

融合多层卷积特征自适应更新的目标跟踪算法

曾梦媛, 尚振宏*, 刘辉, 李健鹏

昆明理工大学信息工程与自动化学院, 云南 昆明 650500

摘要 针对传统手工特征表达能力不足和滤波器模型存在误差累积的影响, 提出一种融合多层卷积特征自适应更新的目标跟踪算法。该算法采用分层卷积神经网络提取图像特征, 利用线性加权的方法融合多层卷积特征预测目标位置; 利用多尺度下目标卷积特征确定目标最佳尺度; 利用平均峰值相关能量评价目标响应的置信度, 根据相邻两帧目标图像的帧差均值和位移评估目标的运动情况, 根据预测位置可信度和目标图像外观变化, 调整滤波器模型的学习率。在 OTB-2013 公开测试集上验证本算法性能, 并与现有基于相关滤波的主流运动目标跟踪算法进行相比, 实验结果表明, 本算法在精度和成功率上表现更优, 且在复杂情况下稳健性更强。

关键词 图像处理; 目标跟踪; 相关滤波; 卷积特征; 多尺度; 自适应更新; 学习率

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP57.021008

Target Tracking Algorithm Based on Adaptive Updating of Multilayer Convolution Features

Zeng Mengyuan, Shang Zhenhong*, Liu Hui, Li Jianpeng

Faculty of Information Engineering and Automation, Kunming University of Science and Technology,
Kunming, Yunnan 650500, China

Abstract Herein, we propose a target-tracking algorithm based on adaptive updating of multilayer convolutional features to address the insufficiency of traditional manual feature expression and the error accumulation of filter models. First, the algorithm uses a layered convolutional neural network to extract the image features, and fuses multi-convolution features through linear weighting to predict the target position. Then, the multiscale target convolution features are used to determine the target optimal scale. Finally, the average peak correlation energy is used to evaluate the confidence of the target response. We evaluate the motion condition of the target according to the frame differential mean and displacement of the two adjacent frames of the target image, and adjust the learning rate of the filter model according to the predicted position credibility and the appearance of the target image. The performance of the algorithm is verified using the OTB-2013 public test set and compared with the existing mainstream moving target tracking algorithm based on correlation filtering. Experimental results show that the proposed algorithm provides higher accuracy and success rate, and is more robust in complex cases.

Key words image processing; target tracking; correlation filtering; convolution features; multi-scale; adaptive updating; learning rate

OCIS codes 100.4550; 100.4999; 150.1135

1 引言

目标跟踪是计算机视觉中的一个重点研究方向, 包含机器学习、事件检测^[1]、信号处理、视频监控^[2]、自动驾驶、统计学等相关方面知识。虽然目标

跟踪这些年取得了很大进展, 但由于跟踪目标存在诸如遮挡、快速运动、模糊、光照及变形等^[3-4]因素, 相关滤波跟踪算法仍面临巨大挑战^[5]。

自 Bolme 等^[6]提出一种具有高效实时性的最小化输出平方误差和(MOSSE)相关滤波算法以来,

收稿日期: 2019-05-28; 修回日期: 2019-06-27; 录用日期: 2019-07-02

基金项目: 国家自然科学基金(61462052)

* E-mail: shangzhenhong@126.com

各种基于 MOSSE 改进的算法层出不穷。特征是决定跟踪效果的重要因素之一, Henriques 等^[7]提出一种基于方向梯度直方图特征(HOG)的目标跟踪算法;受此启发, Danelljan 等^[8]又提出一种基于颜色特征(CN)的目标跟踪算法,使算法精度得到提升。目标跟踪中另外一个重要问题就是尺度^[9], Danelljan 等^[10]在 2017 年提出一种 DSST (Discriminative Scale Space Tracker)跟踪算法,此算法的最大优点在于尺度滤波器移植性良好,可以推广到其他算法中;另一种由 Li 等^[11]提出的融合 HOG 和 CN 特征的 SAMF (Scale Adaptive Kernel Correlation Filter)算法,很大程度上提高了跟踪精度。尽管有些融合传统特征和自适应调节尺度的算法使跟踪性能得到提高,但传统特征存在不能有效捕捉目标物体的语义信息的缺点。

最近,随着深度学习兴起,卷积神经网络在目标跟踪领域得到了很大成功。Ma 等^[12]提出融合多层卷积特征的响应,结合目标的空间和语义信息的 HCFT/CF2 (Hierarchical Convolutional Features for visual Tracking)算法,使跟踪得到较为稳健的结果。Danelljan 等^[13]提出使用灰度特征以及 VGG 网络提取第一和第五层的卷积特征的 C-COT (Continuous Convolution Operators for visual Tracking)算法,通过插值方式得到精细的目标位置,但跟踪速度极慢。为了提高跟踪的速度, Danelljan 等^[14]提出分解卷积,在 C-COT 算法基础上大大减少了模型参数。

除了对特征方面进行改进之外,还有许多算法通过选择性更新滤波模型来提高跟踪精度。Yin 等^[15]针对滤波器难以跟踪快速运动的目标以及存在误差累积问题,提出了一种融合均值漂移法和帧差法的新方法。Wang 等^[16]针对当前帧是否出现遮挡和目标丢失,提出一种平均峰值相关能量(APCE)的方法。基于相关滤波的跟踪算法虽然高效,但在跟踪过程中,目标会发生不同程度的变化,而用于模版更新的学习率是一个固定值,将在这种情况下提取的目标样本用于更新会造成误差累积,影响分类器性能。

尽管从预测目标可靠性和滤波器更新策略方面去改进,能够使滤波器实时获得更新以适应目标变化,使跟踪速度有所提升,但无差别地更新同时也引入了错误信息。因此,本文从特征选取、模版更新和学习率三方面出发,提出一种融合多层卷积特征自适应更新的目标跟踪算法。选取包含目标丰富的空

间信息和语义信息的卷积特征作为图像特征,传统手工特征因缺乏语义信息,不能正确区分目标和背景,卷积特征能克服传统手工特征中语义信息不足的缺点,提高跟踪器的跟踪精度;采用一种多指标学习率自适应调节相关滤波器模版的更新策略,提升跟踪实时性;根据平均峰值相关能量评估预测目标的置信水平,判断当前帧是否出现了遮挡和目标丢失,选择性地更新模型,利用图像帧差均值^[17]和位移^[18]判断目标运动变化情况,选择合适的学习率,提高模型可靠性,解决现有基于卷积特征的相关滤波跟踪算法中滤波器模型误差累积的影响,提高跟踪精度。

2 基于卷积特征及相关滤波的跟踪方法

2.1 卷积特征

特征是影响跟踪效果的关键因素,鉴于传统手工特征缺乏语义信息,而卷积神经网络是一个带有卷积结构的多层神经网络,图像经过预处理提取的卷积特征包含丰富的语义信息,因此本文选取卷积神经网络提取目标特征。

根据文献[7]所述,VGG-19 网络大致分为五层,不同的卷积层包含着不同的信息。从低层到高层,语义信息增加,分辨率减少。鉴于 VGG 网络是金字塔状结构,第三、第四层处于网络中段,可以使用第三、第四层输出的卷积特征辅助第五层卷积特征来逐层精确预测目标的位置。图 1 所示为目标在 VGG-19 网络不同卷积层的输出图像,可以看出,图(b)空间信息丰富,图(d)的语义信息最为丰富,图(c)效果居中。为此设计了三个相关分类器,采用线性加权方法将其融合,以便提取优质的图像特征。

2.2 相关滤波跟踪算法

本文在相关滤波算法框架下,设计了四个部分:分类器训练、位置估计、尺度估计和模型更新。

2.2.1 分类器训练

训练滤波器时,通过建立最小化代价函数为每层神经网络训练一个位置滤波器,得到最优相关滤波器

$$\epsilon = \left\| \sum_{l=1}^d \mathbf{h}^{(l)} * \mathbf{f}^{(l)} - \mathbf{g} \right\|^2 + \lambda \sum_{l=1}^d \left\| \mathbf{h}^{(l)} \right\|^2 \quad (1)$$

式中: $\mathbf{f}^{(l)}$ 为一个 d 维特征向量, d 代表所选特征的维数, $l \in \{1, 2, \dots, d\}$;每个通道对应的滤波器为 $\mathbf{h}^{(l)}$; $*$ 为卷积运算;上标 l 为特征的某一维度; \mathbf{g} 为

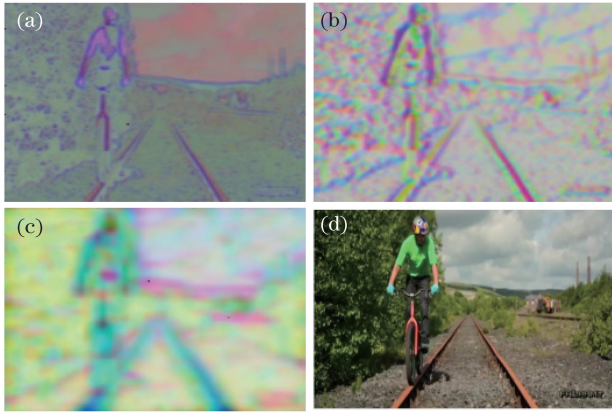


图1 目标在VGG网络中的输出图像。(a) Input;
(b) conv3-4;(c) conv4-4;(d) conv5-4

Fig. 1 Output images of target in VGG network.

(a) Input; (b) conv3-4; (c) conv4-4; (d) conv5-4

理想高斯函数输出; λ 为正则化系数, λ 的作用一方面是消除 f 频谱中零频分量的影响,另一方面是避免滤波器过拟合。利用(1)式可在频域中求解得到

$$H^{(l)} = \frac{\overline{G}F^{(l)}}{\sum_{k=1}^d \overline{F^{(k)}}F^{(k)} + \lambda} = \frac{A_t^{(l)}}{B_t}, \quad (2)$$

式中: $H^{(l)}$ 、 $F^{(l)}$ 、 G 为对应变量的频域形式,上划线表示该变量的共轭复数; $A^{(l)}$ 、 B 为相关滤波器函数的分子和分母;下角标 t 为当前帧数。

2.2.2 位置估计

滤波器训练完成后对下一帧图像中的目标进行位置估计,提取测试样本的多层卷积特征 Z 后,分别与对应层的位置相关滤波器卷积,生成目标位置的空间响应,可表示为

$$y = \mathcal{F}^{-1} \left\{ \frac{\sum_{l=1}^d [\overline{A^{(l)}}Z^{(l)}]}{B + \lambda} \right\}, \quad (3)$$

式中: \mathcal{F}^{-1} 表示傅里叶逆变换, Z 为新一帧图像中输入的候选样本的频域描述。归一化每层空间响应,将空间响应统一归一化到 $[0,1]$ 区间。最后将归一化后的空间响应 y'_c 按照权重 a 线性融合, a_c 表示每层对应的融合权重, c 代表的是卷积层层数。得到最终目标的空间响应 y ,新的位置由响应 y 的最大值确定:响应的表达式为

$$y = \sum_{c=1}^3 a_c y'_c. \quad (4)$$

2.2.3 尺度估计

在对目标进行尺度估计时,由于 VGG-19 的第五层是最后一层卷积层,其输出的特征语义信息最丰富,在目标识别和定位的性能上有较大优势,因此

采用卷积神经网络的第五层提取图像特征。目标尺度估计过程与位置滤波器预测目标位置的过程类似,在位置滤波器的预测位置以当前中心位置为中心点利用卷积层第五层提取图像特征,采集多个尺度候选样本,假设输入样本大小为 $P \times R$,目标最佳尺度由输入样本与一维尺度滤波器进行卷积得到的尺度响应最大值确定,评估目标尺度样本的选择原则为

$$a^n P \times a^n R, n \in \left(\left| -\frac{S-1}{2} \right|, \dots, \left| \frac{S-1}{2} \right| \right), \quad (5)$$

式中: a 为步长因子; P 、 R 分别为目标在前一帧的宽和高; S 为尺度数量; n 为尺度值。

2.2.4 模型更新

对于跟踪任务来说,目标始终处于变化之中,需对位置和尺度滤波器进行更新以适应目标外观及尺度的变化。为解决图像中的每个像素点要计算 $d \times d$ 维的线性方程的问题,可以对(2)式中分子 $A_t^{(l)}$ 和分母 B_t 分别进行更新,即

$$\begin{cases} A_t^{(l)} = (1 - \eta)A_{t-1}^{(l)} + \eta \overline{G}_t F_t^{(l)} \\ B_t = (1 - \eta)B_{t-1} + \eta \sum_{k=1}^d \overline{F_t^{(k)}} F_t^{(k)} \end{cases}, \quad (6)$$

式中: η 为学习率。位置滤波器和尺度滤波器均按照(6)式进行更新。对每一帧图像完成位置预测和尺度估计后,在每一个卷积层均按上述方式更新位置滤波器和尺度滤波器。在出现遮挡、超出视野等情况时,现有的算法会提取目标样本用于更新,而学习率又是一个固定值,会造成误差累积。根据文献[12],平均峰值相关能量能够检测到响应的震荡程度,以此描述目标响应的可靠性,判断当前帧是否出现了遮挡和目标丢失,选择性地更新模型;帧差均值能够评估样本目标图像的运动情况;目标位移能区分目标和背景的运动情况。所以在平均峰值相关能量、帧差均值和目标位移三个条件下,滤波器模版能区分多种跟踪状态,为模型选择合适的学习率 η 。

3 自适应选择学习率的模型更新算法

由于模型更新是适应目标外观变化的重要环节。传统的更新方式在每一帧跟踪完毕后按当前帧内容对滤波器进行线性插值更新,这种更新方式会造成最近几帧内容对滤波器的影响较大。在测试视频中,第一帧中的目标是给定的,后续帧中的目标位置是利用滤波器预测出来的,因此预测目标的可靠性对是否更新滤波模型尤为重要。现有的一些算法

采用隔几帧更新滤波器的策略;还有一些算法根据预测目标的可靠性更新滤波器的策略,目标可靠性高,则更新模型,反之则不更新。但现有算法的学习率都是一个固定值,只能适应某些情况下目标外观的变化,对于一些复杂情况下目标外观发生的变化,滤波器则不能很好地追踪到目标。本文选用平均峰值相关能量评估预测位置的可靠性,平均峰值相关能量小于预设的阈值表示跟踪失败,不需要更新滤波器,学习率保持不变;平均峰值相关能量大于预设阈值则表示定位成功,需更新滤波器。用帧差均值评估目标和背景的运动情况,帧差均值小于预设的阈值表明目标和背景的变化不大,只需设置较小的学习率;反之,则需用目标位移区分是目标发生变化还是背景发生变化,针对不同情况选用不同的学习率。因此本文结合预测响应图的平均峰值相关能量评估模型的可靠性,以及相邻两帧目标图像的帧差均值和位移评估目标外观变化,自适应地为滤波模型选择合适的学习率。具体步骤如下:

Step1 计算当前帧的平均峰值相关能量。第 t 帧目标的平均峰值相关能量 $a_{APCE}^{(t)}$ 计算方法为

$$a_{APCE}^{(t)} = \frac{|F_{\max} - F_{\min}|^2}{\text{mean} \left[\sum_{w,h} (F_{w,h} - F_{\min})^2 \right]}, \quad (7)$$

式中: F_{\max} 和 F_{\min} 分别为空间响应中的峰值和最小值; $F_{w,h}$ 为目标空间响应中位置 (w, h) 上的响应,该指标可以反映目标响应的震荡程度。

Step2 记两帧灰度图像 f_n 和 f_{n-1} , 这两张图像内的像素分别为 $f_n(x, y)$, $f_{n-1}(x, y)$, 则差分图像 D 的计算方式为

$$D(x, y) = |f_n(x, y) - f_{n-1}(x, y)|. \quad (8)$$

帧差分法得到的差分图能够体现目标或背景的变化情况,可用帧差均值衡量该变化的情况。帧差均值 e 的计算式为

$$e = \frac{\sum_{x,y}^{M,N} D(x, y)}{MN}, \quad (9)$$

式中: $D(x, y)$ 为相邻两帧的差分图像; M, N 为差分图像的大小。差分均值越大表明目标或背景变化越大,反之表明目标或背景变化幅度越小。

Step3 帧差均值能够体现样本图像的运动情况,但不能区别目标和背景,因此需要区别目标和背景的运动情况。通过跟踪算法给出目标的目标预测位置以估计目标位移。以 (x_{t-1}, y_{t-1}) 为上一帧的目标位置, (x_t, y_t) 为当前帧预测位置,计算得到两个位置之间的欧氏距离 s 以判断目标的运动状态,

欧氏距离的表达式为

$$s = \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2}. \quad (10)$$

如果相邻两帧之间预测位置的欧氏距离 s 较大,则说明目标可能处于快速运动之中,反之则表示目标运动幅度不大。

Step4 自适应选择学习率。对于预测目标的可靠性可以根据平均峰值相关能量来判断。

$a_{APCE}^{(t)} < \alpha$ 表示目标丢失或定位目标失败,不需要更新滤波器,学习速率为 $\eta = \eta_1$ 。

$a_{APCE}^{(t)} \geq \alpha$ 表示跟踪器定位目标成功,需使用帧差均值对目标及背景的运动状态进行分类以调节学习率:

1) $e < \beta_1$ 表明目标和背景的变化不大,因此只需设置较小的学习率 $\eta = \eta_2$;

2) $\beta_1 \leq e < \beta_2$ 说明目标或背景变化适中,两种情况下均需设置适中的学习率,因此学习率设置为 $\eta = \eta_3$;

3) $e \geq \beta_2$ 说明目标或背景变化较大,这可能是目标变化较大,也有可能是背景发生变化,需要使用目标位移 s 对目标和背景进行区分;

a) $s < \phi$ 表明目标运动较小,背景发生变化,因此学习率设置适中, $\eta = \eta_4$;

b) $s \geq \phi$ 说明目标发生较大变化,需要对滤波器进行较多更新以适应目标外观,因此学习率设置为较大的值 $\eta = \eta_5$ 。

具体设置的学习率为

$$\eta = \begin{cases} \eta_1, & a_{APCE}^{(t)} < \alpha \\ \eta_2, & a_{APCE}^{(t)} \geq \alpha, e < \beta_1 \\ \eta_3, & a_{APCE}^{(t)} \geq \alpha, \beta_1 \leq e < \beta_2 \\ \eta_4, & a_{APCE}^{(t)} \geq \alpha, e \geq \beta_2, s < \phi \\ \eta_5, & a_{APCE}^{(t)} \geq \alpha, e \geq \beta_2, s \geq \phi \end{cases}. \quad (11)$$

4 实验结果与分析

4.1 实验环境与参数设置

为了验证本算法的有效性,实验均在软件为 MATLAB 2018b, CPU 为 Inter(R) Core(TM) i3-8100, GPU 为 NVIDIA GeForce GTX TIX TITAN X, 内存为 12G, 系统为 win10 的计算机上实施。

实验中对所有的测试视频采用相同的参数,具体设置如下:第五,四,三层中卷积特征所占比重为 $\{1, 0.5, 0.6\}$, 尺度步长 $a = 1.02$, 尺度数量 S 为 33, 正则化系数 $\lambda = 0.1$ 。实验阈值设置如下:平均峰值相关能量阈值 $\alpha = 15$, 帧差均值阈值 $\beta_1 = 2.5, \beta_2 =$

5, 目标位移阈值 $s=11$ 。学习率 η_1 至 η_5 的值设置为 0, 0.005, 0.0075, 0.009, 0.02。

4.2 评估方法

为有效地评估本算法性能, 本文利用 OTB-2013^[19] 测试集的 50 个视频进行测试。采用文献^[19]中的评价指标: 重叠精度(OP)、距离精度(DP)、中心位置误差 CLE(Center Location Error); 评价方式为 OPE (One-Pass Evaluation)。其中, CLE 为预测目标的中心位置与标注的准确位置之间的欧氏距离; DP 为 CLE 小于预设阈值 20 pixel 的百分比; OP 为重叠率大于预设阈值 0.5 的百分比。根据距离精度值与阈值之间的距离精度曲线图(Precision Plot)以及目标框重叠率与阈值之间的成

功率曲线图(Success Plot)对算法进行评估。

4.2.1 定量对比实验

将本算法 Ours, 与固定学习率的本算法 Ours1 和主流的 6 种跟踪算法(SiamFC^[20]、HCF/CF2^[12]、SRDCF^[21]、DSST^[10]、SAMF^[11]、KCF^[7])作一次性通过评估。KCF 算法是采用 HOG 特征核函数的目标跟踪算法; DSST 以及 SAMF 算法是利用尺度变化的目标跟踪算法; SRDCF 算法是引入空间正则化的相关滤波改进算法, 可以有效地抑制背景区域的响应, 并在复杂背景等场景下获得更佳的性能; SiamFC 和 HCF/CF2 算法都是近年来利用深度学习进行目标跟踪的算法。图 2 显示了 8 种算法的距离精度跟成功率曲线图。

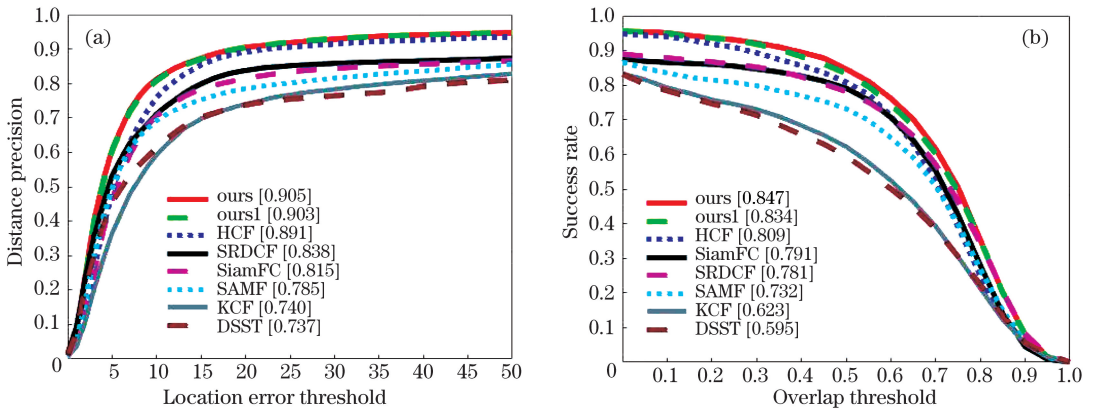


图 2 8 种算法的距离精度曲线图和成功率曲线图。(a)距离精度曲线图;(b)成功率曲线图

Fig. 2 Distance precision and overlap success curves of 8 algorithms. (a) Distance precision curves; (b) overlap success curves

图 2 显示, 本算法的重叠精度为 0.905, 距离精度为 0.847; 比未使用学习率调整的算法的重叠精度只高出 0.002, 距离精度高出 0.013; 但比次优的 HCF 算法, 距离精度高出 0.014, 重叠精度高出 0.038。不管在距离精度上还是在成功率上, 本算法都远超 SRDCF、DSST、SAMF 等算法。这表明使用卷积特征确实能

提高相关跟踪器的性能。同时, 对比同样使用多层卷积特征的 HCF 算法, 本文算法在距离精度和成功率上都更优, 说明本文的尺度处理对跟踪器性能的改善起到了一定的作用。证明基于学习率调整的模版更新方式, 能提高目标跟踪的准确度。

表 1 和表 2 为不同属性下各个算法的距离精度值

表 1 不同算法在不同挑战因素下的距离精度值

Table 1 DPs of different algorithms for different challenge factors

Indicator	Ours	HCF	SRDCF	SiamFC	DSST	SAMF	KCF
FM	0.846	0.851	0.781	0.756	0.690	0.730	0.698
SV	0.889	0.875	0.874	0.825	0.784	0.748	0.770
IV	0.879	0.897	0.811	0.805	0.846	0.836	0.800
OCC	0.862	0.868	0.828	0.770	0.687	0.708	0.712
DEF	0.897	0.926	0.900	0.920	0.696	0.807	0.688
MB	0.855	0.878	0.838	0.735	0.749	0.799	0.713
IPR	0.885	0.853	0.850	0.811	0.770	0.821	0.756
OPR	0.880	0.865	0.869	0.844	0.770	0.815	0.753
OV	0.959	0.935	0.791	0.706	0.699	0.699	0.713
BC	0.905	0.905	0.811	0.901	0.695	0.802	0.706
LR	0.991	0.965	0.958	0.699	0.917	0.800	0.948

和成功率精度值,其中加粗值为最优结果。表中包含超出视野(OV)、光照变化(IV)、低分辨率(LR)、运动模糊(MB)、遮挡(OCC)、尺度变化(SV)、快速移动(FM)、背景干扰(BC)、非刚性形变(DEF)、平面内旋转(IPR)和平面外旋转(OPR) 11个属性。

表2 不同算法在不同挑战因素下的成功率精度值

Table 2 Success rate accuracy values of different algorithms for different challenge factors

Indicator	Ours	HCF	SRDCF	SiamFC	DSST	SAMF	KCF
FM	0.800	0.783	0.757	0.743	0.595	0.713	0.681
SV	0.833	0.779	0.799	0.791	0.605	0.668	0.618
IV	0.833	0.860	0.765	0.783	0.666	0.782	0.647
OCC	0.814	0.798	0.785	0.745	0.578	0.685	0.611
DEF	0.882	0.879	0.830	0.892	0.561	0.770	0.592
MB	0.779	0.786	0.818	0.703	0.646	0.772	0.644
IPR	0.849	0.784	0.819	0.799	0.625	0.750	0.627
OPR	0.856	0.817	0.820	0.822	0.609	0.750	0.619
OV	0.838	0.811	0.752	0.680	0.529	0.663	0.619
BC	0.861	0.873	0.723	0.866	0.567	0.750	0.603
LR	0.809	0.834	0.776	0.640	0.666	0.738	0.686

4.2.2 定性对比实验

为了更好地说明本文算法在复杂环境下的跟踪性能,采用6组典型视频 Shaking, Football, Freeman3,

从表1可以看出,在FM、IV、OCC、DEF和MB情况下,所提算法略低于HCF算法,在其他情况下都优于6种算法。从表2可以看出,所提算法除在IV、DEF、MB、BC和LR情况下不是最优,其余挑战因素下都是最优。此外,在IV、MB和DEF挑战因素下,本算法还需进一步优化。

Jogging2, Lemming 和 MotorRolling 评估本算法与HCF、SiamFC、SRDCF、DSST、SAMF和KCF这7种算法在不同挑战因素下的性能,如表3所示。

表3 测试视频的属性

Table 3 Attributes of test videos

Test video	Frame	Target size	Challenge factor
Football	362	39 × 50	OCC, IPR, OPR, BC
Freeman3	474	12 × 13	SV, IPR, OPR
Lemming	1336	61 × 103	IV, SV, OCC, FM, OPR, OV
Shaking	365	61 × 71	IV, SV, IPR, OPR, BC
Jogging2	307	37 × 114	OCC, DEF, OPR
MotorRolling	164	122 × 125	IV, SV, MB, FM, IPR, BC, LR

图3为7种算法在这6组视频中部分帧的跟踪结果。图3(a)Football视频中,在相似目标情况下,本算法采用的卷积特征,能够正确分辨目标与背景。图3(b)Freeman3视频中,目标出现平面内外旋转和尺度变化,除了本算法提出的自适应更新算法外,其他算法已出现较大偏差。图3(c)Lemming视频中,目标大部分外观被遮挡,本算法能在遮挡情况下识别目标并正确跟踪。图3(d)Shaking视频中,存在相似目标以及光照变化,本算法能分辨背景与目标,有效处理光照变化,完成跟踪。图3(e)Jogging2视频中,本算法能在目标被完全遮挡时跟踪到目标。图3(f)MotorRolling视频中,本算法能在目标旋转以及快速运动情况下正确识别目标。综上所述,本算法在复杂背景、遮挡、尺度等情况下能

很好地追踪到目标。

5 结论

本文在融合多层卷积特征的相关滤波跟踪算法的框架下,提出一种自适应更新学习率的相关滤波跟踪算法。在位置滤波器上融合三层卷积特征预测目标位置,在尺度滤波器上采用单层卷积特征确定目标最佳尺度。采用平均峰值相关能量、帧差均值和位移三种策略相结合的一种多指标选择学习率的更新方式,提高滤波器模型可靠性。将本算法与6种主流算法在OTB-2013公开测试集上进行测试对比,实验结果表明所提算法的平均距离精度达到90.5%,成功精度达到84.6%,都优于其他6种算法。虽然本算法的精度和成功率都有很大提升,但



图3 7种算法在6个视频序列的跟踪结果比较。(a) Football;(b) freeman3;(c) lemming;
(d) shaking;(e) jogging2;(f) motorrolling

Fig. 3 Comparison of tracking results of 7 algorithms for six video sequences. (a) Football; (b) freeman3; (c) lemming;
(d) shaking; (e) jogging2; (f) motorrolling

同时也牺牲了速度。卷积特征的训练和滤波模型的更新十分耗时,故需对速度进行进一步优化,以便更好地提高跟踪算法的实时性和准确性。

参 考 文 献

- [1] Jing W, An J F, Jiang G, et al. Standoff fully-polarimetric holographic-radar-imaging experiments in terahertz regime for concealed target detection[J]. Chinese Journal of Lasers, 2019, 46(6): 0614032. 经文, 安健飞, 江舸, 等. 针对隐藏目标检测的太赫兹波站开式全极化全息雷达成像实验[J]. 中国激光, 2019, 46(6): 0614032.
- [2] Oh S, Hoogs A, Perera A, et al. A large-scale benchmark dataset for event recognition in surveillance video[C] // CVPR 2011, June 20-25, 2011, Colorado Springs, CO, USA. New York: IEEE, 2011: 3153-3160.
- [3] Smeulders A W M, Chu D M, Cucchiara R, et al. Visual tracking: an experimental survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(7): 1442-1468.
- [4] Yang Jianfeng, Zhang Jianpeng. Long Time Target Tracking Based on Kernel Correlation Filtering[J]. Laser & Optoelectronics Progress, 2019, 56(2): 021502. 杨剑锋, 张建鹏. 基于核相关滤波的长时间目标跟踪[J]. 激光与光电子学进展, 2019, 56(2): 021502.
- [5] He X D, Zhou S Z. Fast scale adaptive kernel correlation filtering algorithm for target tracking[J]. Laser & Optoelectronics Progress, 2018, 55(12):

121501.

何雪东, 周盛宗. 快速尺度自适应核相关滤波目标跟踪算法[J]. 激光与光电子学进展, 2018, 55(12): 121501.

- [6] Bolme D, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters[C] // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE, 2010: 2544-2550.
- [7] 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.
- [8] Danelljan M, Khan F S, Felsberg M, et al. Adaptive color attributes for real-time visual tracking[C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 1090-1097.
- [9] Pan Z F, Zhu Y L. Kernelized correlation filters object tracking method with multi-scale estimation[J]. Laser & Optoelectronics Progress, 2016, 53(10): 101501. 潘振福, 朱永利. 多尺度估计的核相关滤波器目标跟踪方法[J]. 激光与光电子学进展, 2016, 53(10): 101501.
- [10] Danelljan M, Häger G, Khan F S, et al. Discriminative scale space tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(8): 1561-1575.

- [11] Li Y, Zhu J K. A scale adaptive kernel correlation filter tracker with feature integration [M] // Agapito L, Bronstein M, Rother C. Computer vision - ECCV 2014 Workshops. Lecture notes in computer science. Cham: Springer, 2015, 8926: 254-265.
- [12] Ma C, Huang J B, Yang X K, et al. Hierarchical convolutional features for visual tracking [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 3074-3082.
- [13] Danelljan M, Robinson A, Khan F S, et al. Beyond correlation filters: learning continuous convolution operators for visual tracking [M] // Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9909: 472-488.
- [14] Danelljan M, Bhat G, Khan F S, et al. ECO: efficient convolution operators for tracking [C] // IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, Hawaii, USA. New York: IEEE, 2017: 6638-6646.
- [15] Yin H P, Chai Y, Yang S X, et al. Fast-moving target tracking based on mean shift and frame-difference methods [J]. Journal of Systems Engineering and Electronics, 2011, 22(4): 587-592.
- [16] Wang M M, Liu Y, Huang Z Y. Large margin object tracking with circulant feature maps [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 4800-4808.
- [17] Xiong C Z, Che M Q, Wang R L, et al. Adaptive model update via fusing peak-to-sidelobe ratio and mean frame difference for visual tracking [J]. Acta Photonica Sinica, 2018, 47(9): 0910001. 熊昌镇, 车满强, 王润玲, 等. 融合峰旁比和帧差均值自适应模型更新的视觉跟踪 [J]. 光子学报, 2018, 47(9): 0910001.
- [18] Hou X D. Algorithm research in the key links of video monitoring [D]. Nanning: Guangxi University, 2012. 侯绪达. 一种改进后的自适应均值位移目标跟踪算法研究 [D]. 南宁: 广西大学, 2012.
- [19] Wu Y, Lim J, Yang M H. Online object tracking: a benchmark [C] // 2013 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2013, Portland, OR, USA. New York: IEEE, 2013: 2411-2418.
- [20] Bertinetto L, Valmadre J, Henriques J F, et al. Fully-convolutional Siamese networks for object tracking [M] // Hua G, Jégou H. Computer vision-ECCV 2016 Workshops. Lecture notes in computer science. Cham: Springer, 2016, 9914: 850-865.
- [21] Danelljan M, Hager G, Khan F S, et al. Adaptive decontamination of the training set: a unified formulation for discriminative visual tracking [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 26-July 1, 2016, Las Vegas, Nevada. New York: IEEE, 2016: 1430-1438.