

自适应特征融合的多尺度核相关滤波目标跟踪

陈法领^{1,2,3,4,5*}, 丁庆海^{1,6}, 常铮^{1,2,4,5}, 陈宏宇^{1,2,3,4,5}, 罗海波^{1,2,4,5}, 惠斌^{1,2,4,5}, 刘云鹏^{1,2,4,5}

¹中国科学院沈阳自动化研究所, 辽宁 沈阳 110016;

²中国科学院机器人与智能制造创新研究院, 辽宁 沈阳 110169;

³中国科学院大学, 北京 100049;

⁴中国科学院光电信息处理重点实验室, 辽宁 沈阳 110016;

⁵辽宁省图像处理与视觉计算重点实验室, 辽宁 沈阳 110016;

⁶航天恒星科技有限公司, 北京 100086

摘要 为了提高复杂场景中目标跟踪的稳健性, 解决由光照变化、目标形变、尺度变化和遮挡等导致的目标跟踪失败问题, 提出一种自适应特征融合的多尺度核相关滤波目标跟踪算法。该算法首先通过 2 种不同的特征分别训练 2 个核相关滤波器, 利用这 2 个滤波器响应的峰值旁瓣比和相邻两帧的响应一致性获得融合权重, 同时采用自适应加权的融合策略将这 2 个滤波器的响应结果进行融合, 完成目标的位置估计; 然后以此为中心进行多尺度采样, 构建尺度金字塔, 并通过贝叶斯估计的方法确定目标的最优尺度; 最后依据目标跟踪的置信度进行跟踪模型更新, 以避免模型退化。选取 51 组视频序列进行测试, 并与近年来性能优异的目标跟踪算法进行对比。实验结果表明, 所提算法能有效降低光照变化、目标形变、尺度变化和遮挡等因素影响, 对测试视频序列取得了较高的跟踪精度和成功率, 整体性能优于对比算法。

关键词 机器视觉; 计算机视觉; 目标跟踪; 核相关滤波; 自适应特征融合; 多尺度估计

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/AOS202040.0315001

Multi-Scale Kernel Correlation Filter Algorithm for Visual Tracking Based on the Fusion of Adaptive Features

Chen Faling^{1,2,3,4,5*}, Ding Qinghai^{1,6}, Chang Zheng^{1,2,4,5},

Chen Hongyu^{1,2,3,4,5}, Luo Haibo^{1,2,4,5}, Hui Bin^{1,2,4,5}, Liu Yunpeng^{1,2,4,5}

¹Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China;

²Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang, Liaoning 110169, China;

³University of Chinese Academy of Sciences, Beijing 100049, China;

⁴Key Laboratory of Opto-Electronic Information Processing, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China;

⁵Key Laboratory of Image Understanding and Computer Vision, Shenyang, Liaoning 110016, China;

⁶Space Star Technology Co., Ltd., Beijing 100086, China

Abstract In this study, we propose a multi-scale kernel correlation filter algorithm for visual tracking based on the fusion of adaptive features to promote the robustness of visual tracking in complex scenarios and tackle the tracking failure problems that can be attributed to illumination variation, target deformation, scale variation, occlusion, etc. First, two kernel correlation filters are separately trained using two different features. Then, the peak side-lobe ratio of the responses and the correlation filter response consistency of two consequent frames are considered to be the weight factors for feature fusion. Meanwhile, an adaptive strategy is adopted to fuse two responses for estimating the position. Next, multi-scale image patches are sampled to construct a scale pyramid based on the estimated position center, and the Bayesian method is employed to estimate the optimal scale of the target. Finally,

收稿日期: 2019-07-25; 修回日期: 2019-09-11; 录用日期: 2019-09-29

基金项目: 国家重点研发计划(2018YFC080192)

* E-mail: chfling@sia.cn

the tracking model is updated according to the confidence of the tracking result to prevent the deterioration of the model. 51 video sequences are selected for conducting tracking evaluation, and the visual tracking algorithms that exhibited excellent performances in recent years are compared with our proposed algorithm. The experimental results demonstrate that the proposed algorithm effectively reduces the interferences, including the illumination variation, target deformation, scale variation, and occlusion. High tracking accuracy and success rate can be achieved using the aforementioned sequences, and the overall performance of our algorithm is observed to be better than those of the comparison algorithms.

Key words machine vision; computer vision; target tracking; kernel correlation filter; adaptive features fusion; multi-scale estimation

OCIS codes 150.0155; 150.1135; 100.4999

1 引 言

目标跟踪是计算机视觉领域一个重要且富有挑战性的研究课题,在智能监控、自动驾驶、人机交互和精确制导等领域具有非常广泛的应用^[1]。在过去的几十年中,国内外学者对目标跟踪进行了广泛而深入的研究^[2-3],并取得了丰硕成果。但在实际应用中,受到场景光照变化、目标形变、尺度变化或遮挡等不利因素影响,实现稳健的目标跟踪仍然面临着非常严峻的挑战^[4]。

相关滤波在频域的快速计算可使目标跟踪的速度达到每秒数百帧,因此其成为当前目标跟踪研究领域的热点之一。Bolme等^[5]提出了最小输出误差平方和滤波(MOSSE)算法,首次将相关滤波方法应用到了目标跟踪领域。在MOSSE算法基础上,Henriques等^[6]提出了循环结构核(CSK)算法,通过循环移位方式完成样本密集采样,解决了分类器无法获取大量训练样本的问题。随后,Henriques等^[7]又提出了核相关滤波(KCF)算法,将单通道灰度特征扩展到多通道梯度方向直方图(HOG)特征,提高了算法的跟踪精度和适应性。Danelljan等^[8]将三维三原色(RGB)特征映射为11维颜色名(CN)特征,也取得了较好的跟踪效果。

上述方法利用单一特征构建目标外观模型,在复杂场景中目标的区分度差并且容易受到干扰。针对此问题:Bertinetto等^[9]提出了Staple算法,将HOG特征和颜色直方图特征进行融合,以提升算法性能;李聪等^[10]将RGB颜色空间映射到Lab颜色空间,再与HOG特征进行融合,完成目标跟踪;沈秋等^[11]将HOG特征、CN特征和局部二值模式(LBP)特征与灰度特征进行串联,选择其中最大响应所对应的位置作为跟踪结果;赵高鹏等^[12]通过计算灰度特征和LBP特征响应的峰值旁瓣比,将这2种特征进行融合。此外,Qi等^[13-14]提出将深度卷积神经网络获取的特征应用于相关滤波目标跟踪中,

也取得了改善效果,但是特征提取的计算开销较大,很难满足目标跟踪的实时性要求。

为了处理相关滤波目标跟踪中的尺度变化问题, Danelljan等^[15]设计了一个三维的位置-尺度相关滤波器,用于适应目标的尺度变化; Li等^[16]在上一帧目标位置进行不同尺度的采样,并将其中最大响应所对应的尺度作为当前目标尺度; Zhang等^[17]通过计算当前帧和前一帧目标位置的置信度来评估相邻两帧之间的尺度变化。这些方法虽然能够适应一定程度的目标尺度变化,但并未有效解决目标跟踪过程中的尺度变化问题。

针对上述问题,本文在KCF算法基础上,提出一种自适应特征融合的多尺度核相关滤波目标跟踪算法。该算法利用目标候选区域的HOG特征和CN特征分别训练2个核相关滤波器,并采用自适应加权的融合策略将这2个滤波器的响应结果进行融合,得到目标中心位置;然后构建尺度金字塔,并采用贝叶斯估计的方法确定目标的最优尺度;最后依据目标跟踪的置信度来更新跟踪模型,从而实现目标的稳定跟踪。

2 核相关滤波目标跟踪原理

核相关滤波目标跟踪通过训练样本寻找一个分类器函数 $f(\mathbf{x}_i, \mathbf{w}) = \langle \mathbf{w}, \mathbf{x}_i \rangle$, 使其在某种决策条件下的损失最小, 其中 \mathbf{x}_i 为训练样本, \mathbf{w} 为该分类器函数的待求参数, $\langle \cdot, \cdot \rangle$ 为内积运算。将训练样本和它对应标签的误差平方和作为损失函数, 可以得到 \mathbf{w} 的求解形式为

$$\min_{\mathbf{w}} \sum_{i=1}^n [f(\mathbf{x}_i) - y_i]^2 + \lambda \|\mathbf{w}\|^2, \quad (1)$$

式中: \mathbf{x}_i 和 y_i 分别为第 i ($i=1, 2, \dots, n$) 个训练样本及其所对应的标签, n 为训练样本的个数; λ 为防止分类器过拟合的正则化系数; $\|\cdot\|$ 为范数。对(1)式求偏导数, 并令其为0, 可以得到(1)式的一般解为

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

式中: \mathbf{X} 是由训练样本组成的样本矩阵, 每行代表一个样本 \mathbf{x}_i ; \mathbf{I} 为与 $\mathbf{X}^T \mathbf{X}$ 同维数的单位矩阵; \mathbf{y} 为每个训练样本 \mathbf{x}_i 所对应的标签 \mathbf{y}_i 的集合。对于核相关滤波目标跟踪, 由于它的样本矩阵 \mathbf{X} 是由初始目标样本 x 经过循环移位得到的, 因此 \mathbf{X} 具有循环结构, 利用循环矩阵的离散傅里叶变换 (DFT) 性质, (2) 式在频域中的表达式为

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

式中: \hat{x}_j 、 \hat{y}_j 和 \hat{w}_j 分别为初始目标样本 \mathbf{x} 、标签集合 \mathbf{y} 和分类器的待求参数 \mathbf{w} 的 DFT 中的一个元素; \hat{x}^* 为 \hat{x} 的复共轭。进一步地, 核相关滤波通过核函数的处理方法将输入样本 \mathbf{x} 映射到高维特征空间中, 那么分类器的待求参数 \mathbf{w} 可以在它的对偶空间中表示为 $\mathbf{w} = \sum \alpha_i \varphi(\mathbf{x}_i)$, 其中 α_i 为对偶空间中的系数, $\varphi(\mathbf{x}_i)$ 为训练样本 \mathbf{x}_i 映射到高维特征空间后的表示, 将 \mathbf{w} 的求解问题转换为在它的对偶空间中求解 α , 并且 α 在频域中的形式可以表示为

$$\hat{a}_j = \frac{\hat{y}_j}{\hat{\kappa}(x, x) + \lambda}, \quad (4)$$

式中: $\hat{\kappa}(x, x)$ 为核函数 $K = \langle \varphi(x), \varphi(x) \rangle$ 的 DFT, \langle, \rangle 为内积运算。对于一帧新的图像 z , 它所对应的分类器在频域中的响应输出为

$$\hat{f}(z) = \hat{\kappa}(x, z) \odot \hat{a}, \quad (5)$$

式中: $\hat{\kappa}(x, z)$ 为核函数 $K = \langle \varphi(x), \varphi(z) \rangle$ 的 DFT。(5) 式的傅里叶逆变换最大值所对应的坐标即为目标在新一帧图像中的位置。

3 自适应特征融合的多尺度目标跟踪

3.1 自适应特征融合

3.1.1 特征选择

在目标跟踪中, 特征的选择及其区分度对目标跟踪的结果具有重要影响。HOG 特征通过计算图像局部区域的方向梯度进行目标表达, 可以有效地描述目标的轮廓和形状信息, 并在一定程度上保持了几何和光照不变性。KCF 算法^[7] 通过采用多通道的 HOG 特征, 能够较好地适应光照变化、平面内旋转等情况。

颜色特征则是从全局角度对目标进行描述, 具有旋转不变性, 同时不易受到目标尺度或形状变化

的影响。CN 特征是颜色特征的一种, 它通过概率映射方式, 将图像从原始的三维 RGB 空间转换到 11 维的颜色特征空间。相对于其他颜色特征, CN 特征具有更好的目标描述能力^[8]。

虽然相对于灰度特征, 采用 HOG 特征和 CN 特征都能够一定程度上提高相关滤波的跟踪性能, 但是受单一特征对目标描述的局限性在复杂场景中目标的部分特性发生变化时单一场景不能对目标进行有效的描述, 进而会影响目标跟踪质量。由于 HOG 特征和 CN 特征分别从不同的角度对目标进行抽象和特征提取, 同时 HOG 特征具有较好的几何和光照不变性, 而 CN 特征则对目标尺度和形状变化不敏感, 两者具有较强的互补性, 因此本文将这 2 种特征进行融合, 以提升目标跟踪的性能。

3.1.2 融合方法

在相关滤波中, 峰值旁瓣比 R_{PS} 代表一个相关滤波响应 (CFR) 的峰值尖锐度, 通常用于衡量目标跟踪的置信度^[5]。对于相关滤波响应, 在峰值位置 x 处的 R_{PS} 可表示为

$$R_{PS}(x) = \frac{\max\{x\} - \mu(x)}{\sigma(x)}, \quad (6)$$

式中: $\max\{x\}$ 为相关滤波响应中 x 的最大值; $\mu(x)$ 和 $\sigma(x)$ 分别为 x 均值和标准差。 R_{PS} 的值越大, 表示目标跟踪的置信度越高; 反之, 则表示目标跟踪的置信度越低。但是仅使用 R_{PS} 指标来表示目标跟踪的置信度是不够的, 特别是在目标被遮挡或者出现相似目标干扰时, 容易导致跟踪漂移甚至失败。文献^[18] 提出将平均峰值相关能量 (EAPC) 作为衡量目标跟踪的置信度评价指标; 文献^[19] 则提出将响应的平滑性约束 (SCCM) 作为衡量各个子块跟踪的置信度评价指标。受文献^[19] 启发, 本文定义相邻两帧目标的相关滤波响应的一致性 C_{CFR} , 表达式为

$$C_{CFR} = \|\hat{f}_t(x, y) - \hat{f}_{t-1}(x + \Delta x, y + \Delta y)\|_2^2, \quad (7)$$

式中: $f_t(x, y)$ 和 $f_{t-1}(x + \Delta x, y + \Delta y)$ 分别为目标在第 t 帧和第 $t-1$ 帧时的相关滤波响应; Δx 和 Δy 为相邻两帧之间目标位置的相对变化; $\|\cdot\|_2$ 为 L2 范数运算。在目标跟踪中, 由于相邻两帧的时间间隔通常只有 20 ms 甚至更短, 因此相邻两帧之间目标和背景的变化是连续的, 它们的相关滤波响应具有较高的相似性。由此, 可以将相邻两帧之间的 C_{CFR} 值作为目标跟踪的置信度评价指标: 当 C_{CFR} 值较小时, 相邻两帧目标的相关滤波响应的相似性较高, 目标跟踪的稳定性较高; 反之, 则说明目标跟踪

的稳定性较低。

通过以上分析,本文将 R_{PS} 和 C_{CFR} 作为目标跟踪的置信度评价指标,构造一个二元函数 $f(R_{PS}, C_{CFR})$ 作为目标跟踪的置信度评价函数,其定义式为

$$f(R_{PS}, C_{CFR}) = (1 - \rho)R_{PS} + \rho \frac{1}{C_{CFR} + \epsilon}, \quad (8)$$

式中: $\rho \in [0, 1]$, 为 R_{PS} 与 C_{CFR} 之间的权重调节系数; ϵ 为 0.01, 用于避免分母为 0。

先利用 HOG 特征和 CN 特征训练 2 个核相关滤波器, 然后分别提取目标候选区域的 HOG 特征和 CN 特征, 并计算二者的核相关滤波响应, 对这 2 个响应进行高斯滤波, 以剔除异常响应值, 最后计算它们滤波后的置信度 f_{HOG} 和 f_{CN} , 计算式分别为

$$f_{HOG} = (1 - \rho)R_{PS,HOG} + \rho \frac{1}{C_{CFR,HOG} + \epsilon}, \quad (9)$$

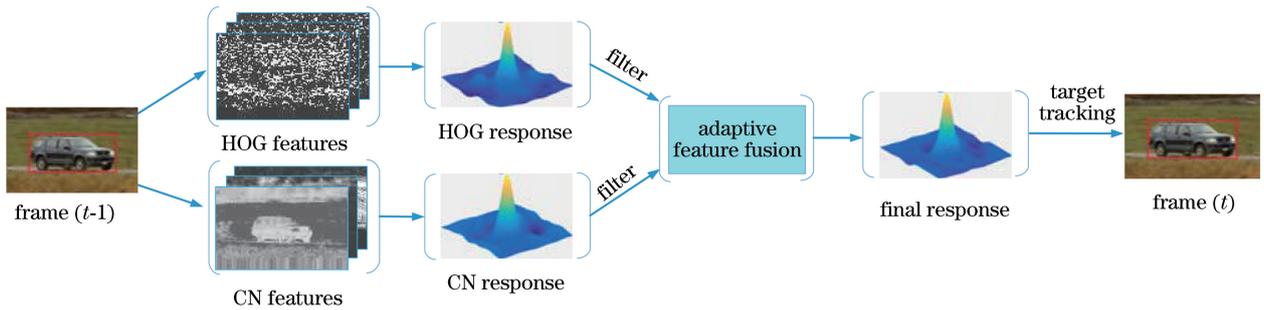


图 1 自适应特征融合过程示意图

Fig. 1 Schematic of adaptive features fusion process

3.2 多尺度估计

为了解决相关滤波目标跟踪中的尺度变化问题, 文献[16-18]分别提出 3 种不同的解决方案, 但是均未考虑相邻两帧之间尺度的变化规律。在目标跟踪中, 相邻两帧的时间间隔通常只有 20 ms 甚至更短, 因此相邻两帧之间目标的尺度变化是微小且是连续的, 可以将这种变化近似为一种高斯分布^[20-21], 即目标在当前帧的尺度变化 s_t 服从高斯分布, 以它在前一帧中的尺度 s_{t-1} 为均值, σ 为方差。

由此获得相邻两帧之间尺度变化的先验分布, 接下来可以寻找一个似然函数 $p(f_t | s_t)$, 并在此基础上按照下式完成对目标尺度变化的贝叶斯估计(最大后验概率)。

$$\max\{p(s_t | f_t)\} = p(f_t | s_t)p(s_t), \quad (12)$$

式中: $p(s_t | \hat{f}_t)$ 为似然概率; $p(s_t)$ 为先验概率。

当给定一个目标的尺度时, 利用核相关滤波器可以求出此尺度下跟踪候选区域与目标的最大相似度, 该最大相似度可以表示目标在该尺度下的概率。

$$f_{CN} = (1 - \rho)R_{PS,CN} + \rho \frac{1}{C_{CFR,CN} + \epsilon}, \quad (10)$$

式中: $R_{PS,HOG}$ 和 $R_{PS,CN}$ 分别为 HOG 特征和 CN 特征所对应的相关滤波响应的 R_{PS} 值; $C_{CFR,HOG}$ 和 $C_{CFR,CN}$ 分别为 HOG 特征和 CN 特征在相邻两帧之间的 C_{CFR} 值。最后将这 2 个置信度作为特征融合的权重因子, 可得到融合后的相关滤波响应为

$$\hat{f}_{final} = \frac{f_{HOG}}{f_{HOG} + f_{CN}} \hat{f}_{HOG} + \frac{f_{CN}}{f_{HOG} + f_{CN}} f_{CN}, \quad (11)$$

式中: \hat{f}_{HOG} 和 \hat{f}_{CN} 分别为 HOG 特征和 CN 特征所对应的核相关滤波响应; $\frac{f_{HOG}}{f_{HOG} + f_{CN}}$ 和 $\frac{f_{CN}}{f_{HOG} + f_{CN}}$ 为它们的权重因子。本文的自适应特征融合过程如图 1 所示。

通过构建尺度金字塔来完成对目标尺度变化的极大似然估计, 即以目标在当前帧的估计位置为中心, 将前一帧的目标尺度 s_{t-1} 作为基准尺度, 进行多尺度采样, 则有

$$s_t = \{s_m = s_{t-1} + md, m \in [-M/2, M/2]\}, \quad (13)$$

式中: m 为多尺度采样的子层数; s_m 为多尺度采样的每个子层的尺度; M 为尺度金字塔的层数; d 为相邻 2 层之间的变化量。通过提取多尺度样本的 HOG 特征, 构建一个尺度滤波器, 完成对目标尺度的极大似然估计, 然后通过(12)式得到每层尺度的最大后验概率, 将具有最大后验概率的尺度 s_m 作为当前帧目标尺度 s_t 的最优估计。

3.3 更新策略

在目标跟踪中, 由于目标所处的场景及目标自身都不可避免地会发生变化, 如果固定采用在起始帧获得的跟踪模型, 则不能准确地描述变化后的状态, 因此在跟踪过程中必须对跟踪模型进行动态更新, 以保证跟踪的稳健性。但是如果跟踪模型的更

新频率过快,则容易将噪声引入到跟踪模型中,反之,则不能描述跟踪过程的动态变化。因此,跟踪模型的更新策略将直接影响目标跟踪的性能。根据特征融合部分的分析,采用融合后相关滤波响应的 R_{PS} 值和 C_{CFR} 值作为目标跟踪的置信度评价指标,同样地,定义特征融合后的置信度为

$$f_{\text{final}} = (1 - \rho)R_{\text{PS,final}} + \rho \frac{1}{C_{\text{CFR,final}} + \epsilon}, \quad (14)$$

式中: $R_{\text{PS,final}}$ 为特征融合后的相关滤波响应的 R_{PS} 值; $C_{\text{CFR,final}}$ 为特征融合后的相关滤波响应在相邻两帧之间的 C_{CFR} 值。 $R_{\text{PS,final}}$ 和 $C_{\text{CFR,final}}$ 的计算方法与(6)式和(7)式给出的计算方法一致,通过特征融合后的相关滤波响应得到。当 f_{final} 值较小时,表示目标跟踪的置信度较低,说明此时目标外观发生了较大变化或者发生了跟踪漂移,不能更新跟踪模型;反之,当 f_{final} 值较大时,则认为此时目标跟踪的置信度较高,可以更新跟踪模型。本文设置一个置信度阈值 f_{th} ,用于判断是否对当前帧的跟踪模型进行更新;另外,为了使更新策略更可靠,在进行模型更新时还参考了历史帧信息,即如果连续多帧目标跟踪的置信度 f_{final} 都大于设定的阈值 f_{th} ,则利用当前帧信息来更新跟踪模型,否则不更新。在核相关滤波目标跟踪中,有 2 个参数需要在频域进行更新^[7],一个是对偶矩阵参数 $\hat{\alpha}$,另一个是目标的外观参数 x 。在本文中有 2 个跟踪模型需要更新,分别是采用 HOG 特征的跟踪模型和 CN 特征的跟踪模型,具体参数更新为

$$\begin{cases} \hat{\alpha}_{\text{HOG}_{t-j}} = (1 - \eta)\hat{\alpha}_{\text{HOG}_{t-1-j}} + \\ \eta \frac{\hat{y}_{\text{HOG}_j}}{\hat{\kappa}(x'_{\text{HOG}}, x'_{\text{HOG}})_j + \lambda} \\ \hat{x}_{\text{HOG}_{t-j}} = (1 - \eta)\hat{x}_{\text{HOG}_{t-1-j}} + \eta\hat{x}'_{\text{HOG}} \end{cases}, \quad (15)$$

$$\begin{cases} \hat{\alpha}_{\text{CN}_{t-j}} = (1 - \eta)\hat{\alpha}_{\text{CN}_{t-1-j}} + \eta \frac{\hat{y}_{\text{CN}_j}}{\hat{\kappa}(x'_{\text{CN}}, x'_{\text{CN}})_j + \lambda}, \\ \hat{x}_{\text{CN}_{t-j}} = (1 - \eta)\hat{x}_{\text{CN}_{t-1-j}} + \eta\hat{x}'_{\text{CN}} \end{cases} \quad (16)$$

式中: $\hat{\alpha}_{\text{HOG}_{t-1}}$ 和 $\hat{\alpha}_{\text{CN}_{t-1}}$ 分别为 2 个核相关滤波器在第 t 帧时跟踪模型更新前的对偶矩阵参数; $\hat{\alpha}_{\text{HOG}_t}$ 和 $\hat{\alpha}_{\text{CN}_t}$ 分别为 2 个核相关滤波器在第 t 帧时跟踪模型更新后的对偶矩阵参数; $\hat{x}_{\text{HOG}_{t-1}}$ 和 $\hat{x}_{\text{CN}_{t-1}}$ 分别为 2 个核相关滤波器在第 t 帧时跟踪模型更新前的目标

外观参数; \hat{x}_{HOG_t} 和 \hat{x}_{CN_t} 分别为这 2 个核相关滤波器在第 t 帧时跟踪模型更新后的目标外观参数;

$\frac{\hat{y}_{\text{HOG}}}{\hat{\kappa}(x', x')_{\text{HOG}} + \lambda}$ 和 $\frac{\hat{y}_{\text{CN}}}{\hat{\kappa}(x', x')_{\text{CN}} + \lambda}$ 分别为在第 t 帧

获得的对偶矩阵参数; \hat{x}'_{HOG} 和 \hat{x}'_{CN} 分别为在第 t 帧获得的目标外观参数; η 为更新率。通过采用(15)和(16)式给出的线性插值方式完成对跟踪模型的更新,这样既能够保留目标在上一帧的相关信息,同时也可以将当前帧的信息更新到跟踪模型中。

3.4 算法流程

本文所提算法的具体实现流程如下。

1) 初始化。利用在起始帧中给出的目标位置和尺度分别初始化 HOG 特征和 CN 特征的核相关滤波器,获得它们的对偶矩阵参数 $\hat{\alpha}_{\text{HOG}_0}$ 和 $\hat{\alpha}_{\text{CN}_0}$ 及其对应的目标外观参数 \hat{x}_{HOG_0} 和 \hat{x}_{CN_0} 。

2) 位置估计。分别在当前帧中提取目标候选区域的 HOG 特征和 CN 特征,计算它们的核相关滤波响应 \hat{f}_{HOG_t} 和 \hat{f}_{CN_t} ,并将这 2 个响应进行高斯滤波,然后根据(9)式和(10)式分别计算它们滤波后的置信度 f_{HOG_t} 和 f_{CN_t} ;最后按照(11)式计算得到特征融合后的相关滤波响应 \hat{f}_{final_t} ,完成当前帧的目标位置估计。

3) 尺度估计。以当前帧估计的目标位置为中心,根据前一帧目标尺度 s_{t-1} ,按照(13)式进行多尺度采样,并提取它们的 HOG 特征,构建尺度相关滤波器,进行目标尺度变化的极大似然估计,然后根据相邻两帧之间尺度变化的先验分布,完成对当前帧目标尺度的最大后验概率估计。

4) 模型更新。根据(14)式计算特征融合后相关滤波响应的置信度 f_{final} ,如果当前帧及其历史帧的 f_{final} 值都大于 f_{th} ,则分别按照(15)和(16)式更新 2 个核相关滤波器的跟踪模型,否则不更新。

5) 保存跟踪结果,并判断跟踪是否完成。未完成则继续执行步骤 2)。

4 实验结果与分析

4.1 实验环境及参数设置

用于完成本文所提算法的计算机硬件配置如下:CPU 型号为 Intel(R) Core(TM) i5-2500,主频为 3.30 GHz,内存为 8 GB;算法的软件开发平台为 MATLAB R2016b。本文选取 OTB50 公开测试

集^[22]中的 51 组视频序列进行跟踪性能测试,按照目标跟踪中常见的挑战因素将该测试集中的视频序列属性分为 11 类,具体包括光照变化(IV)、目标形变(DEF)、尺度变化(SV)、遮挡(OCC)、运动模糊(MB)、快速运动(FM)、平面旋转(IPR)、非平面旋转(OPR)、目标消失(OV)、背景杂乱(BC)和低分辨率(LR),测试集中的每个视频序列都至少包含 1 种上述属性。

本文将文献[22]中给出的中心位置误差 E_{CL} 、距离精度 P_d 和重叠精度 P_o 。这 3 种评价指标作为目标跟踪结果的评判依据。 E_{CL} 为跟踪到的目标中心位置 (x_t, y_t) 与真实的目标中心位置 (x_g, y_g) 之间的欧氏距离:

$$E_{CL} = \sqrt{(x_t - x_g)^2 + (y_t - y_g)^2}。 \quad (17)$$

P_d 评价指标是指 E_{CL} 小于某一阈值的帧数占该视频序列全部帧数的百分比。 P_o 评价指标则是指跟踪到的目标区域 R_t 与真实的目标区域 R_g 之间的重叠率 S_o 。大于某阈值的帧数占该视频序列全部帧数的百分比,其中 S_o 可表示为

$$S_o = \frac{|R_t \cap R_g|}{|R_t \cup R_g|}, \quad (18)$$

式中: $|\cdot|$ 为区域内的像素个数。

基于以上 3 种评价指标,采用一次性通过评估(OPE)的测试方法进行目标跟踪性能测试。首先给出目标在视频序列起始帧中的位置和尺度,然后在后续的每帧中由目标跟踪算法确定目标的位置和尺度。这是一种直观的测试方法,也是贴合实际应用的一种测试方法。本文进行 P_d 指标评价时的 E_{CL} 阈值设置为 20 pixel,进行 P_o 指标评价时的 S_o 阈值设置为 0.5。

本文所提算法中的具体参数设置如下:用于对相关滤波响应进行高斯滤波的模板尺寸为 3×3 ;尺度估计的金字塔层数 $M=17$,变化步长 $d=0.025$;目标跟踪的置信度阈值 $f_{th}=6$,参考历史帧的数量为 3,更新率 $\eta=0.01$;其余参数设置与 KCF 算法一致。另外,用于对 R_{PS} 和 C_{CFR} 的权重进行调节的参数 ρ 在本文中设置为 0.5。

本文在(8)式中通过 R_{PS} 和 C_{CFR} 定义了目标跟踪的置信度评价函数,他们之间的权重调节系数 $\rho \in [0, 1]$,其中 R_{PS} 代表当前帧的相关滤波响应的峰值尖锐度, R_{PS} 值越大表示相关滤波响应的峰值越尖锐,说明目标跟踪的可靠度越高; C_{CFR} 则代表相邻两帧的相关滤波响应的相似程度, C_{CFR} 值越大,表示相邻两帧的相关滤波响应的相似程度越高,说明目

标跟踪的稳定性越高。 R_{PS} 和 C_{CFR} 分别从不同的角度反映了目标跟踪的置信度情况,其中 R_{PS} 代表当前帧的可靠度, C_{CFR} 则代表连续帧的稳定性。 ρ 值增大,说明 C_{CFR} 在目标跟踪置信度评价中的权重增加。如果 $\rho > 0.5$,则表示 C_{CFR} 在目标跟踪的置信度评价中占据主导因素,反之,则表示 R_{PS} 在目标跟踪的置信度评价中占据主导因素。在 OTB 测试集中,选用 E_{CL} 阈值为 20 pixel 时的 P_d 值作为目标跟踪结果的评价指标,获得参数 ρ 与目标跟踪性能之间的关系,如图 2 所示。从图 2 中可以看出,在 $\rho = 0.5$ 时, R_{PS} 与 C_{CFR} 之间取得了较好的平衡,目标跟踪的效果最佳,因此本文将 ρ 的值设置为 0.5。

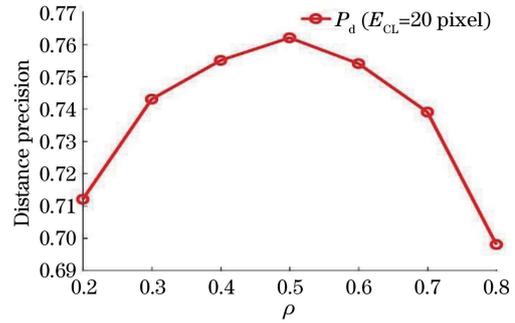


图 2 权重调节系数 ρ 与目标跟踪性能之间的关系
Fig. 2 Relationship between weight adjustment ρ and the target tracking performance

4.2 特征融合性能

为了验证本文提出的自适应权重融合策略的跟踪性能,在 OTB50 测试集中进行 3 种算法的对比实验。3 种算法分别如下:本文提出的使用 HOG 特征和 CN 特征进行自适应特征融合的算法(所提算法),以及分别仅采用 HOG 特征和 CN 特征的 proposed_HOG 算法和 proposed_CN 算法。proposed_HOG 算法和 proposed_CN 算法仅采用单一特征进行目标跟踪,其余参数设置均与所提算法一致。3 种算法的对比实验结果如图 3 所示,其中图 3(a)图例中方括号内的数值为当 E_{CL} 阈值为 20 pixel 时的 P_d 值,图 3(b)图例中方括号内的数值为当 S_o 阈值为 0.5 时的 P_o 值。从图 3 中可以看出,相对于只使用单一 HOG 特征和 CN 特征的算法,采用自适应特征融合后的算法的 P_d 值分别提高了 10.9% 和 15.3%, P_o 值则分别提高了 12.4% 和 19.0%,这说明采用本文提出的自适应特征融合策略可以有效提升目标跟踪的整体性能。

4.3 尺度估计性能

为了验证本文算法的性能,选取 OTB50 测试集中 4 组有代表性的存在尺度变化的视频序列

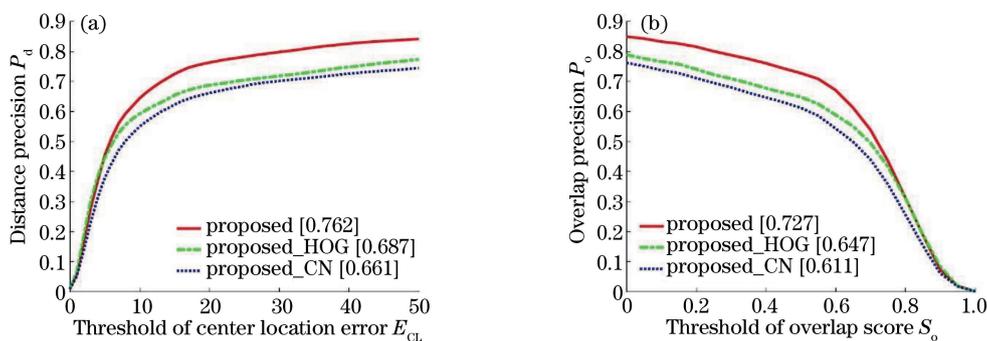


图 3 3 种目标跟踪算法的距离精度曲线和重叠精度曲线。(a)距离精度;(b)重叠精度

Fig. 3 Distance precision curves and overlap precision curves of three target tracking algorithms.

(a) Distance precision; (b) overlap precision

Blurcar2、Dog1、Doll、Carscale 进行测试。实验首先以起始帧的目标尺寸为基准,然后在后续帧中将估计的目标尺寸与起始帧的目标尺寸进行比较,得到估计的目标尺度变化,最后再与真实的目标尺度变化进行比较。在 4 组视频中,视频序列 Blurcar2 的目标尺度变化范围较小,在 0.5 与 3.5 之间;视频序列 Dog1 和 Doll 的目标尺度变化范围较大,相当于起始帧有 15 倍和 18 倍;视频序列 Carscale 的目标尺度变化范围最大,相当于起始帧有 30 多倍。

图 4 所示为估计的目标尺度与真实的目标尺度之间的对比结果。从图 4 中可以看出,本文提出的多尺度估计方法可以较准确地预测目标的尺度变化,即使在目标尺度发生了很大变化的条件下,所提方法仍能较好地预测目标尺度。表 1 所示为所提算法对 4 组视频序列的跟踪结果。其中 Mean E_{CL} 为视频序列所有帧的 E_{CL} 平均值。从表 1 中可以看出,所提算法在这 4 组序列中取得了较优异的跟踪性能。

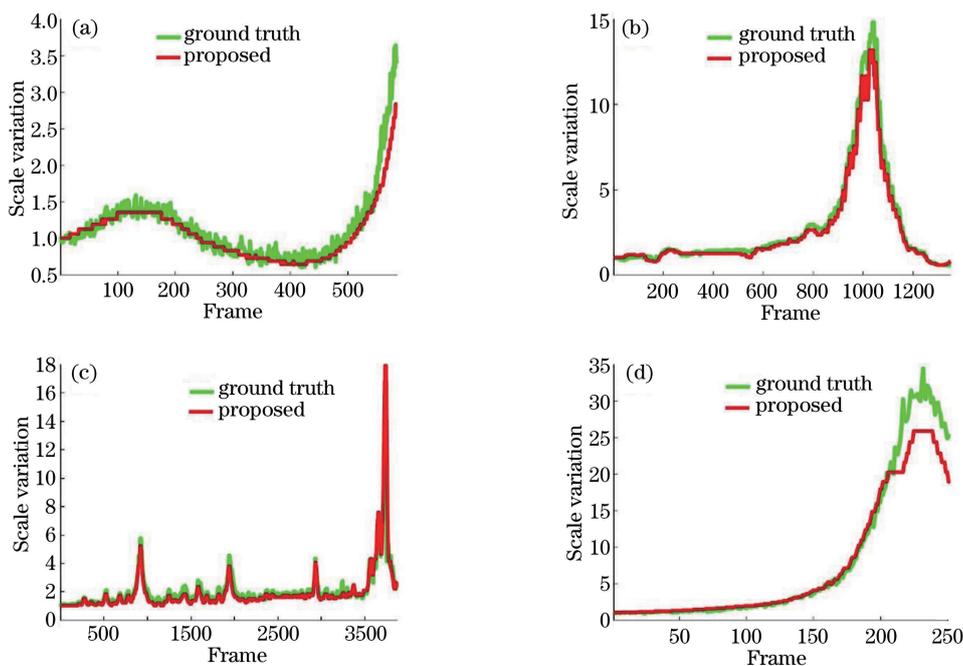


图 4 所提算法对 4 组视频序列进行估计的目标尺度与真实的目标尺度对比。(a) Blurcar2;(b) Dog1;(c) Doll;(d) Carscale

Fig. 4 Comparisons of estimated scale by the proposed algorithm and actual scale on four sequences.

(a) Blurcar2; (b) Dog1; (c) Doll; (d) Carscale

表 1 所提算法对 4 组存在尺度变化的视频序列的目标跟踪结果

Table 1 Tracking results of the proposed algorithm on four scale variation sequences

Sequence	Mean E_{CL}	$P_d/\%$ ($E_{CL}=20$)	$P_o/\%$ ($S_o=0.5$)
Blurcar2	3.42	100.0	100.0
Dog1	3.81	100.0	100.0
Doll	2.26	99.3	99.6
Carscale	3.84	100.0	100.0

4.4 综合跟踪性能

4.4.1 定量分析

为了定量评估所提算法的跟踪性能,利用 OTB50 数据集中的 51 组视频序列对比 33 种算法,这些算法包括所提算法、KCF 算法^[7]、CN 算法^[8]和 DSST 算法^[17],另外还包括在 OTB50 数据集中表

现出色的 29 种算法,如 Struck 算法、TLD 算法、SCM 算法、ASLA 算法、VTD 算法、VTS 算法、CXT 算法和 CT 算法等。

在 OPE 测试方法下,计算这些算法的评价指标 P_d 和 P_o ,对比实验结果如图 5 所示,为了简化显示,这里只列出了在每种指标条件下排名前 10 的算法,其中图 5(a)图例中方括号内的数值为当 E_{CL} 阈值为 20 pixel 时的 P_d 值,图 5(b)图例中方括号内的数值为当 S_o 阈值为 0.5 时的 P_o 值。从图 5(a)中可以看出,当 E_{CL} 的阈值为 20 pixel 时,所提算法的 P_d 值为 0.762,相对于排名第 2 的 DSST 算法,提高了 4.1%;从图 5(b)中可以看出,当 S_o 的阈值为 0.5 时,所提算法的 P_o 值为 0.727,相对于排名第 2 的 DSST 算法,提高了 9.8%。

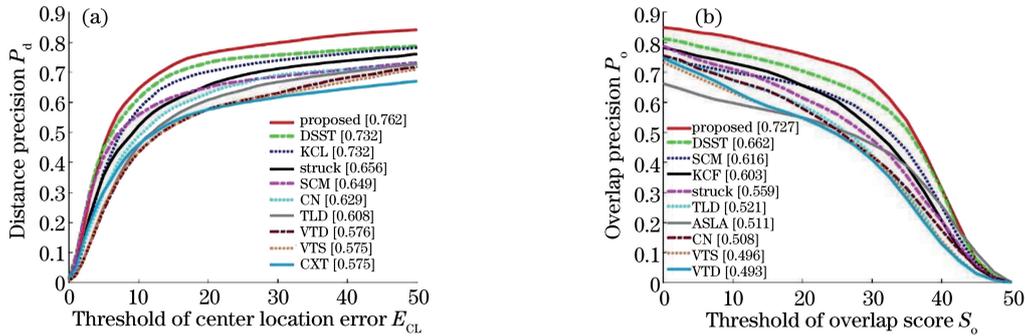


图 5 不同目标跟踪算法的距离精度曲线和重叠精度曲线。(a)距离精度;(b)重叠精度

Fig. 5 Distance precision curves and overlap precision curves of different target tracking algorithms.

(a) Distance precision; (b) overlap precision

为了更全面地评价所提算法的性能,在 OTB50 数据集中,利用图 5 中 P_d 值和 P_o 值排名前 10 的算法分别对比测试 11 种具有不同属性的视频序列,结果

如表 2 和表 3 所示,其中黑体数值为评价指标的最大值。从表 2 和表 3 中可以看出,在 11 种属性中,所提算法的特性大部分超越了其他对比算法。

表 2 排名前 10 的算法对 11 种不同属性的评价指标 P_d

Table 2 P_d scores of the top ten algorithms on eleven attributes

Algorithm	IV	DEF	SV	OCC	MB	FM	IPR	OPR	OV	BC	LR
Proposed	0.780	0.737	0.739	0.761	0.653	0.581	0.704	0.751	0.665	0.714	0.424
DSST	0.730	0.636	0.738	0.692	0.544	0.513	0.768	0.725	0.511	0.694	0.497
KCF	0.657	0.698	0.648	0.695	0.571	0.534	0.691	0.678	0.590	0.676	0.387
Struck	0.558	0.521	0.639	0.564	0.551	0.604	0.617	0.597	0.539	0.585	0.545
SCM	0.594	0.586	0.672	0.640	0.339	0.333	0.597	0.618	0.429	0.578	0.305
CN	0.576	0.607	0.599	0.621	0.551	0.482	0.674	0.645	0.438	0.629	0.408
TLD	0.537	0.512	0.606	0.563	0.518	0.551	0.584	0.596	0.576	0.428	0.349
VTD	0.557	0.501	0.597	0.545	0.375	0.352	0.599	0.620	0.462	0.571	0.168
VTS	0.573	0.487	0.582	0.534	0.375	0.353	0.579	0.604	0.455	0.578	0.187
CXT	0.501	0.422	0.550	0.491	0.509	0.515	0.610	0.574	0.510	0.443	0.371

4.4.2 定性分析

为了更直观地评估所提算法的跟踪性能,选择在定量分析中排名前 5 的算法对 6 组具有不同属性的

视频序列 David (IV)、Basketball (DEF)、Carscale (SV)、Jogging1(OCC)、Trellis(OPR)和 Soccer(BC)分别进行定性分析,结果如图 6~11 所示。

表 3 排名前 10 的算法对 11 种不同属性的评价指标 P_o 。
Table 3 P_o of the top ten algorithms on eleven attributes

Algorithm	IV	DEF	SV	OCC	MB	FM	IPR	OPR	OV	BC	LR
Proposed	0.712	0.733	0.721	0.738	0.591	0.532	0.674	0.702	0.672	0.648	0.419
DSST	0.681	0.610	0.640	0.632	0.528	0.503	0.679	0.632	0.512	0.627	0.437
SCM	0.568	0.565	0.635	0.599	0.339	0.335	0.560	0.575	0.449	0.550	0.308
KCF	0.543	0.628	0.474	0.580	0.561	0.523	0.613	0.579	0.610	0.630	0.355
Struck	0.491	0.473	0.471	0.493	0.518	0.567	0.528	0.506	0.550	0.545	0.410
TLD	0.460	0.456	0.494	0.468	0.482	0.473	0.476	0.497	0.516	0.388	0.327
ALSA	0.503	0.456	0.544	0.451	0.281	0.260	0.488	0.494	0.359	0.468	0.163
CN	0.450	0.511	0.421	0.479	0.480	0.437	0.550	0.501	0.458	0.531	0.399
VTS	0.503	0.441	0.453	0.465	0.328	0.325	0.477	0.496	0.508	0.516	0.183
VTD	0.480	0.443	0.460	0.468	0.320	0.319	0.500	0.510	0.491	0.515	0.170

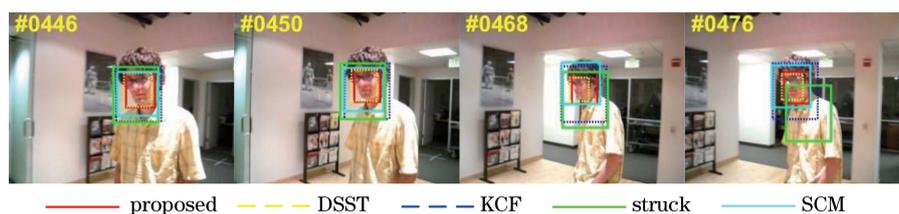


图 6 5 种算法对视频序列 David 的目标跟踪结果对比

Fig. 6 Comparison of tracking results among five algorithms on David sequence

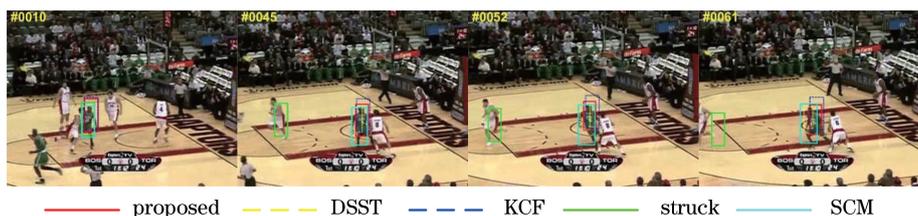


图 7 5 种算法对视频序列 Basketball 的目标跟踪结果对比

Fig. 7 Comparison of tracking results among five algorithms on Basketball sequence

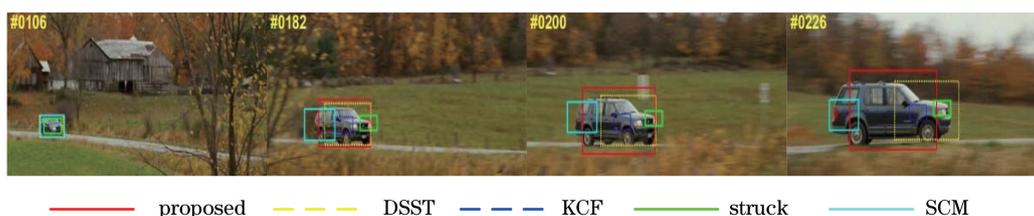


图 8 5 种算法对视频序列 Carscale 的目标跟踪结果对比

Fig. 8 Comparison of tracking results among five algorithms on Carscale sequence

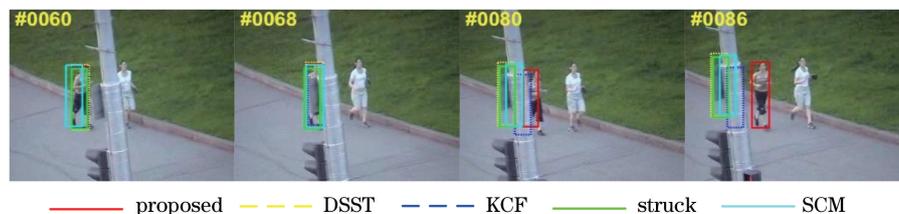


图 9 5 种算法对视频序列 Jogging1 的目标跟踪结果对比

Fig. 9 Comparison of tracking results among five algorithms on Jogging1 sequence

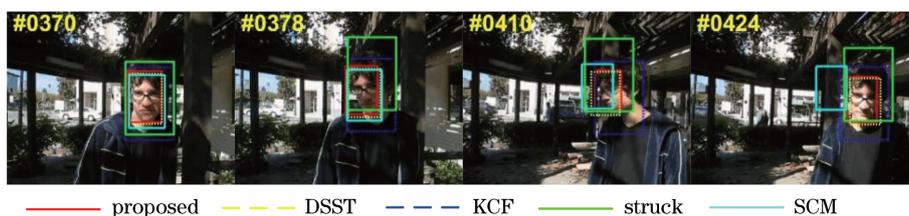


图 10 5 种算法对视频序列 Trellis 的目标跟踪结果对比

Fig. 10 Comparison of tracking results among five algorithms on Trellis sequence

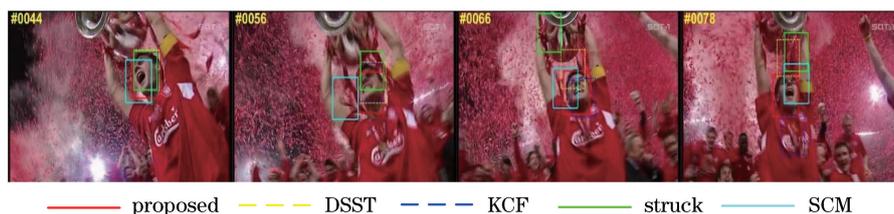


图 11 5 种算法对视频序列 Soccer 的目标跟踪结果对比

Fig. 11 Comparison of tracking results among five algorithms on Trellis sequence

从图 6 中可以看出,目标从第 446 帧开始背景亮度逐渐升高然后降低,存在明显光照变化,只有本文算法和 DSST 算法能准确地跟踪到目标,SCM 算法和 KCF 算法虽然也能跟踪到目标,但给出的跟踪框中包含了大量背景信息,Struck 算法的跟踪结果则出现了明显的跟踪漂移现象。

从图 7 中可以看出,目标在跟踪过程中发生了较大形变,Struck 算法在第 45 帧时丢失了目标,DSST 算法只能定位目标局部,KCF 算法和 SCM 算法给出的跟踪框中则包含了部分背景信息,该情况在第 52 帧和第 61 帧中更明显,只有本文算法在整个过程中能够较准确地跟踪目标。

从图 8 中可以看出,在目标尺度变化较小时,各算法都能较好地跟踪目标,在第 182 帧中当目标尺度发生较大变化时,只有本文算法和 DSST 能够较完整地跟踪目标。在第 200 帧和第 226 帧中,目标尺度进一步变大,DSST 算法只跟踪目标局部,但本文算法仍能跟踪到目标全部。

从图 9 中可以看出,目标在运动过程中逐渐靠近遮挡物,然后被完全遮挡,最后逐步远离遮挡物再次出现在场景中。当目标被严重遮挡后再次出现时,其他算法给出的跟踪框全部停留在了遮挡物上,只有本文算法还能继续准确地跟踪目标。

从图 10 中可以看出,目标在第 378 帧时开始出现了非平面旋转变,此时 KCF 算法和 Struck 算法跟踪框中包含较多背景信息,SCM 算法在第 410 帧时也不能准确跟踪目标,并在第 424 帧时出现了跟踪失败情况,而本文算法和 DSST 算法则能够一

直准确地跟踪目标。

从图 11 中可以看出,目标在跟踪过程中出现了较多背景干扰,在第 44 帧时 SCM 算法的跟踪框包含了许多的背景信息,并在第 56 帧时丢失目标,Struck 算法在第 66 帧时也出现了跟踪失败情况,同时 DSST 算法和 KCF 算法的跟踪框也都包含了大量背景信息,并在第 78 帧时丢失目标,而本文算法则一直稳定地跟踪目标。

5 结 论

本文在核相关滤波基础上,提出了一种自适应特征融合的多尺度目标跟踪算法。首先采用一种自适应权重的融合策略,将 HOG 特征和 CN 特征的核相关滤波输出响应进行融合,用于完成目标跟踪过程中的位置估计;然后以此为中心进行多尺度采样,构建尺度相关滤波器,同时结合相邻两帧的目标尺度变化规律,实现了对目标尺度变化的精确估计;最后根据目标跟踪的置信度,完成对跟踪模型的稳定更新。实验表明,本文算法较好地解决了目标跟踪过程中的光照变化、目标形变、尺度变化和遮挡等问题,有效提高了目标跟踪的成功率,整体性能优于对比算法。同时,在实验中发现,如果在跟踪过程中出现了目标的快速运动情况,由于本文算法暂时没有考虑目标的运动信息,进行目标跟踪时容易出现跟踪漂移甚至失败的情况,因此,如何进一步改进本文算法的特征融合策略,把目标的运动信息特征也进行融合,是下一步研究的重点。

参 考 文 献

- [1] Yilmaz A, Javed O, Shah M. Object tracking: a survey[J]. *ACM Computing Surveys*, 2006, 38(4): 13.
- [2] 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.
- [3] Lu H C, Li P X, Wang D. Visual object tracking: a survey [J]. *Pattern Recognition and Artificial Intelligence*, 2018, 31(1): 61-76.
卢湖川, 李佩霞, 王栋. 目标跟踪算法综述[J]. *模式识别与人工智能*, 2018, 31(1): 61-76.
- [4] Luo H B, Xu L Y, Hui B, et al. Status and prospect of target tracking based on deep learning[J]. *Infrared and Laser Engineering*, 2017, 46(5): 0502002.
罗海波, 许凌云, 惠斌, 等. 基于深度学习的目标跟踪方法研究现状与展望[J]. *红外与激光工程*, 2017, 46(5): 0502002.
- [5] 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.
- [6] Henriques J F, Caseiro R, Martins P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[M] // Fitzgibbon A, Lazebnik S, Perona P, et al. *Computer vision-ECCV 2012. Lecture notes in computer science*. Berlin, Heidelberg: Springer, 2012, 7575: 702-715.
- [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] Bertinetto L, Valmadre J, Golodetz S, et al. Staple: complementary learners for real-time tracking [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 1401-1409.
- [10] Li C, Lu C Y, Zhao X, et al. Scale adaptive correlation filtering tracking algorithm based on feature fusion [J]. *Acta Optica Sinica*, 2018, 38 (5): 0515001.
- [11] Shen Q, Yan X L, Liu L F, et al. Multi-scale correlation filtering tracker based on adaptive feature selection [J]. *Acta Optica Sinica*, 2017, 37 (5): 0515001.
沈秋, 严小乐, 刘霖枫, 等. 基于自适应特征选择的多尺度相关滤波跟踪 [J]. *光学学报*, 2017, 37(5): 0515001.
- [12] Zhao G P, Shen Y P, Wang J Y. Adaptive feature fusion object tracking based on circulant structure with kernel [J]. *Acta Optica Sinica*, 2017, 37(8): 0815001.
赵高鹏, 沈玉鹏, 王建宇. 基于核循环结构的自适应特征融合目标跟踪 [J]. *光学学报*, 2017, 37(8): 0815001.
- [13] Qi Y K, Zhang S P, Qin L, et al. Hedged deep tracking [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 4303-4311.
- [14] Wang X, Hou Z Q, Yu W S, et al. Target scale adaptive robust tracking based on fusion of multilayer convolutional features [J]. *Acta Optica Sinica*, 2017, 37(11): 1115005.
王鑫, 侯志强, 余旺盛, 等. 基于多层卷积特征融合的目标尺度自适应稳健跟踪 [J]. *光学学报*, 2017, 37(11): 1115005.
- [15] Danelljan M, Häger G, Khan F S, et al. Accurate scale estimation for robust visual tracking [C] // *Proceedings of the British Machine Vision Conference 2014*, September 1-5, 2014, Nottingham. UK: BMVA Press, 2014.
- [16] 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.
- [17] Zhang K H, Zhang L, Liu Q S, et al. Fast visual tracking via dense spatio-temporal context learning [M] // Fleet D, Pajdla T, Schiele B, et al. *Computervision-ECCV 2014. Lecture notes in computer science*. Cham: Springer, 2014, 8693: 127-141.
- [18] 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.
- [19] Liu T, Wang G, Yang Q X. Real-time part-based

- visual tracking via adaptive correlation filters [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 4902-4912.
- [20] Bibi A, Ghanem B. Multi-template scale-adaptive kernelized correlation filters [C] // 2015 IEEE International Conference on Computer Vision Workshop (ICCVW), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 613-620.
- [21] Ma J K, Luo H B, Hui B, et al. Robust scale adaptive tracking by combining correlation filters with sequential Monte Carlo [J]. Sensors, 2017, 17(3): 512.
- [22] Wu Y, Lim J, Yang M H. Online object tracking: a benchmark [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2013, Portland, Oregon. New York: IEEE, 2013: 2411-2418.