

DOI: 10.13382/j.jemi.B2104708

基于改进 KCF 算法和多特征融合的车辆跟踪研究

郭秋蕊 李建良 田 焱 刘晓静

(天津科技大学电子信息与自动化学院 天津 300222)

摘要:针对目前车辆跟踪研究算法中,核相关滤波算法(KCF)在复杂背景下存在特征提取单一以及尺度无法自适应的不足,本文提出一种多特征融合的尺度自适应算法。该算法以颜色直方图信息作为颜色特征,将具有更多语义信息的高层卷积特征和拥有较高分辨率的底层卷积特征作为深度特征,并与颜色特征进行自适应特征融合。然后,采用上下文图像对目标背景信息进行约束优化,并通过平均峰值相关能量检测衡量响应置信度,最后利用高置信度的跟踪结果来避免模型易受干扰的问题。通过在 OTB100 数据集上的实验表明,本文算法的精度分别比其他的主流跟踪算法 Staple、SAMF、KCF、TLD、DSST 和 CSK 高出 4.9%, 5.7%, 10.2%, 10.3%, 23.4%, 29.7%。

关键词:核相关滤波;多特征融合;颜色直方图;卷积特征;上下文感知

中图分类号: TP391.4; TN911.7 **文献标识码:** A **国家标准学科分类代码:** 510.40

Research on vehicle tracking based on improved KCF algorithm and multi-feature fusion

Guo Qiurui Li Jianliang Tian Yao Liu Xiaojing

(School of Electronic Information and Automation, Tianjin University of Science and Technology, Tianjin 300222, China)

Abstract: In view of the current vehicle tracking research algorithms, the Kernel correlation filtering algorithm (KCF) has the shortcomings of single feature extraction and the inability to adapt the scale in a complex background, this paper proposes a multi-feature fusion scale adaptive algorithm. The algorithm uses color histogram information as color features, takes high-level convolution features with more semantic information and low-level convolution features with higher resolution as depth features, and performs adaptive feature fusion with color features. Then, the context image is used to constrain and optimize the target background information, and the response confidence is measured by the average peak correlation energy detection, and finally the high-confidence tracking result is used to avoid the problem of the model being vulnerable to interference. In addition, in order to achieve high target tracking accuracy, the algorithm in this paper uses a hierarchical model update strategy to update the extracted features. Experiments on the OTB100 data set show that the accuracy of the algorithm in this paper is better than other mainstream tracking algorithms Staple, SAMF, KCF, TLD, DSST and CSK are 4.9%, 5.7%, 10.2%, 10.3%, 23.4%, 29.7% higher.

Keywords: kernel correlation filtering; multi-feature fusion; color histogram; convolution features; context aware

0 引言

近年来,相关滤波算法因其跟踪精度高,跟踪速度快等优点在跟踪领域占据半壁江山,成为运动车辆跟踪的研究热点之一^[1-3]。2010年,Blome等^[4]首次提出最小误差平方和滤波器(minimum output sum of squared error, MOSSE),该算法是采用密集样本采样的方式去寻找目标

响应最大的滤波器,算法的优点是跟踪速度非常的快,但跟踪精度差强人意。在该算法的基础上 Danelljan等^[5]提出判别尺度空间跟踪(discriminative scale space tracker, DSST)算法,该算法针对目标尺度估计和位置估计分别采用了一维和二维滤波器。接着,针对目标线性不可分的情况,Henriques等^[6]提出一种使用方向梯度直方图(histogram of oriented gradient, HOG)特征的核相关滤波器(kernelized correlation filter, KCF)跟踪算法,该算

法引入循环矩阵采集密集样本,基于核函数提高鲁棒性并用 HOG 特征替代灰度特征,使得目标跟踪精度可以在保持高帧率的状态下得到大幅提升。但是前述几个算法在求解滤波器模板时都会将跟踪目标的历史信息忽略,导致跟踪精度不稳定。在文献[7]中 Mueller 提出上下文感知目标跟踪算法(context-aware correlation filter tracking, CACF),该算法同时利用了目标信息和周围的背景信息,提出一种可以显式的学习目标周围信息的框架,使得目标预测更为准确,其中提出的(context-aware, CA)框架可以与多种相关滤波器集成,体现了算法的高效性,但实时跟踪的效果较差。

针对 KCF 算法在特征提取时方法单一,以及在目标外观建模和尺度自适应方面的缺点,展开对算法基础理论进行推导,并针对其不足进行改进。

本文的创新点在于:

1) 本文算法引入上下文感知模型,使得跟踪目标区域产生高度响应,并通过颜色直方图对上下文模型的目标中心定位,最后在响应层将两者的响应结果融合,从而更好的解决了核相关滤波特征信息采集不足与上下文感知框架特征表达不足的问题。

2) 本文算法为进一步反映出目标跟踪图像的全局信息,将颜色直方图作为颜色特征,该特征对于运动目标模糊和尺度变换等情况不敏感,因此可以得到较好的跟踪效果。

3) 本文算法使用 VGGNet 网络用于深度特征提取,其不同于早期的卷积神经网络。如 AlexNet, GoogLeNet 是通过比较大的卷积核来提取特征的,因此会增加参数数量,提高计算量。本文使用的 VGG-NET-19 通过每个 block 中多个 3×3 的卷积核代替之前的大尺寸卷积核,可以在同等感受野情况下使得参数量减少 1/2,大大降低提取和运算的复杂度,使得跟踪效果不受目标快速运动的影响。

1 基本算法原理

1.1 相关滤波原理

针对相关滤波跟踪算法的提出,其主要是将信号的相关原理应用到运动车辆跟踪中,对运动目标模板与待检测样本进行相关计算,计算结果最大值的点作为目标车辆预测中心点,同时相关理论的应用可以减少计算量,极大的提高目标跟踪的实时性^[8]。

假设两个信号分别为 f 和 g ,则它们的相关性如式(1)所示。

$$f \otimes g = \int_{-\infty}^{+\infty} f^*(t)g(t + \tau) \quad (1)$$

其中, $f^*(t)$ 是 t 时刻信号 $f(t)$ 的复共轭, $g(t + \tau)$

是 $t + \tau$ 时刻的 g 信号。将相关滤波应用到目标跟踪中,事实上就是求解一个模板相关滤波器的问题,将得到的滤波器与待测的目标进行相关运算,并得到最大输出响应。具体过程如式(2)所示。

$$g_1 = f_1 \otimes h_1 \quad (2)$$

其中, f_1 为输入的待检测目标图像, h_1 为相关滤波模板, g_1 为输出响应结果。为提高计算效率,将用傅里叶变换把计算从时域转换到频域上,具体如式(3)所示。

$$G_1 = F_1 \odot H_1^* \quad (3)$$

其中, G_1 是 g_1 的傅里叶形式, F_1 是 f_1 的傅里叶形式, H_1 是 h_1 的傅里叶形式, H_1^* 是 H_1 的复共轭。将式(3)进行简单变形就可求出模板相关滤波器:

$$H_1^* = \frac{G_1}{F_1} \quad (4)$$

1.2 核相关滤波原理

KCF 算法的原理是在第 1 帧图像中获得跟踪车辆的真正位置和其相关的 HOG 特征信息,随后用获取的信息训练跟踪器。当滤波器处在后续帧中对跟踪车辆定位阶段时,通过获得的滤波器模板去预测目标的坐标点以及目标附近采样点坐标的响应值,由此取响应最高的位置,即为目标所在。

以跟踪目标位置的中心点去训练滤波器,截取并通过循环移位^[9]得到 $M \times N$ 运动目标样本,将 $X_i (i \in \{0, 1, \dots, M \times N - 1\})$ 当作训练样本,提取对应的训练样本特征 x_i ,样本标签 y_i 采用二维高斯函数表示,训练的目的是寻找一个 ω 使样本特征和标签特征在给定的情况下误差函数最小,即:

$$\min_{\omega} \sum_i [\omega^T x_i - y_i]^2 - \lambda \|\omega\|^2 \quad (5)$$

其中, λ 是正则化参数,可防止训练过程中出现分类器过拟合情况。接着将 ω 转换为样本特征 x_i 的高维矩阵 $\varphi(x_i)$ 的线性加权:

$$\omega = \sum_i \alpha_i \varphi(x_i) \quad (6)$$

同时引入核函数 $\kappa(x, x') = \varphi^T(x) \varphi(x')$,由此求解 ω 的问题就转化成了求解滤波函数系数 α 的问题:

$$\hat{\alpha} = \frac{\hat{y}}{\hat{\kappa}^{xx} + \lambda} \quad (7)$$

其中,符号 \wedge 表示的是针对变量的傅里叶变换; α 为需要求解的滤波器系数; y 是由标签 y_i 组成的标签向量,其中 $\hat{\kappa}^{xx} = \kappa(x, x')$ 表示的是样本特征 x 的核自相关量。

训练滤波器模板完成后,在新一帧的目标图像中,以上一帧目标位置为中心,计算出检测区域的宽和高来截取检测样本,并提取检测样本特征 z ,得到输出响应:

$$\hat{f}(z) = \hat{k}^{xz} \odot \hat{\alpha} \quad (8)$$

其中, $\hat{\mathbf{k}}^{xz}$ 为目标样本特征 x 和检测样本特征 z 的核互相关向量; \odot 表示点乘。最后输出响应中最大的区域即为当前帧运动目标的中心位置。此外, 利用线性插值更新目标样本特征 \hat{x} 和滤波器系数 $\hat{\alpha}_i$ 可以有效适应目标的变化:

$$\begin{cases} \hat{x}_t = \eta \hat{x}_t + (1 - \eta) \hat{x}_{t-1} \\ \hat{\alpha}_t = \eta \hat{\alpha}_t + (1 - \eta) \hat{\alpha}_{t-1} \end{cases} \quad (9)$$

其中, η 表示学习率; \hat{x}_{t-1} 为 $t - 1$ 帧的样本特征; \hat{x}_t 为第 t 帧的目标样本特征。

1.3 上下文感知滤波原理

上下文感知滤波器的原理是借助了相关滤波算法的原理, 但是该算法与传统相关滤波算法不同, 该算法更注重目标车辆区域周围的样本, 所以不单采集车辆区域作为唯一样本, 同时还需对车辆区域周围的背景信息进行采样, 作为训练滤波器的辅助信息。

针对上下文感知算法分类器的训练^[10-11], 为得到新的分类器, 需要通过岭回归的方法去训练。已知上下文感知滤波器的目标函数为:

$$\min_{\omega} \|A_0 \omega - y\|_2^2 + \lambda_1 \|\omega\|_2^2 + \lambda_2 \sum_{i=1}^k \|A_i \omega\|_2^2 \quad (10)$$

其中, A_0 是对运动目标区域采样后构成的样本集, y 代表的是期望输出, A_i 是对中心区域周围循环采样后的样本集, ω 表示相关滤波器, λ_1, λ_2 为系统的正则项系数。为提高计算效率, 一般在核相关滤波中引入循环矩阵, 设 \mathbf{A} 为循环矩阵, 则 \mathbf{A} 可以由矩阵 \mathbf{F}_1 构建, 公式如下:

$$\begin{aligned} \mathbf{A} &= \mathbf{F}_1 \text{diag}(\hat{a}) \mathbf{F}_1^H \\ \mathbf{A}_T &= \mathbf{F}_1 \text{diag}(\hat{a}^*) \mathbf{F}_1^H \end{aligned} \quad (11)$$

其中, a 是样本集 A_0 的一个样本, \hat{a} 表示样本 a 的傅里叶形式。

上下文感知器也同式(11), 需要构建循环矩阵 \mathbf{B} 如下:

$$\mathbf{B} = \begin{bmatrix} A_0 \\ \sqrt{\lambda_2} A_1 \\ \vdots \\ \sqrt{\lambda_2} A_i \end{bmatrix} \quad (12)$$

将期望输出 y 转换成矩阵形式 \bar{y} 如下:

$$\bar{y} = \begin{bmatrix} y \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (13)$$

利用式(12)和(13)可将目标函数式(10)转换为关

于 ω, \mathbf{B} 的函数为:

$$\min_{\omega} \|\mathbf{B}\omega - \bar{y}\|_2^2 + \lambda_1 \|\omega\|_2^2 \quad (14)$$

通过对式(14)中的 ω 进行求导并令其为 \mathbf{O} , 可以求解出相关滤波器 ω :

$$\omega = (\mathbf{B}^T \mathbf{B} + \lambda_1 \mathbf{I})^{-1} \mathbf{B}^T \bar{y} \quad (15)$$

将式(12)和(13)代到式(15)中, 并将结果频域化可得:

$$\hat{\omega} = \frac{\hat{a}_0^* \odot \hat{y}}{\hat{a}_0^* \odot \hat{a}_0 + \lambda_1 + \lambda_2 \sum_{i=1}^k \hat{a}_i^* \odot \hat{a}_i} \quad (16)$$

设 $\omega = \mathbf{B}^T \alpha$ 得到:

$$\alpha = (\mathbf{B}^T \mathbf{B} + \lambda_1 \mathbf{I})^{-1} \bar{y} \quad (17)$$

同理将式(12)代入式(15)后, 再代入式(17)中, 将结果频域化可得:

$$\hat{\alpha} = \frac{\hat{y}}{\hat{a}_0^* \odot \hat{a}_0 + \lambda_1 + \lambda_2 \sum_{i=1}^k \hat{a}_i^* \odot \hat{a}_i} \quad (18)$$

由公式 $f(z) = r_p(\omega, z) = z \cdot \omega$ 可得此时的输出为 $f(z) = z \cdot \omega$, 将 $\omega = \mathbf{B}^T \alpha$ 代入 $f(z)$ 式可得:

$$f(z) = r_d(\alpha, \mathbf{B}, z) = z \cdot \mathbf{B}^T \cdot \alpha \quad (19)$$

其中, \mathbf{B} 除了运动目标之外还包括上下文补丁信息。因此, α 现在由对偶变量 $\{\alpha_0 \ \alpha_1 \ \dots \ \alpha_k\}$ 串联组成。将 \mathbf{B} 和 α 代入式(19)可得:

$$\hat{r}_d = \hat{z} \odot \hat{a}_0^* \odot \hat{\alpha}_0 + \sqrt{\lambda_2} \sum_{i=1}^k \hat{z} \odot \hat{a}_i^* \odot \hat{\alpha}_i \quad (20)$$

2 多特征融合实现尺度自适应

目前在目标跟踪领域大多数判别类算法都仅限于估计运动对象的位置, 带来的问题就是随着目标车辆的形变, 尺度无法自适应调整大小, 这类算法都是在第1帧选取目标车辆时采用固定的跟踪框跟踪和预测目标, 这就导致无法更好地应对尺度自适应的问题, 还会影响处于更新过程中的模型, 使滤波模板增多或者减少了携带的目标周围背景中原本的信息量^[12-13]。针对上述问题, 本文提出改进核相关滤波(KCF_CACF)和多特征融合尺度自适应的车辆跟踪算法, 改进的核相关滤波算法主要是通过结合上下文感知框架来改进本身算法的不足, 在分类器训练阶段为解决循环移位带来的缺点, 首先对上下文信息进行采样提取背景信息, 然后利用颜色直方图提升信息的表征能力, 由此可以降低背景信息过多的干扰, 减少运动车辆原本信息的流失。最后在改进的算法基础上结合神经网络提取更多目标车辆的外观信息, 使得本文的跟踪算法具有更强的鲁棒性。

2.1 多特征提取

1) 颜色直方图特征

颜色直方图主要是把直方图作为目标图像的颜色特征,在给定的颜色空间基础上,将直方图作为目标车辆的特征,依照其颜色分布进行特征匹配,由于颜色直方图是对车辆表面颜色分布进行统计,所以对运动车辆的尺度,姿态变化影响并不敏感。本文算法的颜色直方图是在 HSV 色系下提取的^[14]。所谓 HSV 空间是一种不同于 RGB 的直观颜色模型,该模型的特点是将色彩的明暗、色调以及鲜艳程度利用六角锥模型直观的表达出来,这样利于进行颜色之间的对比。

按照对颜色变化的敏感程度将 HSV 的 3 个分量量化,设置量化后 3 个分量的取值范围分别是 $\{0, 1, \dots, L_H - 1\}$, $\{0, 1, \dots, L_S - 1\}$, $\{0, 1, \dots, L_V - 1\}$, 按照 $[H, S, V]$ 的形式排列成一个矢量,其范围是 $\{0, 1, \dots, L_H - 1, \dots, L_H + L_S - 1, \dots, L_H + L_S + L_V - 1\}$, 设颜色 i 的像素点个数为 m_i , 则总数为:

$$M = \sum_{i=0}^{L_H+L_S+L_V-1} m_i \quad (21)$$

颜色 i 出现的概率即为颜色直方图,即:

$$p_i = \frac{m_i}{M} \quad (22)$$

本算法基于颜色直方图进行运动目标跟踪时,采用巴氏距离作为两直方图相似度的度量,计算公式为:

$$\rho(p, q) = \sum_{k=1}^m \sqrt{p(k) \cdot q(k)} \quad (23)$$

$$d(p, q) = \sqrt{1 - \rho(p, q)}$$

其中, ρ 为两直方图的巴氏系数, p 为候选目标直方图分布, q 为模板直方图分布, d 为两直方图的巴氏距离,其值大小表示两直方图相似程度的高低。

2) VGG-NET 提取深度特征

本算法采用已经训练过的 VGG-NET-19 用于深度特征提取,该神经网络的扩展性和泛化性都有较强提升,卷积网络在前向传播过程中随着卷积层的深浅变化得到的信息不同,因此采集不同卷积层的特征映射来表示运动车辆的外观信息^[15]。

具体来说,在 VGG-NET 网络模型中,首先输入运动车辆图像,按照其都服从的分布模式对目标车辆做归一化处理,接着经 5 个卷积块提取深度特征,由最大化池化操作后从输出通道个数为 64 一直到 512,最后经过 3 个全连接层后将运动图像变平。从实验结果发现,在跟踪复杂序列时,使用 VGG-19 网络的最高层 Conv5 在运动目标处于遮挡和变形情况下不易跟踪失败;在跟踪简单序列时,Conv5 特征的跟踪效果不如 Conv3 和 Conv4 的特征好^[16],所以本文算法基于 VGG-NET-19 提取所需的深度特征,以满足不同序列的目标跟踪需求。

在利用神经网络提取所需特征前首先对已有图像做前期处理,将目标图像进行缩放至 224×224 的大小以满

足计算需求。接着设提取的 Conv3-4, Conv4-4 和 Conv5-4 的 3 层深度特征为 df_layer_1 , df_layer_2 及 df_layer_3 。在计算特征前,由于本文算法存在边界效应,如果不利用窗函数可能会导致频谱信号泄露,所以本文通过二维汉宁窗来缓解边界效应,那么最终用于描述目标外观的深度特征 f 为:

$$\begin{cases} f_1 = df_layer_1 * hannwin \\ f_2 = df_layer_2 * hannwin \\ f_3 = df_layer_3 * hannwin \end{cases} \quad (24)$$

2.2 自适应特征融合

KCF 跟踪算法只使用单一特征对目标车辆进行建模,不能准确判断车辆位置,本文算法使用两种特征进行训练,其中深度特征编码高层语义信息,对外形变化不敏感,可以用于粗定位;而浅层特征具有更高的细节分辨率,适合精确定位。将两种特征分开处理,鲁棒性方面由深度特征负责,准确性方面由浅层特征负责,两种特征采用自适应融合策略^[17-18],实现特性互补。本文算法将利用神经网络采集的 3 层卷积特征作为深度特征,以颜色直方图特征作为浅层特征,两种特征分别训练相关滤波器,构建两个独立的外观模型,如图 1 是采用自适应特征融合策略来结合两种特征的响应图。

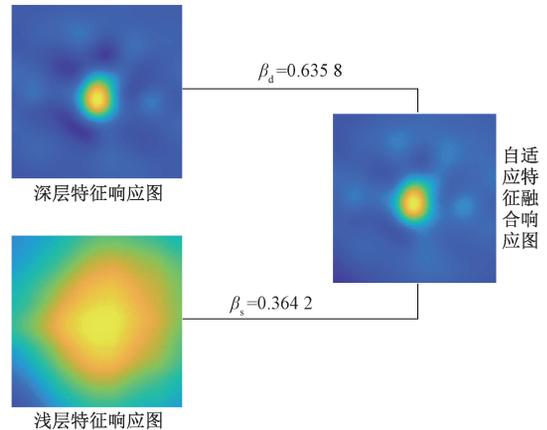


图 1 自适应特征融合图

Fig. 1 Adaptive feature fusion map

$$y_B(t) = \beta_d y_d(t) + \beta_s y_s(t) \quad (25)$$

其中, y_d 表示深层特征分数, y_s 表示浅层特征分数, y_B 表示两种分数加权得到的总分, $\beta = (\beta_d, \beta_s)$ 表示深、浅分数的权重。

用特征响应图反映运动目标定位的精确性和鲁棒性,精确性与预测车辆周围的响应锐利程度有关,主峰越尖表示准精确性越强;鲁棒性与主峰到干扰峰的间隔有关,主峰到次峰的距离越大表示鲁棒性越强^[19]。为了评估预测目标的可靠性,本文采用了一种预测质量评估方法。

$$\xi_t \cdot \{y\} = \min_t \frac{y(t^*) - y(t)}{\Delta(t-t^*)} \quad (26)$$

其中, y 表示目标图像搜索区域的检测分数函数, $y(t) \in \mathbb{R}$ 是位置 $t \in \mathbb{R}^2$ 的目标预测分数, t 表示候选预测目标。 Δ 距离函数定义为:

$$\Delta(\tau) = 1 - e^{-\frac{\kappa}{2}|\tau|^2} \quad (27)$$

引入松弛变量 $\xi = \xi_{t^*} \{y_\beta\}$, 基于等式(25), 共同估计分数权重 β 和目标状态 t^* , 结合最大化质量评估式(26), 经最小化损失函数得到:

$$\begin{aligned} \text{minimize: } & L_t \cdot (\xi, \beta) = -\xi + \mu(\beta_d^2 + \beta_s^2) \\ \text{subject to: } & \beta_d + \beta_s = 1, \beta_d \geq 0, \beta_s \geq 0 \end{aligned} \quad (28)$$

$$y_\beta(t^*) - \xi \Delta(t^* - t) \geq y_\beta(t), \forall t \in \Omega$$

在本实验中, 从深层和浅层分数中分别寻找局部极大值, 将局部极大值按照响应值大小进行排序筛选后作为有限候选状态 Ω , 通过等式(28)优化每个状态 $t^* \in \Omega$, 选择具有最低总体损失的候选状态 t^* 作为最终的预测结果, 并得到相应的权重系数 $\beta = (\beta_d, \beta_s)$, 其中 $\beta_d = 0.6358, \beta_s = 0.3642$ 。

2.3 响应置信度估计

本算法引入 APCE 遮挡判别机制应对目标车辆发生遮挡时会出现跟踪失败的问题, 若响应图只有一个明显峰值, 说明跟踪情况稳定, 反之, 存在波动则说明跟踪发生漂移。假定响应图的最大响应值为 F_{\max} , 最小的响应值为 F_{\min} , 响应图中 w 行 h 列的响应值为 $F_{w,h}$, 则平均峰值相关能量(APCE)^[20]表示如下:

$$APCE = \frac{|F_{\max} - F_{\min}|^2}{\text{mean}\left(\sum_{w,h} (F_{w,h} - F_{\min})^2\right)} \quad (29)$$

3 实验结果

3.1 实验配置

本文算法的硬件实验平台: Intel(R) Core(TM) i5-10300 H CPU @ 2.50 GHz 的 PC, 在软件 MATLAB2019a 上进行仿真实验。在 OTB100^[21] 的数据集上进行实验, 将本文提出的目标跟踪算法与经典的目标跟踪算法作多角度对比。

3.2 定量分析

本文算法采用精度图(precision plot)和成功率图(success plot)来对跟踪效果进行分析, 如图2~6所示; 具体的说, 精确图和成功率图分别与中心位置误差(center location error, CLE)与重叠精度(overlap precision, OP)相关。精确图的绘制是指 CLE 小于阈值的帧数占视频序列总帧数的百分比; 成功率图的绘制是指 OP 小于阈值的帧数占视频序列总帧数的百分比。

$$CLE = \sqrt{(x_c - x_T)^2 + (y_c - y_T)^2} \quad (30)$$

其中, (x_T, y_T) 是真实目标中心坐标; (x_c, y_c) 是跟踪目标中心坐标。

$$OP = \frac{|r_c \cap r_T|}{|r_c \cup r_T|} \quad (31)$$

其中, r_c 为跟踪目标面积, r_T 为目标真实面积。

本实验首先分别在 OPE, TRE 和 SRE 三种不同跟踪策略下进行算法评估, 将3种情况下的精度图与成功率图放在一起进行比较, OPE 是用第1帧的目标位置坐标进行初始化, 然后运行算法得到折线图; TRE 和 SRE 是分别在时间(不同帧开始)和空间(不同位置预测框)上打乱顺序进行算法评估, 如图2所示, 在 OPE 下, 算法的成功率为 0.800, 比 Staple 算法提升 4.6%, 距离精度为 0.842, 比 Staple 提升了 4.9%, 比 SMAF 提升了 5.7%, 比 KCF 提升了 10.2%。在 TRE 实验中, 算法成功率为 0.792, 比 SAMF 算法提升了 1.9%, SAMF 距离精度与本文算法保持较小的精度误差, 但比其他算法提升了很多。在 SRE 实验中本文算法较 Staple 算法成功率略有下降 0.7%, 距离精度与 Staple 算法和 SAMF 算法有较小的差距, 但比其他算法有很大的提升, 比 DSST 算法提升 5.5%。综上, 可以看出本文算法在3种策略下跟踪效果良好。

为了更直观地体现各个跟踪算法在不同属性下的跟踪效果, 采用一次跟踪评估(OPE)作为评估方式, 分别列出几种算法在快速运动、背景杂波、光照变化和遮挡情况下的精度图与成功率图。在快速运动情况下, 本文算法在精确度上较 Staple 算法来说提升了 8.7%, 在成功率上提升了 6.9%, 而与 KCF 算法相比在精确度上提高 12.8%, 在成功率上提高 13.4%, 如图3所示。在背景杂波情况下, 本文算法在精度上较 Staple 和 KCF 算法提升了 9.5%, 在成功率上较 Staple 提升 7.6%, 较 KCF 算法提升 13.4%, 如图4所示。在光照变化情况下, 本文算法在精度上较 Staple 算法提升了 4.2%, 在成功率上提升 3.1%, 而与 KCF 算法相比, 精确度提高 5.5%, 成功率提高 15.6%, 如图5所示。在目标遮挡情况下, 本文算法在精度上较 SAMF 算法提升 0.5%, 在成功率上提升 2.8%, 同 KCF 算法比较, 精确度提升 9.5%, 成功率提升 19.2%, 如图6所示。综上所述, 该算法在相对复杂的运动背景下有较强的鲁棒性。

3.3 定性分析

从前述实验结果图中表明本文算法在目标跟踪时有一定优势, 所以应用该算法和其他主流的6个算法在具有代表性的4个车辆视频中进行定性分析, 选取的4个视频集包含多种不同的属性, 如表1所示。所选的视频序列分别包含两到三种的复杂属性, 其中有光照变化

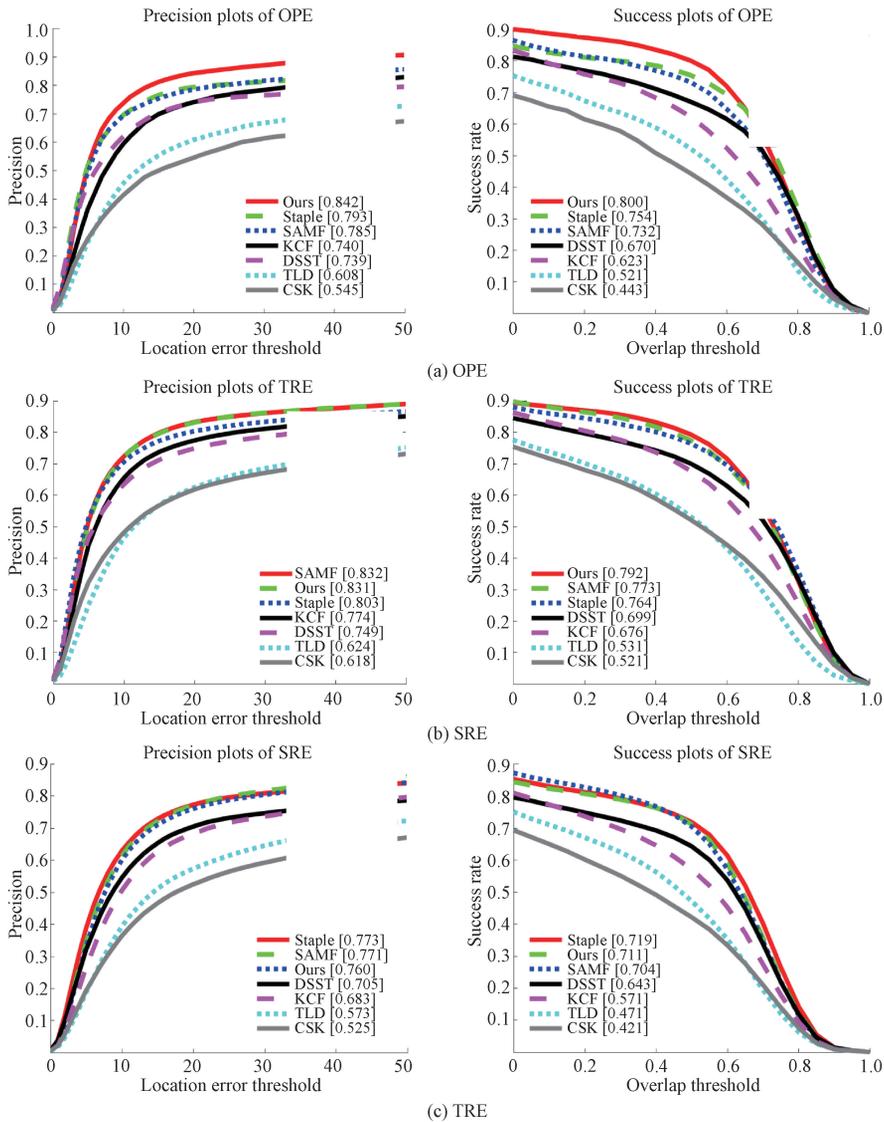


图 2 OPE、SRE、TRE 的精度图与成功率

Fig. 2 Precision and success plots of OPE, SRE, TRE

(illumination variation, IV), 尺度变换 (scale variation, SV), 背景复杂 (background clutters, BC), 遮挡 (occlusion), 快速运动 (fast motion, FM) 和离开视野 (out-of-view, OV)。

表 1 车辆视频序列属性表

Table 1 Vehicle video sequence attribute table

测试视频序列	视频帧数	相关属性
Car4_1	659	IV SV
carDark_1	393	IV BC SV
carScale_1	252	SV OCC FM
Suv_1	945	OCC OV

在 Car4_1 的灰度视频中, 主要体现了视频序列的光照变化, 几种算法在车辆行驶存在尺度变换的情况下跟

踪框保持稳定跟踪, 在视频第 225 帧时, 由于车辆驶入桥下, 光照发生明显变化。车辆从光线充足到光线被遮挡的环境下使得跟预测踪框偏移, 后续跟踪过程中, 除本文之外的跟踪算法外皆有不同程度的漂移, 由图 7(a) 可以看出本文算法跟踪框依旧可以稳定跟踪。

在 carDark_1 视频中包含光照不明, 背景杂乱等属性。背景杂乱是该视频序列的显著特点, 该视频序列是行车记录仪的提取视频, 夜晚会导致视频分辨率降低, 穿行的车辆以及潮湿的地面会引起车灯光线反射或折射从而使图像处理的复杂度增加, 视频前期算法跟踪效果稳定, 在后续视频帧中出现车辆超车, 变道行驶等干扰情况时, TLD 和 Staple 算法跟踪框有明显漂移问题, 本文算法也在复杂背景的跟踪过程中出现短暂时漂移和尺度无法

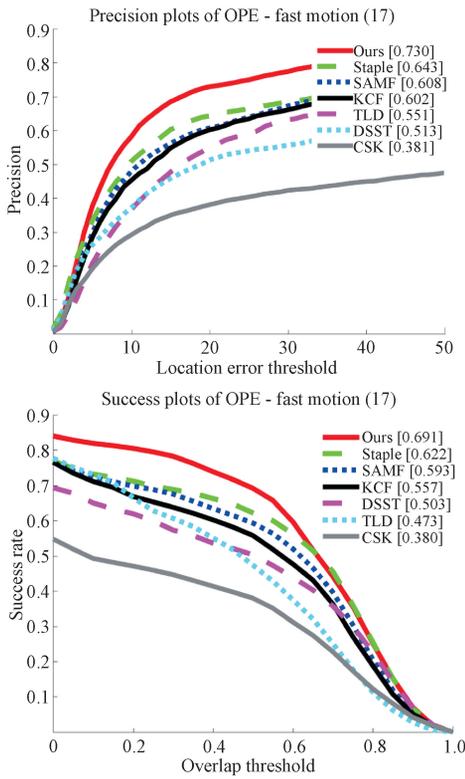


图 3 快速运动精度图与成功率图

Fig. 3 Precision and success plots of fast motion

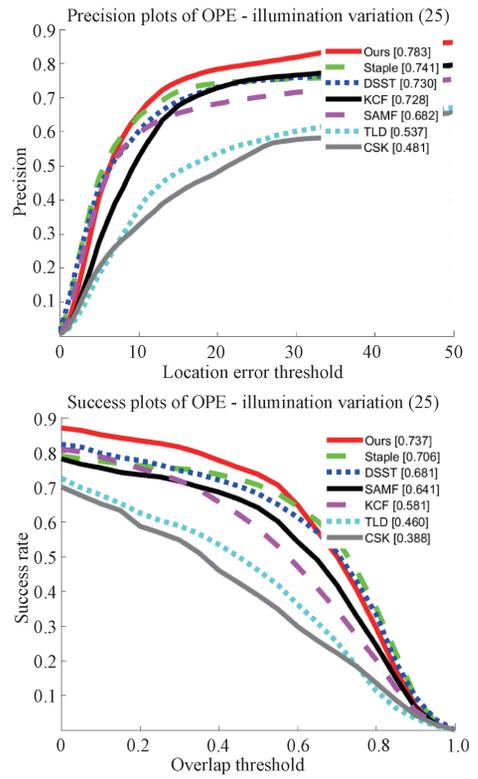


图 5 光照变化精度图 and 成功率图

Fig. 5 Precision and success plots of illumination variation

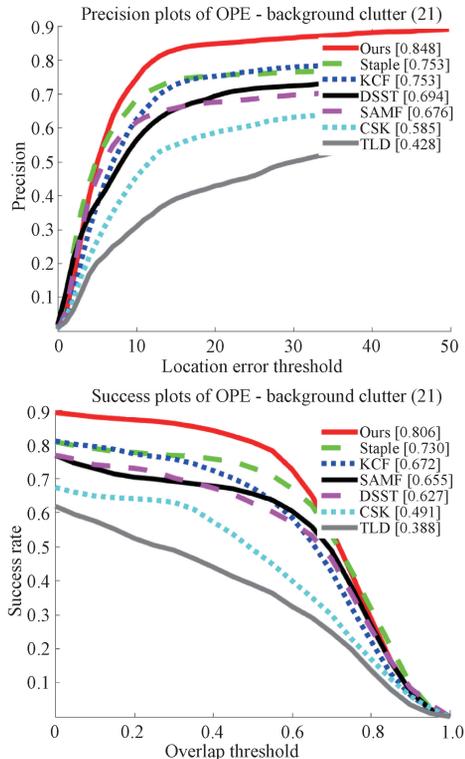


图 4 背景杂波图 and 成功率图

Fig. 4 Precision and success plots of background clutter

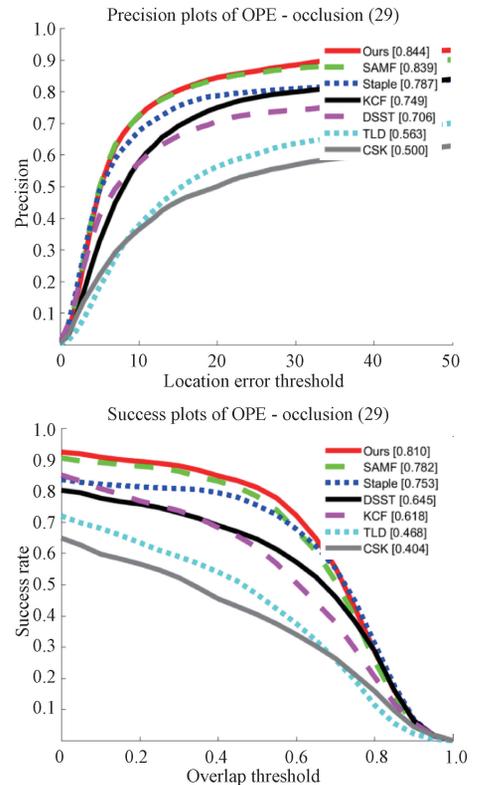


图 6 遮挡时精度图 and 成功率图

Fig. 6 Precision and success plots of occlusion

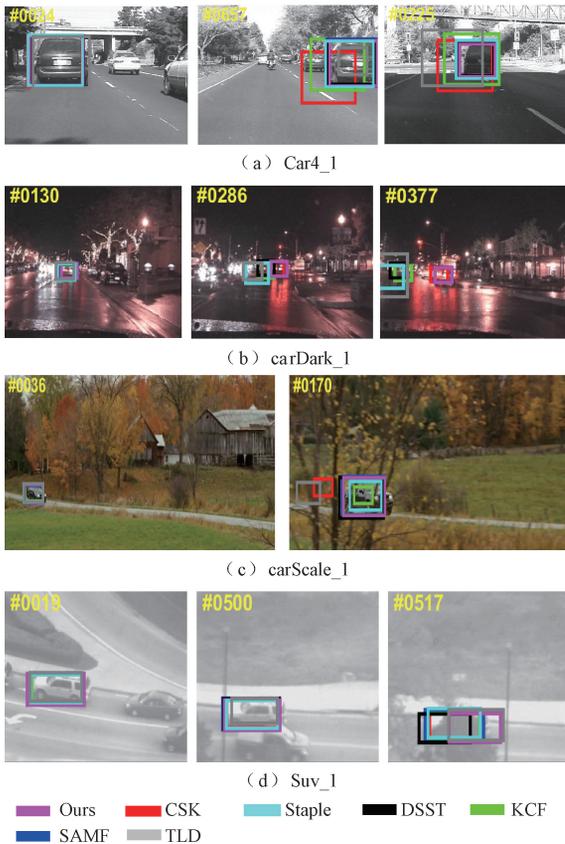


图 7 不同算法在 4 个车辆视频集上的跟踪效果
Fig. 7 Comparison of tracking effects of different algorithms on 4 car video sets

自适应问题,但是总体比其他的主流算法跟踪稳定。

在 carScale_1 视频中主要体现运动车辆快速运动、尺度变换以及部分遮挡属性。由于车辆由远及近行驶,视频中运动车辆在尺度上发生了由小到大的变换,而横向视频序列也使得车辆与背景的相对速度比其他运动目标快,在视频第 36 帧时,各个算法跟踪框与目标车辆匹配,到 170 帧时随着物体临近拍摄位置以及树木的部分遮挡,算法跟踪框皆有变换和漂移,明显看出在跟踪过程中 DSST 和本文算法跟踪框和运动车辆外形跟踪匹配度高。

Suv_1 是一个灰度视频,该视频序列主要属性是有部分遮挡和全遮挡的属性,即运动车辆在视频中发生连续部分遮挡并且有离开视野的情况发生,初始的跟踪效果几种算法都基本稳定,在第 500 帧后出现电线杆、车辆还有树木遮挡,连续出现 3 次的树木遮挡,造成了部分遮挡和目标离开视野情况,这会导致部分算法在进行预测目标位置和计算后续帧的分类器时遇到样本不足或目标选错等问题,部分算法跟踪框有漂移。而后续随着跟踪继续甚至有算法跟踪框出现丢失,总体来看,本文算法具有轻微漂移,跟踪鲁棒性良好。

4 结 论

针对传统 KCF 算法的不足,本文提出一种结合上下文感知框架的改进算法,在此基础上再结合多特征融合自适应算法,改善原本算法无法尺度自适应、特征提取单一以及跟踪效果较差的情况。在训练阶段对目标车辆周围的上下文信息进行采样以完成图像块的约束与优化,利用颜色直方图特征描述运动车辆外观,提高被跟踪区域的信息表达能力,利用结构简洁明了的 VGG-19 网络提取图像深度特征,将提取的特征采用自适应特征融合策略融合,引入 APCE 机制,减少滤波模板不必要的更新,防止模板多次更新有所污染。通过实验和原本的算法进行对比,结果表明虽然 VGG-19 最后的 3 层全连接层导致计算量有所增加,实时处理能力略微下降,但该算法表明运动车辆在快速移动、光照变化、尺度变换和遮挡等复杂背景下具有较强的鲁棒性,较单一算法来说在车辆跟踪的精确度和成功率上都具有较好的综合性能和实用价值。

参考文献

- [1] 张薇,康宝生. 相关滤波目标跟踪进展综述[J]. 中国图象图形报, 2017,22(8):1017-1033.
ZHANG W, KANG B SH. Overview of the progress of correlated filtering target track[J]. Journal of Image and Graphics, 2017,22(8):1017-1033.
- [2] ZHANG B, LI Z, PERINA A, et al. Adaptive local movement modeling for robust object tracking[J]. IEEE T Circ Syst Vid, 2017,27:1515.
- [3] 孟球,杨旭. 目标跟踪算法综述[J]. 自动化学报, 2019,45(7):1244-1260.
MENG L, YANG X. Overview of target tracking algorithms[J]. Acta Automatica Sinica, 2019, 45(7): 1244-1260.
- [4] BLOME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters[C]. IEEE Computer Vision and Pattern Recognition, 2010: 2544-2550.
- [5] DANELLJAN M, HAGER G, KHAN F S, et al. Accurate scale estimation for robust visual tracking[C]. British Machine Vision Conf, 2014, 65: 1-11.
- [6] HENRIQUES J F, RUI C, MARTINS P, et al. High-speed tracking with kernelized correlation filters [J]. IEEE Trans. On Pattern Analysis & Machine Intelligence, 2014,37(3): 583-596.

- [7] MULLER M, SMITH N, GHANEM B. Context-aware correlation filter tracking [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1387-1395.
- [8] 周正松,陈虹君,周红. 基于多特征融合的尺度自适应 KCF 目标跟踪算法[J]. 四川大学学报:自然科学版, 2020,57:697.
ZHOU ZH S, CHEN H J, ZHOU H. Scale-adaptive kernel correlation filtering tracking algorithm based on multi-feature fusion[J]. J Sichuan Univ: Nat Sci Ed, 2020,57:697.
- [9] 范文兵,赵周鼎,王诗. 多特征融合的自适应相关滤波跟踪算法[J]. 计算机工程与应用, 2018, 54(14): 19-25.
FAN W B, ZHAO ZH D, WANG SH. Adaptive correlation filter tracking algorithm based on multi-feature fusion [J]. Computer Engineering and Applications, 2018,54(14):19-25.
- [10] 万欣,张春辉,张琳,等. 加权全局上下文感知相关滤波视觉跟踪算法[J]. 计算机工程与应用, 2018, 54(16):1-12.
WAN X, ZHANG CH H, ZHANG L, et al. Weighted global context-aware correlation filter visual tracking algorithm[J]. Computer Engineering and Applications, 2018,54(16):1-12.
- [11] 茅正冲,陈海东. 自适应尺度的上下文感知相关滤波跟踪算法[J]. 计算机工程与应用, 2021, 57(3): 168-174.
MAO ZH CH, CHEN H D. Adaptive scale context-aware correlation filter tracking algorithm [J]. Computer Engineering and Applications, 2021,57(3):168-174.
- [12] 邓雪菲,彭先蓉,张建林,等. 基于相关滤波的长时目标跟踪算法[J]. 光电技术及应用, 2019, 40(5): 742-748.
DENG X F, PENG X R, ZHANG J L, et al. Long-term target tracking algorithm based on correlation filtering[J]. Optoelectronics Technology and Application, 2019, 40(5):742-748.
- [13] 夏爱明,伍雪冬. 基于上下文感知和尺度自适应的实时目标跟踪[J]. 红外技术,2021,43(5):429-436.
XIA AI M, WU X D. Real-time target tracking based on context awareness and scale adaptation [J]. Infrared Technology, 2021,43(5):429-436.
- [14] 董春燕,刘怀,梁秦嘉,等. 自适应特征融合的核相关滤波目标跟踪算法研究[J]. 南京师范大学学报, 2020,20(3):50-56.
DONG CH Y, LIU H, LIANG Q J, et al. Research on target tracking algorithm based on adaptive feature fusion and kernel correlation filtering [J]. Journal of Nanjing Normal University, 2020,20(3):50-56.
- [15] 陈威,李决龙,刑建春等. 基于核相关滤波器和分层卷积特征的长时间目标跟踪[J]. 计算机科学, 2019, 46(9):272-276.
CHEN W, LI J L, XING J CH, et al. Long-term object tracking based on kernelized correlation filter and hierarchical convolution features[J]. Computer Science, 2019,46(9):272-276.
- [16] 徐晓光,李海. 多尺度特征在 YOLO 算法中的应用研究[J]. 电子测量与仪器学报, 2021, 6(35):96-101.
XU X G, LI H. Application research of multi-scale features in YOLO algorithm [J]. Journal of Electronic Measurement and Instrumentation, 2021,6(35):96-101.
- [17] 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, 53:193.
PAN ZH F, ZHU Y L. Multi-scale estimation kernel correlation filter target tracking method [J]. Progress in Laser and Optoelectronics, 2016,53:193.
- [18] 王敏,陈金勇,孙康,等. 基于神经网络的高光谱图像异常目标检测方法[J]. 国外电子测量技术, 2019, 1(38):96-101.
WANG M, CHEN J Y, SUN K, et al. Anomaly targets detection for hyperspectral imagery based on neural network [J]. Foreign Electronic Measurement Technology, 2019,1(38):96-101.
- [19] 蓝金辉,王迪,申小盼. 卷积神经网络在视觉图像检测的研究进展[J]. 仪器仪表学报, 2020, 41(4): 167-182.
LAN J H, WANG D, SHEN X P. Research progress on visual image detection based on convolutional network [J]. Chinese Journal of Scientific Instrument, 2020,41(4): 167-182.
- [20] WANG M, LIU Y, HUANG Z. Large margin object tracking with circulant feature maps [C]. IEEE Conference on Computer Vision and Recognition, 2013: 2411-2418.
- [21] WU Y, LIM J, YANG M H. Object tracking benchmark[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015,37(9):1834-1838.

作者简介



郭秋蕊, 2019 年于燕山大学里仁学院获得学士学位, 现为天津科技大学硕士研究生, 主要研究方向为自动驾驶、目标跟踪。

E-mail: ruiqiu0318@126.com

Guo Qiurui received her B. Sc. degree from Liren College of Yanshan University in 2019. Now she is a postgraduate at Tianjin University of Science and Technology. Her main research interests include autonomous driving and target tracking.



李建良, 1996 年于天津大学获得学士学位, 2016 年于河北工业大学获得硕士学位现为天津科技大学副教授, 主要研究方向为检测技术及主动隔振。

E-mail: LJL@tust.edu.cn

Li Jianliang received his B. Sc. degree from Tianjin University in 1996 and M. Sc. degree from Hebei University of Technology in 2016. Now he is an associate professor at Tianjin University of Science and Technology. His main research interests include detection technology and active vibration isolation.