

引用格式: 胡昭华, 黄嘉净. 基于快速多尺度估计的重新检测目标跟踪算法[J]. 微电子学与计算机, 2020, 37(9): 11-17. [HU Z H, HUANG J J. A re-detection object tracking algorithm based on fast multi-scale estimation [J]. Microelectronics & Computer, 2020, 37(9): 11-17.]

基于快速多尺度估计的重新检测目标跟踪算法

胡昭华^{1,2}, 黄嘉净¹

(1 南京信息工程大学 电子与信息工程学院, 江苏 南京 210044;

2 南京信息工程大学 江苏省大气环境与装备技术协同创新中心, 江苏 南京 210044)

摘 要: 为了解决跟踪过程中的漂移问题, 本文提出了一种基于快速多尺度估计的重新检测目标跟踪算法。在跟踪过程中, 跟踪器通过目标的最大响应值确定目标的位置, 并且通过一个新的自适应检测指标检测当前位置的可靠性。不同于其他检测指标的是, 新的检测指标减少了对最大响应值的依赖。若检测出当前帧的位置不可靠, 我们的方法可以通过 Edge boxes 方法产生一系列的目标候选框, 并通过非极大值抑制方法和欧几里德度量方法选出最优目标位置。此外, 我们提出了自适应更新的方法来减少由于跟踪失败导致的误差。实验结果表明, 本文提出的算法在精确度和成功率方面都有很好的效果。

关键词: 目标跟踪; 重新检测; 深度特征; 相关滤波; 快速多尺度估计

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2020)09-0011-07

A re-detection object tracking algorithm based on fast multi-scale estimation

HU Zhao-hua^{1,2}, HUANG Jia-jing¹

(1 School of Electronic & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China; 2 Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: To solve the problem of drift in the tracking, in this paper, a re-detection object tracking algorithm based on fast multi-scale estimation is proposed. In the tracking process, the tracker locates the target position through the maximum value of the response and detects the reliability of the current position using a new self-adaptive detection criterion. In contrast to other detection criteria, our new detection criterion reduces the dependence on the maximum response value. If the current location is determined to be unreliable, our method can generate target candidate boxes by using the edge boxes algorithm and select the best target location by applying the non-maximum suppression (NMS) and the Euclidean metric methods. Furthermore, we proposed an adaptive updating method to reduce the errors caused by tracking failure. Experimental results show that our approach has a good performance in terms of precision and success rates.

Key words: object tracking, re-detection; deep feature; correlation filter; fast multi-scale estimation

1 引言

目标跟踪是计算机视觉的一个重要的研究领域,在视频监控、无人驾驶和医疗图像等领域都有着广泛的应用. 尽管这十几年来,目标跟踪的研究已经取得了很大的进步,但在实际跟踪过程中目标可能发生形变、剧烈运动、被遮挡等情况,从而导致目标丢失、跟踪失败,因此,目标跟踪的研究仍是具有挑战性的.

文献[1-3]算法可以有效地解决尺度变化问题,目标特征表示的问题,但不能解决目标丢失并寻回目标的问题. 最小误差平方和跟踪算法通过峰值旁瓣比(Peak to Sidelobe Ratio, PSR)来检测目标是否发生遮挡或跟踪失败;Ting Liu 等^[4]提出置信度平滑约束(Smooth Constraint Confidence Maps, SCCM)来衡量不同子块的跟踪性能. Ma C^[5]等提出长时跟踪算法,结合目标检测的方法,当目标的最

大响应值低于某一阈值时,则认为跟踪失败,采用目标检测的方法继续搜索目标的位置.

因此,建立新的检测指标和重新检测机制是本文关注的重点. 本文在核相关滤波算法的基础上构建了一个新的自适应检测指标,新的检测指标可以更加准确的检测出当前帧是否出现跟踪失败的情况. 与其他指标不同的是,新的检测指标减少了对最大响应值得依赖. 当检测到目标丢失时,通过重新检测的方法重新寻回目标,恢复跟踪. 若当前帧需要进行重新检测,则当前帧的模型拒绝更新,仍采用前一帧的模型.

2 本文算法整体框架

本文提出的跟踪算法的整体框架图如图 1 所示. 主要包括训练阶段、定位阶段、检测阶段、重新检测阶段和更新阶段.

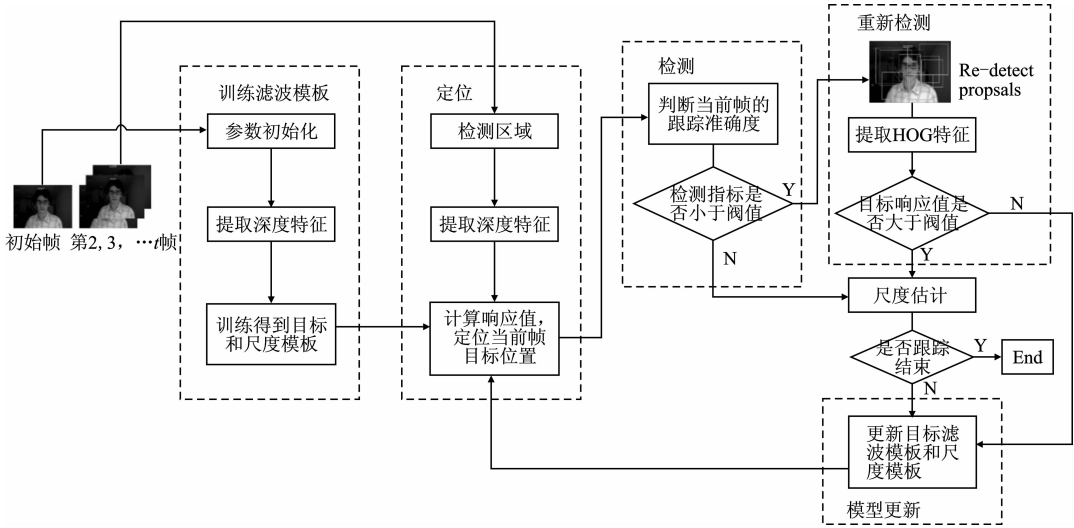


图 1 基于快速多尺度估计的重新检测目标跟踪算法整体框架图

3 重新检测算法

本文算法的具体细节包括滤波模板的学习、新检测指标的构建、重新检测的方法以及在线更新方式.

3.1 核相关滤波算法

核相关滤波算法可以分为三个部分来实现:训练、检测和更新.

(1) 训练:在第 t 帧中,分类器 $f(x)$ 以目标位置为中心进行循环移位采样,产生的样本记为 x_i ,每个样本 x_i 都有一个与其相对应的回归标签 y_i , y_i 由高斯函数得到. 分类器 $f(x)$ 的目的就是找到权重系数 w ,使得样本 x_i 及其回归标签 y_i 有最小的平方误

差,训练分类器的最小代价函数为

$$\min_w \sum_i [f(x_i) - y_i]^2 + \lambda \|w\|^2 \quad (1)$$

式中, λ 是正则化参数,用来防止过拟合. 得到 w 的闭式解为

$$w = (X^H X + \lambda I)^{-1} X^H y \quad (2)$$

利用所有的循环矩阵都能够在傅氏空间使用离散傅里叶矩阵进行对角化的性质进行优化,因此可以优化得到:

$$\hat{w} = \frac{\hat{x} \odot \hat{y}}{\hat{x} \odot \hat{x}^* + \lambda} \quad (3)$$

式中, $\hat{w}$ 、 $\hat{x}$ 、 $\hat{y}$ 分别代表 w 、 x 、 y 的傅里叶变换, $\odot$ 表示元素点乘.

同时,通过一个非线性映射函数 $\varphi(x)$ 将样本映

射到高维空间,并引入高斯核 $k^{xx} = k(x, x')$ 作为核函数,高斯核的计算公式为

$$k^{xx'} = \exp\left(-\frac{1}{\sigma^2} (\|x\|^2 + \|x'\|^2 - 2F^{-1}(\hat{x}^* \odot \hat{x}'))\right) \quad (4)$$

F^{-1} 表示离散傅里叶反变换, σ 表示高斯核函数的带宽, $*$ 表示复共轭。

利用核函数的循环性质和卷积理论,得到 α 的计算公式为

$$\hat{\alpha} = \frac{\hat{\gamma}}{\hat{k}^{xx} + \lambda} \quad (5)$$

(2) 检测:训练阶段完成后,以前一帧的位置为中心提取图像块,根据当前帧的图像块进行检测,计算相关响应 $f(z)$,确定目标的位置。

$$f(z) = F^{-1}((\hat{k}^{xx} \odot \hat{\alpha})) \quad (6)$$

(3) 更新:为了更好地适应目标的外观变化,当 $t > 1$ 时,对目标的外观模型 x 和滤波模板 α 进行更新。

$$\hat{\alpha}^t = (1 - \eta)\hat{\alpha}^{t-1} + \eta\hat{\alpha}^t \quad (7)$$

$$\hat{x}^t = (1 - \eta)\hat{x}^{t-1} + \eta\hat{x}^t \quad (8)$$

式中, η 表示模型的学习率, t 是帧数。

3.2 重新检测

3.2.1 检测指标

在跟踪过程中,可能会出现目标被遮挡、快速运动或者出视野等情况,从而导致目标丢失,整体跟踪效果下降,所以为了保证跟踪的精度,我们需要对跟踪结果进行判断。其中,最简单的就是计算目标的最大响应值 $F_{\max}$,一般而言, $F_{\max}$ 的值越大,则表明跟踪效果较好,当最大响应值 $F_{\max}$ 低于某一个阈值,则认为跟踪失败。但该依据并不可靠,如果在跟踪过程中出现与目标相似的物体或背景,此时依旧会有很高的响应值,但实际却已经出现漂移,如图2所示。图2(c)中“frame”表示帧数,“max response”表示当前帧的最大响应值,“error”表示当前帧的目标框与真实值之间的中心位置距离。以 Tiger2 序列的第67帧和205帧为例,虽然第205帧的目标响应值高于第67帧的目标响应值,但此时已经发生跟踪漂移的情况,由此可以看出,最大响应值检测指标并不可靠,当目标出现与目标相似的背景或物体时,依旧会有很高的响应值,从而导致检测结果出现误差。

故本文提出了一个新的检测指标 DC (Detection Criterion),如公式(9)所示,从而可以更好地度量当前帧的跟踪准确度。

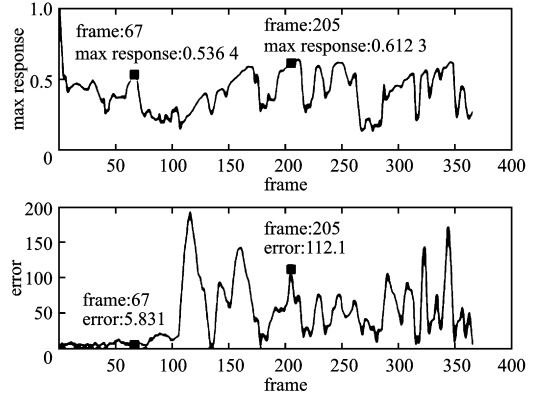
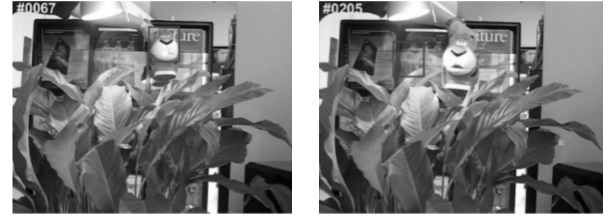


图2 最大响应值检测指标下的中心误差分析

$$DC_t = \gamma_{\text{PSR}}^t * \text{PSR}_t + \gamma_{\text{SCCM}}^t * \frac{1}{\text{SCCM}_t} \quad (8)$$

式中, DC_t 、 PSR_t 和 SCCM_t 分别表示当前帧的 DC、PSR 和 SCCM 值, γ_{PSR}^t 和 γ_{SCCM}^t 表示当前帧 PSR 检测指标和 SCCM 检测指标的权重。

γ_{PSR}^t 和 γ_{SCCM}^t 的计算方式如下:

$$\gamma_{\text{PSR}}^t = \exp(-\beta * (\text{PSR}_t - \text{PSR}_{t-1})^2) \quad (10)$$

$$\gamma_{\text{SCCM}}^t = \exp(-\beta * (\text{SCCM}_t - \text{SCCM}_{t-1})^2) \quad (11)$$

PSR_{t-1} 和 SCCM_{t-1} 分别表示前一帧的 PSR 和 SCCM 值, $\beta = 0.05$ 是惩罚因子。若当前帧的 PSR 值和前一帧的 PSR 值存在显著的差异,则存在目标被遮挡的可能性,此时,我们赋予 PSR 检测指标一个较小的权重,反之,则分配一个较大的权重。同理,SCCM 检测指标也是如此。

PSR 是峰值旁瓣比,其定义为:

$$\text{PSR}_t = \frac{F_{\max}^t - \mu}{\sigma} \quad (12)$$

$F_{\max}^t$ 表示当前帧目标的最大响应值, μ 和 σ 表示响应的均值和标准差。峰值旁瓣比可以量化相关峰的锐利程度, PSR 值越高,说明当前帧的候选目标与前一帧中的目标最接近。一般在正常跟踪的情况下,当 PSR 值的范围在 7~15 之间时,表示目标有较强的峰值^[5],而当 PSR 值在 5 左右时,当前帧的目标检测不太稳定。

SCCM 是置信度平滑约束,其定义为:

$$\text{SCCM}_t = \|F^t - F^{t-1} \oplus \Delta\|_2^2 \quad (13)$$

F^t 表示当前帧的目标响应值, F^{t-1} 表示前一帧的目标响应值, $\oplus$ 表示移位操作, Δ 表示相邻两帧

最大响应值位置的相对位移. 一般情况下, 相邻两帧的目标位置位移较小, 因而 SCCM 值也较小, 若 SCCM 值很大, 则说明该区域变化很大. 一般当 SCCM 值大于 0.5 时, 说明该区域出现遮挡^[5].

当我们同时考虑峰值旁瓣比和置信度平滑约束两个指标, 则可以提高跟踪器的检测能力, 同时减少对最大响应值的依赖. 本文中, 我们设定新的检测指标 DC 的阈值为 7, 当新的检测指标值小于阈值时, 说明当前帧的跟踪效果较差. 具体的跟踪效果将在本文 4.2.1 部分进行分析.

3.2.2 重新检测

若检测阶段检测出当前帧出现目标丢失的情况, 即满足 $DC < 7$ 这个条件时, 目标则进入重新检测阶段, 重新寻回并恢复当前帧的目标.

在重新检测阶段, 首先通过 EdgeBoxes 算法在目标周围产生一系列的候选样本框集合, 然后通过一系列的优化操作, 选出最优目标候选框. 具体的步骤如下:

首先, 提取候选框的 HOG 特征并根据公式(6)计算当前帧生成的候选样本框集合的得分, 当确定了得分后, 对这些候选框的得分进行排序, 选取得分较高的 N 个候选框.

其次, 利用非极大值抑制 (Non-maximum Suppression, NMS) 的方法过滤这些候选框, 将候选框放入列表 M , 选择具有最大得分的候选框 X , 将候选框 X 从 M 集合中移除并加入到最终检测结果 Y 列表中, 计算 M 中剩余的候选框与 X 的重叠区域面积比例 Intersection over Union, IoU, 将 IoU 值大于阈值 N_i 的框从 M 中移除, 重复这个过程, 直到 M 为空. 至此所有的候选框都过滤结束.

最后, 通过欧式定理计算所筛选出的每个候选框与目标框之间的距离, 选出距离最小的候选框即为最优的目标候选框.

至此, 重新检测部分全部结束, 将重新检测的结果反馈至相关滤波继续跟踪.

3.3 快速尺度自适应估计

核相关滤波算法采用固定大小的窗口进行目标的跟踪, 但在跟踪过程中, 目标的尺度可能会发生变化. 如果目标变大, 但跟踪窗口不变, 则会丢失很多目标的边缘信息; 如果目标变小, 跟踪窗口不变, 则又会包含许多背景信息, 从而导致跟踪的精度下降. 因此, 在跟踪的过程中, 根据目标尺度的变化调整跟踪窗口的大小显得尤为重要.

本文采用快速尺度自适应估计的方法, 引入平

移相关滤波器和尺度相关滤波器依次进行目标的位置估计和尺度估计, 两个滤波器是相互独立的, 可以利用不同的特征种类和特征计算方式进行训练, 在提取不同尺度图像块的特征时, 我们添加了主成分分析法, 用以减少维度, 降低计算量, 提高算法的整体计算速度.

3.4 模型更新

在跟踪过程中若是对每一帧都进行更新, 一旦出现跟踪失败的情况, 后面将会无法继续跟踪. 为了避免错误, 本文提出自适应更新的方法, 当检测出当前帧的跟踪效果较差时, 模型拒绝更新, 仍采用上一帧的模板. 具体的更新策略如下:

$$\hat{\alpha} = \begin{cases} (1 - \eta)\hat{\alpha}^{t-1} + \hat{\eta}^t & DC \geq 7 \\ \hat{\alpha}^{t-1} & DC < 7 \end{cases} \quad (14)$$

$$\hat{x} = \begin{cases} (1 - \eta)\hat{x}^{t-1} + \hat{\eta}^t & DC \geq 7 \\ \hat{x}^{t-1} & DC < 7 \end{cases} \quad (15)$$

4 实验结果与分析

4.1 实验评估指标与参数设置

为了验证算法的可行性, 该算法的参数设置如下: 正则化参数 λ 设置为 $1e-4$, 高斯标签函数的空间带宽 σ 设置为 0.5, 公式(14)和(15)中的学习速率为 0.02, 候选目标框的最大值设置为 200.

为了更好地测试算法的性能, 我们从目标跟踪基准 (Object Tracking benchmark, OTB) 数据集中选取了 60 个具有挑战性的视频序列进行试验. 这些序列包括光照变化 (Illumination Variation, IV)、快速运动 (Fast Motion, FM)、平面内旋转 (In-Plane Rotation, IPR)、平面外旋转 (Out-Plane Rotation, OPR)、低分辨率 (Low Resolution, LR) 等 11 种挑战因子. 另外, OTB2013 提供了两种评价指标: 精确度和成功率. 精确度是指目标框与真实值之间的中心距离误差小于某一阈值的帧数占总视频帧数的百分比; 成功率是指目标框与真实值之间的重叠率大于某一阈值的帧数占总视频帧数的百分比.

4.2 实验

为了检验本文算法的性能, 本文进行了以下两组实验: (1) 不同检测指标下算法性能的分析; (2) 本文算法与其他常见算法的比较和分析.

4.2.1 不同检测指标下算法性能的分析

为了验证本文提出的检测指标的有效性, 我们对三种不同检测指标下的算法进行定量分析. 将本文的算法简称为 RDKCF_DC, 将使用 PSR 指标的算法简称为 RDKCF_PSR, 将使用 SCCM 指标的算

法简称为 RDKCF_SCCM.

图 3 是 OTB 数据集中的序列在不同检测指标下的精确度和成功率曲线图. 从图中可以看出,RD-KCF_DC 的成功率达到 0. 569,精确度达到 0. 884,精确度和成功率都优于另外两种检测,证明了我们所提出的检测标准可以更加准确地检测当前目标跟踪的准确度,从而提高跟踪的准确率,并且新的检测指标对于大多数的序列都有良好的适应性.

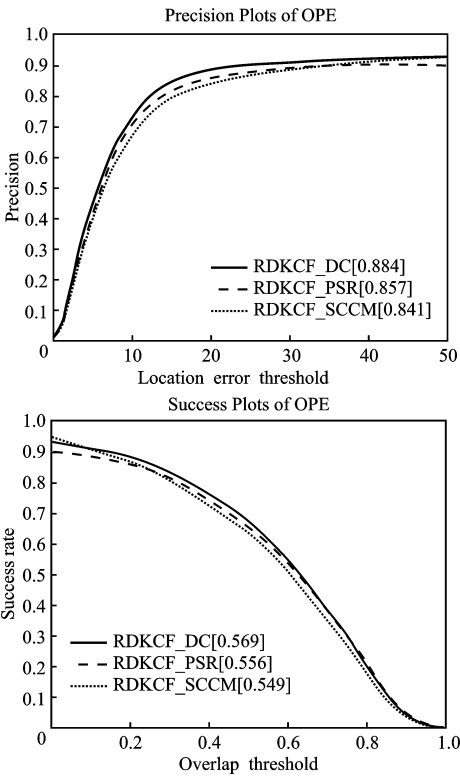


图 3 不同检测指标下的跟踪精确度和成功率曲线图

4. 2. 2 本文算法与其他常见算法的比较和分析

为了验证本文所提出的算法的有效性,同时对比和评估不同算法的性能,本文选取了其他 7 种算法进行实验的对比,包括 SAMF^[6]、KCF^[2]、CSK^[7]、DFT^[8]、CT^[9]、CACF^[10]、DSST^[3]等.

该部分实验先对算法做定性分析,从实验的 60 个序列中选取 6 个包含不同挑战因子的视频序列,分别为 (a) Bird1, (b) Blurowl, (c) Bolt2, (d) Couple, (e) DragonBaby, (f) Human7. 图 4 是本文算法与其他常见算法的跟踪结果比较. 从图 4(a) 中可以看出,第 20 帧时,CSK 算法和 CT 算法已经出现漂移,其他算法依旧可以有效的跟踪目标;第 191 帧是鸟经过云层之后的图片,此时大部分算法 (CSK, SAMF, CT, DFT, KCF, DSST) 都已经出现漂移,但是本文的算法依旧可以准确的跟踪,并且

本文算法可以根据目标大小变化相应的调整目标框的大小;在第 287 帧前,由于目标发生了快速的运动,所有算法都出现了漂移;然而从第 345 帧可以看出,本文算法通过重新检测机制又恢复了准确的跟踪,而其他算法都已经跟踪失败,说明了本文所提出的重新检测方法的有效性. 在 Blurowl, Couple, DragonBaby 的视频序列中,尽管本文算法和其他算法一样,在遇到一些较难的挑战因子时会出现漂移的情况,但只有本文算法可以重新寻回目标,并长时间有效的跟踪目标. 从这些序列的跟踪结果中可以看出,我们的重新检测方法可以克服跟踪过程的一些复杂的挑战.

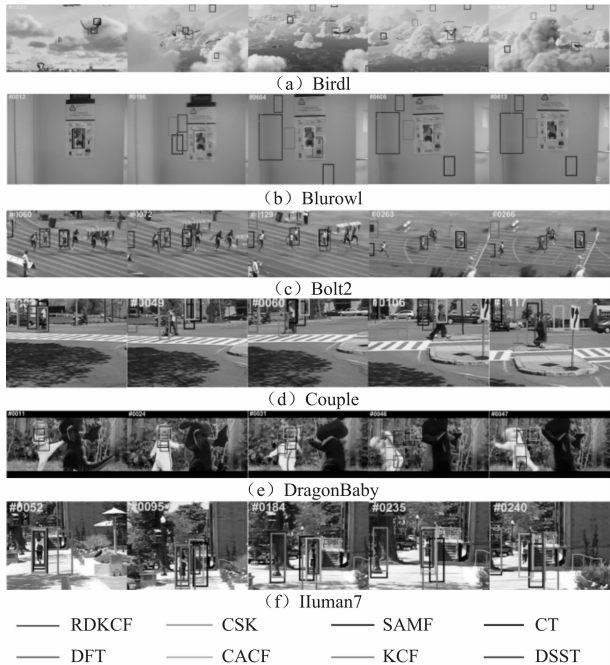


图 4 本文算法与其他常见算法的跟踪结果比较

其次,为了评估本文算法的整体性能,我们对实验算法进行定量分析. 分别绘制实验算法的精确度曲线图和成功率曲线图. 从图 5 中可以看出,本文算法 RDKCF 的成功率达到 0. 586,精确度达到 0. 874,算法的整体性能优于另外几种算法. 本文算法与 SAMF 算法都引入了尺度估计,在成功率上这两个算法都优于其他非尺度估计的算法. 本文算法之所以表现良好,主要有三个原因:首先,本文算法使用了目标的深度特征;其次,算法引入了快速多尺度估计的方法;最后,新的检测指标 DC 和重新检测机制被用来检测和恢复丢失的目标.

为了分析本文算法面对不同挑战因子的跟踪性能,我们选取了三组不同挑战因子的精确度曲线图和成功率曲线图来进行分析,结果如图 6 所示. 在这些

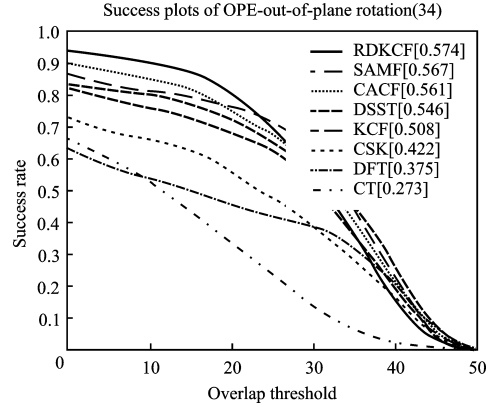
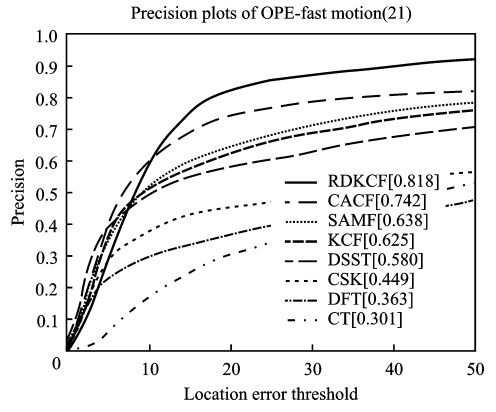
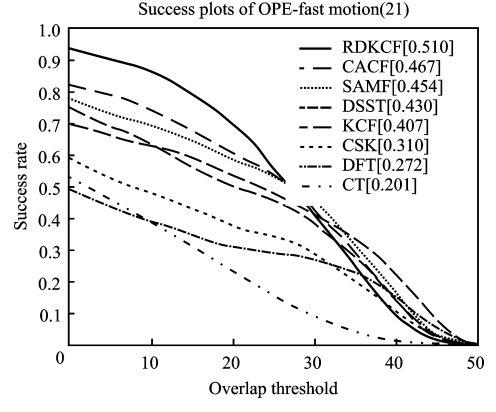
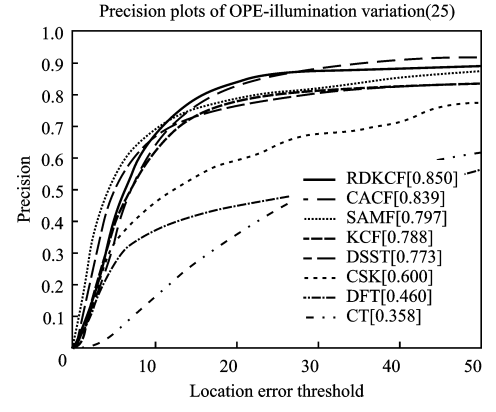
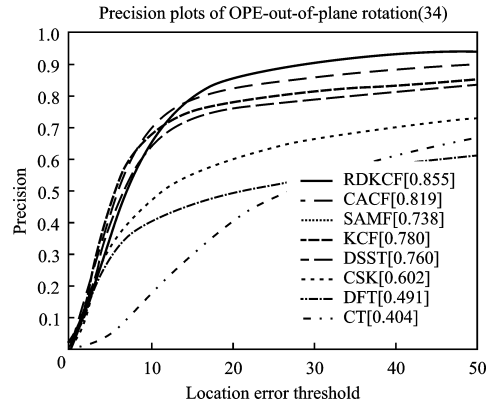
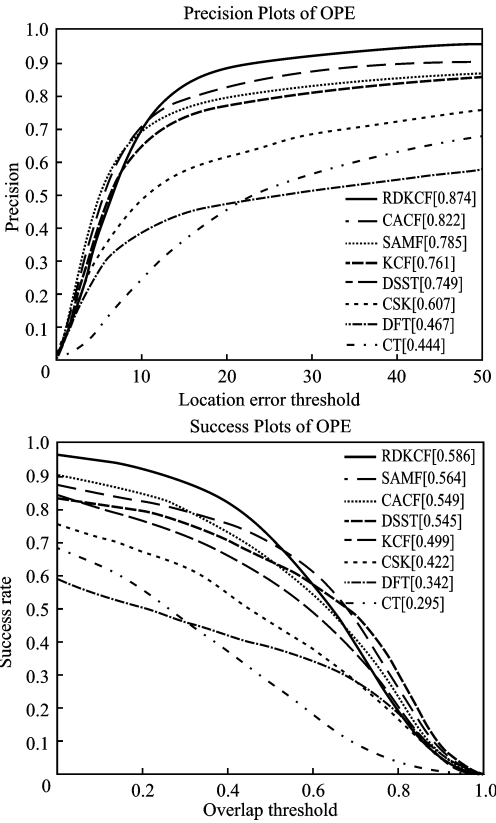


图 5 不同算法的精确度与成功率曲线图

算法中,CACF 算法在快速运动(74.2%)、平面外旋转(81.9%)的精度方面表现良好,而本文的 RDKCF 算法的精度分别达到了 81.8%、85.5%。SAMF 算法在快速运动(46.7%)、平面外旋转(56.7%)的成功率方面表现良好,而 RDKCF 算法的成功率分别为 51.0%、57.4%。但是,每种算法都有其局限性或不足之处,从图 6(c)和图 6(f)可以看出,RDKCF 算法在面对光照变化时的成功率低于 SAMF 和 DSST 算法,其精度也低于 CACF 算法,所以本文的算法不太擅长跟踪光照变化明显的这类视频序列。

(a)

(e)

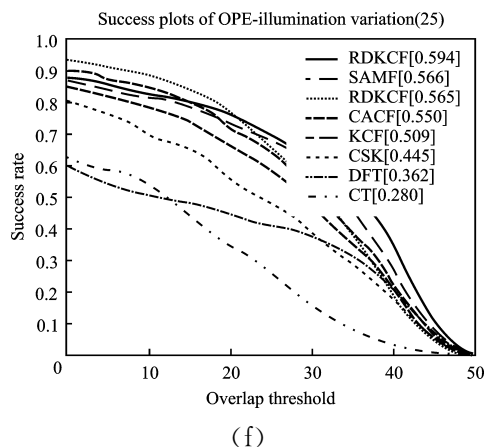


图 6 不同算法对不同挑战因子的性能评估

5 结束语

本文在原有的核相关滤波算法的基础上添加了检测机制和重新检测机制,我们构建了一种新的检测指标,从而可以更加准确的判断当前帧目标跟踪的准确度.若当前帧的跟踪结果不可靠,则通过重新检测的方法重新寻回目标,继续跟踪.另外,本文使用了深度特征,提高了目标特征的表达能力.并在原有平移滤波器的基础上添加了尺度滤波器,从而解决目标的尺度变化问题.本文算法在精确度和成功率上都取得了不错的效果,但在解决光照变化方面仍面临着一定的挑战.在今后的工作中,将主要解决这方面的跟踪问题.

参考文献:

- [1] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters[C]// The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010. San Francisco, CA, USA. IEEE, 2010: 2544-2550.
- [2] HENRIQUES J F, RUI C, MARTINS P, et al. High-Speed Tracking with Kernelized Correlation Filters[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(3):583-596.

- [3] DANELLJAN M, HÄGER G, KHAN F, et al. Accurate scale estimation for robust visual tracking[C]// British Machine Vision Conference. Nottingham, Nottingham, UK, 2014. BMVA Press, 2014.
- [4] LIU T, WANG G, YANG Q. Real-time part-based visual tracking via adaptive correlation filters[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, Massachusetts, 2015: 4902-4912.
- [5] BOSTAN MASADUSETTS. Long-term correlation tracking[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Boston, Massachusetts, 2015: 5388-5396.
- [6] LI Y, ZHU J. A Scale Adaptive Kernel Correlation Filter Tracker with Feature Integration[C]// European Conference on Computer Vision. Springer, Cham, 2014: 254-265.
- [7] JOÃO F HENRIQUES, CASEIRO R, MARTINS P, et al. Exploiting the Circulant Structure of Tracking-by-Detection with Kernels[C]// Proceedings of the 12th European conference on Computer Vision - Volume Part IV. Springer, Berlin, Heidelberg, 2012: 702-715.
- [8] SEVILLA LARA L, LEARNED MILLER E. Distribution fields for tracking[C]//2012 IEEE Conference on computer vision and pattern recognition. Providence, Rhode Island, USA, IEEE, 2012: 1910-1917.
- [9] ZHANG K, ZHANG L, YANG M H. Real-time compressive tracking [C]//European conference on computer vision. Springer, Berlin, Heidelberg, 2012: 864-877.
- [10] MUELLER M, SMITH N, GHANEM B. Context-aware correlation filter tracking[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, Hawaii cite, 2017: 1396-1404.

作者简介:

胡昭华 女, (1981-), 博士, 副教授. 研究方向为视觉跟踪、模式识别. E-mail: zhaohua_hu@163.com.

黄嘉净 女, (1994-), 硕士研究生. 研究方向为目标跟踪、机器学习.