

DOI: 10.11992/tis.201806038

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.tp.20180928.1338.004.html>

一种自适应模板更新的判别式 KCF 跟踪方法

宁欣^{1,2}, 李卫军^{1,2,3}, 田伟娟², 徐驰², 徐健¹

(1. 中国科学院半导体研究所 高速电路与神经网络实验室, 北京 100083; 2. 威富集团 形象认知计算联合实验室, 北京 100083; 3. 中国科学院大学 微电子学院, 北京 100029)

摘要:为了解决单目标跟踪算法中存在的目标旋转、遮挡和快速运动等挑战,提出了一种基于自适应更新策略的判别式核相关滤波器(kernelized correlation filter, KCF)目标跟踪新框架。构建了外观判别式模型,实现跟踪质量有效性的评估。构造了新的自适应模板更新策略,能够有效区分目标跟踪异常时当前目标是否发生了旋转。提出了一种结合目标检测的跟踪新构架,能够进一步有效判别快速运动和遮挡状态。同时,针对上述 3 种挑战,分别采用模板更新、目标运动位移最小化以及目标检测算法实现目标跟踪框的恢复,保证了跟踪的有效性和长期性。实验分别采用 2 种传统手动特征 HOG 和 CN(color names)验证提出的框架鲁棒性,结果证明了提出的目标跟踪新方法在速度和精度方面的优越性能。

关键词:目标跟踪;目标检测;高速核相关滤波算法;模板更新;卷积神经网络

中图分类号:TP391 **文献标志码:**A **文章编号:**1673-4785(2019)01-0121-06

中文引用格式:宁欣,李卫军,田伟娟,等.一种自适应模板更新的判别式 KCF 跟踪方法[J].智能系统学报,2019,14(1):121-126.

英文引用格式:NING Xin, LI Weijun, TIAN Weijuan, et al. Adaptive template update of discriminant KCF for visual tracking[J]. CAAI transactions on intelligent systems, 2019, 14(1): 121-126.

Adaptive template update of discriminant KCF for visual tracking

NING Xin^{1,2}, LI Weijun^{1,2,3}, TIAN Weijuan², XU Chi², XU Jian¹

(1. Laboratory of Artificial Neural Networks and High-speed Circuits, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China; 2. Image Cognitive Computing Joint Lab, Wave Group, Beijing 100083, China; 3. School of Microelectronics, University of Chinese Academy of Sciences, Beijing 100029, China)

Abstract: To solve the challenges of in-plane/out-of-plane rotation (IPR/OPR), fast motion (FM), and occlusion (OCC), a new robust visual tracking framework of discriminant kernelized correlation filter (KCF) based on adaptive template update strategy is presented in this paper. Specifically, the proposed discriminant models were first used to determine the tracking validity and then a new adaptive template update strategy was introduced to effectively distinguish whether or not the object has rotated when the object tracking was abnormal. Furthermore, a new visual tracking framework combining object test is presented, which could further effectively distinguish FM and OCC. Meanwhile, to overcome the above-mentioned challenges, three measures were taken to recover the object tracking frame: template updating, object movement displacement minimization, and use of an object detection algorithm ensuring validity and long-term visual tracking. We implemented two versions of the proposed tracker with representations from two conventional hand-actuated features, histogram of oriented gradient (HOG), and color names (CN) to validate the strong compatibility of the algorithm. Experimental results demonstrated the state-of-the-art performance in tracking accuracy and speed for processing the cases of IPR/OPR, FM, and OCC.

Keywords: visual tracking; object detection; high-speed kernelized correlation filters; template update; convolution neural network

收稿日期: 2018-06-22. 网络出版日期: 2018-09-30.

基金项目: 国家自然科学基金项目(61572458).

通信作者: 李卫军. E-mail: wjli@semi.ac.cn.

近年来,计算机视觉作为一门新兴学科发展十分迅速,目标跟踪作为视频监控、分析和理解

的关键技术,由于其重要的学术价值、潜在的商业价值和广泛的应用前景迅速成为计算机视觉领域研究的热点和难点,已成功应用于智能监控、机器人、医学成像及人体运动分析等人机交互领域^[1-2],越来越多的学者和机构在相关方面进行了大量的研究工作,同时取得了较好的成就。然而,目标形状变化、快速运动和遮挡一直是限制目标跟踪在实际场合应用的瓶颈。

最近,KCF^[3]以其高效的计算能力成功引起了众多研究者的重视和青睐,同时在目标跟踪中取得了较好的性能。该方法首次将快速 Fourier 变换成功应用到目标跟踪中,达到了实时跟踪的效果。然而该算法未对遮挡和快速运动中的目标跟踪作进一步处理;同时为了减少训练样本数目少的缺陷,每一帧均用于模板更新,以增强模板的检测性能,但这也使得大量的与目标无关的背景信息被加入到模板更新中,导致跟踪框漂移,目标跟踪性能下降。针对经典的 KCF 算法存在的问题,最新的跟踪算法主要从 2 个方面作改进:一种是目标特征表示的优化设计,另一种是鲁棒的分类器的设计。

1) 目标新特征的优化设计方法。特征的优化设计是目标跟踪领域一个非常重要的研究线索,该方法主要致力于目标表示的设计,即通过采用机器学习方法或者深度学习方法设计表达能力更好的特征,实现目标更详细的描述。文献[4-8]对传统特征进行改进;考虑到深度学习在特征表示中的优越性,文献[9-13]采用深度卷积神经网络挖掘高层卷积特征,以增强跟踪算法的性能。文献[9]直接从预训练的深度 CNN 中提取若干层特征并用于学习自适应相关滤波器,保证了跟踪准确率和鲁棒性。进一步地,文献[10]提出了一种基于 CNN 的序列训练方法,该方法把每个通道的特征输出作为单独的跟踪器,最后将分类器的结果级联作为最后的决策输出。目标特征的优化设计方法,虽取得了一定成绩,但目标跟踪算法旨在区分出背景和区域,该方法所学习到的特征只表示了目标视频的底层视觉信息,未对背景作进一步约束,导致学习到的模板判别能力非常有限。因此,获取判别性好的目标特征仍是一个比较棘手的难题。

2) 鲁棒的分类器的设计方法。鲁棒的分类器设计方法主要是通过建立优化机制,对需要进行模板更新的样本作限制,进一步保证了模板对于目标的检测效果。为了充分利用结构化输出的 SVM 能够处理复杂输出,如树、序列和集合等优势,Hare 等^[14]首次将该算法应用到跟踪中,大大

提高了跟踪准确率。为降低计算复杂度,Ning 等^[15]提出了一种基于双线性结构化的 SVM 模型,该模型采用的特征图可达到近似非线性核的效果,因此跟踪性能得到大幅度提升,然而该算法的跟踪速度依然限制了其实际应用。因此,Wang 等^[16]提出了一种高置信度的模板更新策略,能尽可能地避免模板更新引入的背景信息,但目标跟踪过程中,目标往往不是一成不变的,目标状态的变化会引起当前跟踪框的置信度较小,从而导致学习到的模板无法拟合目标的变化。

以上 2 种方法致力于设计更复杂的特征或者分类器,在提高跟踪精度的同时,无疑大大增加了算法的复杂度,从而限制了视觉跟踪的实时性能,同时,上述方法并没有对目标的快速运动和遮挡作进一步的处理。

1 跟踪算法新构架

本文提出了一种自适应模板更新的判别式核相关滤波跟踪新构架,如图 1 所示。能够判断当前跟踪状态是否正常,及时恢复丢失的目标,同时能够判断当前跟踪的目标是否发生形状变化,并制定出不同的机制恢复丢失的目标,实现了高效的目标跟踪。

1.1 基于判别式模型的跟踪质量判定

为了实时监督跟踪目标是否发生形状变化或者目标丢失的情况,本文提出一种基于外观判别式分类器的跟踪质量判定标准,以判断跟踪结果的准确性,实现目标的有效跟踪。本文跟踪结果判定标准的定义和实现如下:

1) 获取初始帧中目标框的初始信息。采用目标检测算法对输入帧执行目标检测,获取目标框的初始信息,包括目标所在位置以及尺寸大小。

2) 构建判别式分类器的样本训练集并分配样本标签。以一定步长在目标框的周围进行稠密采样,获取与目标框大小相同的训练样本;假设样本图像所在框与目标框之间的重叠阈值为 T_{ovp} ,分别计算所有样本图像与目标框的重叠值;当重叠值大于 T_{ovp} 时,设置该样本图像的标签为 1,否则为 0。

3) 提取样本的 HOG 特征在低维空间的特征表示,即 HOG 描述子。首先提取所有样本的 HOG 特征,然后采用 PCA-net^[17]的思想对 HOG 特征作进一步的降维处理,旨在获取更抽象且判别性更好的特征表示 f 。

4) 对所有正负样本重复 3),得到所有样本的低维特征表示,为当前目标学习判别模型 svm_1 ,并加入到集合 SVM_{obj} 中, SVM_{obj} 为跟踪目标出现的所有外观状态对应的分类器集合。

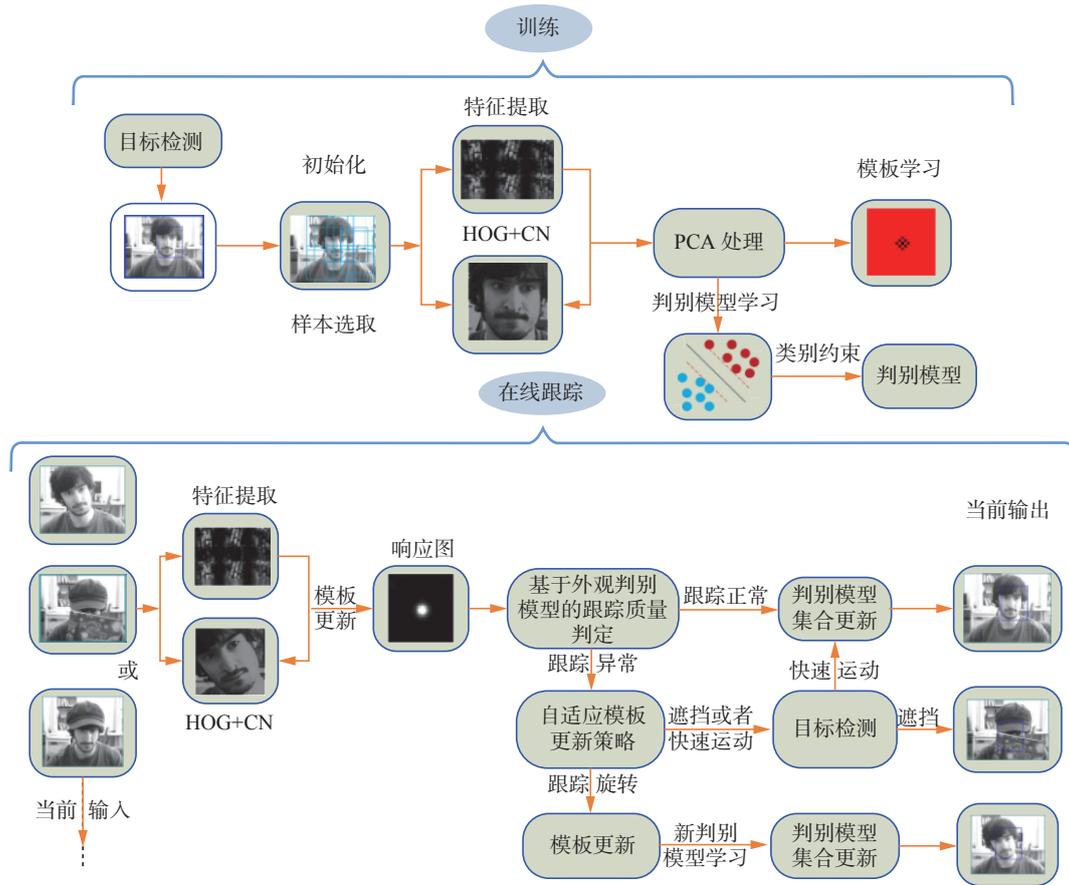


图 1 目标跟踪新框架图示

Fig. 1 Overview of the proposed framework for visual tracking

5) 针对易混淆的训练样本, 在模型学习的过程中加入类别约束, 对易错分的样本进行惩罚, 使得特征的类别信息更明显, 判别能力更强。针对出现错分的样本: 若隶属于目标的样本特征 f_{obj} 被错分到背景类别中, 即 f_{obj} 在目标中的类别判分小于在背景中的判分; 若隶属于背景的样本特征 f_{bgd} 被错分到目标类别中, 即 f_{bgd} 在目标中的类别判分大于在背景中的判分。因此, 引入下列类别约束条件:

$$s(f_{obj}, svm_1) - s(f_{bgd}, svm_1) > 0 \quad (1)$$

6) 采用 KCF 获取测试帧的跟踪结果, 并计算跟踪区域对应的特征 f 在集合 SVM_{obj} 下的判分 $s(f, SVM_{obj})$, 以判断当前目标的跟踪状态 trk_sta 是否正常, 计算如式 (2)、式 (3):

$$s(f, SVM_{obj}) = \max(s(f, svm_i), i = 1, 2, \dots, n) \quad (2)$$

$$trk_sta = \begin{cases} \text{normal}, & s(f, SVM_{obj}) > T_{sco} \\ \text{abnormal}, & \text{其他} \end{cases} \quad (3)$$

式中: n 表示外观判别模型的数目; T_{sco} 表示跟踪结果在判别模型中的判分阈值。若 $trk_sta = \text{normal}$, 表明跟踪结果可信; 否则, 表明目标状态变化, 可能发生旋转、目标快速运动或者被遮挡等情况。

1.2 自适应模板更新策略

本节给出了一种基于极小值的模板更新策略, 旨在针对性地筛选出不同形状以及与背景相似度较大的目标样本, 用于对当前模板进行更新, 以增强模板对于不同形状和与背景相似的正样本的泛化能力。

目标跟踪过程中, 可观察到置信度呈现先下降后上升的趋势, 有 2 种情况可以解释: 1) 目标形状从偏离初始模板到逐渐恢复至与模板相似, 如目标的平面内旋转; 2) 目标形状发生变化的过程中, 目标形状逐渐偏离初始模板, 而背景中存在与初始模板相近的区域, 导致目标所在位置框朝着背景方向漂移, 如目标平面外的旋转。为了能够让模板具有较好的鲁棒性, 需要同时保证图像标签为 1 的正样本的数量和多样性, 当上述 2 种情况出现时表示目前样本的状态发生了变化, 需要对模板及时进行更新, 使得跟踪器具有更好的检测性能。因此, 模板的更新条件设置为: 连续多帧目标跟踪框的置信度呈现先增后减的趋势, 即伴随有极小值的出现。基于这样的思想, 本文基于极值的模板更新策略具体过程如下。

1) 输入下一帧图像, 采用 KCF 算法进行跟

踪, 获取跟踪结果的置信度。

2) 重复步骤 1), 获取连续 n 帧的置信度集合 $C = \{c_1, c_2, \dots, c_n\}$, 观察置信度的变化情况。设 F 表示更新标志位, 定义为

$$F = \begin{cases} 1, & \text{argmin}(C) == n - 1 \\ 0, & \text{其他} \end{cases} \quad (4)$$

3) 若 F 为 1, 则极小值对应的跟踪结果用于模板更新, 同时考虑到模板更新具有一定的记忆功能, 为了能包含更完整的样本信息, 对接下来的 3 帧, 直接设置标志位 F 为 1, 并执行模板更新操作。

4) 模板未被更新时, 若极小值出现在帧 Fra_n 时, 保留最后的 $n-1$ 帧, 否则仅保留最后一帧; 模板被更新时, 剔除所有 n 帧的置信度, 得到更新后的置信度集合 C'_n , 计算式为

$$C'_n = \begin{cases} \{\}, & \text{argmin}(C) == n - 1 \\ \{c_2, c_3, \dots, c_n\}, & \text{argmin}(C) == n \\ \{c_n\}, & \text{其他} \end{cases} \quad (5)$$

5) 将用于模板更新的样本作为训练样本, 学习新的外观判别式模型 svm , 并加入到 SVM_{obj} 中。

1.3 目标跟踪新构架

本文提出了一种自适应模板更新的判别式核相关滤波跟踪新构架, 能够实时监督目标的跟踪情况, 捕捉发生旋转或者处于遮挡和快速运动的目标, 并进一步设计跟踪丢失恢复机制, 以保证目标实时且长期有效地跟踪, 具体实施如下。

1) 输入图像, 并执行 1.1 节提出的跟踪质量判定方法, 捕捉跟踪异常的情况。

2) 获取跟踪异常的图像帧, 利用 1.2 节提出的模板更新策略获取模板更新标志 F 。 $F = 1$ 表明目标发生旋转, 有必要进行模板更新; 否则, 表明目标形状未发生较大的变化, 此时目标可能被干扰物遮挡或者处于快速运动的状态。

3) 对当前帧执行目标检测算法, 若丢失目标被成功检测, 则利用检测结果继续跟踪, 否则认为目标被干扰物遮挡, 在这种情况下, 假设遮挡过程具备时长短期性和目标运动位移最小化的性质, 因此, 保持目标丢失前的位置不变, 直到跟踪状态 trk_sta 正常或者目标被成功检测。

2 实验与分析

2.1 实验设置和数据库

为验证提出的新框架的有效性, 首先从公共数据集 OTB-100^[18] 中选取 50 组视频进行测试。然后, 分别选取具有旋转、快速运动和遮挡属性的视频对本文新构架的鲁棒性进行评估。图 2 给出了 OTB-100 中的部分视频帧。一次成功率和帧速率分别用于跟踪评估。一次成功率用于评估

跟踪的准确率, 包含跟踪精度和成功率两个指标; 帧速率用于评估跟踪速度的快慢。



图 2 OTB-100 数据库
Fig. 2 OTB-100 dataset

本文实验参数设置如下: 1) 本文提出的基于极值的自适应更新策略中, 最小值存在性判断需要的帧数目 n 设置为 4, 重叠率阈值 T_{sco} 设置为 0.7; 2) 本文用于跟踪质量判定的判别模型学习中, 训练样本标签确定需要的重叠率阈值 T_{ovp} 设置为 0.7。

2.2 新构架的目标跟踪

本节采用基于 HOG 和 CN 特征结合的跟踪新构架, 即 AU_DKCF 进行测试, 通过比较不同特征以验证提出的方法的有效性。表 1 分别列出了新构架 AU_DKCF 与对比算法的跟踪结果。由表 1 可以看出, 新构架在跟踪速度和精度方面取得了较高的性能, 均优于对比算法。其原因在于: 新构架能够较准确判断目标是否跟踪异常, 同时可进一步区别目标跟踪异常所产生的原因, 并利用不同机制恢复丢失的目标, 实现目标长期且准确地跟踪。AU_DKCF 在采用 2 种手工特征结合时, 跟踪速度能达到 105.2 f/s, 可有效实现目标的实时跟踪。

表 1 本文新构架和对比算法的目标跟踪结果

Table 1 Object tracking results of proposed method and contrastive methods on OTB-50

方法	跟踪准确率/%		帧速率/(f/s)
	精确度	成功率	
KCF ^[3]	73.20	52.00	172.0
DCF ^[3]	72.80	50.40	292.0
LMCF ^[16]	83.90	62.40	85.2
TLD ^[19]	60.80	33.40	28.0
ORIA ^[20]	45.70	23.10	9.0
Ours (HOG)	85.40	64.30	143.0
Ours (CN)	86.70	52.40	121.0
AU_DKCF	87.50	63.10	105.2

2.3 不同属性的视频跟踪

本节针对具有目标形状变化、目标被遮挡和快速运动等属性的视频序列, 分别进行目标跟踪测试, 以验证提出的跟踪算法对于上述挑战的鲁棒性。表 2 给出了本文新构架和对比算法在 OTB-100 数据集上, 对于不同属性视频的跟踪结果。由表 2 可以看出, 本文提出的新构架对于上述挑战的有效性。

表 2 本文新构架对不同属性视频的跟踪结果

Table 2 Tracking results on videos with sequence attributes %

方法	精确度		
	遮挡	快速运动	平面内/外旋转
KCF ^[3]	74.90	65.00	74.00
Struck ^[14]	58.40	53.90	52.10
TLD ^[19]	58.30	57.60	51.20
LMCF ^[16]	62.70	61.60	58.85
AU_DKCF	75.80	72.10	79.90

2.4 跟踪结果可视化

本节给出了本文新框架的跟踪可视化效果, 以展示本文方法的有效性。图 3~5 分别给出了本文 AU_DKCF 和经典滤波算法 (线性相关滤波 DCF 和核相关滤波 KCF) 对旋转属性的视频, 快速运动属性的视频和遮挡属性的视频的跟踪结果。从图中可以观察得到, 新框架在处理具备旋转挑战的视频时, 获得了更优的性能。图 3~5 中, 绿色框表示 Groundtruth, 紫色框表示 DCF, 蓝色框表示 KCF, 红色框表示 AU_DKCF。



图 3 足球视频的可视化跟踪结果

Fig. 3 Visualizing tracking result of soccer



图 4 行人视频的可视化跟踪结果

Fig. 4 Visualizing tracking result of boy



图 5 篮球视频的可视化跟踪结果

Fig. 5 Visualizing tracking result of basketball

3 结束语

本文在充分分析、研究和实验的基础上, 取得了以下几点结论: 1) 提出一种基于外观判别式分类器的跟踪质量判定方法, 旨在判断当前跟踪状态是否正常, 以便及时恢复丢失的目标; 2) 提出了一种新的自适应模板更新策略, 在提高跟踪速度的同时, 能够判断当前跟踪的目标是否发生形状变化, 以增强模板的泛化能力; 3) 提出一种融合跟踪质量判定和模板自适应更新的跟踪新构架, 旨在实时判断当前跟踪目标所处状态, 同时制定出不同的机制恢复丢失的目标, 实现了高效的目标跟踪。后续作者将致力于将现有跟踪算法和深度学习结合用于安防监控中的异常行为分析、客流量统计和警戒区域报警等应用中。

参考文献:

- [1] YILMAZ A, JAVED O, SHAH M. Object tracking: a Survey[J]. ACM computing surveys, 2006, 38(4): 45–50.
- [2] CANNONS K. A review of visual tracking[R]. Technical Report CSE 2008-07. York University, Canada, 2008.
- [3] HENRIQUES J F, CASEIRO R, MARTINS P, et al. High-speed tracking with kernelized correlation filters[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(3): 583–596.
- [4] DANELLJAN F M, KHAN F S, FELSBURG M, et al. Adaptive color attributes for real-time visual tracking[C]//Proceedings of the 20014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 1090–1097.
- [5] JIA Xu, LU Huchuan, YANG M H. Visual tracking via adaptive structural local sparse appearance model[C]//Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA, 2012: 1822–1829.
- [6] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 34(7): 1409–1422.
- [7] KWON J, LEE K M. Visual tracking decomposition[C]//Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA, 2010: 1269–1276.
- [8] ROSS D A, LIM J, LIN R S, et al. Incremental learning for robust visual tracking[J]. International journal of computer vision, 2008, 77(1/2/3): 125–141.
- [9] HONG S, YOU T, KWAK S, et al. Online tracking by learning discriminative saliency map with convolutional neural network[C]//Proceedings of the 32nd International

- Conference on Machine Learning. Lille, France, 2015: 597–606.
- [10] MA Chao, HUANG Jiabin, YANG Xiaokang, et al. Hierarchical convolutional features for visual tracking[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 3074–3082.
- [11] QI Yuankai, ZHANG Shengping, QIN Lei, et al. Hedged deep tracking[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 4303–4311.
- [12] WANG Lijun, OUYANG Wanli, WANG Xiaogang, et al. Visual tracking with fully convolutional networks[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 3119–3127.
- [13] WANG Lijun, OUYANG Wanli, WANG Xiaogang, et al. Stct: sequentially training convolutional networks for visual tracking[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 1373–1381.
- [14] HARE S, SAFFARI A, TORR P H S. Struck: structured output tracking with kernels[C]//Proceedings of the 2011 International Conference on Computer Vision. Barcelona, Spain, 2011: 263–270.
- [15] NING Jifeng, YANG Jimei, JIANG Shaojie, et al. Object tracking via dual linear structured SVM and explicit feature map[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 4266–4274.
- [16] WANG Mengmeng, LIU Yong, HUANG Zeyi. Large margin object tracking with circulant feature maps[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA, 2017: 4800–4808.
- [17] CHAN T H, JIA Kui, GAO Shenghua, et al. PCANet: a simple deep learning baseline for image classification?[J]. *IEEE transactions on image processing*, 2015, 24(12): 5017–5032.
- [18] WU Yi, LIM J, YANG M H. Online object tracking: a benchmark[C]//Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA, 2013: 2411–2418.
- [19] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2012, 34(7): 1409–1422.
- [20] WU Yi, SHEN Bin, LING Haibin. Online robust image alignment via iterative convex optimization[C]//Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Rhode Island, USA, 2012: 1808–1814.

作者简介:



宁欣,男,1989年生,博士,IEEE、中国计算机学会、中国人工智能学会、中国自动化学会会员,主要研究方向为计算机视觉、类脑视觉计算、形象认知计算和神经网络。发表学术论文15篇,被SCI收录5篇。



李卫军,男,1975年生,教授,博士生导师,博士,主要研究方向为机器视觉、模式识别与智能系统、仿生图像处理技术、仿生模式识别理论与方法、近红外光谱定性分析技术、高维信息计算。主持国家自然科学基金项目2项,企业合作研究项目3项。发表学术论文30篇。



田伟娟,女,1990年生,硕士研究生,主要研究方向为视频、图像处理和目标跟踪相关算法。