

文章编号:1671-6833(2020)03-0032-05

基于改进核相关滤波的长时目标跟踪算法

陈丽萍, 王铭羽, 杨文柱, 王思乐, 陈向阳

(河北大学 网络空间安全与计算机学院, 河北 保定 071002)

摘要:长时目标跟踪研究在当前的目标跟踪研究领域占有主导地位。在长时目标跟踪过程中,由于目标快速移动和物体遮挡等造成的模板漂移是研究难点。为解决该问题,提出了一种基于改进核相关滤波的长时目标跟踪算法。该算法采用核相关滤波跟踪框架,联合高置信度的模版更新机制和目标重检机制解决模板漂移问题。实验结果表明,本文算法在长时目标跟踪过程中相对于传统算法更加稳定可靠。

关键词:长时目标跟踪; 模板漂移; 核相关滤波; 置信度

中图分类号: TP391 **文献标志码:** A **doi:**10.13705/j.issn.1671-6833.2019.02.001

0 引言

目标跟踪是计算机视觉领域的重要研究内容,也是计算机视觉领域的难点之一,除了光照变化、障碍物遮挡等环境影响,目标本身的旋转、尺度变化及出视野情况都对进行准确的目标跟踪造成挑战^[1]。根据观测模型的不同,可将视觉目标跟踪方法分为生成式方法和判别式方法。比较著名的生成式方法有 Meanshift 算法^[2-3]、粒子滤波算法^[4]。生成式方法重点关注对目标本身的刻画,不考虑背景的影响,在目标形变剧烈或遮挡严重时容易产生目标漂移;判别式方法同时考虑了目标和背景信息,在现有跟踪方法中具有明显的优势,其中比较经典且性能突出的判别类方法有 FCT 算法(fast compressive tracking)^[5]和相关滤波算法(correlation filter-based tracking, CFTs)。

相关滤波算法是由 Bolme 最早引入的,该算法通过相关滤波器(minimum output sum of squared error, MOSSE)^[6]来构造一种自适应的训练策略。由于该算法在实时目标跟踪上的优秀表现,此后基于该算法框架的改进算法被相继提出^[7-10],其中最具有影响力的为 Henriques 等^[7]提出的核相关滤波算法(kernelized correlation filters, KCF)。KCF 不仅能以每秒数百帧的速度

运行,且实现简单,在 VOT(visual object tracking)竞赛测评中展现出明显的高效性和鲁棒性。

传统的相关滤波跟踪算法均是针对短时视频的跟踪研究,而在实际应用中往往需要长时间对目标进行准确跟踪。在进行长时目标跟踪过程中,由于目标的外观变化、周围环境造成的目标遮挡以及光照变化等,传统滤波跟踪算法很容易发生模板漂移现象。

为解决上述问题,提出了一种基于改进核相关滤波的长时目标跟踪算法 LKCF(long-term kernelized correlation filter)。该算法以 KCF 算法为基本跟踪框架,为避免模板漂移,采用了一个高置信度的模版更新机制。另外,通过构造目标重检机制来解决由于遮挡造成的目标丢失问题,以便实现鲁棒的目标跟踪。

1 LKCF 算法

1.1 相关滤波原理

设 f 表示输入图像; h 表示相关滤波器; g 表示响应输出,则

$$g = h \otimes f, \quad (1)$$

式中: g 由高斯函数产生; h 为滤波器,在跟踪时,将下一帧的搜索窗口与滤波器进行相关计算,目标的新位置根据响应最大值所对应的位置确定。

收稿日期:2019-06-10;修订日期:2019-12-12

基金项目:河北省自然科学基金资助项目(F2015201033, F2017201069)

作者简介:陈丽萍(1974—),女,河北保定人,河北大学讲师,主要从事图像处理 and 机器视觉研究。

通信作者:杨文柱(1968—),男,河北保定人,河北大学教授,博士,主要从事机器视觉研究与智能系统研究, E-mail: wenzhuyang@163.com。

通过快速傅里叶变换,在频率域计算相关性可加快运算速度。另外,由卷积定理可知:

$$F\{g\} = (F^*(h)) \odot F(f), \quad (2)$$

式中: F 表示傅里叶变换; \odot 表示点乘; $*$ 表示复数共轭,则有

$$H^*(f) = \frac{G(f)}{F(f)}. \quad (3)$$

1.2 滤波器训练

LKCF 采用了 KCF 的训练机制。在 KCF 中,对相关滤波器的求解实质是一个岭回归问题,即

$$\min_w \sum_i (\varphi(x_i) - y_i)^2 + \lambda \|w\|^2, \quad (4)$$

式中: λ 为正则化参数,用来控制训练的过拟合; w 为权重参数矩阵,用以描述样本 x 和评分 $f(x)$ 之间的关系。每一个样本 x_i 对应一个标记 $y_i \in [0, 1]$ 。通过各种垂直、水平的循环移位来构建训练样本矩阵 X ,并采用核函数 K 将目标 x 的低维特征空间映射到高维空间。滤波器 w 的求解空间则被转化到对偶空间 α 中,核化的岭回归表示为^[11]:

$$\alpha = (K + \lambda I)^{-1}y, \quad (5)$$

式中: I 为单位矩阵; K 是核矩阵,也是一个循环矩阵,由循环矩阵的性质可以继续化解为:

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}^{xx'} + \lambda}, \quad (6)$$

式中: $\hat{k}^{xx'}$ 为 K 的第一行元素组成的向量;所求的 $\hat{\alpha}$ 即为用于跟踪图像位置的相关滤波器系数。选用高斯函数作为核函数,对于核函数有 $k(x, x') = g(x^T, x')$, 则

$$k^{xx'} = g(F^{-1}(\hat{x}^* \odot \hat{x}')), \quad (7)$$

由高斯核函数的性质可以得到:

$$k^{xx'} = \exp\left(-\frac{1}{\sigma^2}(\|x\|^2 + \|x'\|^2 - 2F^{-1}(\hat{x}^* \odot \hat{x}'))\right). \quad (8)$$

1.3 模板更新机制

传统的相关滤波算法采用的是线性插值方法,该方法能够将最新的特征信息添加到训练模型中,从而简单有效地进行在线跟踪。但是对于实际应用中的复杂环境,相关滤波类算法仍存在着很大挑战,如长时跟踪产生的错误累积问题、目标遮挡时的模板漂移问题等。

针对以上问题,通过分析模板更新时对跟踪阶段的影响,提出了对长时目标跟踪的基本解决方案:①设计了最大响应值和响应模式的稳定度两个指标来反映跟踪器的跟踪置信度,以此避免目标在遇到遮挡和严重形变时的跟踪失败;②运用 TLD 算法^[12]思想将跟踪任务分解为跟踪、学习和检测三部分,以减轻在线模版更

新时的错误累积。检测器负责在检测到跟踪失败时重新初始化跟踪器,把跟丢的目标找回并将产生的误差清零,而跟踪器的结果则提供训练数据来更新检测器。

当新一帧进来时, s 为上一帧目标位置,当前目标位置为 $y = \varphi(s; w)$ 。其中, w 的求解转化到对偶空间 α 中,通过式(6)来计算。通过循环采样获得的整个响应图按下式计算^[13]:

$$F(s, y; w) = F^{-1}(\hat{\psi}^* \odot \hat{w}) = F^{-1}(\hat{k}^{y, s, 0} \odot \hat{\alpha}), \quad (9)$$

式中: \hat{k} 为岭回归所产生的封闭解的核矩阵; $\hat{\alpha}$ 为回归函数的向量。

为判断跟踪目标位置的精确度,设定了两个基准。第一个基准是响应图 $F(s, y; w)$ 中最大响应分数 F_{\max} , 定义为:

$$F_{\max} = \max F(s, y; w), \quad (10)$$

如果当前帧的 F_{\max} 小于第一个阈值 T_a (外观阈值),则调用重检测机制,得到新的响应图 F' 和相应的 F'_{\max} 。这时如果 $F'_{\max} < 1.5 F_{\max}$,则固定目标位置框,提取出历史模版的信息。在接下来的视频流中增加搜索窗口的范围,逐帧查找目标位置,直到历史信息的目标框与当前帧的搜索框之间的目标响应大于 T_a 为止。

高置信度更新机制的第二个判断基准为平均峰值相关能量 (average peak-to correlation energy)^[13],即 APCE 为

$$APCE = \frac{|F_{\max} - F_{\min}|^2}{\text{mean}(\sum (F_{i,j} - F_{\min})^2)}, \quad (11)$$

式中: F_{\max} 、 F_{\min} 以及 $F_{i,j}$ 分别为 $F(s, y; w)$ 的最大值,最小值和指定的第 i 行第 j 列的元素。如果当前帧中的 APCE 大于算法设定的第二个阈值 T_i (运动阈值),则认为当前帧中的跟踪结果是高置信度的。此时,相关滤波模版将会在固定的学习率参数 μ 下进行在线更新,更新方式如下:

$$\hat{x}' = (1 - \mu) \hat{x}^{t-1} + \mu \hat{x}'; \quad (12)$$

$$\hat{\alpha}' = (1 - \mu) \hat{\alpha}^{t-1} + \mu \hat{\alpha}', \quad (13)$$

式中: t 为当前帧的索引。

1.4 LKCF 算法流程

步骤 1 在首帧中标出欲跟踪目标的位置与尺寸,利用式(6)训练核相关滤波器中的参数;

步骤 2 提取下一帧的图像特征 f_t ,并用式(9)计算响应图 $F(s, y; w)$;

步骤 3 通过式(10)和(11)计算两个判断基准 F_{\max} 和 APCE;

步骤 4 如果 F_{\max} 小于第一个阈值 T_a 则调

用检测器 D_f 实现重检测机制,得到新的响应图 F' 和相应的 F'_{\max} 。这时如果 $F'_{\max} < 1.5 F_{\max}$, 则固定目标位置框,调用历史更新中最近的模版信息 f'_i , 加大目标搜索框的 padding 并重复步骤 2、3,计算得 F'_i 和相应的 F'_{\max} ,直到 $F'_{\max} > T_a$ 时,判断为找回跟踪目标 $f_i = f'_i$, 其中 $i = \operatorname{argmax}_i f'_i$;

步骤 5 保存目标位置信息 $s=f_i$;

步骤 6 采用 APCE 基准判断跟踪的置信度,若 $APCE \geq T_i$,则用式 (12) 和 (13) 来更新滤波模版 H_i , 否则转到步骤 2;

步骤 7 更新检测器 D_f , 转到步骤 2 直到视频结束。

2 实验与分析

实验采用 13 个目标跟踪基准数据库 OTB50 和 OTB100 中的具有挑战性的视频序列对提出的算法进行评测,序列涉及了运动模糊 (MB)、快速运动 (FM)、遮挡 (OCC)、背景杂乱 (BC) 和光照变化 (IV) 等多种因素。实验结果采用两种度量评估标准:

(1)成功重叠率 (OS),即预测边框与实际边框重叠超过阈值的帧占比,可利用有效帧的重叠率均值来计算;

(2)距离精度 (DP),即跟踪到的位置与目标实际位置的距离小于指定阈值的帧占比。

2.1 实验配置

实验用电脑的配置:CPU 为 Intel Core i7,内

存 4 GB,选择 Windows 10 操作系统,使用 MATLAB 7.0 来实现和验证算法。正样本空间 α 的窗口大小设置为目标大小的 1.8 倍,搜索窗口的大小设置为目标框空间 α 的 1.5 倍。训练过程

中的正则化参数 $\lambda = 10^{-4}$ 。高斯核的 $\sigma = \frac{\sqrt{mn}}{10}$,

其中 n 和 m 是指在特征单元中测量目标的宽度和高度。式 (12) 和 (13) 中学习率 $\mu = 0.01$ 。检测阶段中的阈值 $T_a = 0.25, T_i = 0.5$ 。

2.2 实验比较

在公共基准数据集 OTB100 上将 LKCF 与 7 种典型的跟踪算法进行了对比实验。7 个对比算法包括 Struck (structured output tracking with kernels)^[14]、TLD (tracking-learn-detection)^[12]、在线 AdaBoost 方法 (OAB)^[10]、KCF^[7] 以及依据在线分类器分类的跟踪算法 MIL (multiple instance learning)^[15]、CT (compressive tracking)^[8] 和 ASLA (adaptive structural local sparse appearance model)^[16]。采用距离精度和重叠成功率评估标准^[17],以 0.5 为重叠阈值的重叠率、20 个像素内的精度分数以及平均帧速来定量比较,如表 1 所示。结果显示,LKCF 算法的成功重叠率 (OS) 为 78.9%、距离精度 (DP) 为 86.6%,优于其他 7 个对比算法。虽然 KCF 的平均 OS 为 62.2%,DP 为 74.1%,但是 KCF 的帧速率 172 Hz 明显比 LKCF 的 34 Hz 快。这是由于 LKCF 算法的重新检测机制增长了检测时间,但在实际应用中仍可满足工程的实时要求。

表 1 与其他算法的跟踪结果对比
Table 1 Comparison with other algorithms

参数	LKCF	Struck ^[14]	OAB ^[10]	MIL ^[15]	CT ^[8]	ASLA ^[16]	KCF ^[7]	TLD ^[12]
DP/%	86.6	65.6	50.6	47.5	39.4	53.2	74.1	60.8
OS/%	78.9	55.9	39.4	37.3	34.1	51.1	62.2	52.1
帧速率/Hz	34	20	42	38	64	16	172	28

不同环境下目标跟踪效果的分析。将 LKCF 与其他 4 个目前最好的追踪器 KCF、TLD、IVT (incremental visual tracking) 和 Struck 在公共数据集 OTB50 上进行了对比,实验效果如图 1 所示。

在图 1(a)中,目标 David 经历光照变化,所有被测试的算法都可以稳定地跟踪目标,原因在于上述算法均使用了基于 PCA 的外观模型,可适应目标的外观变化。在图 1(b)和图 1(c)中,目标对象遭受了重度遮挡,此时只有 TLD 和

LKCF 方法可以完成跟踪,因为它们使用了重新检测机制。在图 1(d)的快速运动场景下,KCF、TLD 和 LKCF 方法表现良好,而 Struck 和 IVT 算法无法在帧#297 之后跟踪目标。在图 1(e)中,目标的运动方向和背景环境发生了很大变化,大多数算法均丢失了目标,只有 LKCF 算法能正确跟踪。在图 1(f)的杂波背景下,KCF 跟踪器由于仅使用了 HOG 特征,使跟踪目标产生了漂移。相比之下,LKCF 和 TLD 跟踪器由于采用了重新检测机制使得跟踪效果更好。

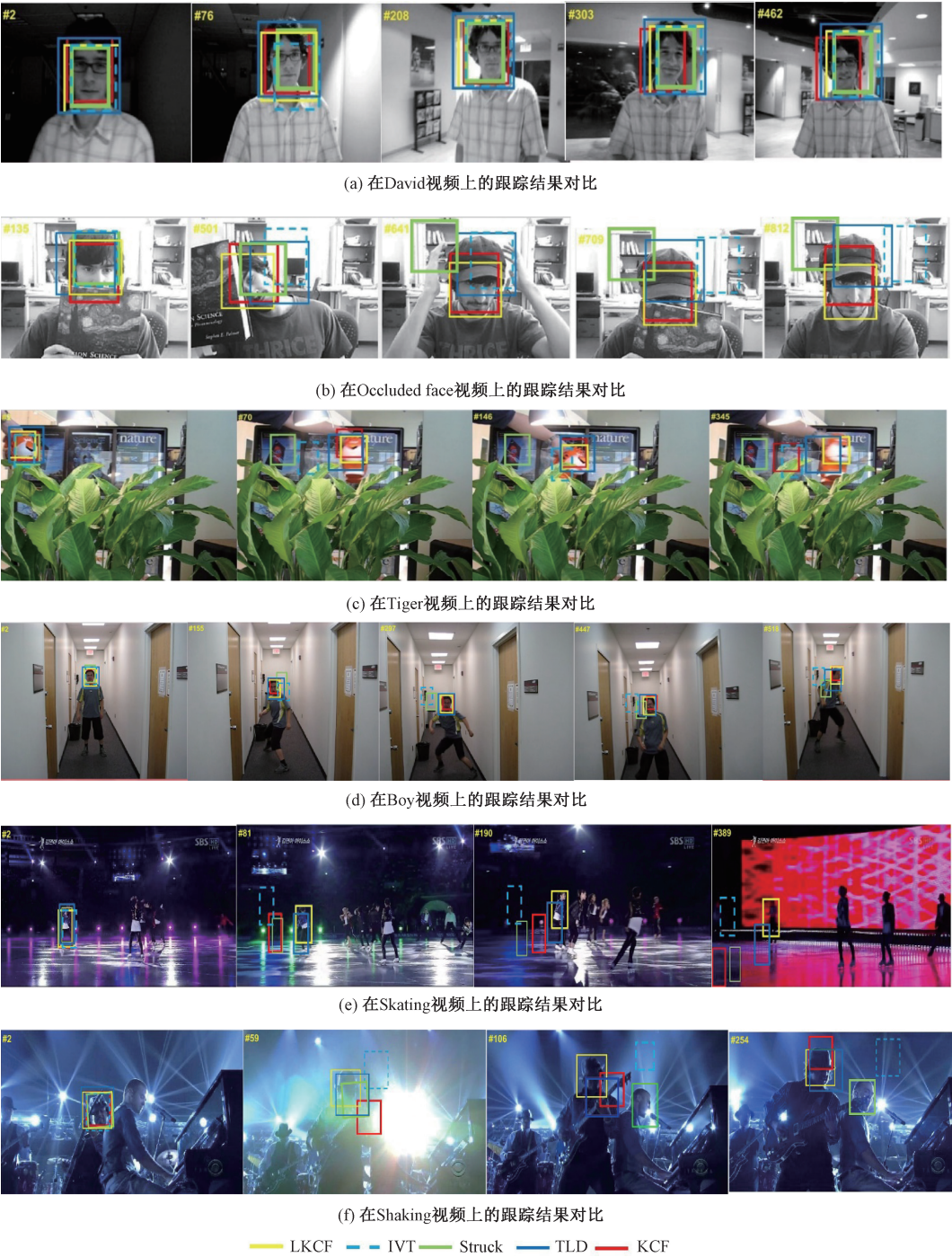


图 1 不同场景下的跟踪结果对比

Figure 1 Tracking results in different video scenes

总之,由于采用了重新检测机制和有条件的模板更新机制,笔者提出的 LKCF 算法有效避免了检测中的误差累积和模板漂移问题,完整地完成了跟踪任务。

3 结论

通过改进核相关滤波跟踪算法,实现了一种长时核相关滤波目标跟踪算法。①构建了高置信度模版更新机制,使长时目标跟踪中的模板漂移

问题得到较好解决;②构造了目标重检机制,以此解决由于遮挡造成的目标丢失问题。实验结果表明,所提算法在处理长时目标跟踪时,相较于传统算法在效率、准确性和鲁棒性方面表现更好。下一步工作将探索更富表现力的特征来描述目标对象,并扩展尺度变换模块来完善目标跟踪器。

参考文献:

[1] 毛晓波,周晓东,刘艳红.基于 FAST 特征点改进的

- TLD 目标跟踪算法[J].郑州大学学报(工学版), 2018,39(2):1-5.
- [2] 毛晓波,郝向东,梁静. 基于 ELM 与 Mean Shift 的抗遮挡目标跟踪算法[J].郑州大学学报(工学版), 2016,37(1):1-5.
- [3] HOU Z Q, LIU X, YU W S, et al. Mean-shift tracking algorithm with improved background-weighted histogram[C]//International Conference on Intelligent Systems Design & Engineering Applications.Zhangjiajie, China:IEEE, 2014:597-602.
- [4] 宋策. 基于粒子滤波的目标跟踪技术研究[D].长春:中国科学院大学,2014.
- [5] ZHANG K H, ZHANG L, YANG M H. Fast compressive tracking[J]. IEEE transactions on pattern analysis & machine intelligence, 2014, 36(10):2002-2015.
- [6] 张微,康宝生.相关滤波目标跟踪进展综述[J].中国图象图形学报, 2017,22(8):1017-1033.
- [7] HENRIQUES J F, CASEIRO R, MARTINS P, et al. High-speed tracking with kernelized correlation filters[J]. IEEE transactions on pattern analysis & machine intelligence, 2014, 37(3):583-596.
- [8] ZHANG K H, ZHANG L, YANG M H. Real-time compressive tracking[C]//European Conference on Computer Vision. Florence, Italy: Springer, 2012:864-877.
- [9] DANELLJAN M, HAGER G, KHAN F S, et al. Discriminative scale space tracking[J]. IEEE transactions on pattern analysis & machine intelligence, 2017, 39(8):1561-1575.
- [10] GRABNER H, GRABNER M, BISCHOF H. Real-time tracking via on-line boosting[C]//British Machine Vision Conference 2006. Edinburgh, England: British Machine Vision Association, 2006:47-56.
- [11] 任维雅,李国辉. 面向监督学习的稀疏平滑岭回归方法[J]. 国防科技大学学报, 2015, 37(6):121-128.
- [12] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection[J]. IEEE transactions on pattern analysis & machine intelligence, 2012, 34(7):1409-1422.
- [13] WANG M M, LIU Y, HUANG Z. Large margin object tracking with circulant feature maps[C]//IEEE Conference on Computer Vision. Honolulu, Hawaii, USA: IEEE, 2017:4800-4808.
- [14] HARE S, SAFFARI A, TORR P H S. Struck: Structured output tracking with kernels[C]//IEEE International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011:263-270.
- [15] BABENKO B, YANG M H, BELONGIE S. Visual tracking with online multiple instance Learning[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, Florida, USA: IEEE, 2009:983-990.
- [16] JIA X, LU H C, YANG M H. Visual tracking via adaptive structural local sparse appearance model[C]//IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA:IEEE, 2012:1822-1829.
- [17] VOJIR T, NOSKOVA J, MATAS J. Robust scale-adaptive mean-shift for tracking[J]. Pattern recognition letters, 2014, 49:250-258.

An Improved Kernelized Correlation Filter for Long-term Target Tracking

CHEN Liping, WANG Mingyu, YANG Wenzhu, WANG Sile, CHEN Xiangyang

(School of Cyber Security and Computer, Hebei University, Baoding 071002, China)

Abstract: At present, the research of long-term target tracking played a leading role in the field of target tracking. Due to the target fast moving and object occlusion, the template drift became one of the research issues in this field. To solve the above-mentioned problem, a long-term target tracking algorithm based on the improved Kernelized Correlation Filter was proposed. The kernelized correlation filter algorithm as the tracking framework was used. Furthermore, a highly reliable template update strategy as well as a conditional target re-detection mechanism was constructed to avoid the problem of template drift. Experimental results indicated that the proposed algorithm was more stable and reliable than the traditional algorithm in the long-term target tracking process.

Key words: long-term target tracking; template drift; kernelized correlation filter; confidence