

文章编号 1004-924X(2013)12-3191-07

梯度特征稀疏表示目标跟踪

孙晓燕^{1,2}, 常发亮^{1*}

- (1. 山东大学 控制科学与工程学院, 山东 济南 250061;
2. 山东建筑大学 计算机科学与技术学院, 山东 济南 250101)

摘要:传统的压缩感知目标跟踪算法在光照变化剧烈、目标与背景存在一定相似性的情况下容易产生跟踪偏差,故本文提出了一种基于梯度方向直方图特征进行压缩感知跟踪方法。该方法用梯度方向直方图特征替换原来的广义类 Haar 特征进行压缩感知跟踪。首先,将梯度方向直方图作为原始特征,并利用压缩感知理论得到稀疏表示的特征子空间;然后,在后续帧中用朴素贝叶斯分类器进行目标位置的搜索;最后,对分类器进行在线更新。由于梯度特征能更稳定地表示目标,所以这种跟踪方法具有更好的鲁棒性;另外在计算时采用了积分直方图技术,有效克服了计算量大的问题。对不同视频的实验结果表明,该方法在实验环境 Intel Core2 2.93 GHz, matlab R2010a, 图像大小 320×240 下,跟踪速率可达到 10 frame/s。在目标姿态、环境光照变化剧烈,背景中存在与目标有一定相似性的物体等情况下跟踪准确。

关键词:目标跟踪;梯度方向直方图;稀疏表示;压缩感知

中图分类号:TP391 **文献标识码:**A **doi:**10.3788/OPE.20132112.3191

Object tracking based on sparse representation of gradient feature

SUN Xiao-yan^{1,2}, CHANG Fa-liang^{1*}

(1. School of Control Science and Engineering, Shandong University, Jinan 250013, China;

2. School of Computer Science and Technology, Shandong Jianzhu University, Jinan 250101, China)

* Corresponding author, E-mail: flchang@sdu.edu.cn

Abstract: As traditional compressive sensing tracking algorithm will produce tracking errors in circumstances when illumination has dramatic change or there exists a object similar to the target in background, this paper proposes a sparse representation object tracking algorithm by taking the histogram of gradient feature to replace the generalized Haar feature. The algorithm uses the histogram of gradient feature as an original feature firstly, and gets the sparse representation of object feature subspace by using compressive sensing theory. In the subsequent frames, the naive Bayes classifier is used to search the target location and the classifier is online updated finally. As the histogram of gradient feature can represent the target more stably, this algorithm is more robust than original compressive tracking algorithm. Furthermore, the integral histogram is adapted to effectively reduce computational load when the gradient feature is computed. Experiments on different videos show that the tracking

收稿日期:2013-08-06; **修订日期:**2013-08-21.

基金项目:国家自然科学基金资助项目(No. 61273277);教育部留学回国人员科研启动基金资助项目(No. 20101174);山东省自然科学基金资助项目(No. ZR2011FM032);济南市高校自主创新计划资助项目(No. 201004002);高等学校博士学科点专项科研基金资助项目(No. 2013031110038)

algorithm can reach the tracking rate of 10 frames per second in an experimental environment of Intel Core2 2.93 GHz, matlab R2010a, image size 320×240 , and it achieves stable tracking in some special conditions as mentioned above.

Key words: object tracking; histogram of gradient feature; sparse representation; compressive sensing

1 引言

目标跟踪是计算机视觉领域一项重要的任务,有着非常广泛的应用^[1]。理想的目标跟踪算法一定要在实时性与鲁棒性两方面达到平衡^[2]。早期的目标跟踪算法主要有:模板匹配法、光流法、均值漂移^[3]和粒子滤波^[4]。目前实时跟踪的热点方法是将机器学习的各种在线学习方法应用到目标模型的表示与更新中,以适应目标的内在变化和环境的在外变化,这一类方法也被称为基于检测的跟踪算法(tracking by detection)^[5]。Avidan 的组合跟踪器(ensemble tracker)^[6]使用 Adboost 方法将一组能够动态更新的弱分类器组合成强分类器,进行像素是目标还是背景的分类判断。David 等^[7]用增量学习方式得到目标表观的低维子空间表示,该方法在复杂情况下的目标跟踪取得了不错的效果。Grabber 和 Bischof 提出在线提升(online boosting)算法^[8]从特征池中选择弱分类器组合成强分类器,后来又进一步提出一种半监督学习的跟踪方法以克服模型更新带来的漂移问题^[9]。颜佳等将在线提升算法进行改进以解决目标被严重遮挡时出现的漂移甚至丢失问题^[10]。Babenko 等提出一种在线多示例学习(MIL)方法以解决正样本的不确定性带来的漂移问题^[11]。Saffari 等^[12]提出一种在线随机森林学习方法并将其用于目标跟踪取得了较好的实时跟踪效果。Zhang 等提出一种简单有效的压缩感知跟踪方法^[13],通过随机投影从多尺度的图像特征空间中得出原始目标的稀疏表示表观模型,并用朴素贝叶斯分类器进行目标与背景的分类。该方法使用特征简单,且用特别稀疏的投影矩阵压缩特征空间来表示目标,速度可达到 40 frame/s。但是也正因为使用了简单的广义类 Haar 特征,在光照变化剧烈,目标与背景存在一定相似性的情况下容易跟偏^[14]。

为了解决压缩感知目标跟踪中由于特征简单而引起的在光照变化剧烈、目标与背景存在一定

相似性的情况下容易跟偏的问题,本文将原来的广义类 Haar 特征替换为目标的梯度方向直方图特征构建原始高维特征空间,然后采用极度稀疏的投影矩阵压缩特征空间建立目标的稀疏表观模型,最后用朴素贝叶斯分类器找到目标位置并进行在线更新。梯度方向直方图特征虽然计算量比灰度特征大,但对光照变化和局部区域的微小形变比较鲁棒,采用投影矩阵对原始高维特征空间压缩可以大大降低其计算量,因此将灰度特征替换为梯度特征可使得跟踪在实时性与鲁棒性之间达到平衡。实验结果表明,即使是在目标纹理或光照变化剧烈,目标与背景存在一定相似性等情况下,基于梯度方向直方图特征的压缩感知跟踪也能够准确实时地跟踪目标。

2 压缩感知跟踪

压缩感知理论是 Donoho 于 2006 年提出的^[15],只要信号是稀疏的或在某个变换域是稀疏的,就可以用一个测量矩阵将原始高维信号投影到一个低维空间上,并且可以从这些少量的投影中以高概率重构出原始信号。也就是说,通过一个满足约束等距性(RIP)条件的稀疏测量矩阵对原图像特征空间做投影,得到一个低维压缩子空间,该低维压缩子空间可以很好地保留高维图像特征空间的信息。文献[13]中根据这一理论提出一种基于压缩感知的跟踪算法(Compressive Tracking,简称 CT),该算法利用压缩感知从高维的特征空间中提取低维特征:

$$\mathbf{V} = \mathbf{R}\mathbf{X}, \quad (1)$$

其中: $\mathbf{X} \in \mathbf{R}^n$ 为高维原始信号,在文献[13]中为各种不同尺度的矩形滤波器与原图像卷积后的结果组合。 $\mathbf{V} \in \mathbf{R}^m$ ($m \ll n$) 为压缩后的低维特征, $\mathbf{R} \in \mathbf{R}^{m \times n}$ 为满足 RIP 条件的投影矩阵,其形式为

$$r_{i,j} = \sqrt{s} \times \begin{cases} +1, & \text{概率为 } 1/2s \\ 0, & \text{概率为 } 1/s \\ -1, & \text{概率为 } 1/2s \end{cases}, \quad (2)$$

其中: s 是在 2~4 间的随机取值。实际特征提取

为^[14]:

$$v_i = \sum_{k=1}^{NR} \frac{p_{i,k} \cdot \text{rect}_{i,k}}{\sqrt{NR}}, \quad (3)$$

式中: $p_{i,k}$ 值在 1 与 -1 等概率随机选取。实际就是在候选区域内随机选取 $NR(NR=2\sim 4)$ 个不同大小的矩形图像块 $\text{rect}_{i,k}$, 求这几个小块的灰度加权(若系数 $p_{i,k}$ 为 -1, 则权重为负)。特征的压缩生成如图 1 所示, 其中矩阵 R 中黑色、灰色和白色小块分别代表系数为 -1、+1 和 0。

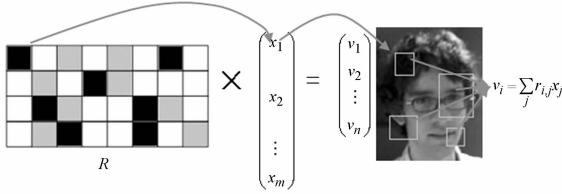


图 1 特征压缩生成示意图

Fig. 1 Diagram of feature generation by compression

后续的跟踪过程就是在后续图像帧中遍历上一次目标位置的邻近区域, 通过如下朴素贝叶斯分类器计算候选区域的分类分数, 并找到分类分数最大的候选区域作为目标位置:

$$H(v) = \log\left(\frac{\prod_{i=1}^m p(v_i | y=1)p(y=1)}{\prod_{i=1}^m p(v_i | y=0)p(y=0)}\right) = \sum_{i=1}^m \log\left(\frac{p(v_i | y=1)}{p(v_i | y=0)}\right), \quad (4)$$

$$H(*) = \max(H(v)), \quad (5)$$

其中: m 为降维后的特征维数, $p(v_i | y=1)$ 和 $p(v_i | y=0)$ 满足标准正态分布, 其均值和标准差分别为: μ^1, σ^1 和 μ^0, σ^0 。

在确定了最大的 $H(v)$ 为目标之后, 最后对相关参数 μ^1, σ^1 和 μ^0, σ^0 进行更新, 以适应目标外观的变化。

文献[13]中基于压缩感知的跟踪算法每个弱分类器对应提取的图像特征是图像的灰度信息, 在环境光照变化剧烈, 或者目标与背景存在一定相似性的情况下, 图像灰度信息并不可靠, 会带来跟踪的偏离(参见 4.1 部分)。

3 基于梯度的稀疏表示跟踪

基于图像灰度信息的压缩感知跟踪在一些自然条件下会带来跟踪的偏离, 故考虑将其替换为基于图像梯度信息的梯度方向直方图(HOG)。

HOG 特征是目标检测方面应用较多的特征。Dollar 等已经通过实验证实, 在行人检测效果上, HOG 特征是效果最好的一种特征^[16]。将原来简单的广义类 haar 特征, 改为更能准确表示目标的 HOG 特征, 可使得目标的表示更加稳定。

3.1 梯度方向直方图压缩特征空间生成

采用梯度方向直方图特征后, 原始高维信号 X 为各种不同尺度的矩形滤波器与原图像卷积后的结果提取梯度方向直方图特征的组合, 压缩后的低维特征 $V \in R^m$ 中每个特征就是在候选区域内随机选取 2~4 个不同大小的矩形图像块 $\text{rect}_{i,k}$, 求这几个小块 HOG 特征的加权和。具体的生成过程如下:

3.1.1 梯度的计算

候选区域内一点 (x, y) 的梯度计算为:

$$\begin{cases} G_x(x, y) = I(x+1, y) - I(x-1, y) \\ G_y(x, y) = I(x, y+1) - I(x, y-1) \end{cases}, \quad (6)$$

则该点的梯度幅值和梯度方向角度分别为:

$$\begin{cases} G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \\ \alpha(x, y) = \arctan^{-1}\left(\frac{G_y(x, y)}{G_x(x, y)}\right) \end{cases}. \quad (7)$$

3.1.2 构建每一子块的梯度方向直方图

随机选取 2~4 个不同大小的矩形图像子块 $\text{rect}_{i,k}$, 统计每个矩形图像块的梯度方向直方图 $H_{i,k}$ 。将梯度方向分为 bin 块, 块内每个像素都为某个方向的直方图通道投票, 投票采取加权投票方式, 权值根据该像素点的梯度幅度计算(计算时可采用积分直方图^[17]快速得到)。由于每个矩形图像块 $\text{rect}_{i,k}$ 的大小不一, 直接统计得到的直方图需按照 $\text{rect}_{i,k}$ 内像素点个数进行归一化。

3.1.3 最终特征表示

按下式求 2~4 个小块的特征加权和:

$$v_i^b = \sum_{k=1}^{NR} \frac{p_{i,k} \cdot h_{i,k}^b}{\sqrt{NR}}, b = 1, \dots, bin. \quad (8)$$

图 2 显示为梯度方向直方图特征提取的示意图。

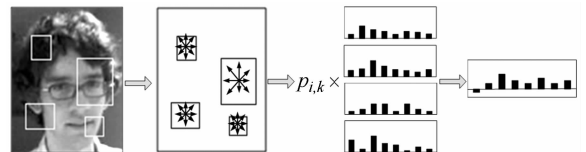


图 2 梯度方向直方图特征提取示意图

Fig. 2 Feature extraction of histogram of gradient orientation

3.2 基于梯度特征的稀疏表示跟踪

采用梯度方向直方图特征的稀疏表示跟踪算法步骤如下:

1) 初始化: 初始帧选定目标(手动选取), 利用投影矩阵提取目标样本和背景样本的特征表示;

a. 手动选取目标;

b. 采集目标样本和背景样本。目标样本 $\|l(x) - l_0\| < \alpha$ 在内采集, 其中 l_0 为初始选定的目标位置, 背景样本在目标周围 $\beta < \|l(x) - l_0\| < \gamma$ 范围内采集。

c. 计算 $\|l(x) - l_0\| < \gamma$ 图像区域的梯度方向积分直方图;

d. 利用积分直方图, 快速计算目标样本和背景样本的特征。并利用投影矩阵 R , 按照式(8)分别计算目标和背景稀疏表示特征, 代入式(9)更新参数 μ^1, σ^1 和 μ^0, σ^0 。

$$\begin{cases} \mu_i^1 = \lambda \mu_i^0 + (1 - \lambda) \mu_i^1 \\ \sigma_i^1 = \sqrt{\lambda (\sigma_i^0)^2 + (1 - \lambda) (\sigma_i^1)^2 + \lambda (1 - \lambda) (\mu_i^1 - \mu_i^0)^2} \end{cases}, \quad (9)$$

式中: λ 为更新率, λ 越小表示更新速度越快, 原来的特征保留越少。

2) 计算: 读取新一帧图像, 在上一帧目标位置 l_{t-1} 附近 $\|l(x) - l_{t-1}\| < \gamma$ 范围内采集样本, 并计算所有样本涉及图像区域的 HOG 积分直方图。通过积分直方图和式(8)计算候选区域内每个样本的特征, 代入(4)中计算 $H(v)$ 并按式(5)选取最大的 $H(v)$ 为该帧中的目标位置;

3) 更新: 在目标位置附近采集目标样本和背景样本, 分别计算目标样本和背景样本的特征(计算时利用步骤 2)中已经计算出的 HOG 积分直方图以降低计算量), 利用式(9)进行参数的更新。

4) 跳转到步骤 2) 循环处理下一帧。

4 实验结果及分析

4.1 与 CT 的跟踪结果比较

实验采用文献[7]提供的 david_indoor 视频序列进行基于梯度方向直方图的稀疏表示跟踪(HOGCT)与 CT 跟踪方法^[13]的比较。该视频序列一开始光照特别弱, 后来越来越强, 中间还存在有与目标相似性较大的背景物体干扰。实验过程中参数的取值如下: $\alpha=4$, 取得的目标样本个数为

35 个, $\beta=8, \gamma=20$, 在 $\beta < \|l(x) - l_0\| < \gamma$ 范围内随机取 50 个背景样本。更新率 $\lambda=0.85$ 时的实验结果如图 3 所示, 通过计算两者跟踪结果左上角坐标与人工标定真理图的欧氏距离, 得到两者的误差曲线如图 4 所示。从图 3 和图 4 可以看出, 在跟踪一开始的第 4 帧中, HOGCT 与 CT 跟踪到的目标位置基本差不多。随着跟踪的进行, 在 196 帧和 383 帧时, 由于光照的变化越来越大, 人脸部的灰度特征变化较大, CT 出现了偏移 to 灰度较低区域的误差。而由于 HOGCT 是基于目标梯度特征表示的, 仍能够准确跟踪目标。在 468 帧时, 背景中存在与原始目标有一定灰度分布相似性的物体, CT 直接从目标偏移 to 该物体, 出现较大误差, 而 HOGCT 的跟踪误差较小。531 帧目标与相似性物体重叠, CT 的跟踪又回到目标上。后来随着目标特征的更新, CT 又偏向于下部(578 帧), 而 HOGCT 一直能够较好的跟踪目标。虽然 CT 跟踪中可调小更新率 λ , 使得更新速率变快, 理论上能使跟踪适应目标外观变化的能力更强, 但势必对跟踪算法的稳定性有很大影响。

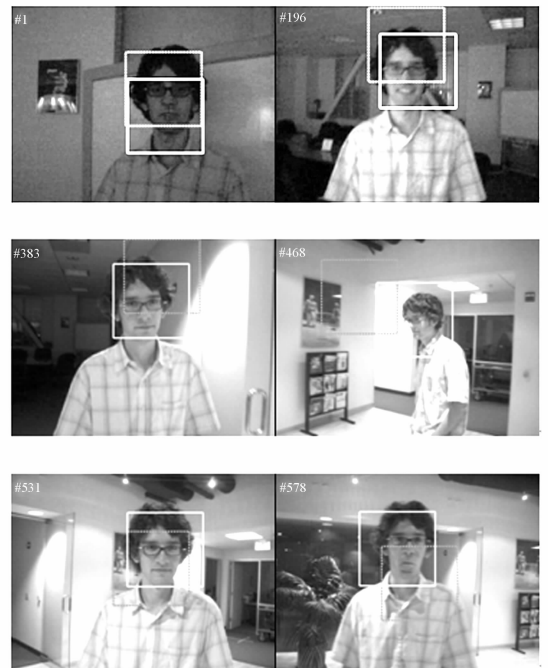


图 3 HOGCT 和 CT 的跟踪比较, 亮实线框为 HOGCT 的跟踪结果, 暗虚线框为 CT 的跟踪结果

Fig. 3 Comparison of HOGCT with CT (bright solid boxes for HOGCT, and dark dashed boxes for CT)

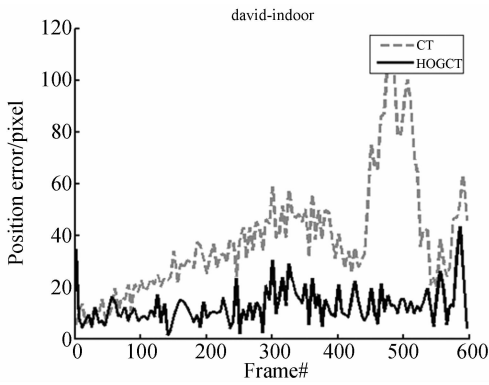


图 4 david_indoor 视频跟踪误差比较

Fig. 4 Error comparison of david_indoor sequence

4.2 与几种在线学习跟踪算法的比较

本实验选取了当前比较流行的几种在线学习跟踪方法,使用自己拍摄的行人视频序列与本文提出的 HOGCT 跟踪方法进行比较。视频中背景较复杂,行人目标也经历了从正侧面到背面的姿态变化。比较的方法包括:基于多示例学习的跟踪方法 MIL Tracker^[11], Online boosting 方法^[8] (OAB), Fragtrack^[18], CT 方法^[13]。实验过

程中这些方法均采用原作者提供的参数, HOGCT 中更新率 $\lambda=0.65$,其他参数取值与 4.1 中相同。跟踪结果比较如图 5 所示,误差曲线参见图 6。从中可以看出,随着目标姿态的变化,其他跟踪方法都出现了不同程度的偏移, HOGCT 跟踪比起其他方法具有较好的鲁棒性。

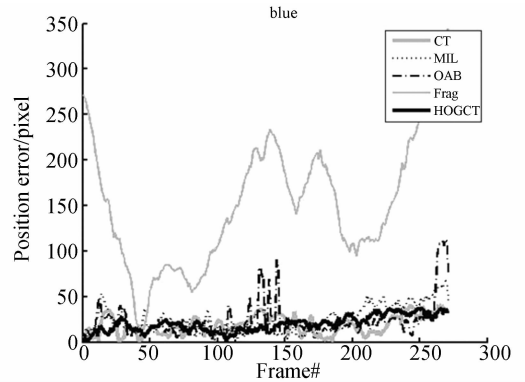


图 6 HOGCT 与其他跟踪方法误差比较

Fig. 6 Error comparison of HOGCT with other on-line learning methods

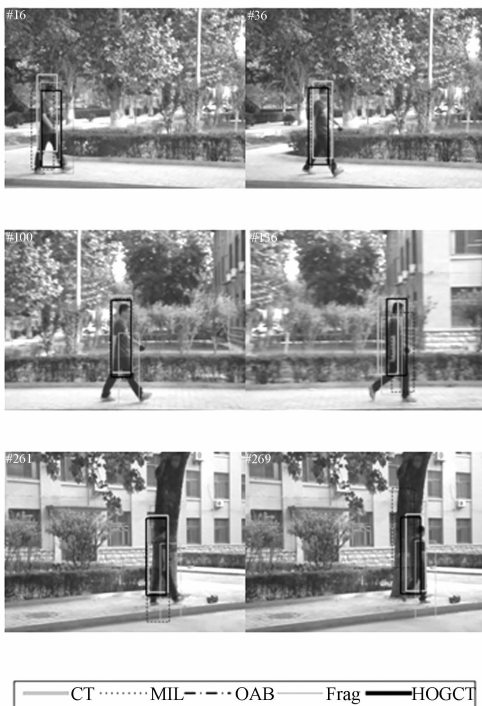


图 5 HOGCT 与其他在线学习跟踪方法比较

Fig. 5 Comparison of HOGCT with other online learning tracking methods

5 结 论

本文提出一种基于梯度方向直方图特征的压缩感知跟踪方法,实验结果表明,在实验环境 Intel Core2 2.93 GHz, matlab R2010a, 图像大小 320×240 下该方法的跟踪速率可达到 10 frame/s,并且在背景复杂,光照剧烈变化,背景中存在与目标有一定相似性物体等复杂情况下,跟踪效果较好。因为 HOG 特征的计算量比起原来 CT 中采用的广义类 Haar 特征大,至少是原来的 8 倍,因此基于 HOG 特征的压缩感知跟踪比起原来的基于广义类 Haar 特征的 CT 跟踪速度要慢一些。通过积分直方图技术并只计算采样窗口内的 HOG 特征直方图可降低部分计算量,如何提高该方法的速度是未来需要进一步研究的问题。同文献^[13]中的 CT 一样, HOGCT 暂时没有解决目标尺度变化问题,在这方面可以在搜索候选区域时将窗口大小分别缩放 10% 进行搜索^[19],但这样势必带来更多的计算量。因此,如何使跟踪能适应目标的尺度变化也是一个需要继续研究的问题。

参考文献:

- [1] YILMAZ A, JAVED O, SHAH M. Object tracking: a survey [J]. *ACM Computing Surveys*, 2006, 38(4): 1-45.
- [2] 杨戈, 刘宏. 视觉跟踪算法综述[J]. *智能系统学报*, 2010, 5(2): 95-105.
YANG G, LIU H. Survey of visual tracking algorithms [J]. *CAAI Transactions on Intelligent Systems*, 2010, 5(2): 95-105. (in Chinese)
- [3] 常发亮, 赵瑶, 陈振学, 等. 基于 Mean Shift 和自适应预测的非刚性目标跟踪算法[J]. *控制与决策*, 2009, 24(12): 1821-1825.
CHANG F L, ZHAO Y, CHEN Z X, *et al.*. Non-rigid object tracking algorithm based on Mean Shift and adaptive prediction [J]. *Control and Decision*, 2009, 24(12): 1921-1825. (in Chinese)
- [4] 王国良, 刘金国. 基于粒子滤波的多自由度运动目标跟踪[J]. *光学精密工程*, 2011, 19(4): 864-869.
WANG G L, LIU J G. Moving object tracking with multi-degree-of-freedom based on particle filters [J]. *Opt. Precision Eng.*, 2011, 19(4): 864-869. (in Chinese)
- [5] YANG H X, SHAO L, ZHENG F, *et al.*. Recent advances and trends in visual tracking: A review [J]. *Neurocomputing*, 2011, 74: 3823-3831.
- [6] AVIDAN S. Ensemble tracking [C]. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*, 2005: 261-271.
- [7] ROSS D, LIM J, LIN R S, *et al.*. Incremental Learning for Robust Visual Tracking [J]. *International Journal of Computer Vision*, 2008, 77: 125-141.
- [8] GRABNER H, BISCHOF H. On-line boosting and vision [C]. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*, 2006: 260-267.
- [9] GRABNER H, LEISTNER C, BISCHOF H. Semi-supervised on-line boosting for robust tracking [C]. *ECCV*, 2008: 234-247.
- [10] 颜佳, 吴敏渊. 遮挡环境下采用在线 boosting 的目标跟踪[J]. *光学精密工程*, 2012, 20(2): 439-446.
YAN J, WU M Y. On-line boosting based target tracking under occlusion [J]. *Opt. Precision Eng.*, 2012, 20(2): 439-446. (in Chinese)
- [11] BABENKO B, YANG M H, BELONGIE S. Robust object tracking with online multiple instances learning [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(8): 1619-1632.
- [12] SAFFARI A, LEISTNER C, SANTNER J, *et al.*. On-line random forests [C]. *IEEE 12th International Conference on Computer Vision*, 2009: 1393-1400.
- [13] ZHANG K H, ZHANG L, YANG M H. Real time compressive tracking [J]. *ECCV*, 2012: 864-877.
- [14] 朱秋平, 颜佳, 张虎, 等. 基于压缩感知的多特征实时跟踪[J]. *光学精密工程*, 2013, 21(2): 437-444.
ZHU Q P, YAN J, ZHANG H, *et al.*. Real-time tracking using multiple features based on compressive sensing [J]. *Opt. Precision Eng.*, 2013, 21(2): 437-444. (in Chinese)
- [15] DONOHO D L. Compressed sensing [J]. *IEEE Transactions on Information Theory*, 2006, 52(4): 1289-1306.
- [16] DOLLAR P, WOJEK C, SCHIELE B, *et al.*. Pedestrian detection; An evaluation of the state of the art [J]. *IEEE Transaction of Pattern Analysis and Machine Intelligence*, 2012, 34(4): 743-761.
- [17] PORKILI 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.
- [18] ADAM A, RIVLIN E, SHIMSHONI I. Robust Fragments-based Tracking using the Integral Histogram [C]. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*, 2006: 798-805.
- [19] BIRCHFIELD S, RANGARAJAN S. Spatiograms vs. histograms for region based tracking [C]. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference*, 2005: 1158-1163.

作者简介:



孙晓燕(1978—),女,山东淄博人,博士研究生,2000年于山东工业大学获得学士学位,2003年于山东大学获得硕士学位,主要从事机器视觉及运动目标跟踪方面的研究。E-mail: sxy@sdjzu.edu.cn

导师简介:



常发亮(1965—),男,山东潍坊人,教授,博士生导师,主要从事机器视觉和模式识别方向的研究。E-mail: flchang@sdu.edu.cn

(版权所有 未经许可 不得转载)