

文章编号:1674-2974(2016)10-0102-08

# 基于支持向量机优化的行人跟踪学习检测方法<sup>\*</sup>

孙 炜,薛 敏<sup>†</sup>,孙天宇,胡梦云,吕云峰

(湖南大学 电气与信息工程学院,湖南 长沙 410082)

**摘 要:**提出一种基于 SVM(Support Vector Machine)优化的 TLD(Track-Learning-Detection)行人检测跟踪算法.将行人作为正样本,背景作为负样本,提取出行人的 HOG 特征并投入线性 SVM 中进行训练,得到行人检测分类器,并标定出目标区域,实现行人自动识别;然后在 TLD 算法的基础上对行人进行跟踪和在线学习,估计检测出的正负样本并实时修正检测器在当前帧中的误检,利用相邻帧间特征点配准剔除误配点,同时更新跟踪器数据,以避免后续出现类似错误.实验表明,该算法能够适应遮挡变化且自动识别并稳定跟踪目标行人,较传统跟踪算法具有更强的鲁棒性.

**关键词:**支持向量机;行人检测跟踪;TLD

**中图分类号:**TP391.4

**文献标识码:**A

## The Optimized Pedestrian Tracking-Learning-Detection Algorithm Based on SVM

SUN Wei, XUE Min<sup>†</sup>, SUN Tian-yu, HU Meng-yun, LV Yun-feng

(College of Electrical and Information Engineering, Hunan Univ, Changsha, Hunan 410082, China)

**Abstract:** A new method based on optimized TLD (Track-Learning-Detection) and SVM (Support Vector Machine) for tracking pedestrian was proposed. First, with pedestrians as positive samples and the background as negative samples respectively, HOG (Histogram of Oriented Gradient) descriptor of pedestrian was extracted and combined with linear SVM to train the pedestrian classifier, which was used to obtain the calibrated pedestrian area accurately. Then, adaptive tracking and online learning on the pedestrians on the basis of TLD were integrated to estimate the reliability of the positive and negative samples, to rectify error existing in the current frame caused by detection and to update the tracking data simultaneously to avoid subsequent similar mistakes. The experiment results demonstrate that, compared with the conventional tracking algorithm, the proposed algorithm can not only significantly adapt to occlusions and appearance changes but also automatically identify and track pedestrian targets at arbitrary position, manifesting stronger robustness.

**Key words:** support vector machine; pedestrian detection and tracking; TLD

<sup>\*</sup> 收稿日期:2015-05-21

基金项目:高等学校博士学科点专项科研基金资助项目(20130161110009);湖南省自然科学基金资助项目(14JJ1011)

作者简介:孙 炜(1975-),男,河南淮滨人,湖南大学教授,博士生导师

<sup>†</sup> 通讯联系人, E-mail: xueminhnu@hotmail.com

在计算机视觉应用中,长期稳定实时检测跟踪运动物体已经成为一个重要的研究课题,随着技术的不断成熟,该领域的应用也相当广泛,比如:工业生产、实时监控、自动目标定位、自动导航、人机交互、增强现实技术、SLAM、游戏开发等. 研究人员根据实际应用的需要提出不同的跟踪方案,其中 D. Comaniciu 等<sup>[1]</sup>利用改进的 mean-shift 来跟踪运动物体,该方法可通过迭代步骤找到离跟踪目标最近的位置,但其不能解决目标被遮挡或发生形变等问题,在前后背景颜色相似的环境中,容易发生目标跟丢的情况. Martinez 等<sup>[2]</sup>用背景差分法提取目标,根据轨迹建立数学模型实现行人跟踪,该方法可取得较好的检测效果,但由于计算时间长,跟踪的实时性不能得到保证. 季玉龙等<sup>[3]</sup>提出的对运动目标建模的方法需要大量的先验知识,对视频帧中出现的相似目标干扰没有很好的鲁棒性,不能保证长期的跟踪性和适应性.

TLD 算法<sup>[4-7]</sup>将检测过程和跟踪过程融合起来并引入学习机制,当目标发生形变或被遮挡时,可以实时更新目标模型,实现在线学习和评估,在一定程度上可以克服目标外观变化及运动模糊,具有较

强的稳定性. 但是 TLD 算法中的跟踪目标并不能自动识别,当目标移动过快时,跟踪准确度也有所下降,因此泛化性仍有待提高. 本文在其基础上提出了一种基于支持向量机分类优化的行人跟踪学习检测方法,该方法利用 SVM 分类器检测到待跟踪目标,保留 TLD 算法在复杂背景下对目标区域实时学习与跟踪的基础上在检测模块中加入图像特征点配准<sup>[8]</sup>以实现更高的跟踪精度.

### 1 方法概述

本文利用视频初始帧中行人方向梯度直方图和支持向量机检测出行人在视频帧中的特征及位置,将目标行人提取出来,记录行人位置左上角坐标与行人的宽和高. 确认目标后将视频帧输入到检测模块与跟踪模块并产生相应的实时正负样本,学习模块根据捕捉到的正负样本不断将信息反馈给检测模块与跟踪模块,利用图像特征点配准进一步剔除图像元中的误配点,达到良好的检测跟踪效果. 算法具体框架如图 1 所示.

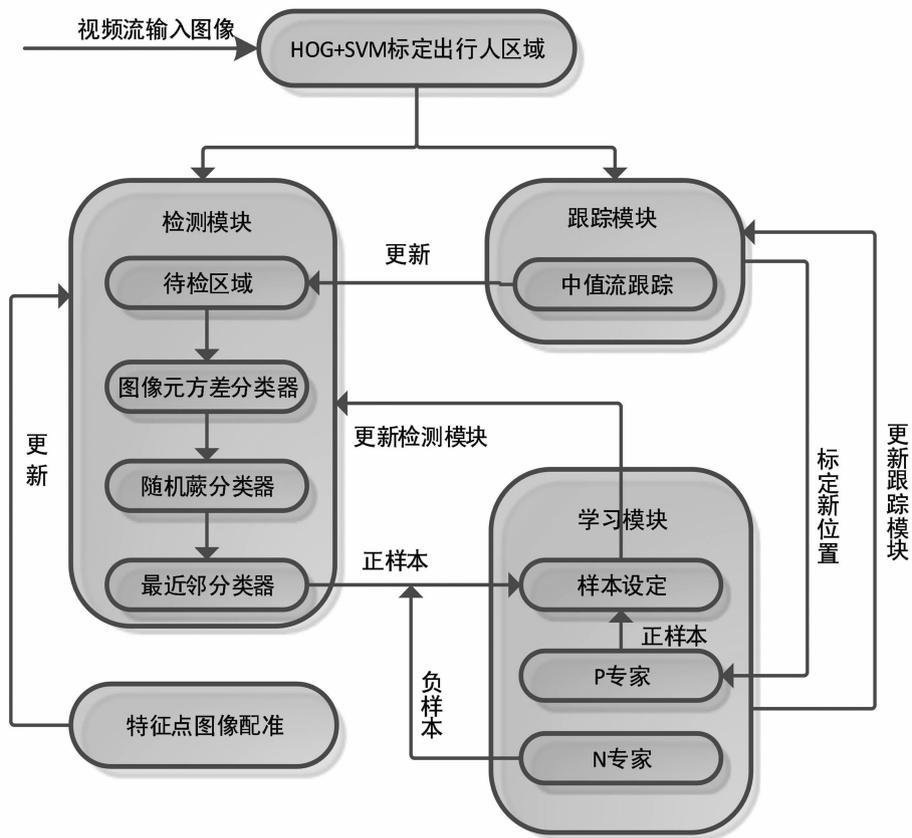


图 1 改进算法框架图  
Fig. 1 The block diagram of improved algorithm

## 2 基于 HOG 与线性 SVM 的行人检测方法

### 2.1 行人特征提取

行人特征可以用灰度、边缘、SIFT 特征、Haar 特征等来描述. 由于 HOG<sup>[9]</sup> 属于对局部目标进行检测, 对阴影、光照条件的变化、小角度旋转以及微小行人动作有较好的鲁棒性. 因此本文采用 HOG 来提取行人特征. 目标的 HOG 提取过程<sup>[10]</sup> 如下:

1) 对图像进行 gamma 校正以实现标准化, 降低噪声干扰及光照影响, 如图 2 所示.

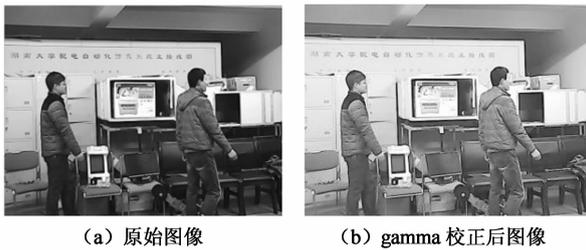


图 2 校正前后对比图

Fig. 2 The original image and the image after gamma

2) 对各像素点求梯度大小及方向, 捕捉目标区域的轮廓.

$$G_x(x, y) = H(x + 1, y) - H(x - 1, y), \quad (1)$$

$$G_y(x, y) = H(x, y + 1) - H(x, y - 1), \quad (2)$$

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}, \quad (3)$$

$$\alpha(x, y) = \tan^{-1}\left(\frac{G_y(x, y)}{G_x(x, y)}\right). \quad (4)$$

式中:  $G_x(x, y)$ ,  $G_y(x, y)$ ,  $H(x, y)$ ,  $G(x, y)$ ,  $\alpha(x, y)$  分别为图像中  $x, y$  处的水平方向梯度、垂直方向梯度、像素值、梯度大小及方向. 本文分别采用  $[1, 0, -1]^T$ ,  $[-1, 0, 1]^T$  梯度算子与原图像进行卷积求得垂直方向与水平方向梯度, 最终求得梯度的大小与方向.

3) 构建细胞单元梯度方向直方图.

本文中行人检测样本取自 Navneet Dalal, Bill Triggs 建立的 INRIA 数据库, 库中图片像素为  $64 \times 128$ . 以  $8 \times 8$  像素作为一个 cell 单元, 对单元内每个像素梯度进行统计并投影到 9 个 bin 中. 将  $2 \times 2$  个 cell 作为一个 block, 扫描步长为单个 cell, 依次对图像进行遍历, 则共有 105 块 block, 每个 block 中有  $2 \times 2 \times 9 = 36$  个特征描述子, 最终整个图片包括  $36 \times 105 = 3780$  个 HOG 特征描述子.

4) 根据式(5)对 HOG 特征描述子进行  $L_2$  范数归一化后用作后续分类识别的特征向量:

$$v \rightarrow v / \sqrt{\|v\|_2^2 + \epsilon^2}. \quad (5)$$

式中:  $v$  为未进行归一化的描述子;  $\|v\|_2$  为  $v$  的 2 阶范数;  $\epsilon$  为常数.

### 2.2 使用支持向量机(SVM)检测行人具体位置

在机器学习中, SVM<sup>[11-12]</sup> 被广泛应用于训练感知器与统计分析及分类中. 该算法可将低维向量投影到高维空间中, 并在此空间中形成一个最大分离超平面, 使其距离平面两端的数据间隔最大, 如图 3 所示. 最大分离超平面如式(6)表示:

$$y = \omega^T x + b. \quad (6)$$

两平面间的几何间隔如式(7)表示:

$$\eta = \frac{\omega^T x + b}{\|\omega\|}. \quad (7)$$

超平面目标函数转化为在  $y_i(\omega^T x + b)$  约束条件下求  $\frac{1}{\|\omega\|}$  的最大值,  $x$  为平面上一点,  $\omega$  为垂直于超平面的向量;  $\|\omega\|$  为  $\omega$  的 2-范数;  $b$  为位移因子;  $y$  为对应的类别因子;  $i$  为样本序列号, 两条实线间距是  $\frac{2}{\|\omega\|}$ , 进而转化为求  $\frac{1}{2} \|\omega\|^2$  的最小值. 线性约束条件作用于二次目标函数可转化为更易求解的拉格朗日对偶问题, 如式(8)表示:

$$\xi = \frac{1}{2} \|\omega\|^2 + \sum_{i=1}^n \lambda_i (y_i (\omega^T x + b) - 1). \quad (8)$$

式中:  $\lambda$  为拉格朗日乘子, 问题变成求解目标函数  $\mathcal{J}(\omega) = \max_{x \geq 0} \xi(\lambda)$  的最小值, 通过拉格朗日计算可求得  $\omega$  及  $b$ .

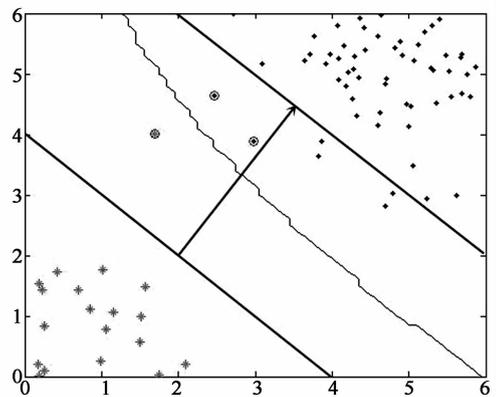


图 3 SVM 示意图  
Fig. 3 Diagram of SVM

为训练分类器, 本文选择包含不同姿势及形态的行人图像作为正样本, 不包含行人的任意背景图片作为负样本. 本文选择 3 000 张正样本(如图 4(a)所示), 2 000 张负样本(如图 4(b)所示)并将其标准

化为  $64 \times 128$  像素作为训练样本.



(a) 正样本



(b) 负样本

图 4 部仇正负样本图例

Fig. 4 Sub-sample of positive and negative images

用初次训练好的分类器检测不包含人体的图像,有时会得到错误的目标区域<sup>[13-14]</sup>,将这些错误区域(Hard Example)归到负样本中,再次进行训练,迭代多次,可明显改善分类器效果.最终训练好的分类器可将行人区域很好地划分出来,效果图如图 5(a),(b)所示.



(a) 目标行人 a



(b) 目标行人 b

图 5 检测到的目标行人 a,b  
Fig. 5 The detected pedestrian a,b

### 3 利用改进的 TLD 对行人进行跟踪

在初始帧利用 SVM 分类器得到待跟踪行人位置后,将位置信息传递给 TLD 中的 bounding box 以对该算法进行目标区域的初始化. TLD(Tracking-Learning-Detection)是一种对特定目标进行自适应检测与在线学习并实现实时跟踪的算法<sup>[15]</sup>,该算法将目标跟踪任务分成 3 部分:跟踪模块、学习模块和检测模块.3 部分协同组合,并以并行方式运行发挥各自优势<sup>[16]</sup>,具有较高的可靠性和鲁棒性.

1) 基于中值光流跟踪法的跟踪模块,选定  $t$  时刻图像帧中的跟踪点  $x_1(t), x_2(t) \cdots x_n(t)$ , 利用光流法求得  $t+1$  时刻跟踪点位置  $x_1(t+1), x_2(t+1) \cdots x_n(t+1)$ , 再反向追踪  $t$  时刻跟踪点预测位置  $\hat{x}_1(t), \hat{x}_2(t) \cdots \hat{x}_n(t)$ ,  $\hat{x}_i$  与  $x_i(t)$  间的欧氏距离定义为 FB 误差,FB 误差最小的一半跟踪点作为下一时刻跟踪点,依次进行下去,得到动态跟踪的特征点.

2) 检测模块由图像元方差分类器、在线随机森林分类器和 KNN(K-Nearest Neighbor)最近邻分类器级联组成.当前灰度值方差小于原始图像元灰度值方差 0.5 的图像元标记为负样本.随机森林分类器由若干个基本分类器构成,每个分类器都产生一个后验概率,后验概率的平均值大于经验阈值时则被认为含有目标图像.最近邻分类器 KNN 用存储的正负样本与当前区域进行相似度匹配,匹配相似度用  $S(p_i, p_j)$  度量,公式为式(9),式(10):

$$NCC(p_i, p_j) = \frac{\sum_{i=1}^n (p_i(x, y) \cdot p_j(x, y))}{(\sqrt{\sum p_i(x, y)^2} \cdot \sqrt{\sum p_j(x, y)^2})} \quad (9)$$

$$S(p_i, p_j) = 0.5(NCC(p_i, p_j) + 1). \quad (10)$$

式中:  $NCC(p_i, p_j)$  为归一化互相关系数;  $p_i, p_j$  分别为两个不同的图像元. 令  $M = \{q_1^+, q_2^+ \cdots q_{n_1}^+, q_1^-, q_2^- \cdots q_{n_2}^-\}$  为正负样本集合,  $q_i^+$  为目标行人,  $q_i^-$  为背景图像元, 正样本最近邻相似度为  $S^+(p, M) = \max_{p_i^+ \in M} S(p, p_i^+)$ , 负样本最近邻相似度为  $S^-(p, M) = \max_{p_i^- \in M} S(p, p_i^-)$ , 相对相似度为  $S^r(p, M) = \frac{S^+(p, M)}{S^+(p, M) + S^-(p, M)}$ ,  $S^r(p, M)$  值越大则代表图像元相似度越高, 当  $S^r(p, M)$  大于更新阈值时, 则认为此图像元存在行人目标.

检测模块默认每一帧图像互不关联<sup>[17]</sup>, 用滑动的扫描窗口对当前帧全局搜索以确定目标位置, 由于相邻帧间存在相同的目标区域, 令前一帧行人目标图像  $img1$  为参考图像, 当前帧  $img2$  为待配准图像, 对  $img2$  进行空间变换和灰度变换后得到的图像为  $img2^*$ ,  $img1$  和  $img2^*$  间的相似度函数为  $\Theta(\theta) = \Gamma(img2^*, img1)$ , 式中  $\Gamma()$  为两图像间的相似性度量, 检测出  $img1$  和  $img2$  的 Harris 角点, 以角点位置为坐标中心, 将与其相邻的窗口  $w$  内的像素点归一化后的像素值作为该点特征点匹配向量, 将特征点向量匹配后  $\Theta(\theta)$  小于阈值的点作为图像配准点, 然后通过分析配准点周围 8 邻域相似像素点数目来剔除误配点. 选取图像标准差作为 8 邻域相似像素检测阈值  $t$  以实现不同图像自适应阈值选取, 计算配准点像素与 8 邻域像素的绝对值. 用函数

$$Num(i, j) = \sum_{(-1 \leq x \leq 1, -1 \leq y \leq 1, \text{且 } x, y \neq 0)} R(i+x, j+y). \quad (11)$$

统计在配准点的 8 邻域范围内与之相似的像素点的个数, 其中

$$R(i+x, j+y) = \begin{cases} 1, & \text{绝对值} \leq \text{阈值 } t; \\ 0, & \text{绝对值} > \text{阈值 } t. \end{cases} \quad (12)$$

当  $Num(i, j) = 8$  时, 8 邻域像素点与目标像素点相似, 故该像素点为区域内部点, 应将此特征点剔除; 当  $Num(i, j) = 0$  时, 表示 8 邻域范围没有与特征点相似的点, 故此特征点为零散分布的孤立点或噪声点, 也应予以剔除. 但在灰度差异较大的图像中, 可能会把局部灰度变化较小的少量角点也筛选掉. 故在此基础上将局部角点响应函数最大值的像素点作为初始角点, 计算该角点一定半径内各角点对此点的权重, 并以各角点响应函数作为各自权值, 采用最小二乘法进一步修正角点坐标. 如式(13)表示:

$$J = \sum_{j=1}^n p_j s_j. \quad (13)$$

式中:  $s_j = \|\hat{x} - x_j\|^2 + \|\hat{y} - y_j\|^2$ .

令  $R_j$  为此初始角点  $j$  的角点响应函数,  $p_j$  为此角点在半径  $r$  内角点集合的权重,  $n$  为角点集合中像素个数. 由于角点响应函数值越大则其权重越大, 则  $p_j$  可由式(14)表示:

$$p_j = \frac{R_j}{\sum_{i=1}^n R_i}. \quad (14)$$

其中角点响应函数如式(15)表示:

$$R = \det(M) - k \text{tr}^2(M). \quad (15)$$

式中:  $k$  为常数;  $M = \sum G \begin{bmatrix} I_x^2 & I_{xy} \\ I_{xy} & I_y^2 \end{bmatrix}$ ,  $G$  为抗噪高斯窗口函数;  $I_x, I_y$  和  $I_{xy}$  分别为像素点在  $x, y$  方向上的偏导及二阶混合偏导;  $n$  为邻域内像素点个数. 令  $\frac{\partial J}{\partial \hat{x}}, \frac{\partial J}{\partial \hat{y}}$  分别为 0, 可得  $\hat{x}, \hat{y}$  的值, 最终确定精确的角点坐标  $(\hat{x}, \hat{y})$ , 经过多次筛选, 剔除非角点与伪角点. 将剔除误配点后的图像反馈到检测器中, 用目标特征不断调整跟踪器. 图 6(a), 图 7(a) 为两组移动行人测试集相邻两帧图片, 图 6(b), 图 7(b) 为这两组图片剔除误配点后的配准结果.



(a) 待配准图像



(b) 剔除误配点后的图像

图 6 剔除误配点前后比较图  
Fig. 6 The registration and rejected registration points of image



(a)待配准图像



(b)剔除误配点后的图像

图 7 剔除误配点前后比较图

Fig. 7 The registration and rejected registration points of image

3)PN 学习过程中<sup>[18]</sup>,先给定一个视频和扫描框,同时用分类器标定出视频区域中的目标和背景,用跟踪器来提供正的训练样本,检测器提供负的训练样本.跟踪器跟踪目标后,靠近目标窗口的扫描窗口通过 P 约束来更新检测器,若扫描窗口中不存在目标物体,则 N 约束更新检测器. P 约束利用时间连续性,根据前一帧目标出现的位置预测目标轨迹,估计目标在当前帧的位置,若检测器将此位置定义为负样本,则 P 约束生成正样本;N 约束分析出当前帧目标可能出现的最优位置,与此位置重叠度低的区域标记为负样本,继续更新跟踪器. PN 约束通过在线处理视频帧逐步提高检测器正确率,相互补偿来确保学习模块的稳定性与可靠性,利用误检来提高学习性能.

## 4 实验结果与分析

本实验编程平台:VS2010 与 opencv2.4.3 开发库;MATLAB R2013a.

微机环境配置:CPU Intel(R)Core(TM)i3-4150 3.50 GHz,内存为 4 G.

### 4.1 实验测试集说明

为了更好地验证本算法检测跟踪目标的鲁棒性和准确性,将实验结果与文献[19]中高斯混合模型

改进的 meanshift 跟踪算法<sup>[19]</sup>,文献[20]中 kalman 滤波和模板匹配跟踪算法<sup>[20]</sup>及基本 TLD 算法分别在 4 种不同的数据集中进行验证比较,并分析各算法的准确性、实时性及鲁棒性.其中 shooting 和 football 序列来自网上公开测试集 VTD\_data\_images,pedestrian1 及 pedestrian2 数据集为实时采集.测试集序列内容如表 1 所示.

表 1 测试集序列及其描述

Tab. 1 The description of the testing sequences

测试序列	部分截图	测试集描述
Shooting		目标在运球及投篮过程中姿势发生了较大变化
Football		目标绕竖直方向旋转(倾斜),腿部姿态也会有较大变化
Pedestrian1		目标迅速移动,并伴随侧身、转身、姿态变化及被遮挡等情况
Pedestrian2		目标侧身、转身、消失然后重新回到视频帧中以及相似目标干扰等情况

## 4.2 性能分析

### 4.2.1 算法准确性比较

将本文算法、改进的 meanshift 算法、kalman 滤波和模板匹配算法、基本 TLD 算法在不同的测试集中进行实验,对比跟踪效果.用式(16)来表示跟踪成功率:

$$\text{score} = \frac{\text{area}(R_T \cap R_G)}{\text{area}(R_T R_G)}. \quad (16)$$

式中: $R_T$  表示跟踪算法跟踪到的目标区域, $R_G$  表示目标实际所在区域.若  $\text{score} > 0.5$ ,则认为该算法正确跟踪此帧.测试集实验结果如表 2 所示.由表 2 可知,在未遮挡情况下,各算法均可以获得较准确的跟踪结果,但在有遮挡时,本文算法正确率均高于其他各算法.

本文算法、文献[19]改进的 mean-shift 算法、文献[20]kalman 滤波和模板匹配算法及基本 TLD 算法在 pedestrian1 测试集中跟踪部分结果(分别选取测试集第 30 帧、第 53 帧、第 174 帧、第 350 帧)如图 8 所示.从图 8 可知,当目标未被遮挡时,3 种算法都可以准确跟踪行人,目标被遮挡后,基本 TLD 算法、文献[19]与文献[20]中的跟踪算法均出现目标跟丢甚至误检现象,而本文算法可始终追踪到行人,对目标定位跟踪表现出良好的抗干扰性和准确性.

表 2 测试集测试结果  
Tab. 2 The results of testing sequences

测试序列	总帧数	目标存在帧数	本算法正确处理帧数/正确率	改进的 meanshift 算法正确处理帧数/正确率	kalman 滤波和模板匹配算法正确处理帧数/正确率	TLD 算法正确处理帧数/正确率
Shooting	434	434	408 / 94.01%	411 / <b>94.70%</b>	387 / 89.17%	403/92.86%
Football	800	800	732 / <b>91.50%</b>	726 / 90.75%	697 / 87.13%	729/91.13%
Pedestrian1	397	389	352 / <b>90.49%</b>	344/ 88.43%	326 /83.80%	347/89.20%
Pedestrian2	2 000	1 863	1 564 / <b>83.95%</b>	1 421 / 76.27%	1 298 /69.67%	1 511/81.10%

注:正确率最高者用黑体加粗,次高者用斜体标出



图 8 各算法对 pedestrian1 序列跟踪结果截图  
Fig. 8 The tracking results of pedestrian1 sequence using each algorithm

4.2.2 算法实时性比较

本文采用平均帧率(average frame per second)作为算法实时性分析的评价指标,各算法在测试集中的平均帧率如表 3 所示.由表 3 可知,基本 TLD 算法与本文算法在测试集的处理上速度高于其他两种算法,而由跟踪准确性分析可知,本文算法比基本 TLD 算法获得了更准确的跟踪结果,表现出较强的跟踪性能.

表 3 测试集平均帧率测试结果  
Tab. 3 The average FPS results of each sequences

测试序列	平均帧率/(帧·s <sup>-1</sup> )			
	本文算法	改进的 meanshift 算法	kalman 滤波和模板匹配算法	TLD 算法
Shooting	<b>22.5</b>	18.1	16.2	<i>21.3</i>
Football	<b>21.1</b>	17.6	14.1	<i>19.2</i>
Pedestrian1	<i>15.7</i>	13.1	10.9	<i>16.6</i>
Pedestrian2	<b>19.5</b>	18.1	15.8	<i>18.9</i>

注:平均帧率最高者用黑体加粗,次高者用斜体标出

4.2.3 算法鲁棒性分析

在 pedestrian2 序列中(如图 9 所示),行人检测分类器确定目标位置后,当目标姿势变化较大时,学习模块通过在线学习不断更新跟踪模块和检测模块样本,确定新的正样本.本算法在目标侧身,转身,消失然后重现时均可以准确跟踪行人,当相似目标出现时则不返回跟踪框,表现出良好的抗干扰性.



图 9 Pedestrian2 序列部分截图  
Fig. 9 Snapshots of pedestrian2 sequence

## 5 结 论

检测跟踪一直是计算机视觉研究的热点和难点,而 TLD 算法通过在线学习与检测,根据产生地正负样本不断更新跟踪结果,具有较好的实时性与准确性.本文在其基础上引入 SVM 行人检测分类器,确定并读取目标区域位置信息,实现行人的自动识别,同时联合图像配准技术,将误配点排除,跟踪精度得到进一步提高.最后进行实验验证,通过与当前两种主流跟踪算法及基本 TLD 算法在不同测试集上进行试验比较,可知本算法在长时间跟踪过程中能够不断适应目标形变带来的干扰,表现出更好的跟踪效果与鲁棒性能.

## 参考文献

[1] CQMANICIU D, RAMESH V, MEER P. Real-time tracking of nonrigid objects using mean shift[C]//Proceedings in Computer Vision and Pattern Recognition, Hilton Head Island, SC; IEEE Conference,2000;2,142-149.

[2] ANTONINI G, MARTINEZ S V, BIERLAIRE M, *et al.* Behavioral priors for detection and tracking of pedestrians in video sequences[J]. International Journal on Computer Vision, 2006,69(2):159-180.

[3] 季玉龙,董天罡,梁海军,等.基于粒子滤波的多区域辨识性建模跟踪算法[J].四川大学学报:工程科学版,2013,45(1):79-83.  
JI Yu-long, DONG Tian-gang, LIANG Hai-jun, *et al.* Robust particle tracking via multi-region based discriminative modeling[J]. Journal of Sichuan University: Engineering Science Edition, 2013,45(1):79-83. (In Chinese)

[4] KALAL Z, MATAS J, MIKOLAJCZYK K. Online learning of robust object detectors during unstable tracking[C]//Proceedings of 12th International Conference on Computer vision workshops. New York: IEEE Xplore,2009:1417-1424.

[5] KALAL Z, MATAS J, MIKOLAJCZYK K. Pn learning: bootstrapping binary classifiers by structural constraints[C]// Proceedings of Conference on Computer Vision and Pattern Recognition. New York: IEEE Press,2010:49-56.

[6] ALALZ K, MIKOLAJCZYK K, MATAS J. Forward-backward error: automatic detection of tracking failures[C]//Proceedings of International Conference on Pattern Recognition. New York: IEEE Press,2010:23-26.

[7] KALAL Z, MIKOLAJCZYK K, M ATAS J. Tracking-learning-detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2012,34(7):1409-1422.

[8] 魏国剑,侯志强,李武,等.融合光流检测与模板匹配的目标跟踪算法[J].计算机应用研究,2014,31(11):3498-3501.  
WEI Guo-jian, HOU Zhi-qiang, LI Wu, *et al.* Object tracking algorithm fused with optical flow detection and template matching[J]. Application Research of Computers, 2014,31(11):3498-3501. (In Chinese)

[9] DALAL N, TRIGGS B. Histograms of oriented gradients for

human detection[C]//Computer Vision and Pattern Recognition, International Conference. Beijing;2005;886-893.

[10] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//Proceedings of Computer Vision and Pattern Recognition, IEEE Computer Society Conference on. San:CA;IEEE,2005,1:886-893.

[11] 曹文明,王耀南,文益民.输电线路除冰机器人基于小波矩及 SVM 的障碍物识别研究[J].湖南大学学报:自然科学版,2012,39(9)33-38.  
CAO Wen-ming, WANG Yao-nan, WEN Yi-ming. Research on obstacle recognition based on wavelet moments and SVM for deicing robot on high voltage transmission line[J]. Journal of Hunan University: Natural Sciences,2012,39(9)33-38. (In Chinese)

[12] 于海滨,刘济林.应用于公交客流统计的机器视觉方法[J].中国图像图形学报,2008,13(4):716-722.  
YU Hai-bin, LIU Ji-lin. A vision-based method to estimate passenger flow in bus[J]. Journal of Image and Graphics, 2008,13(4):716-722. (In Chinese)

[13] HUANG Dong-ze, CAI Zhi-hao, XIANG HE Xiang, *et al.* A SVM embedded particle filter for multi-object detection and tracking [C]//Proceedings of 2014 IEEE Chinese Guidance, Navigation and Control Conference. Yantai;IEEE Press,2014:2094-2099.

[14] 方卫宁.基于 RBF 神经网络的复杂场景人群目标的识别[J].北京交通大学学报,2009,33(4):29-33.  
FANG Wei-ning. Recognizing the passenger number in complex scenes by RBF neural network[J]. Journal of Beijing Jiaotong University,2009,33(4):29-33. (In Chinese)

[15] ANDRILUKA M, ROTH S, SCHIELE B. People-tracking-by-detection and people-detection-by-tracking [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press,2008:1-8.

[16] COLLINS R T, LIUY, LEORDEANU M. Online selection of discriminative tracking features [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005,27(10):1631-1643.

[17] YAMAZAWA K, YOKOYA N. Detecting moving objects from omnidirectional dynamic images based on adaptive background subtraction[C]// Proceedings of International Conference on Image processing. Institute of Electrical and Electronics Engineers Computer Society, Barcelona, Spain;2003,953-956.

[18] FENARI V, TUYTELAARS T, VAN G L. Real-time affine region tracking and coplanar grouping[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai, HI; Institute of Electrical and Electronics Engineers Computer Society, 2001;226-233.

[19] QUAST K, KAUP A. Shape adaptive mean shift object tracking using Gaussian mixture models[M]. Lecture Notes in Electrical Engineering. Heidelberg: Springer, 2013, 158: 107-122.

[20] 陈金令,苗东,康博,等.基于 Kalman 滤波和模板匹配的目标跟踪技术研究[J].光学与光电技术,2014,12(6):9-12.  
CHEN Jin-ling, MIAO Dong, KANG Bo, *et al.* Research of target tracking based on kalman filtering and template matching[J]. Optics & Optoelectronic Technology, 2014,12(6):9-12. (In Chinese)