

DOI: 10.19650/j.cnki.cjsi.J1905219

基于双相关滤波器的多通道尺度自适应目标跟踪^{*}

韩晓微¹, 王雨薇¹, 谢英红¹, 高源², 鲁正²

(1. 沈阳大学信息工程学院 沈阳 110044; 2. 东北大学信息科学与工程学院 沈阳 110004)

摘要:针对目标跟踪中尺度变化和表观剧烈变化导致的目标跟踪不稳定问题,设计了多通道特征融合的尺度估计策略,提出了基于双相关滤波器的多通道尺度自适应目标跟踪算法。考虑到CN特征对姿态及尺度不敏感,以及HOG特征对光照变化和移动都有着较好稳定性,将CN、HOG以及灰度特征进行了特征融合,提升了对于目标表观变化的跟踪鲁棒性。在保证误差风险最小的前提下使用岭回归进行滤波器求解,同时建立了尺度滤波器实现了目标的多尺度判断,使得目标发生尺度变化时能保持稳定的跟踪。使用TB-100数据集在多场景下对算法进行性能测试,证明该算法在目标表观变化、尺度变换、背景干扰等情况下有良好的跟踪效果。

关键词: 目标跟踪; 相关滤波; 尺度变换; 多通道; 特征融合

中图分类号: TP391 TH741 **文献标识码:** A **国家标准学科分类代码:** 460.40

Multi-channel scale adaptive target tracking based on double correlation filter

Han Xiaowei¹, Wang Yuwei¹, Xie Yinghong¹, Gao Yuan², Lu Zheng²

(1. School of Information Engineering, Shenyang University, Shenyang 110044, China;

2. College of Information Science and Engineering, Northeastern University, Shenyang 110004, China)

Abstract: Aiming at the instability problem of target tracking caused by scale change and severe apparent change in target tracking, in this paper a multi-channel feature fused scale estimation strategy is designed, and a multi-channel scale adaptive target tracking algorithm based on double correlation filter is proposed. Considering that the features of CN are insensitive to attitude and scale, and the features of HOG have good stability to illumination change and target movement, the features of CN, HOG and gray scale are fused to improve the tracking robustness to target apparent change. On the premise of ensuring minimum error risk, the ridge regression is used to solve the filter. The scale filter is established to realize the multi-scale judgment of the target, so that when the target scale changes the tracking of the target remains stable. The TB-100 data set was used to test the performance of the algorithm in multiple scenarios. The experiment results show that the algorithm has good tracking effect under the conditions of target apparent change, scale transformation and background interference.

Keywords: target tracking; correlation filtering; scale transformation; multi-channel; feature fusion

0 引言

目标跟踪是计算机视觉领域研究的重要部分,其广泛的应用在安防监控和军工等领域^[1-3]。现如今目标跟踪领域已经取得了非常丰富的研究成果,但由于尺度变换和遮挡等不确定因素,建立既高精度又高速的跟踪算

法仍存在严峻的挑战。

在当前的目标跟踪领域,已经有很多学者对其进行了研究,并且获得了一定的研究成果^[4-6]。Babenko等^[7]提出了多示例学习目标跟踪算法,成功减少了跟踪的漂移现象。Kalal等^[8]将跟踪器、检测器、学习模块3部分相结合,实现了在遮挡情况下的有效跟踪。但上述算法在目标表观剧烈变化时可能会产生漂移现象,进而导致

收稿日期:2019-06-03 Received Date:2019-06-03

^{*} 基金项目:国家自然科学基金(61503274)、沈阳市科技计划(18013015)、沈阳市双百工程计划(Z18-5-013)项目资助

跟踪误差的产生。

2010年,通过相似性判断求解目标响应峰值来确定目标中心位的相关滤波跟踪方法,因其运行效率高并且有较好的跟踪效果而逐渐得到应用。Bolme等^[9]提出了一种最小化平方和误差相关滤波器,实现了对于目标表观变化的有效跟踪。Henriques等^[10]进一步提出了一种核化的循环结构相关滤波器,通过加入正则项建立模型,利用循环矩阵进行稠密采样实现了高效的目标跟踪。之后提出了一种核相关滤波器^[11],在之前的基础上进行了多通道特征的提取,实现了较为鲁棒的跟踪。Danelljan等^[12]通过引入颜色名称特征(color names, CN),实现了对颜色变化的稳定跟踪。Li等^[13]提出了加入尺度池的方法使目标跟踪适应一定的尺度变化。Mueller等^[14]提出了上下文感知滤波跟踪算法,实现了对于背景干扰的有效跟踪。但以上算法多数无法满足尺度变换或未能实现尺度的实时更新。

近年来在相关滤波广泛应用的同时,基于深度学习的目标跟踪算法逐步发展起来。Nam等^[15]通过卷积的多域学习实现目标信息的分解,从而实现有效的跟踪。之后其又提出了树形深度卷积神经网络算法^[16],通过利用多个卷积网络实现对目标的表示,使得跟踪的鲁棒性更强。Ma等^[17]提出了分层卷积特征的跟踪算法,其通过对多层卷积特征的分别提取实现目标位置的逐步确认,显著提高了跟踪性能。Danelljan等^[18]通过加入空间正则项,降低了距目标中心较远特征的影响,解决了跟踪的边缘效应。Sun等^[19]将核相关滤波器与卷积神经网络结合实现了更有效的跟踪。这些基于深度学习的目标跟踪算法在跟踪性能方面有了进一步的提升,但其执行效率较低。

为了同时解决以上的尺度变换、目标表观变化和运行效率问题,本文以双相关滤波器为基础,在保证其高效运行的同时提出了一种基于双相关滤波器的尺度自适应跟踪算法。相较于其他算法,本文算法通过将灰度、方向梯度直方图特征(histogram of oriented gradients, HOG)和CN特征进行融合实现了目标表观特征的提取。建立了尺度滤波器,通过对目标预估位置的确定实现了目标尺度的实时更新,完成了目标跟踪的尺度自适应。

1 相关滤波器跟踪模型

1.1 相关滤波器原理

近年来,由于运行效率高和较好的跟踪效果等优势,相关滤波器被广泛地应用于目标跟踪领域,其工作原理是通过判断选定目标间的相似程度来确定二者是否为同一目标,根据相似性与卷积的转化原理,能将目标图像间的相似性问题转化为对图像和分类器的卷积操作。

式(1)中 p 为输入的目标, h 为相关滤波器, g 为经过判断后的输出目标。

$$g = p \otimes h \quad (1)$$

式中: $\otimes$ 表示卷积。根据卷积定理,时域的卷积等于频域的乘积,可将式(1)问题转化为式(2)。

$$g = F^{-1}(\hat{p} \odot \hat{h}^*) \quad (2)$$

式中: $\hat{p}$ 为 p 的傅里叶变换; $\hat{h}$ 为 h 的傅里叶变换; $\odot$ 表示各元素的对应相乘; F^{-1} 表示傅里叶逆变换; $*$ 表示复共轭,通过式(2)将相关问题最终转化到频域的乘积问题,大大降低了运算的复杂度。

通过式(2)可求得最终的 $\hat{h}$ 为:

$$\hat{h}^* = \frac{\hat{g}}{\hat{p}} \quad (3)$$

式中: $\hat{g}$ 为 g 的傅里叶变换。

1.2 双相关滤波器

双相关滤波器将相关性问题转化为正则化的最小二乘法,其通过求解极值实现滤波器的求解。通过对每张图片中目标及其周围一定范围的图像块进行滤波器的训练,其中目标位置图像块 x 的尺寸为 $M \times N$,在滤波器训练过程中通过循环矩阵实现密集采样,图像块 x_i 中 $i \in \{0, \dots, M-1\} \times \{0, \dots, N-1\}$ 。滤波器通过回归最小化进行求解,具体表示为:

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

式中: $\langle \cdot, \cdot \rangle$ 表示点积; w 为滤波器的参数; y 为期望的目标输出; λ 为正则化参数,能够防止模型过拟合现象,使模型得到更好的泛化能力。通过求解式(4)的极值得到 w 为:

$$w = (x^T x + \lambda I)^{-1} x^T y \quad (5)$$

通过核化的方法^[20],将高维非线性变换到低维函数实现加速计算,其中核函数。

$$k(x, y) = \langle \varphi(x), \varphi(y) \rangle \quad (6)$$

其中, φ 表示将特征向量映射到核空间的函数。

此时可将分类器 $f(x_i) = \langle w, x_i \rangle + b$ 映射到更高维,利用映射函数 φ ,得到分类器为:

$$f(x_i) = w^T \varphi(x_i) \quad (7)$$

由于 w 是 $\varphi(x) = [\varphi(x_1), \varphi(x_2), \dots, \varphi(x_n)]^T$ 中向量空间中的一个向量,根据支持向量机核定理得到的 w 为:

$$w = \sum_i \alpha_i \varphi(x_i) \quad (8)$$

其中 α 是 w 的对偶空间变量,即为滤波器系数。

因此通过式(5)、(7)、(8)最终将 w 的求解转化为 α 的求解问题,求解到 α 为:

$$\alpha = (\varphi(x) \varphi(x)^T + \lambda I)^{-1} y \quad (9)$$

由核函数性质可以得到 $k(x, x) = \varphi(x) \varphi(x)^T$, $K_{ij} = k(x_i, x_j)$, K 表示核空间的核矩阵,由此可以求得 α 为:

$$\alpha = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{y} \quad (10)$$

因此此时高维空间的回归函数为:

$$f(z) = \sum_{i=1}^n \alpha_i k(z, x_i) \quad (11)$$

通过循环矩阵性质可知 $\mathbf{K}$ 仍是循环矩阵,结合循环矩阵的性质求解到傅里叶域中的 α 为:

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}^{xx} + \lambda} \quad (12)$$

其中 $\hat{\alpha}$ 和 $\hat{y}$ 表示他们各自的傅里叶变换, $\hat{k}^{xx}$ 是核矩阵 $\mathbf{K}$ 的第一行,为多通道的线性核函数,其计算公式为:

$$\hat{k}^{xx'} = \mathbf{F}^{-1} \left(\sum_c \hat{x}_c^* \odot \hat{x}'_c \right) \quad (13)$$

通过上述训练过程,根据循环矩阵 $\mathbf{K}$ 构建训练样本和测试样本的核函数,将分类器的求解转化到傅里叶域中,对目标下一帧位置进行求解,其目标响应 $f(z)$ 为:

$$f(z) = \mathbf{F}^{-1}(\hat{k}^{xz} \odot \hat{\alpha}) \quad (14)$$

最后通过求取目标响应最大值求取目标位置,其预测中心位置为:

$$P_t = \operatorname{argmax}_f(z) \quad (15)$$

2 多通道尺度自适应跟踪算法设计

2.1 多通道特征融合设计

本文在双相关滤波器单独使用 HOG 特征基础上,基于式(13)通过对保持目标原始信息的灰度特征、描述颜色的 CN 特征^[21]和能够表现目标方向梯度的 HOG 特征进行融合,来实现多特征的提取工作。

其中 CN 颜色特征对于目标的形变和尺度等具有较好的稳定性。CN 特征相比于常用的颜色特征,其将原有的 RGB 的三通道特征映射到 11 维,实现了将原有的 3 种颜色细化成 11 种颜色,此特征对颜色有着更丰富的表示。因此 CN 特征相比于 RGB 颜色直方图特征在光照、阴影等变化造成的颜色失真能保持较好的稳定性。CN 颜色属性对于一张输入的目标图像 I ,将 a 位置的像素值 $I(a)$ 映射到 CN 空间,得到共 11 维的颜色概率特征图 $[q_1, q_2, \dots, q_{11}]$ 。进一步在保留图片有效信息确保算法精度的基础上,为了提升运行效率,将 11 维的颜色容器中各减去 $1/11$,变为 10 维。

HOG 特征^[22]对光照变化和移动有较好的效果。HOG 特征与光流方向直方图特征(histograms of oriented optical flow, HOF)相比,HOG 特征用于描述不同行为条件下人体在空域的姿态特征,即表示目标运动时的不同动作姿态的变化,HOF 特征用于描述人体在时域的变化,即表示目标运动方向和速度的变化。因此,HOG 特征能够更加明确地获取目标移动时的空间位置。

获取 HOG 特征首先需进行灰度处理,然后利用

Gamma 校正法对目标进行颜色空间归一化,以降低光照变化的影响,实现降噪效果,其压缩公式为:

$$I(x, y) = I(x, y)^{\text{gamma}} \quad (16)$$

之后对每个像素点进行梯度方向和大小的求解,具体表示为:

$$G_x(x, y) = H(x+1, y) - H(x-1, y) \quad (17)$$

$$G_y(x, y) = H(x, y+1) - H(x, y-1) \quad (18)$$

式中: $G_x(x, y)$, $G_y(x, y)$ 分别表示点 (x, y) 处水平和垂直方向梯度, $H(x, y)$ 表示点 (x, y) 的像素值。计算此点的梯度幅值和方向,它们分别表示为:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (19)$$

$$\alpha(x, y) = \tan^{-1} \left(\frac{G_y(x, y)}{G_x(x, y)} \right) \quad (20)$$

最后进行细胞单元、块直方图的构建,通过对块的 HOG 特征进行整体串联,实现目标框 HOG 特征的提取。

CN 特征对姿态及尺度不敏感,但其受光照变化的影响,当目标表观颜色较近时容易产生背景干扰。HOG 特征则适应了此部分干扰,其对光照变化和移动都有着较好的稳定性,但其对目标形变较敏感。因此两个特征有着良好的互补效果。此外,再利用能够描述目标原始信息的灰度特征,通过 3 种特征的有效结合,使得进行融合后的特征对大幅度的表观变化不敏感,因此能提升目标表观剧烈变化时的跟踪准确性。

本文所采用的 HOG 特征的每个 cell 为 4×4 像素,其获取到的特征为 31 维, CN 为 10 维特征,灰度特征为一维,由于单一的目标特征提取会导致跟踪不够准确的情况,因此本文提出了一种可以将这 3 种特征进行融合的方法。由于核相关函数只需进行向量间的点乘运算,因此对于 3 种特征的融合工作就是将这些特征进行维度的串联,所得的最终目标的特征维数为 42 维,其融合方式具体表示为:

$$\mathbf{f} = [f_g, f_c, f_h] \quad (21)$$

其中 f_g, f_c, f_h 分别表示灰度、CN 和 HOG 特征。

2.2 尺度自适应设计

在跟踪过程中,待跟踪目标经常发生尺度变化,设计高效的尺度自适应算法是实现目标稳定跟踪的关键之一。传统的多尺度跟踪算法是通过建立尺度池的方法,对多个固定目标尺度分别计算目标图像的响应,最终通过计算峰值最大的尺度确定目标预估尺度,并且无法实现更新。本文通过建立尺度滤波器^[23],通过双线性滤波器的位置预测,对于每一帧图像均进行尺度的估计,实现实时更新。此外此方法能够在空间中进行多尺度的详尽搜索,使其尺度确定更加精确。

本文使用的尺度变换滤波器结合相关滤波器的预估目标位置,实现对目标尺度的调节。在 t 帧获取到预估

的输出目标框,其大小为 $M \times N$,在其目标框周围获取多个不同尺度的图像块样本,图像块的尺度选择为:

$$a^r w_{t-1} \times a^r N_{t-1} \quad (22)$$

其中 a 为尺度系数, r 的取值为 $r \in \left\{ \left[-\frac{S-1}{2} \right], \dots, \left[\frac{S-1}{2} \right] \right\}$, S 为所取图像块数量。

通过特征提取后利用插值变换实现图像块尺寸的统一,将原始的 $M \times N \times 31$ 的特征转换成 $Q \times SX$ 的结构,位置滤波器实现了对目标后续帧的位置确定,而尺度滤波器是在此基础上进一步精确目标的尺度大小。由于两帧间的尺度变化小于位置变化,并且为了保证运行效率,最终尺度滤波器选择了单一的 HOG 特征,因此 Q 则为 31 维特征,转换后的结构是将每个尺度特征各自构成单独向量进行整体求解,之后利用汉宁窗进行滤波,实现图像边缘处理,最终通过式(12)、(13)和(15)求解尺度滤波器的响应最大值得到最终精确尺度下对应的目标位置。

通过更新滤波器系数 α 和每一帧的目标模型 f ,使得跟踪算法能够更加有效的适应目标尺度变化。其更新公式为:

$$\hat{\alpha}^t = (1 - \beta) \hat{\alpha}^{t-1} + \beta \hat{\alpha} \quad (23)$$

$$\hat{f}^t = (1 - \beta) \hat{f}^{t-1} + \beta \hat{f} \quad (24)$$

其中, β 为学习率, t 为图片帧数。

2.3 算法流程

本文对灰度、HOG 和 CN 特征进行特征融合实现特征提取,使用双相关滤波器确定目标初始位置,利用尺度滤波器进行目标框尺度确定,最终实现尺度自适应的目标跟踪,具体跟踪算法流程如下。

输入:视频的所有图像序列。

输出:目标估计位置 pos 。

初始化:令 $t = 1$,初始化首帧目标位置信息,利用其周围特征信息建立初始滤波器模型。

1) 读取下一帧图像序列,根据上一帧所得目标位及双滤波器模型利用公式(14)计算目标响应,通过式(15)获取最大响应得到目标位置信息。

2) 根据所求得的目标位置信息及尺度滤波器,利用式(15)计算得到其最大响应值从而确定目标最终的位置及尺度信息。

3) 通过求解得到的新的估计目标位置,利用式(23)、(24)进行参数更新。

4) 令 $t = t + 1$,转至步骤1)。

本文算法流程如图1所示。

3 实验结果和分析

本文算法使用的平台为 MATLAB R2016b,计算机为

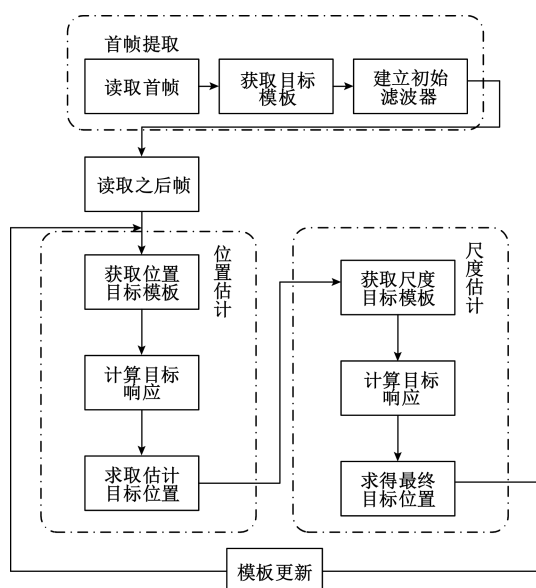


图1 算法流程

Fig.1 Algorithm flow chart

主频 3.40 GHz 的 CPU 以及 8 GB 的内存。为了验证本文算法的有效性,在 TB-100 数据集^[24]中选用了 30 个视频序列进行测试。TB-100 数据集是 Wu 等^[25]于 2015 年提供的,此数据集中共包含 100 个不同场景、不同挑战因素的视频序列,其中的序列数据均已经过人工标注,标注的信息为真实目标框中的原点坐标及目标框的宽、高信息,因此以这些数据集作为实验数据能够更加方便且公正的进行算法测试。

3.1 实验结果定性分析

为了验证算法的有效性,本文对 30 个视频序列进行了实验,本文将选取测试视频序列中 BlurBody、Bolt2、Car1 等 3 个视频序列进行算法定性分析。

第 1 组测试序列 Car1 如图 2 所示,此测试序列存在着尺度变换和背景干扰的影响情况。对于第 3 帧目标距离较近时所有算法均能准确跟踪。220 帧开始,随着目标的逐渐远去和背景干扰的出现,判别尺度空间跟踪 (discriminative scale space tracking, DSST) 算法将干扰误识别为目标,核相关滤波 (kernel correlation filter, KCF) 算法因无尺度变换,最终导致干扰也被识别为目标,分层卷积特性 (hierarchical convolutional features, HCF) 算法对目标的框选也有了一定的偏差,误差最小平方和滤波器 (minimum output sum of squared error filter, MOSSE) 算法则直接跟踪失败。到了 782 帧目标进行了转弯,此时除了本文算法在稳定跟踪,其余算法均跟踪失败。这是由于本文算法使用的双相关滤波器在基准定位的基础上加以尺度滤波器,克服了其它算法中定位与尺度标定的准确性不足现象,使目标框能克服背景干扰,进而实现准

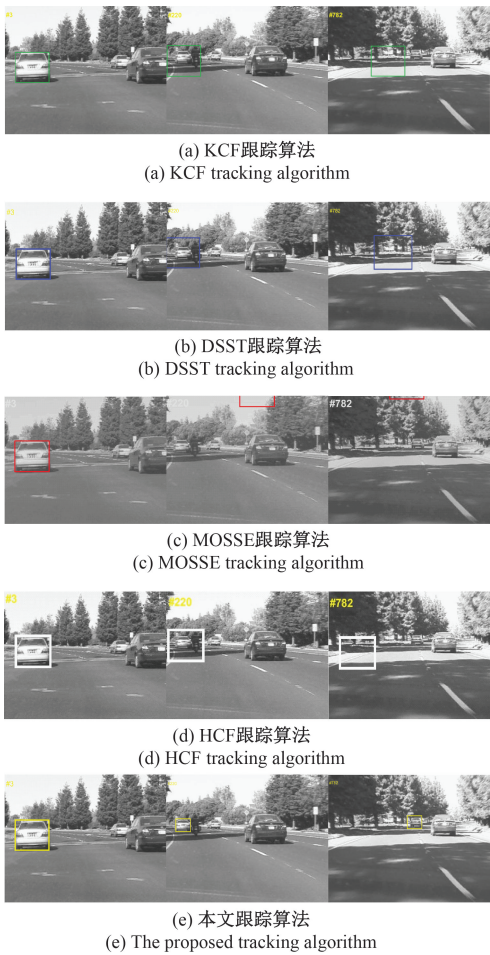


图 2 Car1 测试序列实验结果

Fig.2 Experiment results of Car1 test sequence

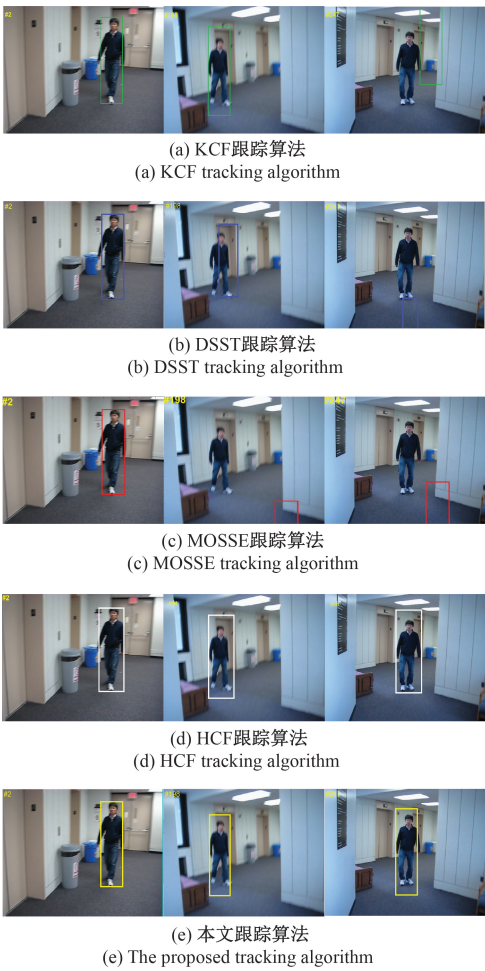


图 3 BlurBody 测试序列实验结果

Fig.3 Experiment results of BlurBody test sequence

确的目标跟踪。

第 2 组测试序列 BlurBody 如图 3 所示,此测试序列存在着强烈的晃动,使得目标被动的产生快速运动。第 2 帧时目标跟踪准确,到 198 帧时随着目标侧身和快速运动的情况,本文算法和 KCF 依旧保持着准确跟踪,而 DSST 对目标产生了一定偏移, MOSSE 算法则直接丢失目标,到了 247 帧随着镜头的强烈晃动,只剩下 HCF 算法和本文算法还保持着稳定的跟踪,其余 3 个算法均丢失目标。这是因为本算法的相关滤波器相较于其它滤波器增加了对特征的描述,使其更适应目标的表现变化,并且本文通过利用线性核降低了模板更新的复杂度,保证了在目标快速运动时跟踪的鲁棒性。

第 3 组测试序列 Bolt2 如图 4 所示,此测试序列存在着快速运动和不断遮挡的情况。通过图 4 中所列出的结果看到序列在第 2 帧时是准确跟踪的,但是到 50 帧时,由于相互交替遮挡的情况,DSST 和 MOSSE 算法错误的跟踪了目标, KCF 则直接丢失目标, HCF 算法有一定偏差但仍保持着跟踪。到了 105 帧时, DSST 和 MOSSE 算

法由于目标的快速移动对后续的跟踪产生了漂移现象, HCF 依旧保持跟踪但有了一定的误差,此时只剩下了本文算法还在准确跟踪。这是由于本文算法通过灰度、CN 和 HOG 特征的融合丰富了目标表现特征描述,能够应对表现剧烈变化情况,并且双滤波器和尺度滤波器的结合使得目标最终跟踪成功。

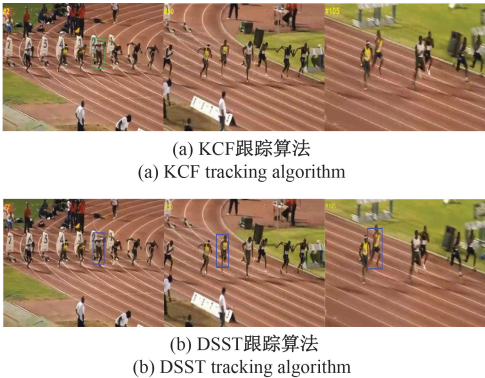




图 4 Bolt2 测试序列实验结果

Fig.4 Experiment results of Bolt2 test sequence

3.2 实验结果定量分析

本小节通过对本文方法和其他 4 个方法 MOSSE、KCF、DSST、HCF 的对比实验来评估这 5 种算法的性能。本文在相同环境下对 5 种算法进行实现,并通过中心位置误差 (center location error, CLE)、距离精度 (distance precision, DP)、重叠精度 (overlap precision, OP) 以及运行速度 (frames per second, FPS) 这 4 种性能评价标准^[26],实现对各个算法的性能比较。

其中 CLE 判断的是算法估计出的目标中心位置与实际标注的目标中心位置的欧氏距离,其单位为像素,CLE 值越小说明跟踪过程中估计目标与实际目标间的距离越小,跟踪效果越好。DP 表示的是在中心位置误差小于一定值时的视频序列的帧数占视频总帧数的百分比,若在中心误差值确定的情况下,百分比越大则说明小于此误差的视频帧的数量所占的比例越高,代表跟踪效果更好。OP 则表示估计位置与实际位置的重叠面积所占两者的覆盖总面积的百分比,具体计算公式为:

$$S = \frac{R_i \cap R_a}{R_i \cup R_a} \quad (25)$$

式中: R_i 表示估计目标所占面积; R_a 表示实际目标所占面积; $\cap$ 表示二者进行交集操作的重叠部分面积, $\cup$ 表示二者并集操作的全部覆盖面积, S 越大则说明预估目标与实际目标的重叠比例越大,跟踪效果越好。FPS 用来衡量算法的运行速度,代表算法每秒运行的帧数,其值越大说明算法运行速度越快,算法的实时性越好。

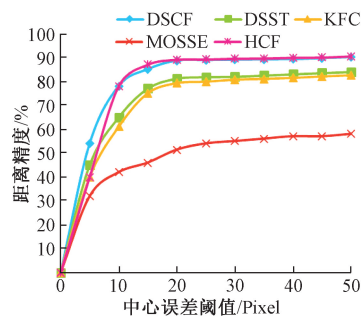
表 1 所示的评价数据是通过对各视频序列的结果取均值所得,其中将指标最好的数据进行加粗显示,排名第二的用下划线进行标注。本文算法与其他算法相比在 CLE 和 OP 两个评价标准下得到了最好的效果。其中对于平均 CLE,本文算法比排名第二位的 HCF 算法提升了 5.4 pixel。对于平均 OP,本文算法比排名第二的 HCF 算法提升了 2.9%。对于平均 DP,本文算法比 HCF 算法低了 0.7%,这是由于 HCF 算法使用了卷积神经网络对特征进行了多层的获取,使得对特征的描述更加丰富,但其运行效率远低于本文算法。

表 1 5 种算法性能评估

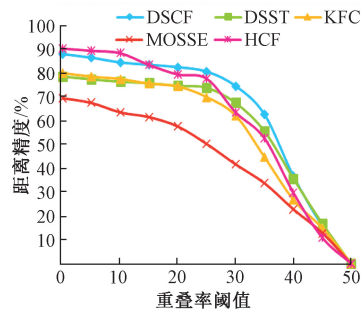
Table 1 Performance evaluation of the five algorithms					
评价标准	本文算法	KCF	DSST	MOSSE	HCF
平均 CLE/pixel	10.3	22.3	21.1	71.5	<u>15.7</u>
平均 DP	<u>88.6%</u>	79.2%	81.2%	51.4%	89.3%
平均 OP	81.1%	70.3%	74.4%	50.6%	<u>78.2%</u>
平均 FPS	29.4	<u>185</u>	29.8	663	10.4

对于平均 FPS,MOSSE 算法使用最小二乘法对滤波器进行求解,且仅使用了单通道的灰度图像,因此 MOSSE 有着相当高的运行速度。KCF 算法通过增加正则项进行滤波器的计算,使用 HOG 特征进行特征提取,滤波器计算方式及特征获取方式的改变使其运算效率低于 MOOSE,但其仍保持着高效的运行速度。DSST 算法因对每一帧图像进行了尺度判断,导致计算量增大,使得运行速度降低。HCF 算法采用了卷积神经网络进行特征的获取,使得其运行效率最低。而本文算法由于多特征融合和尺度滤波器的增加使得运算量增加,但由于使用了线性核的方法,减少了对求解滤波器时的计算量,因此本文算法最终的平均 FPS 相较于 DSST 算法仅降低了 0.4。

本文使用一次通过成功率 (one-pass evaluation, OPE) 方法^[27],此方法是通过初始化首帧,对目标进行一次跟踪,最终求得算法的平均距离精度和成功率。距离精度的计算方法是首先求取估计目标中心点与实际目标中心点的距离,再通过计算两者的距离小于给定阈值的视频帧占总帧数的百分比实现求解,距离精度值越高则说明跟踪效果越好,由于不同的阈值得到的百分比不一样,因此可以获得一条曲线,本文将阈值设为 20 pixel 实现对算法的评估。当某一帧的重叠精度 OP 的值大于设定阈值时,说明此帧是跟踪成功的,成功的帧占所有帧的百分比即为成功率,成功率的值越大则说明跟踪效果越好,因重叠精度的取值范围为 0~1,因此可以绘制出一条曲线,本文将 OP 阈值设定为 0.5 来评估算法成功率。通过图 5 所示曲线可以看出本文算法性能与其他算法相比



(a) 距离精度曲线
(a) Distance accuracy curves



(b) 成功率曲线
(b) Success rate curves

图 5 5 种跟踪算法的距离精度曲线和成功率曲线
Fig.5 Distance accuracy curves and success rate curves of five tracking algorithms

均有较好的表现。

此外本文列出了 6 个干扰情况不同的视频序列在 5 种算法下的距离精度曲线,根据图 6 所示的距离精度曲线可以清晰的看到本文算法在尺度变化、快速移动、背景干扰等多种情况下均有良好的效果。

4 结 论

本文提出了基于双相关滤波器的多通道尺度自适应目标跟踪算法。算法通过灰度特征、对姿态及尺度不敏感的 CN 特征以及对光照变化和移动都有着较好稳定性的 HOG 特征进行特征融合,丰富了对于目标表观特征描述,使得在目标表观剧烈变化时能保持稳定的跟踪。通过岭回归实现滤波器的求解,同时建立了尺度滤波器,实现了目标跟踪的尺度自适应,此外经过实验表明算法在背景干扰的情况下也能实现准确的跟踪。本文使用了 TB-100 数据集对不同的算法分别进行了性能测试,结果显示本文算法的 DP 和 OP 分别为 88.6% 和 81.1%。而本文也存在着一定的局限性,当目标被长时间的完全遮挡时滤波器通过在线更新进行跟踪存在着不确定性,因此在后续工作中需继续对遮挡问题进行研究工作。

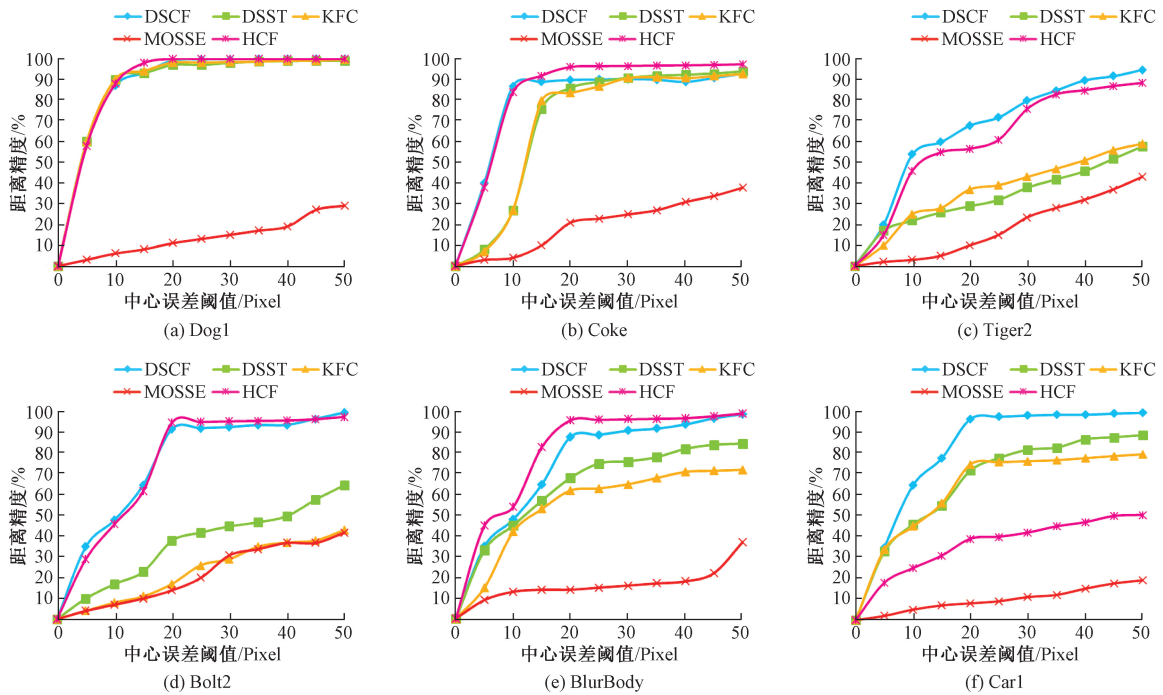


图 6 距离精度曲线
Fig.6 Distance accuracy curves

参考文献

[1] 苑晶, 李阳, 董星亮, 等. 基于运动模式在线分类的 移动机器人目标跟踪[J]. 仪器仪表学报, 2017, 38(3): 568-577.

- YUAN J, LI Y, DONG X L, et al. Target tracking with a mobile robot based on on-line classification for motion patterns [J]. Chinese Journal of Scientific Instrument, 2017, 38(3): 568-577.
- [2] 江南. 基于动态自适应相关滤波的复杂场景目标跟踪算法[J]. 电子测量与仪器学报, 2019, 33(2): 140-147.
- JIANG N. Visual target tracking algorithm based on dynamic adaptive filtering [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(2): 140-147.
- [3] 孙宇嘉, 于纪言, 王晓鸣. 适用于复杂场景的多目标跟踪算法[J]. 仪器仪表学报, 2019, 40(3): 126-137.
- SUN Y J, YU J Y, WANG X M. Multiple object tracking algorithm for the complex scenario [J]. Chinese Journal of Scientific Instrument, 2019, 40(3): 126-137.
- [4] 熊丹, 卢惠民, 肖军浩, 等. 具有尺度和旋转适应性的长时间目标跟踪[J]. 自动化学报, 2019, 45(2): 289-304.
- XIONG D, LU H M, XIAO J H, et al. Robust long-term object tracking with adaptive scale and rotation estimation [J]. Acta Automatica Sinica, 2019, 45(2): 289-304.
- [5] 韦智勇, 杨晓武. 模型自适应分割的人脸跟踪方法[J]. 电子测量与仪器学报, 2018, 32(11): 172-177.
- WEI ZH Y, YANG X W. Face tracking method based on model adaptive threshold segmentation [J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(11): 172-177.
- [6] 汪济洲, 鲁昌华, 蒋薇薇. 一种基于随机场多运动目标跟踪算法[J]. 电子测量与仪器学报, 2017, 31(6): 909-913.
- WANG J ZH, LU CH H, JIANG W W. New multi-target tracking algorithm based on conditional random field [J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(6): 909-913.
- [7] BABENKO B, YANG M H, BELONGIE S. Visual tracking with online Multiple Instance Learning [C]. 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2009, 983-990.
- [8] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-Learning-Detection [J]. IEEE Trans. Pattern Anal. Mach. Intell., 2012, 34(7): 1409-1422.
- [9] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters [C]. 23rd IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010, 13-18.
- [10] HENRIQUES J F, CASEIRO R, MARTINS P, et al. Exploiting the circulant structure of tracking-by-detection with kernels [C]. European Conference on Computer Vision (ECCV 2012), Berlin Heidelberg: Springer, 2012: 702-715.
- [11] 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.
- [12] DANELLJAN M, SHAHBAZ Khan F, FELSBERG M, et al. Adaptive color attributes for real-time visual tracking [C]. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014: 1090-1097.
- [13] LI Y, ZHU J. A scale adaptive kernel correlation filter tracker with feature integration [C]. European Conference on Computer Vision (ECCV), 2014: 254-265.
- [14] MUELLER M, SMITH N, GHANEM B. Context-aware correlation filter tracking [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1387-1395.
- [15] NAM H, HAN B. Learning multi-domain convolutional neural networks for visual tracking [J]. Computer Science, 2015: 4293-4302.
- [16] NAM H, BAEK M, HAN B. Modeling and propagating CNNs in a tree structure for visual tracking [C]. European Conference on Computer Vision, 2016: 1-10.
- [17] MA C, HUANG J B, YANG X K, et al. Hierarchical convolutional features for visual tracking [C]. International Conference on Computer Vision, 2015: 3038-3046.
- [18] DANELLJAN M, HAGER G, KHAN F S, et al. Convolutional features for correlation filter based visual tracking [C]. Proceedings of IEEE International Conference on Computer Vision Workshop, IEEE, 2015: 621-629.
- [19] SUN C, WANG D, LU H, et al. Learning spatial-aware regressions for visual tracking [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 8962-8970.
- [20] 黄宏图, 毕彦彦, 高山, 等. 基于局部敏感核稀疏表示的视频跟踪[J]. 电子与信息学报, 2016, 38(4): 993-999.
- HUANG H T, BI D Y, GAO SH, et al. Visual tracking via locality-sensitive kernel sparse representation [J]. Journal of Electronics and Information Technology, 2016,

38(4): 993-999.

- [21] BERTINETTO L, VALMADRE J, GOLODETZ S, et al. Staple: complementary learners for real-time tracking[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2016: 1401-1409.
- [22] 马浩, 景军锋, 苏泽斌, 等. 基于 LBP 和 HOG 特征的图案织物缺陷检测[J]. 电子测量与仪器学报, 2018, 32(04): 95-102.
MA H, JING J F, SU Z B, et al. Patterned fabric defect detection based on LBP and HOG feature [J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(04): 95-102.
- [23] DANELLJAN M, HAGER G, KHAN F S, et al. Accurate scale estimation for robust visual tracking[C]. Proceedings of 2014 British Machine Vision Conference, 2014: 65.1-65.11.
- [24] 孙锐, 王旭, 张东东, 等. 采用局部线性嵌入的稀疏目标跟踪方法[J]. 电子测量与仪器学报, 2017, 31(8): 1312-1320.
SUN R, WANG X, ZHANG D D, et al. Sparse object tracking method using local linear embedding [J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(8): 1312-1320.
- [25] WU Y, LIM J, YANG M H. Object tracking benchmark[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834-1848.
- [26] 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.
ZHANG W, KANG B SH. Recent advances in correlation filter-based object tracking: A review [J]. Chinese Journal of Image and Graphics, 2017, 22(8): 1017-1033.
- [27] 余铎, 王耀南, 毛建旭, 等. 基于视觉的移动机器人目标跟踪方法[J]. 仪器仪表学报, 2019, 40(1): 227-235.
YU D, WANG Y N, MAO J X, et al. Vision-based object tracking method of mobile robot [J]. Chinese Journal of Scientific Instrument, 2019, 40(1): 227-235.

作者简介



智能、计算机视觉。

E-mail: hwx69@163.com

Han Xiaowei received his B.Sc. degree from Dalian University of Technology in 1991, received his M.Sc. and Ph.D. degree both from Northeastern University in 1997 and 2005, respectively, and was engaged in post-doctoral research at Peking University from October, 2005 to September, 2007. Now, he is a professor in Shenyang University. His main research interests include artificial intelligence and computer vision.



王雨薇, 2017 年于沈阳大学获得学士学位, 现为沈阳大学硕士研究生, 主要研究方向为模式识别、图像处理。

E-mail: 1270506190@qq.com

Wang Yuwei received her B.Sc. degree from Shenyang University in 2017. Now, she is a M.Sc. candidate in Shenyang University. Her main research interests include pattern recognition and image processing.



高源, 2012 年于中国石油大学胜利学院获得学士学位, 2016 年于沈阳大学获得硕士学位, 现为东北大学的博士研究生, 主要研究方向为模式识别、图像处理。

E-mail: gaoyuan012@163.com

Gao Yuan received his B.Sc. degree from Shengli College, China University of Petroleum in 2012, and received his M.Sc. degree from Shenyang University in 2016. Now he is a Ph.D. candidate at Northeastern University. His main research interest includes pattern recognition and image processing.