

尺度自适应与抗遮挡的KCF方法研究

闫焯堃, 刘天波

沈阳建筑大学计算机科学与工程学院, 辽宁 沈阳

收稿日期: 2023年3月11日; 录用日期: 2023年4月11日; 发布日期: 2023年4月19日

摘要

随着当前社会智能化的快速发展以及计算机视觉技术的兴起, 目标跟踪技术在日常生活的诸多领域都得到了广泛应用。自相关滤波思想于2010年被引入目标跟踪领域后, 基于此思想的跟踪器在性能方面的优良表现引起了国内外研究学者的关注。经典的核相关滤波目标跟踪算法(KCF)在较简单的场景下表现了较好的跟踪效果, 但面对跟踪目标尺度变化及目标被遮挡的问题时, KCF算法无法进行稳定跟踪。本文提出了一种基于KCF的尺度自适应及抗遮挡的跟踪方法, 利用尺度系数组对跟踪窗口进行尺度缩放, 通过计算响应值确定跟踪窗口大小。同时利用图像响应值对目标进行遮挡判断, 若目标未遮挡, 则按原KCF算法执行; 若目标被遮挡, 则对图像提取ORB特征, 在后续帧中重新确定目标位置, 完成后续跟踪。OTB100数据集的实验表明本文方法对目标跟踪的可行性和有效性。

关键词

目标跟踪, 尺度自适应, 抗遮挡

Research on KCF Method of Scale Adaptation and Anti-Occlusion

Xiaokun Yan, Tianbo Liu

School of Computer Science and Engineering, Shenyang Jianzhu University, Shenyang Liaoning

Received: Mar. 11th, 2023; accepted: Apr. 11th, 2023; published: Apr. 19th, 2023

Abstract

With the rapid development of social intelligence and the rise of computer vision technology, target tracking technology has been widely used in many areas of daily life. After the idea of autocorrelation filtering was introduced into the field of target tracking in 2010, the excellent performance of the tracker based on this idea has attracted the attention of researchers at home and abroad. The classic kernel correlation filtering target tracking algorithm (KCF) performs better tracking effect in relatively simple scenes, but the KCF algorithm cannot track stably in the face of

the scale change of the tracking target and the occlusion of the target. In this paper, a KCF-based scale adaptive and anti-occlusion tracking method is proposed. The tracking window is scaled using the scale system array, and the response value of the specified region and template is calculated, and then the tracking window scale is determined. At the same time, the image response value is used to judge the occlusion of the target. If the target is not occluded, the original KCF algorithm is used; if the target is occluded, the ORB feature is extracted from the image, and the target position is redefined in the subsequent frame to complete the follow-up tracking. Experiments on OTB100 data set show that the proposed method is feasible and effective for target tracking.

Keywords

Target Tracking, Scale Adaptation, Anti-Occlusion

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

目标跟踪是计算机视觉领域的一个重要研究方向, 因此受到了众多研究人员的广泛关注, 目标跟踪技术的核心是从视频序列中提取目标在时间和空间上的相关性, 并应用算法获得目标的运动轨迹及运动状态。目前目标跟踪技术在生活中的很多领域发挥着重要作用, 比如视频监控、人机交互、智能交通、无人驾驶等。

运动目标跟踪作为计算机视觉领域中一个基础且重点的热门话题, 得到了国内外研究人员的密切关注与研究, 引入各方面的知识并总结出许多算法。基于相关滤波思想的目标跟踪算法在跟踪效果上得到了显著提升。2010 年, Bolme 等人[1]首次将相关滤波思想引入目标跟踪领域, 提出了一种最小输出平方误差和滤波器(MOSSE), 这一结果引起了众多学者的广泛关注, 随后, 使用循环结构的检测和跟踪算法(CSK)[2]和核相关滤波的目标跟踪算法(KCF)[3]相继出现。在此之后, KCF 算法良好的跟踪性能使其成为了公认的跟踪算法之一, 对它的深入研究也越来越多。Ting Liu 等人[4]提出使用结构化方法将目标区域划分为具有给定权重的子块, 并最终使用这些权重计算目标位置。Akin 等人[5]提出了一种基于全局跟踪器和多个局部跟踪器之间的耦合和交互的跟踪器。Ma 等人[6]提出了一种长期目标跟踪算法(LCT), 该算法使用平移相关滤波、尺度相关滤波和置信度检测相关滤波来实现目标尺度自适应, 并对遮挡具有良好的处理效果。Bibi 等人[7]将多特征与多模板进行融合, 同时结合后验概率来获得目标的尺度信息。Danelljan 等人[8]提出了 SRDCF 算法, 通过扩大搜索范围及引入正定因子消除了边界效应。Danelljan 等人[9]提出了 DSST 算法, 通过两个滤波器完成位置估计和尺度估计。Wang M M 等人[10]提出了结合多层次特征和多峰前向检测判定的方法, 从特征描述和模型更新两方面来解决目标跟踪中的遮挡问题。

2. 理论基础

核相关滤波算法(Kernelized Correlation Filters, KCF)于 2014 年由 Joao F. Henriques, Rui Caseiro, Pedro Martins, and Jorge Batista 提出, 相较于当时的其他算法, KCF 在跟踪效果和跟踪速度方面均有较为亮眼的表现, 同时被广泛应用到生活中的实际场景中, 大批学者也开始了对该算法的深入研究。

传统的 KCF 算法由四部分构成, 分别是: 目标初始化、外观建模、运动估计和目标定位。目标初始化阶段即在视频初始帧中主动选择感兴趣的目标, 常用的方法有手动选择、自动检测及读取初始信息等。

建模过程通常包括外观特征提取和建模两方面,外观特征提取即对目标的一种或多种特征如颜色、纹理、边缘信息等进行提取,得到的结果称为特征描述子;建模即将特征描述子以数学统计和计算的方式转化为数学模型,用于后续的目标检测与跟踪目标跟踪。运动估计即对运动中目标的位置进行预测,根据当前帧目标的信息,通过计算下一帧中目标运动至不同位置的概率值来确定目标最有可能到达的位置。

2.1. 循环采样

KCF 算法引入了循环矩阵,利用目标区域循环移位取得的数据来近似模拟训练样本集合,我们通过公式(1)定义基础样本 x ,此时 x 为向量,由图像的第 n 行中的像素值组成。

$$x = [x_1, x_2, x_3, \dots, x_n]^T \quad (1)$$

通过该样本,可以得到循环采样后的循环矩阵,如公式(2)。

$$X = C(x) = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ x_n & x_1 & x_2 & \cdots & x_{n-1} \\ x_{n-1} & x_n & x_1 & \cdots & x_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_2 & x_3 & x_4 & \cdots & x_1 \end{bmatrix} \quad (2)$$

每个循环矩阵都可以利用傅里叶变换实现对角化,即循环矩阵的特征值分解,这个性质可以表示如下:

$$X = F \text{diag}(\hat{x}) F^H \quad (3)$$

其中, $\hat{x} = F(x) = Fx$, 表示 x 的傅里叶变换。

2.2. 脊回归

KCF 算法采用最小二乘法训练分类器,根据线性回归函数 $f(x_i) = w^T x_i$ 可以得到求解 w 的目标回归函数:

$$\min \sum_i (f(x_i) - y_i)^2 + \lambda w^2 \quad (4)$$

在计算过程中,利用公式(5)将计算转化至傅里叶域内进行,可以大大简化计算量。同时,在实际处理问题的过程中,大多为二维图像,因此需要借助核函数将问题转化为非线性回归问题,非线性回归方程如公式(6):

$$X = F \text{diag}(\hat{x}) F^H \quad (5)$$

$$w = \sum_i \alpha_i \varphi(x_i) \quad (6)$$

最终可将问题转化为求解 α , 如公式(7):

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}^{xx} + \lambda} \quad (7)$$

其中 $\hat{\alpha}$ 为 α 的傅里叶变换, k^{xx} 表示核矩阵 K 的第一行元素。

2.3. 快速检测

响应值的计算可以将上述 $\hat{\alpha}$ 带入公式(8)得到:

$$\hat{f}(z) = \hat{k}^{xz} \odot \hat{\alpha} \quad (8)$$

其中 z 表示待检测图像, x 表示已知目标图像, K^{xz} 表示 x 和 z 的核相关。由此求得的响应值最大的位置

即可确定为当前目标位置。

3. 算法改进

3.1. 尺度自适应

通过读取初始信息且 KCF 算法第一帧运行结束后, 位置预测模块定位出的目标位置有一个初始跟踪窗口, 定义其大小为 s_0 , 同时定义一个包含若干参数的尺度系数组 $S = \{a_1, a_2, \dots, a_n\}$, 在进行下一帧跟踪任务时, 对初始的跟踪窗口用 $S_i = a_i s_0$ 进行缩放, 利用不同大小的窗口对样本进行不同尺度的采集, 调整至固定大小后分别与模板进行响应值计算, 得到获得最大响应值的尺度系数, 进而确定当前帧跟踪窗口的大小。流程如图 1 所示。

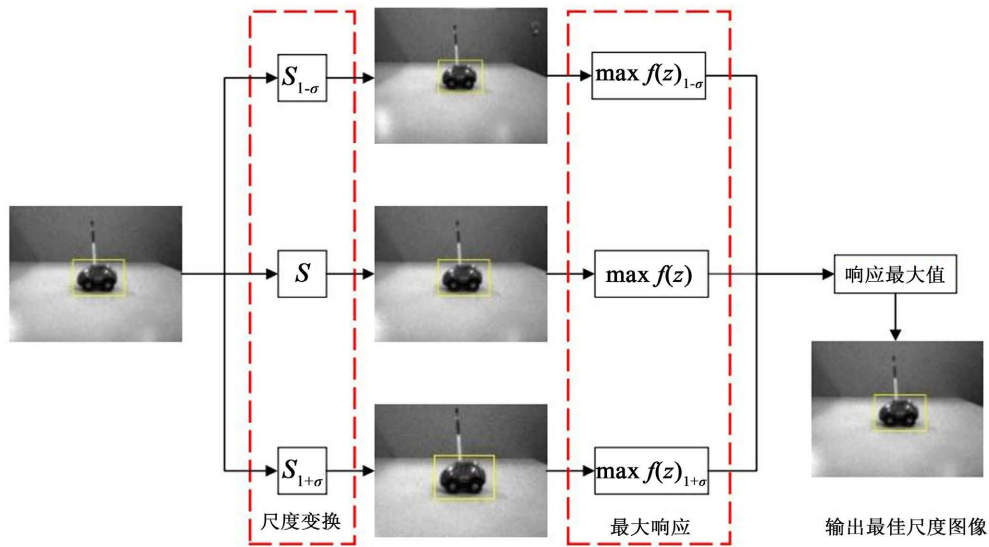


Figure 1. Flow chart of scale adaptation
图 1. 尺度自适应流程图

3.2. 遮挡判断

在对图像进行卷积运算来确定目标位置的过程中, 会涉及到响应值的计算, 通过计算可以得到图像各位置的响应值大小, 当目标周围环境较好, 即没有外部因素影响时, 得到的响应图的峰值是十分尖锐的, 但当目标受到遮挡等外界环境因素影响时, 响应图的峰值会比较小且峰也会比较平坦, 这样得到的目标定位就是不准确的。

本位采用响应峰值和平均峰值相关能量(average peak-to correlation energy, APCE)指标对跟踪结果进行判断, 定义如下:

$$APCE = \frac{|F_{\max} - F_{\min}|^2}{\text{mean}\left(\sum_{i,j} (F_{i,j} - F_{\min})^2\right)} \quad (9)$$

其中: $F_{\max}$ 表示响应峰值, $F_{\min}$ 表示响应最小值, $F_{i,j}$ 表示在第 i 行、第 j 列位置的响应值。当响应峰值越大时, 目标位置的检测的可信度就越高, 进而计算得到的平均峰值能量值也就越大。当发生遮挡或者目标消失时, 平均峰值相关能量会明显减小。当响应峰值和平均峰值相关能量值均大于其一定比例的历史平均值时, 我们就认为跟踪结果的可靠性高, 这样可以有效防止目标模型被污染, 提高跟踪鲁棒性。

3.3. 遮挡再检测

对图像是否遮挡进行判断后, 若目标未被遮挡, 则正常进行; 若目标被遮挡, 则需对训练图像和下一帧图像提取 ORB 特征, 并对二者进行特征匹配, 重新检测出目标位置。

ORB 是 Oriented Fast and Rotated Brief 的简称, 是一种对图像特征进行描述的方法, 主要包括特征点检测以及特征向量创建。首先使用 FAST 算法对图像进行关键点提取, 关键点即图像中较为明显的点, 例如轮廓点、较暗区域中的亮点、较亮区域中的暗点等。然后使用 BRIEF 算法对关键点进行特征描述, 最终形成二进制的特征描述子, 并通过计算特征描述子之间的汉明距离来进行特征点匹配。

4. 实验结果与分析

4.1. 数据集与实验环境

本文选用的数据集为 OTB 数据集, 分为 OTB50 [11] 和 OTB100 [12], 我们选用视频序列更加完整的 OTB100 数据集进行实验分析。数据集内包含三部分内容, 一是图片形式的视频帧序列, 二是每副图像中目标中心位置的坐标及目标的高和宽, 三是以 .mat 文件存储的视频基本信息, 包括视频总帧数, 背景信息, 以及该视频的标记属性等。实验环境为 Windows10 操作系统, CPU 为 i7-6700HQ, 运行内存 8 GB, 在电脑客户端 MATLAB 进行实验, 版本为 MATLAB R2019a。

4.2. 评价指标

通常情况下, 对于目标跟踪效果有两个指标, 即精确度和成功率。精确度体现对目标中心位置的预测是否准确, 成功率体现跟踪窗口是否准确的覆盖目标。同时, 实时性也是对算法进行评价的指标之一。

1) 精确度

精确度表示误差在阈值范围内的帧数占所有帧数的百分比。误差是衡量跟踪精度的指标, 指目标的预测位置与实际位置之间的欧氏距离, 将其与指定阈值(0 px~50 px)结合, 即可得到精确度曲线, 曲线图横坐标为阈值, 纵坐标为精确度, 表示的是误差在横坐标阈值范围内的帧数占所有帧数的百分比, 由此我们可以推断, 位于曲线下方的面积越大, 说明该算法的精确度越高。

2) 成功率

成功率指算法预测到的目标位置区域与目标实际真实位置区域的重叠程度, 可以通过公式(10)求得:

$$S = \frac{|\alpha_p \cap \alpha_t|}{|\alpha_p \cup \alpha_t|} \quad (10)$$

其中: α_p 为算法对目标的预测区域, α_t 为目标的真实位置区域。与精确度一样, 将成功率与阈值(0~1)结合, 即可得到成功率曲线图, 表示的是成功率在横坐标阈值内的帧数占所有帧的百分比。

3) 实时性

实时性指算法在跟踪过程中每秒可以处理的视频帧数, 单位是 fps, 也是判断算法是否优异的重要指标。

4.3. 结果对比分析

本文从 OTB100 数据集中选择 3 组具有代表性的视频序列进行不同属性实验, 除本文改进算法外, 参与实验的算法还有基础算法核相关滤波算法 KCF、长期相关滤波跟踪算法 LCT 及基于循环架构的核检测算法 CSK, 通过定性与定量两方面对不同算法效果进行对比分析。

4.3.1. 定性分析

针对目标尺度变化问题, 选择 Car 4 视频序列进行对比实验验证。视频记录的是一辆汽车, 在行驶过

程中汽车不断改变位置，并与录像设备进行距离的变化。图 2 展示了不同算法在视频序列中的表现效果，可以看出本文算法表现出了较好的跟踪效果。其中算法与颜色的对应关系如表 1 所示。

Table 1. Corresponding table of different algorithms and colors
表 1. 不同算法与颜色对应表

本文算法	KCF	LCT	CSK
Color	Color	Color	



Figure 2. Car 4 video sequence tracking effect
图 2. Car 4 视频序列跟踪效果

针对目标遮挡问题，选择 Walking 2 视频序列进行对比实验验证。视频包含 500 帧图像，拍摄的是一位女士在行走，在这个过程中，有其他人短暂出现在镜头中并对目标进行了遮挡。图 3 展示了不同算法在视频序列中的表现效果，可以看出本文算法表现出了较好的跟踪效果。

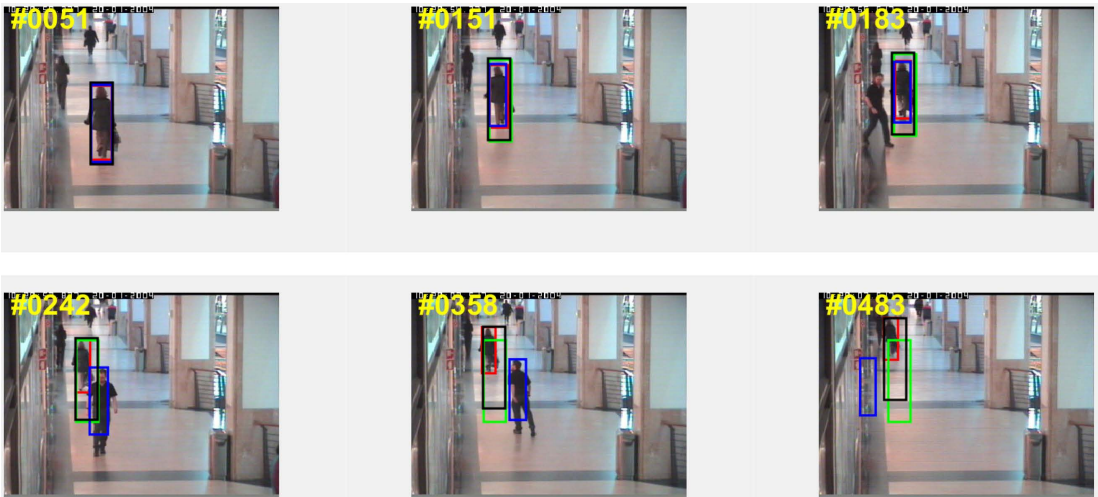


Figure 3. Walking 2 video sequence tracking effect
图 3. Walking 2 视频序列跟踪效果

除上述两种情况外, 另选择 Singer 1 视频序列对综合性场景进行对比实验验证, 该视频中不仅涉及到目标遮挡及尺度变化问题, 同时还存在光照变化、目标旋转等问题, 其记录的是舞台上的歌手, 在镜头逐渐移动的过程中周围环境也发生了变化。图 4 展示了不同算法在视频序列中的表现效果, 可以看出本文算法表现出了较好的跟踪效果。

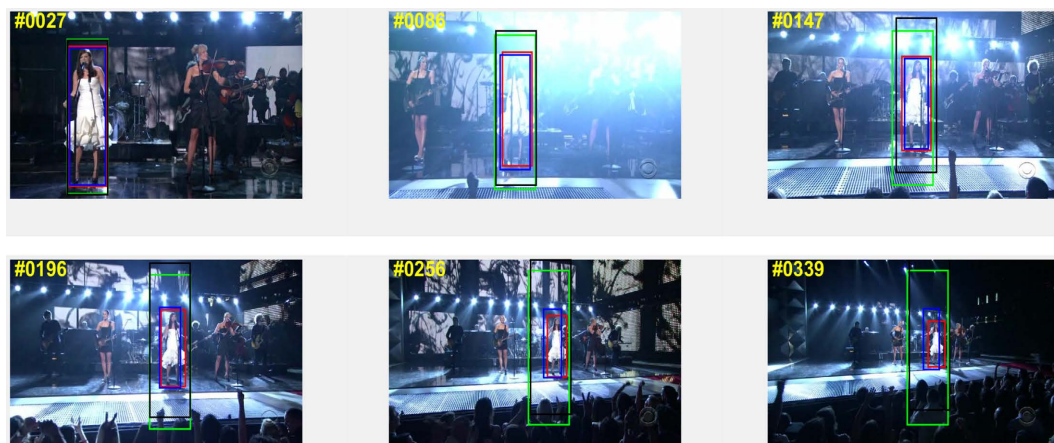


Figure 4. Singer 1 video sequence tracking effect

图 4. Singer 1 视频序列跟踪效果

4.3.2. 定量分析

选择上一小节进行实验的 3 个视频序列, 通过平均精确度和平均成功率进行直观的量化对比, 同时也对实时性进行考量, 表格如表 2 所示。

Table 2. Performance comparison of different algorithms

表 2. 不同算法性能对比

算法	精确度	成功率	实时性
本文算法	0.957	0.983	8.53
KCF	0.735	0.340	10.18
LCT	0.798	0.707	6.31
CSK	0.498	0.320	17.72

通过表格可以看出, 本文算法在精确度和成功率方面均表现良好, 当解决尺度及遮挡问题后, 跟踪窗口内的冗余信息减少, 避免了后续模板更新过程中的污染, 相比于基础算法 KCF 得到了较大提高。由于算法步骤相对增加, 因此实时性降低在所难免, 但仍可以有效地实施跟踪。

5. 结论

本文针对传统 KCF 算法在目标跟踪过程中无法处理目标尺度变化及目标遮挡情况的问题, 提出了一种尺度自适应与抗遮挡的跟踪方法。对于尺度变化场景, 算法通过引入尺度系数组对原算法始终保持不变的跟踪窗口进行缩放, 通过计算响应值筛选出最佳尺度。对于遮挡场景, 首先对目标进行遮挡判断, 若目标被遮挡, 则对图像提取 ORB 特征, 进行前后帧的特征匹配, 最终完成目标再检测。从实验结果看, 本文算法可以较好地处理上述两种问题, 同时在精确度和成功率方面也表现较好。本文算法在处理部分

目标快速移动及目标模糊问题时的效果仍需加强, 在未来将继续探索改进方法, 提高算法性能。

参考文献

- [1] Henriques, J.F., Caseiro, R., Martins, P., *et al.* (2014) High-Speed Tracking with Kernelized Correlation Filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**, 583-596. <https://doi.org/10.1109/TPAMI.2014.2345390>
- [2] Bolme, D.S., Beveridge, J.R., Draper, B.A. and Lui, Y.M. (2010) Visual Object Tracking Using Adaptive Correlation Filters. 2010 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, 13-18 June 2010, 2544-2550. <https://doi.org/10.1109/CVPR.2010.5539960>
- [3] Henriques, J.F., Caseiro, R., Martins, P., and Batista, J. (2012) Exploiting the Circulant Structure of Tracking-by-Detection with Kernels. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y. and Schmid, C., Eds., *Computer Vision—ECCV 2012. Lecture Notes in Computer Science* (Vol. 7575), Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-33765-9_50
- [4] Liu, T., Wang, G. and Yang, Q.X. (2015) Real-Time Part-Based Visual Tracking via Adaptive Correlation Filters. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 4902-4912. <https://doi.org/10.1109/CVPR.2015.7299124>
- [5] Akin, O., Erdem, E., Erdem, A. and Mikolajczyk, K. (2016) Deformable Part-Based Tracking by Coupled Global and Local Correlation Filters. *Journal of Visual Communication and Image Representation*, **38**, 763-774. <https://doi.org/10.1016/j.jvcir.2016.04.018>
- [6] Ma, C., Yang, X.K., *et al.* (2015) Long-Term Correlation Tracking. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 5388-5396. <https://doi.org/10.1109/CVPR.2015.7299177>
- [7] Bibi, A. and Ghanem, B. (2015) Multi-Template Scale-Adaptive Kernelized Correlation Filters. 2015 *IEEE International Conference on Computer Vision Workshop (ICCVW)*, Santiago, 7-13 December 2015, 613-620. <https://doi.org/10.1109/ICCVW.2015.83>
- [8] Danelljan, M., *et al.* (2015) Learning Spatially Regularized Correlation Filters for Visual Tracking. 2015 *IEEE International Conference on Computer Vision (ICCV)*, Santiago, 7-13 December 2015, 4310-4318. <https://doi.org/10.1109/ICCV.2015.490>
- [9] Danelljan, M., *et al.* (2016) Discriminative Scale Space Tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1561-1575. <https://doi.org/10.1109/TPAMI.2016.2609928>
- [10] Wang, M., Liu, Y. and Huang, Z. (2017) Large Margin Object Tracking with Circulant Feature Maps. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 4800-4808. <https://doi.org/10.1109/CVPR.2017.510>
- [11] Wu, Y., Lim, J. and Yang, M.H. (2017) Online Object Tracking: A Benchmark. 2013 *IEEE Conference on Computer Vision and Pattern Recognition*, Portland, 23-28 June 2013, 2411-2418. <https://doi.org/10.1109/CVPR.2013.312>
- [12] Wu, Y., Lim, J. and Yang, M.H. (2015) Object Tracking Benchmark. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **37**, 1834-1848. <https://doi.org/10.1109/TPAMI.2014.2388226>