrian detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2000, 1(3): 148-154.
- [8] KIM K, CHALIDABHONGSE T H, HARWOOD D, et al. Real-time foreground-background segmentation using codebook model [J]. Real-Time Imaging, 2005, 11(3): 172-185.
- [9] KIM K, CHALIDABHONGSE T H, HARWOOD D, et al. Background modeling and subtraction by codebook construction [C]// Proceedings of the 2004 International Conference on Image Processing. Piscataway, NJ: IEEE, 2004, 5: 3061-3064.
- [10] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]// Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2005, 1: 886-893.
- [11] KÉGL B. The return of AdaBoost. MH: multi-class Hamming trees [EB/OL]. [2016-12-20]. <https://core.ac.uk/download/pdf/24989526.pdf>.
- [12] KELLY A. A 3D state space formulation of a navigation Kalman filter for autonomous vehicles [EB/OL]. [2016-12-20]. [http://fric.ri.cmu.edu/users/alonzo/pubs/reports/kalman\\_V2.pdf](http://fric.ri.cmu.edu/users/alonzo/pubs/reports/kalman_V2.pdf).
- [13] POWERS D M W. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation [J]. Journal of Machine Learning Technologies, 2011, 2(1): 37-63.

This work is partially supported by the Projects on the Integration of Industry, Education and Research of Guangzhou (201604010114), the Special Funds for Frontier and Key Technology Innovation of Guangdong (2016B090912001), the International Cooperation Projects of Science and Technology Information Bureau of Guangzhou (2012J5100001).

**DU Qiliang**, born in 1980, Ph. D., associate professor. His research interests include robot, machine vision.

**LI Haozheng**, born in 1994, M. S. candidate. His research interests include computer vision, machine learning.

**TIAN Lianfang**, born in 1969, Ph. D., professor. His research interests include pattern recognition, artificial intelligence.