

基于自适应特征选择的多尺度相关滤波跟踪

沈 秋, 严小乐, 刘霖枫, 孔繁镛, 王丹丹

南京航空航天大学航天学院, 江苏 南京 210016

摘要 近年来, 基于相关滤波的目标跟踪算法因其具有很好的跟踪精度和明显的速度优势, 引起了研究人员的极大关注。提出一种基于自适应特征选择的多尺度相关滤波跟踪算法。首先, 提取三种互补特征, 通过相关滤波响应图评估各特征的跟踪性能, 自适应选择最优特征进行位置跟踪; 其次, 预设响应图阈值作为位置相关滤波模型更新的判断条件, 优化模型更新方式; 最后, 引入尺度相关滤波跟踪器, 进一步提高了算法的尺度适应性和跟踪精度。实验部分将该算法和近年来流行的相关滤波及非相关滤波类跟踪算法进行了对比, 结果表明, 该算法在精度上优于其他算法, 同时具有 53.12 frame/s 的实时跟踪速度。

关键词 机器视觉; 目标跟踪; 相关滤波; 自适应特征选择; 自适应尺度

中图分类号 TP391 **文献标识码** A

doi: 10.3788/AOS201737.0515001

Multi-Scale Correlation Filtering Tracker Based on Adaptive Feature Selection

Shen Qiu, Yan Xiaole, Liu Linfeng, Kong Fanqiang, Wang Dandan

College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 210016, China

Abstract Recently, the correlation filter-based trackers have aroused increasing interest because of their good performance and high efficiency. A multi-scale correlation filtering tracker based on adaptive feature selection is presented. Firstly, we extract three complementary features to learn three independent filter models. By comparing the response maps, we evaluate the tracking performance of each feature, and then adaptively select the most representative feature for tracking. Secondly, to better handle occlusions and drifts, we improve the online model update strategy by setting peak response threshold as a criterion. Furthermore, we learn a separate filter model for scale estimation. The experimental results show that the proposed tracker achieves better accuracy compared with state-of-the-art correlation filter-based trackers and other popular trackers when running at 53.12 frame/s.

Key words machine vision; visual object tracking; correlation filtering; adaptive feature selection; adaptive scale

OCIS codes 150.1135; 150.0155; 100.4999

1 引 言

目标跟踪是计算机视觉领域中的一个基础问题, 在视频监控、运动分析、人机交互等方面均有广泛的应用。近十年来, 目标跟踪算法研究从传统的光流法、均值漂移、粒子滤波跟踪算法, 逐渐转向基于学习的跟踪算法。和传统跟踪算法相比, 基于学习的方法在复杂场景下表现出了更好的跟踪性能^[1-2]。

通常, 基于学习的跟踪算法将跟踪问题转化成分类问题, 利用在线学习的判别分类器将当前帧的目标和背景分类, 将分类结果置信度最大的位置定位为目标位置。在这些方法中, 基于相关滤波的目标跟踪算法在兼顾跟踪精度的同时具备最高的速度, 从而引起了研究人员的极大关注^[3]。相关滤波跟踪算法利用最小化误差模型和循环矩阵理论建立目标背景分类器, 学习一个相关滤波模型, 在后续跟踪中通过寻找预测分布的响应峰值来定位目标位置, 并在计算过程中巧妙地使用快速傅里叶变换来提升速度。早期的相关滤波算法如最小均方误差和滤波跟踪器 (MOSSE)^[4] 和核循环结构跟踪器 (CSK)^[5] 仅利用灰度特征进行跟踪, 随后

收稿日期: 2016-10-21; **收到修改稿日期:** 2017-01-02

基金项目: 国家自然科学基金(61201365, 61401200)、江苏省普通高校研究生科研创新计划(SJLX15_0138)

作者简介: 沈 秋(1982—), 女, 博士, 讲师, 主要从事视频压缩与视频处理方面的研究。E-mail: shenqiu@nuaa.edu.cn

Henriques 等^[6]将 CSK 算法从单通道扩展至多通道,研究人员也愈发认识到特征表达在相关滤波跟踪中的重要性。核相关滤波器(KCF)利用梯度方向直方图(HOG)特征代替原始灰度特征^[6],使跟踪精度得到提升,自适应颜色名跟踪(CNT)将颜色名(CN)特征很好地应用于相关滤波框架中^[7],文献[8]结合卷积神经网络特征对相关滤波跟踪进行了改进。另一种思路是从多特征融合入手,尺度自适应多特征跟踪器(SAMF)采用灰度、CN 和 HOG 特征串联^[9],结合模板和像素特征的分类器(STAPLE)分别训练颜色直方图和 HOG 特征分类器并在决策层进行融合^[10],文献[11]提出了一种基于多特征相关滤波的红外目标跟踪方法。以上研究表明,更稳健的特征能够有效提高跟踪精度,但如何寻求特征表达能力和计算复杂度之间的平衡仍然有待研究。此外,针对传统相关滤波跟踪无法处理尺度变化的问题,研究人员提出了多尺度自适应跟踪方法^[9,12-13],但这些方法存在着一定的稳健性问题。

本文在相关滤波框架的基础上,提出了一种基于自适应特征选择的多尺度相关滤波跟踪方法。首先分析不同特征的跟踪性能表现,然后选用三组互补特征训练三组独立的相关滤波模型,通过相关滤波响应图评估各特征的跟踪能力,进而实现跟踪过程中的自适应特征选择和自适应模型更新。此外,本文算法引入了文献[12]的尺度估计方法,实现了目标的尺度自适应跟踪。实验采用标准跟踪数据集 OTB (Object Tracking Benchmark)中公开的 74 个彩色跟踪视频序列对本文算法进行验证^[1],并和近年来的其他流行跟踪方法进行对比,结果表明,本文算法在平均跟踪精度上优于其他方法,同时达到了 53.12 frame/s 的平均跟踪速度。

2 相关滤波跟踪器

2.1 KCF 跟踪器

在相关滤波理论的基础上,Henriques 等^[6]提出了一种在核空间使用相关滤波器的跟踪方法 KCF。在 KCF 跟踪器中,以目标为中心,选取目标周围大小为 $m \times n$ 的矩形区域图像块 \mathbf{x} 来训练线性分类器 $f(\mathbf{x}) = \langle \mathbf{w}, \varphi(\mathbf{x}) \rangle$ 。和传统的移动窗口获取样本方法不同,分类器将 \mathbf{x} 的所有循环移位图像块 $\mathbf{x}_i (i \in \{0, \dots, m-1\} \times \{0, \dots, n-1\})$ 作为训练样本,进行一维或多维特征提取,对应的标签数据 y_i 用一个高斯函数来描述。首先,建立最小化目标函数模型:

$$\min_{\mathbf{w}} \sum_i L[y_i, f(\mathbf{x}_i)] + \lambda \|\mathbf{w}\|^2, \quad (1)$$

采用正则化最小二乘(RLS)分类器模型,并引入核函数,模型变为

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

式中 $\varphi(\mathbf{x})$ 表示原始输入空间到 Hilbert 特征空间的映射,此时目标解可表示为 $\mathbf{w} = \sum_i \alpha_i \varphi(\mathbf{x}_i)$,并可使用核函数 $\kappa(\mathbf{x}, \mathbf{x}') = \langle \varphi(\mathbf{x}), \varphi(\mathbf{x}') \rangle$ 求解。

利用循环矩阵和离散傅里叶变换(DFT)变换,得到核化正则最小二乘(KRLS)的最优解为

$$\boldsymbol{\alpha} = \mathcal{F}^{-1} \left[\frac{\mathcal{F}(\mathbf{y})}{\mathcal{F}(\mathbf{k}^{xx}) + \lambda} \right], \quad (3)$$

式中 \mathcal{F} 表示离散傅里叶变换, $\mathbf{k}^{xx} = \kappa(\mathbf{x}, \mathbf{x})$ 为核 κ 的输出。采用高斯核作为核函数,计算公式为

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

在新的一帧中,通过获取候选窗口图像块 \mathbf{z} 来检测目标的位置,分类器的输出响应为

$$\hat{\mathbf{y}} = \mathcal{F}^{-1}[\mathcal{F}(\mathbf{k}^{xz}) \odot \mathcal{F}(\hat{\boldsymbol{\alpha}})], \quad (5)$$

式中 $\hat{\mathbf{x}}$ 和 $\hat{\boldsymbol{\alpha}}$ 分别表示学习得到的目标外观模板和相关滤波模板,更新方式为

$$\hat{\boldsymbol{\alpha}}_i = \eta \boldsymbol{\alpha}_i + (1 - \eta) \hat{\boldsymbol{\alpha}}_{i-1}, \quad \hat{\mathbf{x}}_i = \eta \mathbf{x}_i + (1 - \eta) \hat{\mathbf{x}}_{i-1}. \quad (6)$$

找到使 $\hat{\mathbf{y}}$ 最大的位置即为待跟踪目标的位置。

根据上述描述,KCF 算法具有以下特点:1)通过循环偏移构建分类器的训练样本图像块,依据循环矩阵结构实现训练样本的密集采样,利用循环矩阵特性将问题的求解变换到傅里叶变换域,从而避免了矩阵求逆

过程,大大降低了算法的复杂度;2)利用核矩阵的循环矩阵性质,将相关滤波器成功扩展到核空间上,并能够处理多通道特征计算,同时达到和原始图像空间相同的计算速度。因此,KCF跟踪器具有灵活的相关滤波框架和明显的速度优势,本文算法以其为基础进行扩展。

2.2 多尺度跟踪器

针对目标运动过程中出现的尺度变化问题,Danellian等^[12]提出了一种基于相关滤波的尺度自适应目标跟踪算法,即多尺度空间跟踪器(DSST),尺度预测结果比较准确。本文借鉴DSST多尺度跟踪方法,通过学习一个独立的一维相关滤波器去检测目标尺度的变化,称之为尺度相关滤波器。尺度估计方法与位置估计类似:首先在检测出的目标周围采集一系列不同尺度的图像块 \mathbf{x}_{s_i} ,提取 d 维HOG特征进行图像块表达,对应的高斯函数标签为 \mathbf{y}_{s_i} ;利用这些图像块训练RLS分类器,并在线性空间进行求解,获得一个一维的尺度相关滤波器;最后通过寻找这个滤波器的最大输出响应来完成尺度检测。

按照2.1节中位置相关滤波器计算的方法来训练RLS分类器,得到尺度相关滤波器模板 \mathbf{B}_{num} , \mathbf{B}_{den} 的计算和更新公式为

$$\mathbf{B}_{\text{num}i} = (1 - \eta_s) \mathbf{B}_{\text{num}i-1} + \eta_s \mathcal{F}(\mathbf{y}_{s_i}) \odot \mathcal{F}^*(\mathbf{x}_{s_i}), \quad (7)$$

$$\mathbf{B}_{\text{den}i} = (1 - \eta_s) \mathbf{B}_{\text{den}i-1} + \eta_s \sum_{k=1}^d \mathcal{F}(\mathbf{x}_{s_i}^k) \odot \mathcal{F}^*(\mathbf{x}_{s_i}^k). \quad (8)$$

尺度分类器的输出响应为

$$\hat{\mathbf{y}}_s = \mathcal{F}^{-1} \left[\frac{\sum_{k=1}^d \mathbf{B}_{\text{num}}^k \odot \mathcal{F}(\mathbf{z}_s)}{\mathbf{B}_{\text{den}} + \lambda} \right], \quad (9)$$

找到使 $\hat{\mathbf{y}}_s$ 最大的尺度即为新一帧中目标的尺度。

3 自适应特征选择

在相关滤波跟踪框架中,特征表达有效与否对于跟踪性能有着重要的影响。首先,受到相关滤波框架的限制,所提取的特征应属于基于图像块的全局特征,而非基于特征点描述的特征;其次,不同特征在不同跟踪场景中对目标的描述能力不同,若仅选择固定的一个或两个特征进行跟踪,当后续场景变换后,特征有效性可能发生改变,从而导致跟踪失败;最后,由于跟踪算法实时性的要求,特征提取算法本身不宜有较高的计算复杂度^[14]。针对以上问题,分别选取三种代表性视觉特征——形状、纹理、颜色特征,通过相关滤波响应图评估各特征的跟踪性能,实现跟踪过程中的自适应特征选择和自适应模型更新。

3.1 特征分析

3.1.1 特征的适用性分析

HOG特征是一种常用的形状和边缘描述算子,具有一定的平移、旋转不变性和光照不变性^[15],在处理形变、光照变化和局部遮挡类跟踪场景中表现良好。然而,当背景出现较多噪声或目标运动模糊时,目标的梯度特征并不明显,此时HOG特征的描述能力变弱。局部二值模式(LBP)是一种常用纹理特征,侧重于提取纹理信息,具有计算简单、旋转不变性和灰度不变性等优点,而LBP等价模式能够更有效地滤除高频噪声,很好地处理快速运动和运动模糊情况,在一定程度上弥补了HOG特征不适用的跟踪场景^[16]。CN特征利用文献^[17]中描述的映射方法将RGB空间转化为11维颜色特征空间,能够很好地表示目标的颜色信息,在彩色视频序列上得到了良好的跟踪性能。此外,归一化灰度特征(GRAY)作为一种最简单的一维特征,可以很方便地和任意其他特征进行简单串联,补充和提高特征表达能力。图1所示为单独使用HOG_GRAY、LBP_GRAY、CN_GRAY特征在部分视频序列上的跟踪结果,可以看出,三种特征的描述能力在不同场景下有明显的区别,单独使用某一种特征的跟踪算法在应对实际问题时稳健性较差,但可以寻求一种策略,最大化地发挥三种特征各自的优势。

3.1.2 特征的性能表现分析

针对上述三种特征,通过KRLS分别独立训练位置相关滤波器并计算候选图像块的响应图,可以发现,每一帧的相关滤波器响应最大值和跟踪结果之间存在一定的关联。以视频Basketball为例,图2(a)表示三

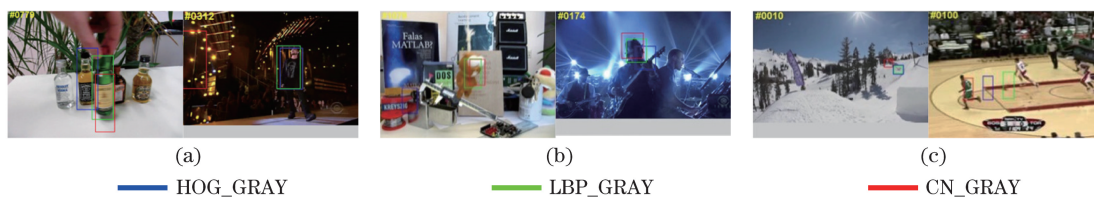


图 1 三组特征跟踪结果比较。(a)相似目标遮挡和光照变化;(b)运动模糊和背景复杂;(c)颜色明显和低分辨率
 Fig. 1 Tracking results comparison of three features. (a) Occlusion and illumination variation;
 (b) motion blur and background clutter; (c) vivid color and low resolution

种特征下位置相关滤波器在每一帧的响应最大值,图 2(b)为对应帧的跟踪结果,图 2(c)和图 2(d)画出了三种特征跟踪的距离精度曲线和成功率曲线。在其他条件相同的情况下,选择不同特征进行目标表达,相关滤波器的输出响应值大小表征目标跟踪结果的可靠性,响应值越大说明跟踪结果越精确。当某一帧的响应得分突然低于一定阈值时,有很大的可能性在该帧的目标已经跟丢;当目标跟丢之后,跟踪框在后续帧中的位置将会保持不变,得分可能会再次升高,但此时的高得分已经没有实际意义。

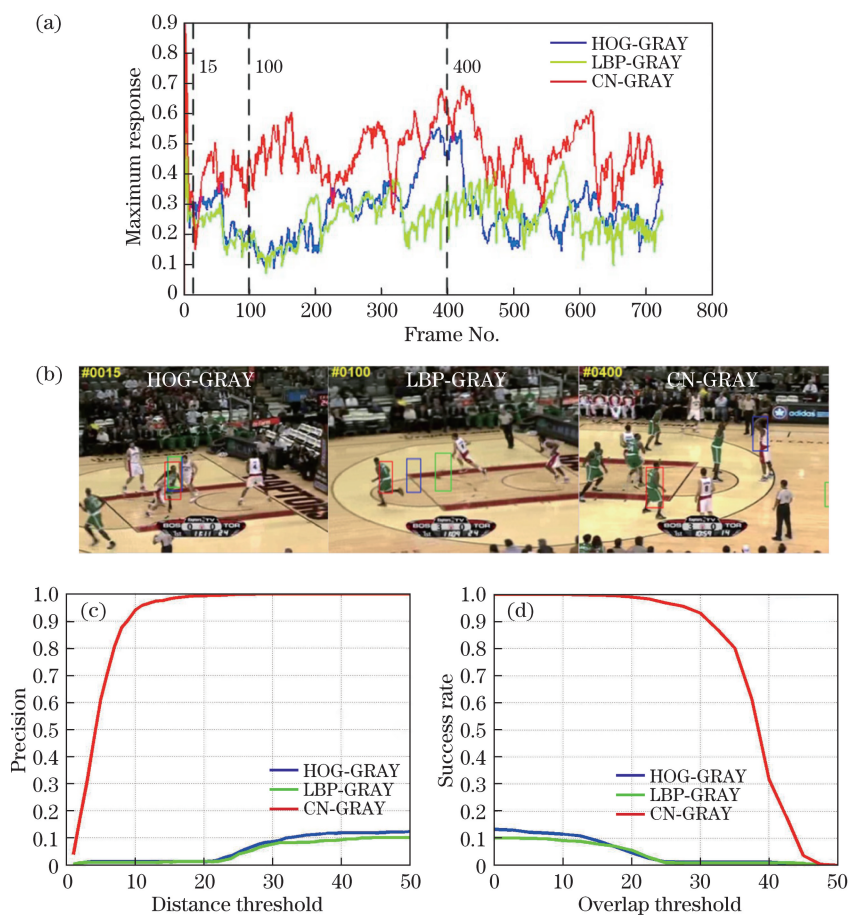


图 2 不同特征响应最大值与跟踪结果的关系。(a)响应最大值曲线;(b)相应帧跟踪结果;(c)跟踪精度曲线;(d)成功率曲线
 Fig. 2 Illustration of maximum response and tracking performance for different features. (a) Maximum response in each frame;
 (b) tracking results; (c) tracking precision plot; (d) success rate plot

每一帧的相关滤波响应分布和跟踪结果之间也存在密切的关联。仍以 Basketball 序列为例,采用 CN_GRAY 进行候选图像块特征表达,每一帧的响应图和该帧的跟踪结果如图 3 所示。当响应峰值位置集中且突出时,表明在该帧中所用特征对目标和背景的区别能力很强,跟踪框精度很高;反之,当滤波的响应峰值较小,响应图分布相对较分散,可能会在背景区域出现伪峰值,此时跟踪框由于干扰或遮挡出现偏差。

3.2 自适应特征选择方法

基于以上结果提出了一种自适应特征选择方法,将滤波响应图作为特征选择及滤波器模型更新的判断

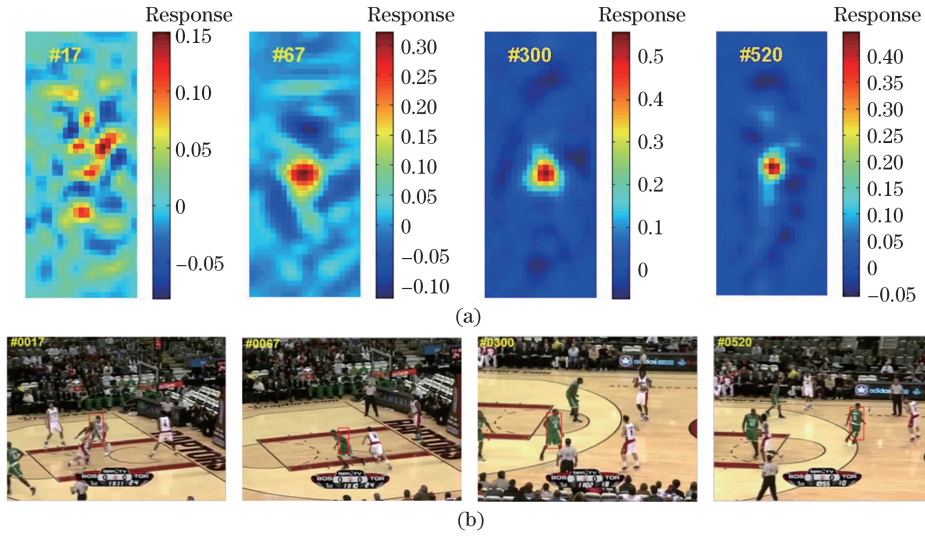


图 3 部分帧(a)相关滤波响应图与(b)跟踪结果的关系(Basketball 序列, CN_GRAY 特征)

Fig. 3 Illustration of (a) response map and (b) tracking performances (Basketball, CN_GRAY feature)

标准。分别用 HOG_GRAY、LBP_GRAY 和 CN_GRAY 特征对图像块进行特征提取,为三组特征设置各自的响应阈值,记为 T_{HOG} , T_{LBP} 和 T_{CN} 。利用第 2 节所述方法,训练三组独立的位置相关滤波器模型 $\{\hat{\mathbf{x}}_{\text{HOG}}, \hat{\boldsymbol{\alpha}}_{\text{HOG}}\}$, $\{\hat{\mathbf{x}}_{\text{LBP}}, \hat{\boldsymbol{\alpha}}_{\text{LBP}}\}$ 和 $\{\hat{\mathbf{x}}_{\text{CN}}, \hat{\boldsymbol{\alpha}}_{\text{CN}}\}$,并计算相应的滤波响应输出 $\hat{\mathbf{y}}_{\text{HOG}}$, $\hat{\mathbf{y}}_{\text{LBP}}$ 和 $\hat{\mathbf{y}}_{\text{CN}}$,计算每一帧的响应最大值,当最大值低于设定阈值时,判断为该特征跟踪将失效。此时重新选取特征,特征选取策略为

$$\hat{\mathbf{y}}_{\text{select}} = \max[\max(\hat{\mathbf{y}}_{\text{HOG}}), \max(\hat{\mathbf{y}}_{\text{LBP}}), \max(\hat{\mathbf{y}}_{\text{CN}})]. \quad (10)$$

根据(10)式选择的特征结果计算当前帧的目标位置,并相应更新相关滤波器模型,用于下一帧跟踪。

当特征固定后,在跟踪过程中还需要对相关滤波器模型进行实时更新。目前的模型更新通常采用(6)式的线性插值方法,利用固定权值进行逐帧更新,但这种方法存在一定的不足。当某一帧跟踪框出现偏差或遮挡时,(6)式会引入该帧的错误模型信息,使跟踪误差不断累积,最终导致跟踪失败。利用图 3 分析的响应图和跟踪结果关系,对模型更新方法进行改进。如图 4 所示,对于某一帧的候选图像块,红色矩形框内为目标区域,矩形框外为背景区域,分别计算目标区域和背景区域的响应峰值 $\max(\hat{\mathbf{y}}_t)$ 和 $\max(\hat{\mathbf{y}}_b)$,模型更新公式为

$$X_{\text{diff}} = \max(\hat{\mathbf{y}}_t) - \max(\hat{\mathbf{y}}_b), \quad (11)$$

$$\hat{\boldsymbol{\alpha}}_i = \begin{cases} \eta \boldsymbol{\alpha}_i + (1 - \eta) \hat{\boldsymbol{\alpha}}_{i-1}, & X_{\text{diff}} > T_{\text{diff}} \\ \hat{\boldsymbol{\alpha}}_{i-1}, & X_{\text{diff}} \leq T_{\text{diff}} \end{cases}, \quad (12)$$

$$\hat{\mathbf{x}}_i = \begin{cases} \eta \mathbf{x}_i + (1 - \eta) \hat{\mathbf{x}}_{i-1}, & X_{\text{diff}} > T_{\text{diff}} \\ \hat{\mathbf{x}}_{i-1}, & X_{\text{diff}} \leq T_{\text{diff}} \end{cases}, \quad (13)$$

式中 T_{diff} 表示本文设置的目标响应峰值和背景响应峰值间的差值阈值,在每一帧的跟踪过程中,只有当目标峰值和背景峰值的差大于该阈值时,才会更新滤波器模型 $\{\hat{\mathbf{x}}, \hat{\boldsymbol{\alpha}}\}$, 否则不进行更新。

综上,本文跟踪算法总体流程如图 5 所示。在位置检测模块,利用 HOG_GRAY 特征、LBP_GRAY 特征和 CN_GRAY 特征进行自适应最优特征选择,并相应地在线更新模型;在尺度检测模块,当目标的中心位置确定后,从目标周围选取一系列尺度的图像块作为样本提取 HOG 特征,采用线性核函数训练 RLS 分类器,获得尺度相关滤波器,通过寻找尺度滤波器输出响应最大值,完成对目标的尺度检测。

4 实验结果与分析

4.1 实验环境及设置

实验硬件环境为 Intel Core i7-4790 CPU,主频 3.60 GHz,内存 8 GB 配置的计算机,算法开发平台为

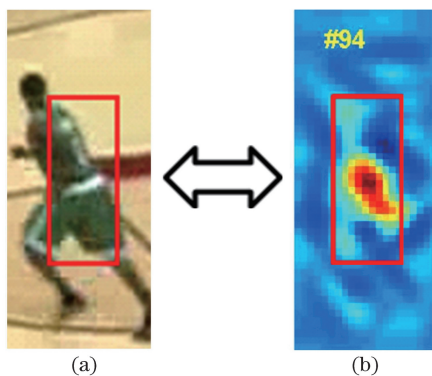


图 4 (a) 候选图像块和(b)对应目标背景区域响应图

Fig. 4 (a) Image block for tracking and (b) corresponding response map with target and background region

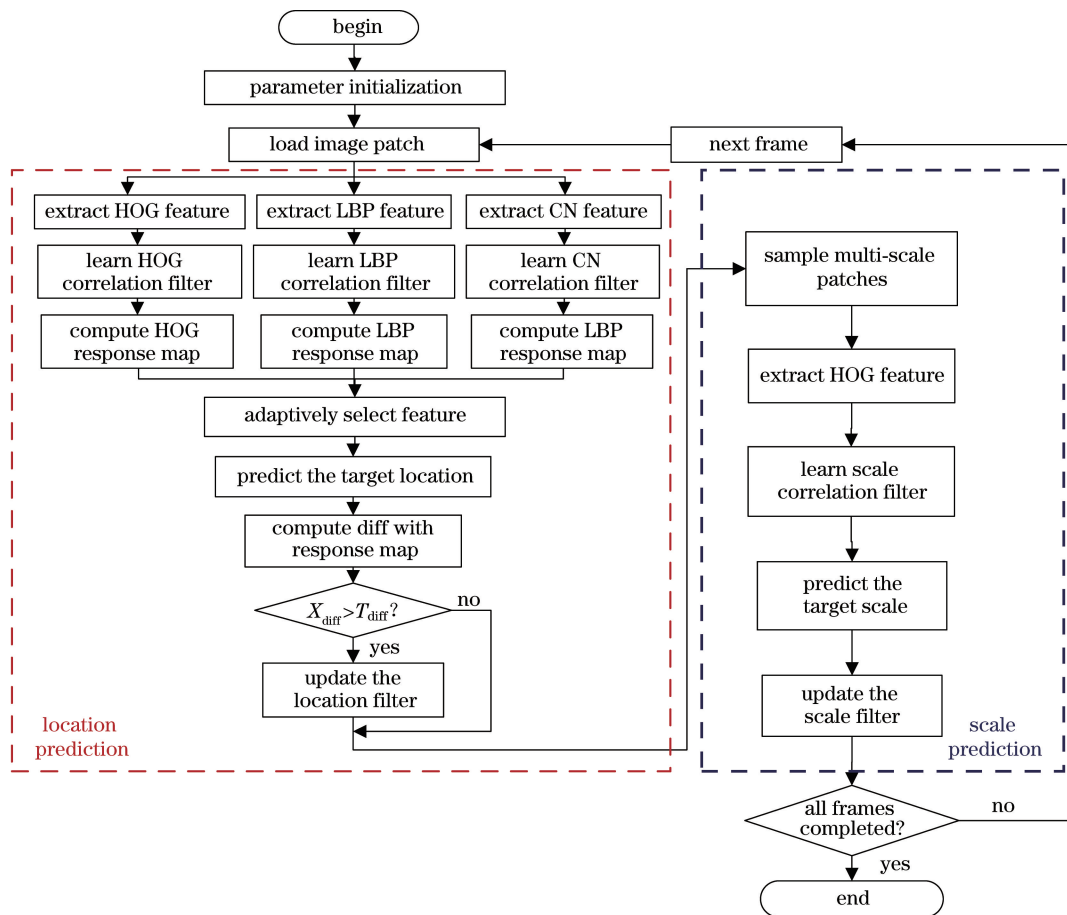


图 5 本文算法流程图

Fig. 5 Overall flowchart of the proposed method

Matlab R2014a。实验测试序列来源于标准跟踪数据集 OTB^[1]，视频序列中包含了跟踪中常见的挑战性场景，包括光照变化、运动模糊、快速运动、目标遮挡、尺度变化、目标形变、背景复杂、目标旋转、低分辨率等。遵循 OTB 的评估方法，使用一次通过评估 (OPE) 模式，计算平均距离精度曲线和成功率曲线，进行跟踪性能评估。距离精度曲线计算中心位置误差 (CLE) 低于某个阈值的帧数占所有帧的百分比，其中 CLE 指跟踪目标中心位置与目标中心真值之间的平均欧氏距离；成功率曲线计算跟踪重叠率大于某个阈值的帧数占所有帧的百分比，其中重叠率是指目标跟踪框与真实目标框的重叠面积与总面积之比。此外，以每秒平均帧数 (FPS) 来报告所有跟踪器的速度。

本文算法位置相关滤波器涉及的参数值与 KCF^[6] 相同,尺度相关滤波器涉及的参数值与 DSST^[12] 相同。除此之外,算法还设置了四个阈值参数,依次为 T_{HOG} 、 T_{LBP} 、 T_{CN} 和 T_{diff} 。首先将 OTB-2013 数据集(36 组彩色序列)作为训练集进行阈值参数遍历,实验发现,不同特征的响应峰值分布在不同的范围内,而 X_{diff} 值低于 0.02 时跟踪结果通常不可靠。在该数据集上训练出一组经验阈值,随后在扩充后的 OTB-2015 数据集(74 组彩色序列)上验证了这组阈值的适用性和有效性。最终实验选取参数值如下: $T_{HOG} = 0.25$, $T_{LBP} = 0.20$, $T_{CN} = 0.35$, $T_{diff} = 0.02$ 。

4.2 实验一:特征选择结果分析

图 6 以 Deer 序列为例,给出了本文算法的特征选择结果。图 6(a)中三条颜色不同的曲线表示不同特征的响应峰值,红色十字点表示本文算法的响应峰值;图 6(b)表示不同特征对应的重叠率结果和本文算法的重叠率结果。可以看到,本文算法以 HOG 作为初始选择特征,开始时得到了很高的重叠率;在第 22 帧中,HOG 的响应峰值低于 0.25,相应重叠率结果在接下来几帧中突然下降为 0,表明此时 HOG 特征跟踪失败;本文算法根据阈值判断及时进行了特征调整,在第 22 帧中选择 CN 特征进行后续跟踪,从始至终重叠率均保持在 0.6 以上,得到了可靠的跟踪结果。该实验表明,本文特征选择方法以及选取的经验阈值是合理有效的。

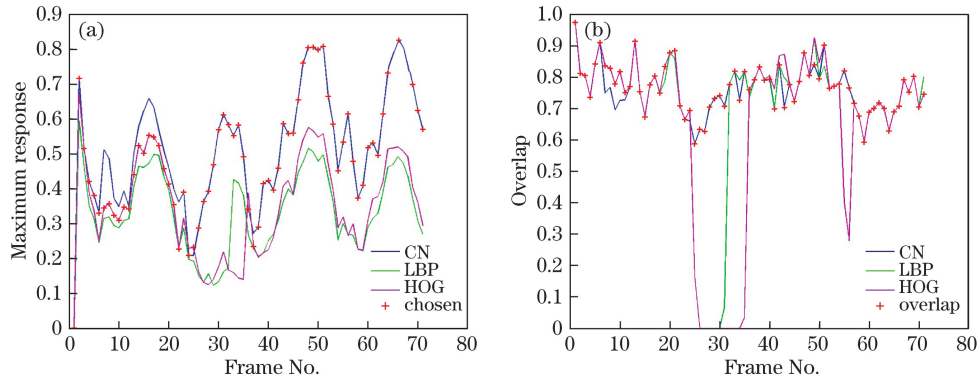


图 6 Deer 序列的特征选择结果分析。(a)响应峰值结果;(b)重叠率结果

Fig. 6 Analysis of feature selection results in Deer sequence. (a) Maximum response results; (b) overlap results

4.3 实验二:与单特征的相关滤波跟踪算法比较

为了评估不同特征对相关滤波跟踪算法性能的影响,验证本文提出的自适应特征选择方法(的有效性,在基于 KCF 和 DSST 结合的多尺度相关滤波跟踪框架下,分别采用 HOG_GRAY、LBP_GRAY、CN_GRAY 和本文方法进行了四组实验。为保证算法比较的公平性,在特征提取时在特征提取时均设为 4。表 1 列出了四种方法的平均 FPS、平均距离精度(DP)和平均重叠率精度(OP),相应的距离精度曲线和成功率曲线如图 7 所示。

表 1 不同特征的平均跟踪性能指标比较(74 组彩色序列)

Table 1 Average tracking performance comparison of different features on 74 color sequences

Algorithm	HOG_GRAY	LBP_GRAY	CN_GRAY	Proposed
Mean FPS	103.67	90.50	96.54	53.12
Mean DP (20 pixel)	0.645	0.583	0.642	0.702
Mean OP (0.5 overlap)	0.558	0.518	0.568	0.631

以上结果表明,和采用单特征的相关滤波跟踪算法相比,本文算法的确能够充分利用各特征的优势,在平均跟踪性能上有了很大的提升。本文算法的平均 DP 和平均 OP 分别为 0.702 和 0.631,比单特征中最优的 HOG_GRAY(0.645 和 0.558)提高了 5.7% 和 7.3%。

4.4 实验三:与 8 种流行跟踪算法比较

为了对比算法的性能,选取近年来出现的 8 种跟踪方法 CSK^[5]、KCF^[6]、DSST^[12]、SAMF^[9]、STAPLE^[10]、TLD^[18]、STRUCK^[19]、SCM^[20] 和本文方法进行比较,9 种方法的简介如表 2 所示。

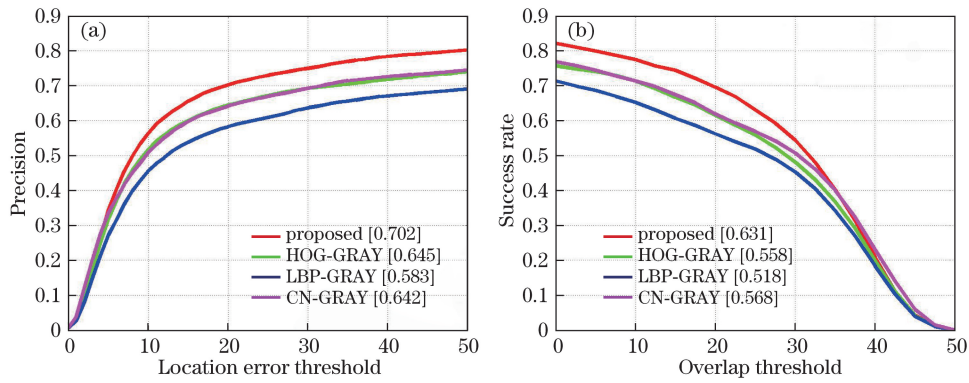


图 7 不同特征的(a)距离精度曲线和(b)成功率曲线(36组彩色序列)

Fig. 7 (a) Precision plots and (b) success rate plots of different features over the 36 color sequences

表 2 9种跟踪算法简介

Table 2 Introduction of nine trackers

Tracker name	Feature	Scale adaptive	Correlation filter-based
CSK	GRAY	No	Yes
KCF	HOG	No	Yes
DSST	HOG, GRAY	Yes	Yes
SAMF	HOG, CN, GRAY	Yes	Yes
STAPLE	HOG, CN	Yes	Yes
TLD	BINARY	Yes	No
STRUCK	HAAR	No	No
SCM	GRAY	Yes	No
Proposed	HOG, CN, LBP	Yes	Yes

表 3 列出了 9 种跟踪算法的平均速度、平均 DP 和平均 OP,其中排名前三的算法依次为 STAPLE、SAMF 和本文算法,相应的距离精度曲线和成功率曲线如图 8 所示。本文算法和原始 KCF 相比,平均 DP 和 OP 分别提高了 5.4%和 11.4%;和 DSST 相比,平均 DP 和 OP 分别提高了 8.3%和 8.4%。此外,本文算法跟踪精度远高于其他非相关滤波类方法,能够接近最前沿的多尺度相关滤波跟踪方法 STAPLE 和 SAMF。最后,本文算法继承了相关滤波本身的速度优势,平均速度为 53.12 frame/s,与 STAPLE、DSST 接近,远远快于其他多尺度跟踪方法,能够满足实时跟踪的需求。

表 3 9种跟踪算法的平均跟踪性能指标比较(74组彩色序列)

Table 3 Average tracking performance comparison of nine trackers on 74 color sequences

Algorithm	TLD	STRUCK	SCM	CSK	KCF	DSST	SAMF	STAPLE	Proposed
Mean FPS	28.36	18.21	0.53	346.91	339.25	50.34	14.58	50.49	53.12
Mean DP	0.534	0.568	0.497	0.461	0.648	0.619	0.731	0.747	0.702
Mean OP	0.454	0.467	0.440	0.361	0.517	0.547	0.655	0.665	0.631

图 9 为本文算法与 STAPLE、DSST、KCF、TLD、SCM 这 5 种算法在部分跟踪序列上的结果对比。图 9(a)目标存在快速运动和运动模糊,只有本文算法和 STAPLE 算法能够始终抓住有效特征,保持准确稳定的跟踪,其他方法均出现了漂移现象;图 9(b)背景复杂且干扰较多,本文算法一直准确跟踪到了目标,并能够根据目标的尺度变化及时调整跟踪窗口大小;在图 9(c)中,当电线杆对目标产生了遮挡之后,除了本文算法和 TLD 算法,其他算法均跟踪失败,说明本文算法具有很好的抗遮挡能力和稳健性;图 9(d)则说明了本文算法在尺度跟踪上的有效性。

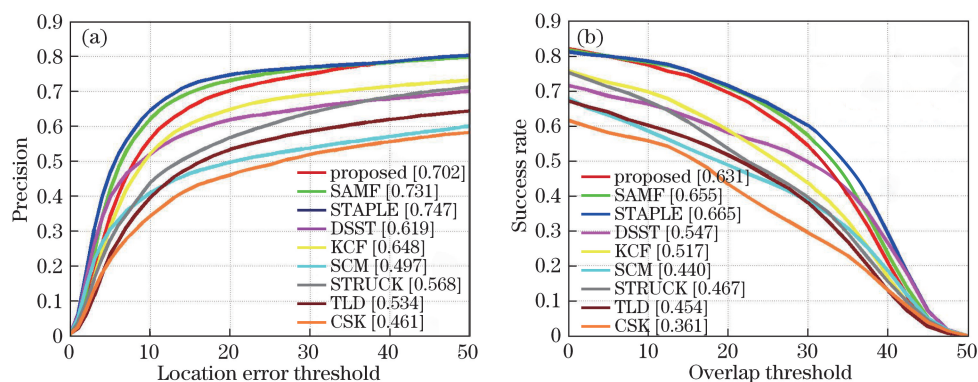


图 8 本文算法和其他 8 种跟踪算法的(a)距离精度曲线和(b)成功率曲线(36 组彩色序列)

Fig. 8 (a) Precision plots and (b) success rate plots of the proposed method and another eight trackers over the 36 color sequences

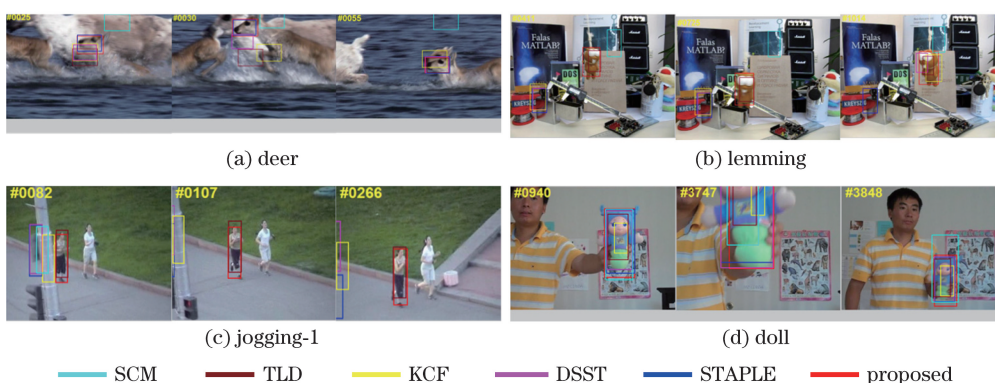


图 9 本文算法和其他 5 种跟踪算法在部分序列上的跟踪结果显示

Fig. 9 Tracking results of the proposed method and another five trackers

5 结 论

在相关滤波框架的基础上,提出了一种基于自适应特征选择的多尺度相关滤波跟踪方法。首先提取三组互补特征 HOG_GRAY、LBP_GRAY 和 CN_GRAY,通过比较各特征的输出响应最大值评估其跟踪性能,自适应地选择最优特征形成位置相关滤波跟踪器;其次,利用响应图作为位置相关滤波模型更新的评估判断条件,优化模型更新方式;此外,引入尺度相关滤波跟踪器,实现目标的尺度自适应跟踪。实验表明,本文算法提高了目标跟踪的稳健性和准确性,并达到了 53.12 frame/s 的跟踪速度。另外,目前本文算法的特征提取和自适应选择采用串行的方式实现,若进一步采用并行计算优化,算法的速度仍有很大提升空间。

参 考 文 献

- [1] Wu Y, Lim J, Yang M H. Object tracking benchmark[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834-1848.
- [2] Liu Wei, Zhao Wenjie, Li Cheng. Long-term visual tracking based on spatio-temporal context[J]. Acta Optica Sinica, 2016, 36(1): 0115001.
刘 威, 赵文杰, 李 成. 时空上下文学习长时目标跟踪[J]. 光学学报, 2016, 36(1): 0115001.
- [3] Chen Z, Hong Z B, Tao D C. An experimental survey on correlation filter-based tracking[Z/OL]. (2015-09-18) [2016-09-20]. <https://arxiv.org/abs/1509.05520>.
- [4] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2010: 2544-2550.
- [5] 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, 2012: 702-715.
- [6] 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.

- [7] Danelljan M, Shahbaz Khan F, Felsberg M, *et al.* Adaptive color attributes for real-time visual tracking[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1090-1097.
- [8] Qi Y, Zhang S, Qin L, *et al.* Hedged deep tracking[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [9] Li Y, Zhu J K. A scale adaptive kernel correlation filter tracker with feature integration[C]. European Conference on Computer Vision, 2014: 254-265.
- [10] Bertinetto L, Valmadre J, Golodetz S, *et al.* Staple: complementary learners for real-time tracking[C]. International Conference on Computer Vision and Pattern Recognition, 2016: 1401-1409.
- [11] He Yujie, Li Min, Zhang Jinli, *et al.* Infrared small target detection method based on correlation filter[J]. Acta Optica Sinica, 2016, 36(5): 0512001.
何玉杰, 李 敏, 张金利, 等. 基于相关滤波器的红外弱小目标检测算法[J]. 光学学报, 2016, 36(5): 0512001.
- [12] Danelljan M, Häger G, Khan F, *et al.* Accurate scale estimation for robust visual tracking[C]. British Machine Vision Conference, Nottingham, 2014.
- [13] Xu Yulong, Wang Jiabao, Li Yang, *et al.* Scale-adaptive tracking based on kernelized correlation filter[J]. Application Research of Computers, 2016, 11: 1-6.
徐玉龙, 王家宝, 李 阳, 等. 基于相关滤波的尺度自适应目标跟踪[J]. 计算机应用研究, 2016, 11: 1-6.
- [14] Guo Pengyu, Su Ang, Zhang Hongliang, *et al.* Online mixture of random Naïve Bayes tracker combined texture with shape feature[J]. Acta Optica Sinica, 2015, 35(3): 0315002.
郭鹏宇, 苏 昂, 张红良, 等. 结合纹理和形状特征的在线混合随机朴素贝叶斯视觉跟踪器[J]. 光学学报, 2015, 35(3): 0315002.
- [15] Felzenszwalb P F, Girshick R B, Mcallester D, *et al.* Object detection with discriminatively trained part-based models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627-1645.
- [16] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.
- [17] van de Weijer J, Schmid C, Verbeek J. Learning color names from real-world Images[C]. IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, Minnesota, USA, 2007.
- [18] Kalal Z, Matas J, Mikolajczyk K. Pn learning: bootstrapping binary classifiers by structural constraints[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2010: 49-56.
- [19] Hare S, Golodetz S, Saffari A, *et al.* Struck: structured output tracking with kernels[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(10): 2096-2109.
- [20] Zhong W, Lu H, Yang M H. Robust object tracking via sparsity-based collaborative model[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2012: 1838-1845.