

特征融合的尺度自适应相关滤波跟踪算法

李聪, 鹿存跃, 赵珣, 章宝民, 王红雨

上海交通大学电子信息与电气工程学院, 上海 200240

摘要 为提高相关滤波(CF)跟踪算法的稳健性,并克服传统 CF 方法无法处理目标尺度变化以及未利用图像颜色特征等问题,提出了一种基于融合颜色特征的尺度自适应相关滤波改进跟踪算法。首先,将目标搜索区域从 3 原色(RGB)颜色空间转换到 Lab 颜色空间,提取搜索区域的 Lab 3 通道颜色特征;然后,融合 Lab 颜色特征与方向梯度直方图(HOG)特征得到多通道特征,利用核相关滤波(KCF)计算输出响应图并寻找图中最大响应位置即目标位置;最后,基于 Lab 颜色特征建立尺度模型,从当前帧的目标位置处截取不同尺度图像块,通过将其与尺度模型比较得到目标尺度最优估计。实验选取 35 段公开彩色视频序列进行测试,并将所提算法与其他 5 种跟踪性能较好的跟踪方法进行对比。实验结果表明,所提方法对彩色视频序列中的目标遮挡、变形、尺度变化等现象具有良好的适应性,其平均性能优于对比方法,同时具有 $76 \text{ frame}\cdot\text{s}^{-1}$ 的实时跟踪速度。

关键词 机器视觉; 目标跟踪; 相关滤波; 特征融合; 尺度估计

中图分类号 TP391.41

文献标识码 A

doi: 10.3788/AOS201838.0515001

Scale Adaptive Correlation Filtering Tracing Algorithm Based on Feature Fusion

Li Cong, Lu Cunyue, Zhao Xun, Zhang Baomin, Wang Hongyu

School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract In order to improve the robustness of correlation filtering (CF) tracking algorithm, and overcome the problems that the traditional CF method cannot handle target scale change and does not use image color feature, a scale adaptive tracking algorithm is proposed based on correlation filtering improvement with fused color features. Firstly, the target searching area of the image is transferred from the color space of the three primary colors (RGB) to the Lab color space to obtain the Lab three channel features of the search area. Then, Lab color features and histogram of oriented gradients (HOG) feature are fused to obtain the image feature of multi-channel. The kernelized correlation filtering (KCF) is used to get the output response chart and find the position of maximum response, namely target location. Finally, the scale model is established through the Lab color feature, and the different scale image blocks are intercepted from the current frame target position. Optimal estimation of the target scale is obtained when we compare the scale image blocks with scale models. 35 pieces of open color video sequences are selected in experiments for testing, and the proposed method is compared with five other tracking methods with excellent performance. Experimental results show that the proposed method is well adapted to the phenomena of target occlusions, deformation and scale change in color video sequences, and its average performance outperforms the other compared methods. At the same time, the real-time tracking speed of the proposed method is $76 \text{ frame}\cdot\text{s}^{-1}$.

Key words machine vision; object tracking; correlation filtering; feature fusion; scale estimation

OCIS codes 150.0155; 150.1135; 100.4999; 110.1085

1 引 言

目标跟踪一般指在给定目标初始位置状态的条

件下,在视频或图像序列中检测目标位置状态的过程。目标跟踪作为机器视觉领域的重要分支,近年来发展迅猛,在众多领域都有着广泛的应用,如在国

收稿日期: 2017-11-06; 收到修改稿日期: 2017-11-28

基金项目: 国家自然科学基金(11174206,61471237)、爱生创新发展基金(ASN-IF2015-1302)

作者简介: 李聪(1992—),男,硕士研究生,主要从事机器视觉与机器学习方面的研究。E-mail: leecc@sjtu.edu.cn

导师简介: 鹿存跃(1971—),男,博士,副教授,主要从事精密驱动以及超声电机和压电制动器方面的研究。

E-mail: lucunyue@sjtu.edu.cn(通信联系人)

防领域中应用于导弹精确制导定位、海空域可疑目标监控等,在民用领域中应用于智能机器人、家宅视频监控等。通常目标在视频或图像序列中存在着外观变形、尺度变化、光照变化、快速旋转运动、背景相似干扰和视野丢失等现象^[1-2],这些现象的存在给目标跟踪带来了很大的困难,甚至会导致跟踪失败,是目前目标跟踪领域亟待解决的问题。研究人员针对各种现象与问题提出的跟踪方法层出不穷,但目前尚未出现通用解决方法,总体而言这些方法可以分为两大类:生成类方法^[3-4]和判别类方法^[5-6]。生成类方法一般对前一视频帧的目标区域进行建模,在当前帧查找与模型最相似的区域即目标位置;判别类方法一般从前一视频帧的目标区域与背景区域中提取正负训练样本,用机器学习的方法训练一个目标分类器,在当前帧用训练好的分类器检测目标位置。

基于相关滤波(CF)的跟踪方法属于判别类方法^[7-11],CF方法以其良好的跟踪性能和极高的计算效率,成为近年来目标跟踪领域除深度学习方法外最大的研究热点。CF方法一般在跟踪过程中利用从前一帧采集的训练数据集训练分类器,利用分类器去检测目标在当前帧的位置,之后更新训练数据集为当前帧采集的数据集,进而更新分类器。Bolme等^[7]将相关滤波用于目标跟踪,提出了误差最小平方和滤波(MOSSE)算法,利用图像灰度特征对分类器进行训练和目标检测;Henriques等^[8-9]引入核函数改进MOSSE算法,提出了循环结构核(CSK)算法和基于CSK算法改进的核相关滤波(KCF)算法,其中KCF算法使用多通道方向梯度直方图(HOG)特征替代灰度特征,跟踪精度得到显著提高;Danelljan等^[10]基于英语语境中常用的颜色名称将颜色分为11类,利用多通道颜色名(CN)特征去扩展CSK算法,提取了CN跟踪算法,也取得了不错的效果;Danelljan等^[11]通过引入目标位置与尺度滤波器实现目标尺度自适应变化,在目标尺度变化跟踪问题上取得了一定的成果。上述研究表明图像特征的选择对跟踪算法的性能影响巨大,然而这些方法仅利用图像的单一特征,利用单一特征解决跟踪问题时,当目标外观表现变化时容易积累模型误差,进而影响跟踪的稳健性,并且这些方法在跟踪过程中没有有效解决目标尺度变化的问题。

针对上述问题,本文在KCF算法框架的基础上,提出了一种基于融合颜色特征的尺度自适应相关滤波跟踪方法。首先在提取到图像HOG特征的

基础上融合Lab颜色空间下的颜色特征,然后训练相关滤波分类器并用训练完成的分类器对每一帧的目标位置进行预测,最后利用Lab颜色特征构建目标尺度模型,实时更新目标尺度,进而实现目标尺度自适应跟踪。实验选取文献[1]中公开的彩色视频序列数据集对算法进行实验验证,并将本文算法与近年来出现的其他主流跟踪方法进行对比。结果表明,本文算法对彩色视频序列的平均跟踪精度优于其他方法,并能有效自适应目标尺度变化,其平均跟踪速度达到了76 frame·s⁻¹。

2 特征融合的尺度自适应跟踪

2.1 KCF跟踪算法

KCF跟踪算法首先在前一帧图像中以跟踪所得目标中心位置及其扩展一定范围大小的窗口经循环移位采集训练图像^[9],若采集得到的训练图像块为 x ,进行特征提取后对应的标签(回归)函数为 y ,设线性分类器模型 $f(x) = \langle w, \varphi(x) \rangle$,在正则化最小二乘并引入核函数的条件下,有

$$w = \arg \min_w \sum |\langle w, \varphi(x) \rangle - y|^2 + \lambda \|w\|^2, \quad (1)$$

式中: w 为分类器系数; $\langle \cdot, \cdot \rangle$ 表示内积; $\varphi(x)$ 为内核 κ 从原始空间到Hilbert特征空间的映射函数,定义核空间内积 $\kappa(x, x') = \langle \varphi(x), \varphi(x') \rangle$; $\lambda \geq 0$ 为正则化参数。

利用循环矩阵的性质以及离散傅里叶变换,求得最优解 $w = \sum \alpha_x \varphi(x)$,其中系数 α_x 可表示为

$$\alpha_x = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}(y)}{\mathcal{F}[\kappa(x, x)] + \lambda} \right\}, \quad (2)$$

式中 \mathcal{F} 和 \mathcal{F}^{-1} 表示离散傅里叶变换和离散傅里叶逆变换。

KCF跟踪所采用的核为高斯核,高斯核函数为

$$\kappa(x, x') = \exp\left(-\frac{|x - x'|}{\sigma^2}\right), \quad (3)$$

式中: $\kappa(x, x') = k^{xx'}$; σ 为高斯核函数带宽参数。对于所有的循环位移块, $k^{xx'}$ 的计算为

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

式中 \odot 表示元素点积。

KCF算法的跟踪过程就是根据上一帧分类器训练的结果,在当前帧以相同大小窗口截取图像块 z ,计算这些图像块在分类器中的输出响应值,响应值最大的位置即为被跟踪目标在当前帧的位置。响

应值的计算表示为

$$\hat{y} = \mathcal{F}^{-1}[\mathcal{F}(k^{\hat{x}}) \odot \mathcal{F}(\hat{a})], \quad (5)$$

式中： \hat{y} 为输出响应值； \hat{x} 和 \hat{a} 分别表示学习得到的外观模板和目标外观模型参数。

2.2 融合 Lab 颜色特征与 HOG 特征

对于提取图像多通道特征作为输入的情况， $k^{xx'}$ 的计算公式可以改写为

$$k^{xx'} = \exp\left\{-\frac{1}{\sigma^2}[\|x\|^2 + \|x'\|^2 - 2\mathcal{F}^{-1}(F)]\right\}, \quad (6)$$

式中： $F = \mathcal{F}(x) \odot \mathcal{F}^*(x')$ ； $x = [x_1, x_2, \dots, x_i]$ 为图像的 i 通道特征。多通道特征条件下，算法的跟踪过程同前述过程一样。

根据(6)式，在 HOG 特征的基础上引入了 Lab 颜色特征。同最常见的 RGB 颜色空间不同，Lab 颜色空间既不依赖光线，也不依赖颜料，它是国际照明委员(CIE)组织确定的一个颜色空间，理论上能描述自然界的所有色彩。Lab 颜色空间 3 通道分量中 L 代表亮度，值域为 $[0, 100]$ ，亮度随数值递增； a 和 b 通道分量的值域都为 $[-128, 127]$ ，当接近 0 时，它们的颜色变化较为显著，靠近值域上下限的颜色变化差异较小，其中 a 的正端代表红色，负端代表绿色， b 的正端代表黄色，负端代表蓝色。RGB 颜色空间(R 代表红色分量， G 代表绿色分量， B 代表蓝色分量)无法直接线性变换到 Lab 颜色空间，需要以 XYZ 颜色空间为中介进行变换。RGB 空间到 XYZ 空间的转换关系为^[12]

$$\begin{cases} X = 0.490 \times R + 0.310 \times G + 0.200 \times B \\ Y = 0.177 \times R + 0.812 \times G + 0.011 \times B \\ Z = 0.010 \times G + 0.990 \times B \end{cases} \quad (7)$$

XYZ 空间到 Lab 空间的转换关系^[12]

$$\begin{aligned} L &= 116f(Y) - 16, a = 500\left[f\left(\frac{X}{0.982}\right) - f(Y)\right], \\ b &= 200\left[f(Y) - f\left(\frac{Z}{1.183}\right)\right], \end{aligned} \quad (8)$$

式中： $X \leq 0.008856$ 时， $f(X) = 7.787X + 0.138$ ； $X > 0.008856$ 时， $f(X) = X^{\frac{1}{3}}$ 。

HOG 特征是机器视觉领域最常用的图像特征之一，它将图像划分为较小的局部细胞元，通过在每一个细胞元上提取梯度信息，绘制梯度方向直方图。图像的几何和光照的变化只会出现在较大的空间领域，因此 HOG 对这两种形变都能保持良好的不变性，适合应用于图像检测领域。根据文献[13]，采用 31 通道 HOG 特征描述目标。

该算法在基于 HOG 特征的基础上引入 Lab 空间下图像的颜色特征，将其与 HOG 特征融合后得到图像的多通道特征，将该多通道特征作为输入，目标外观在多特征融合下得到更好的描述，经过训练的分类器对目标的检测准确性也得到了提高。如图 1 所示，以 Lemming 视频序列中的第 80 帧(F80)图像为例，其中虚线框框定的范围为被跟踪目标，多个实线框框定的范围是扩展后的训练数据采集区域，将这些用于训练的图像块提取出的归一化 Lab 3 通道颜色特征和 31 通道 HOG 特征进行融合，得到 34 通道特征，并对分类器进行训练。分类器训练好后，同前述过程一样，依然利用循环位移得到搜索图像块，在多通道特征条件下利用核相关滤波算法计算第 81 帧的输出响应图，寻找图中的最大响应值，对应图像块的中心位置即为目标在当前帧的位置。之后的每一帧跟踪过程中心位置(P81)都经过相同的训练与检测方法完成。

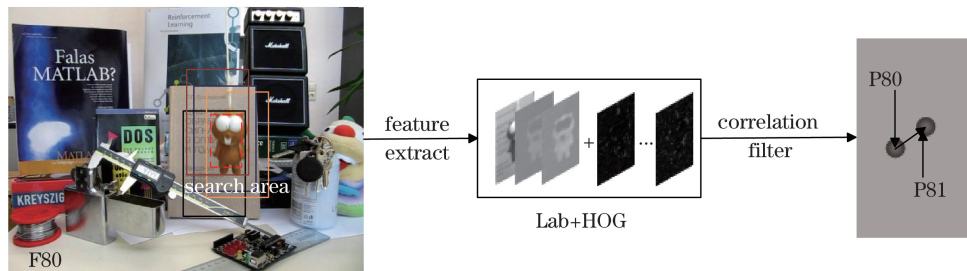


图 1 第 t 帧跟踪过程示意图

Fig. 1 Tracking process diagram of frame t

2.3 目标尺度估计与更新

目标尺度的变化容易导致跟踪漂移现象的产生，KCF 算法依赖循环矩阵，每次产生的训练样本以及之后跟踪过程截取的图像块大小都固定不变，

因此它对于多尺度变化的目标跟踪效果并不理想。如图 2 所示，其中图 2(a)为 Carscale 视频序列，序列从第 80 帧(F080)到 200 帧(F200)出现了目标尺度逐渐变大的现象，若没有自适应尺度变化，基于固

定大小窗口提取的训练图像中包含了目标大量的局部特征,目标外观没有得到有效描述;图 2(b)为 David 视频序列,序列从第 300 帧(F300)到 400 帧(F400)出现了目标尺度逐渐变小的现象,若没有自适应尺度变化,基于固定大小窗口提取的训练图像中则包含了过多的背景特征。以上两种情况都会导致目标跟踪精度的降低甚至导致跟踪失败,因此跟踪过程中自适应目标的尺度变化有助于提升算法的跟踪精度。

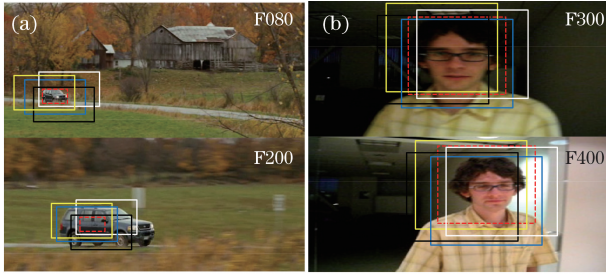


图 2 目标尺度变化。(a) Carscale; (b) David
Fig. 2 Scale change of object. (a) Carscale; (b) David

所提算法通过引入尺度金字塔模型^[11,14]并结合 Lab 颜色直方图特征来估计目标尺度。首先,由前一帧的跟踪结果获取目标的尺度模板:Lab 颜色直方图特征模板 L_{t-1} 和尺寸模板 S_{t-1} 。对于 Lab 颜色空间, a 和 b 通道的数值在接近 0 时即使出现微小的变化也会导致颜色的巨大差异,故通过非均匀量化的方式将 a 和 b 通道的各个数值量化区分为 22 个数值区间,其中正负两端各 11 个区间: $[-127, -70], (-70, -60], (-60, -50], (-50, -40], (-40, -30], (-30, -25], (-25, -20], (-20, -15], (-15, -10], (-10, -5], (-5, 0]$, 正端数值划分方式与负端相同。统计 a 和 b 通道分量量化后的每个区间的数值作为图像的 Lab 直方图特征 L 。以当前帧的跟踪结果得到的目标位置为中心,截取 22 个不同尺度的图像 S_i ($S_i = i \times S_{\min}, S_{\min} = S_{t-1}/11, i \in [1, 22]$), 经尺度变换为 S_{t-1} , 提取 S_i 的 Lab 直方图特征 L_i 。对 L_i 与 L_{t-1} 进行相似度计算, 得到 22 个特征相似度值 $f(i) = f(L_i, L_{t-1})$, 其中最大相似度值为 f_{\max} , 以该值对应的图像尺寸大小作为当前帧的最优尺寸, 并估计 S_t , 对应的直方图特征为 L_t 。下一帧的更新尺度模板为 L_t 与 S_t , 并重复上述过程, 即为尺度自适应过程。

如图 3 所示, 图 3(a) 中实线为尺寸模板 S_{t-1} 大小, 实线内部截取尺度小于 S_{t-1} 大小的图像, 同理实线外部截取大于 S_{t-1} 大小的图像; 采样得到的 22 幅图像经变换与 S_{t-1} 大小相同, 如图 3(b) 所示。

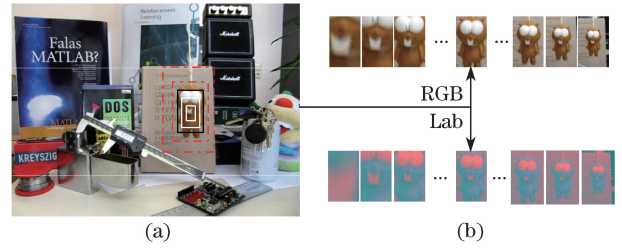


图 3 多尺度特征金字塔。(a) 目标图像; (b) 多尺度图像
Fig. 3 Multi-scale feature pyramid. (a) Object image; (b) multi-scale image

当获得的最优尺寸 S_t 与模板尺寸 S_{t-1} 之间的 Lab 直方图特征相似度值小于阈值 λ_s 时, 采用线性加权法来更新第 t 帧的目标尺度模板:

$$\begin{cases} L_t = \epsilon L_t + (1 - \epsilon) L_{t-1} \\ S_t = (1 - \epsilon) S_t + \epsilon S_{t-1} \end{cases} \quad (9)$$

式中 ϵ 为尺度模型更新学习率, 反之不更新。

3 实验结果与分析

3.1 实验环境与参数设置

实验所用开发平台为 MATLAB R2015b, 所有实验均在硬件配置为 Intel i7-4790 CPU(3.6 GHz)、8 GB 内存的计算机上完成。实验中其他参数均使用 KCF 原作者提供的代码默认参数^[9], 设置尺度更新阈值 $\lambda_s = 0.8$, 尺度模型更新学习率 $\epsilon = 0.2$ 。因该方法涉及到图像颜色特征, 因此仅选取文献[1]中公开的 35 段彩色视频序列(剔除其余 15 段灰度视频序列), 视频序列中囊括了背景遮挡、尺度变化、光照变化、目标旋转变形和运动模糊等各种现象。

选用最常用的 3 种指标评价跟踪算法性能: 中心位置误差 (CLE)、距离精度 (DP) 和重叠精度 (OP)。其中, CLE 评价指标指目标中心 (x_T, y_T) 和目标真实中心 (x_G, y_G) 之间的平均欧氏距离值, 即 $\sqrt{(x_T - x_G)^2 + (y_T - y_G)^2}$; DP 评价指标指 CLE 小于某一个阈值的百分比, 选择阈值为 20 pixel; 对于 OP 评价指标, 设 R_T 为第 t 帧跟踪结果, R_G 为第 t 帧标注框区域, 定义每一帧的跟踪得分为 s_t , 则

$$s_t = \frac{\text{area}(R_T \cap R_G)}{\text{area}(R_T \cup R_G)} \quad (10)$$

式中 \cap 表示取二者的重叠区域; \cup 表示取二者的覆盖总区域。根据 PASCAL 评价指标^[15], 选择重叠率阈值为 0.5, OP 指标的数值即跟踪得分 s_t 大于 0.5 的帧数与序列总帧数的百分比值。

为验证该算法的性能, 进行以下 3 种实验: 1) 进行不同颜色空间下颜色特征对算法跟踪性能的影响

响实验,对比研究不同颜色特征的选择对跟踪性能的影响;2)进行该算法与其他4种流行算法 DSST (Discriminative Scale Space Tracker)^[11]、RPT (Reliable Patch Trackers)^[16]、SAMF (Scale Adaptive Kernel Correlation Filter Tracker with Feature Integration)^[17]、TGPR (Transfer Learning Based Visual Tracking with Gaussian Processes Regression)^[18]以及基准算法 KCF^[9]的整体性能对比实验,这4种流行方法均在相同配置的计算机上调用作者开源的代码进行实验,最后通过评价标准评价各个算法的优劣性;3)进行该算法与整体性能次好算法的定性、定量对比实验。

3.2 不同颜色特征性能对比

为研究不同颜色空间下的颜色特征对算法跟踪性能的影响,分别在5种不同颜色空间(Lab、RGB、YCbCr、HSV和rg颜色空间)下进行对比实验。分

表1 不同颜色空间特征下所提算法与KCF算法的性能对比

Table 1 Comparison of performance of proposed algorithm and KCF algorithm under different color space features

Evaluation index	Lab	RGB	YCbCr	HSV	rg	KCF
Average CLE /pixel	19.4	39.3	46.3	24.6	38.5	42.7
Average DP /%	81.2	73.1	72.6	76.4	74.6	71.8
Average OP /%	74.3	68.3	65.6	71.3	64.5	60.5

表2 所提算法与其他算法的跟踪性能对比

Table 2 Comparison of performance of proposed algorithm and other algorithms

Evaluation	Proposed	SAMF	RPT	TGPR	DSST	KCF
Average CLE /pixel	19.4	34.4	48.5	56.3	50.3	42.7
Average DP /%	81.2	76.8	75.1	70.2	69.8	71.8
Average OP /%	74.3	70.4	64.6	63.4	61.8	60.8
Average speed /($\text{frame}\cdot\text{s}^{-1}$)	76.0	18.7	5.3	0.7	41.7	258.0

从表2数据可以看出,对于其他5种跟踪算法,在相同实验条件下所提出的方法在平均CLE、DP和OP3个指标上均取得了最好的结果,排名第二的SAMF算法与所提算法在这3个指标上差距较小,但其跟踪速度仅为 $18.7 \text{ frame}\cdot\text{s}^{-1}$,无法满足实时性需求。跟踪速度作为基准方法KCF的最大优势,高达 $258 \text{ frame}\cdot\text{s}^{-1}$ 的速度远远超过其他比较方法,但受限于其跟踪精度,综合性能仍显不足。所提算法在引入特征融合与尺度自适应处理的情况下取得了 $76 \text{ frame}\cdot\text{s}^{-1}$ 的平均跟踪速度,可以满足实时性的需求。

图4展示了几种方法的DP曲线与成功率曲线。其中,图4(a)为DP曲线,显示了不同CLE阈值下的DP评价指标数值,图中标注的数值代表阈

别提取每种颜色空间下的图像颜色特征,将其与HOG特征融合后进行实验,基准方法KCF仅利用图像的灰度特征进行实验。各种颜色特征下的算法跟踪性能如表1所示。

从表1数据可以看出,采用Lab颜色特征时相对其他颜色特征取得了最好的跟踪性能,采用HSV颜色特征时的跟踪性次之。相对于基准算法KCF,所提算法在Lab颜色空间下,CLE指标降低了54.6%,DP指标提升了13.1%,OP指标提升了22.8%,算法性能得到了明显的提升。

3.3 不同算法性能对比

表2记录了所提算法与其他5种算法在35段彩色视频数据集上的平均跟踪结果,其中包括CLE、DP、OP评价指标和平均跟踪速度指标(每秒处理的视频序列帧数,单位为 $\text{frame}\cdot\text{s}^{-1}$)。表格从左到右依据OP指标数值降序排列。

值为20 pixel时的DP评价指标数值;图4(b)为成功率曲线,图中标注的数值代表每种方法成功率曲线与坐标轴围成的区域面积(AUC),AUC是机器学习算法的重要评价指标,其值越大表示分类器的分类效果越好,反映了算法的跟踪精度。从图4的两种曲线可以看出,所提算法对比其他算法均取得了最好的跟踪性能。

3.4 不同算法定性、定量对比

图5为实验的部分跟踪结果,其中不同颜色实线框表示SAMF算法、基准算法KCF以及所提算法的跟踪结果,选取每段视频序列中两帧的结果进行展示。这6段视频序列分别为basketball[图5(a)]、carscale[图5(b)]、David[图5(c)]、girl[图5(d)]、shaking[图5(e)]和singer2[图5(f)],它们包含了

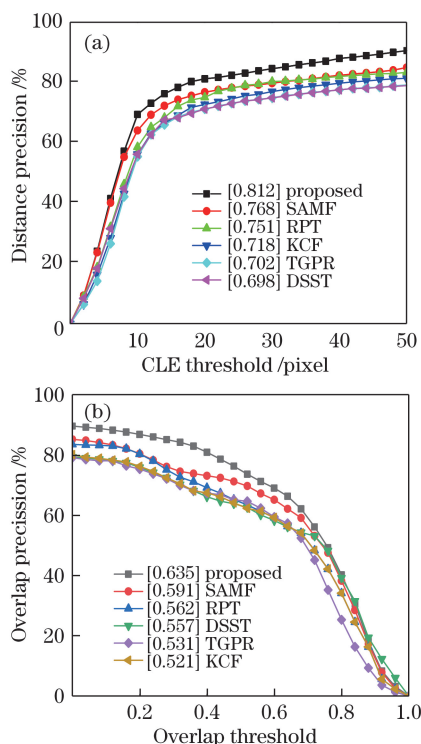


图 4 (a) DP 曲线; (b) 成功率曲线

Fig. 4 (a) DP plot; (b) success rate plot

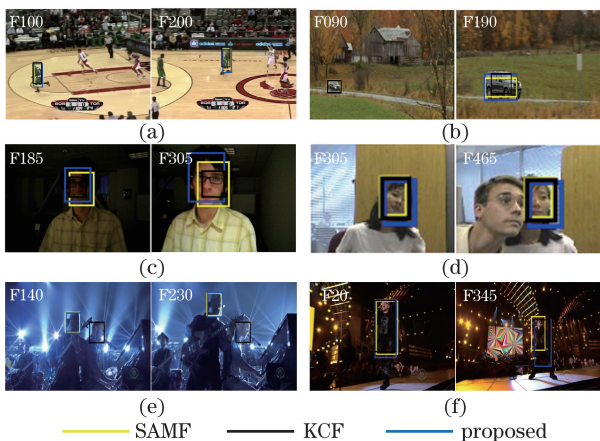


图 5 不同方法的跟踪结果。(a) Basketball; (b) carscale; (c) David; (d) girl; (e) shaking; (f) singer2

Fig. 5 Tracking results of different approaches.

(a) Basketball; (b) carscale; (c) David; (d) girl; (e) shaking; (f) singer2

常见的跟踪问题,如目标遮挡[图 5(a)、(b)、(d)],尺度变化[图 5(a)、(c)、(b)、(d)、(e)],光照变化[图 5(c)、(f)],旋转变换[图 5(a)、(b)、(c)、(d)、(e)],目标变形[图 5(a)、(c)、(f)],运动模糊[图 5(a)、(c)]等。从图 5(b)、(c)可以明显看出,对于目标尺度出现变化的跟踪问题,KCF 算法不能完整地标注目标,而 SAMF 算法表现稍好,但对比所提算法仍显不足,所提算法能够准确完整地标注出目标位置。

图 5(e)的跟踪结果中,KCF 算法出现了跟踪漂移现象,通过查看完整视频序列,发现跟踪漂移现象出现的原因是目标快速运动、旋转变换,KCF 跟踪器误将目标周围的物体检测成目标,导致之后的跟踪一直处于跟踪错误目标的状态,而 SAMF 算法和所提算法则比较准确地跟踪上了目标。3 种算法在多数序列中都能有效地跟踪目标,具有一定的稳健性,所提算法表现最好,KCF 算法表现最差,并且通过前面分析可以知道,所提算法能有效地解决尺度变化的跟踪问题。

图 6 为 3 种算法在 6 段视频序列下的 DP 曲线,每种算法在阈值为 20 pixel 时的 DP 数值如表 3 所示。可以看出,所提算法的 DP 表现最优。从图 6(e)中可以明显看出,KCF 算法出现跟踪漂移现象,故其 DP 极差。

表 3 各个算法的 DP

Table 3 DP of different algorithms

Sequence	Proposed	SAMF	KCF
Basketball	1.00	0.97	0.92
Carscale	0.98	0.85	0.80
David	0.93	0.85	0.84
Girl	0.95	0.94	0.86
Shaking	0.94	0.91	0.03
Singer2	1.00	1.00	0.94

所提算法虽然对大部分视频序列都有较好的跟踪效果,但依然存在跟踪漂移的现象,如图 7 所示。

图 7(a)、(b)对应的视频序列分别为 Lemming 和 MotorRolling,选取其中成功与失败后的各一帧结果。可以看出,对于这两段视频序列,所提方法和 KCF 方法均发生跟踪漂移现象,跟踪失败。对于 Lemming 视频序列,两种方法对视频前 300 帧的跟踪表现良好,跟踪到 300~450 帧时,目标物大范围隐藏在背景干扰物后面,即形成了长时间大范围的遮挡现象,而这段时间分类器提取的训练样本没有包含正样本,经过错误的训练后,分类器无法对下一帧的目标进行正确的检测,于是跟踪失败。对于 MotorRolling 视频序列,目标快速运动产生了运动模糊现象,并且目标经过了大于 360°的旋转,目标两帧位置差异过大的时候,分类器提取的训练样本仅一部分包含正确的目标,同样导致跟踪失败。对于这种现象,可以增大搜索框搜索范围,但随之有可能带来跟踪速度下降的问题。

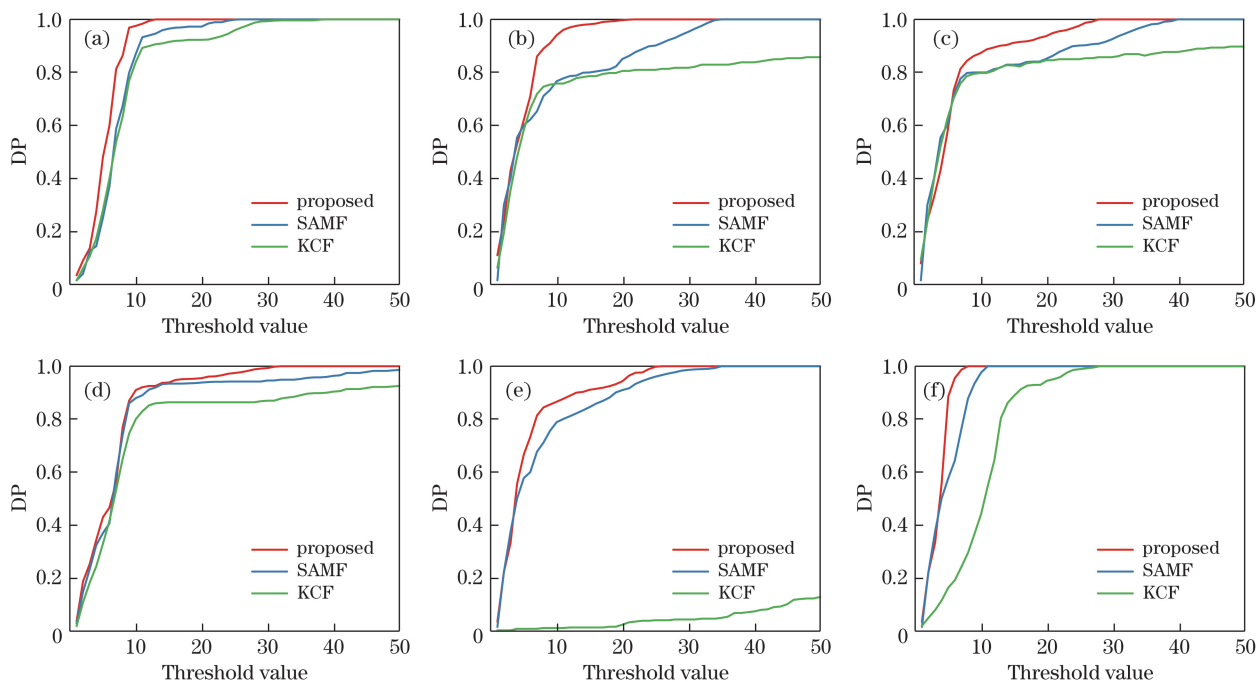


图 6 各个算法 DP。(a) Basketball; (b) carscale; (c) David; (d) girl; (e) shaking; (f) singer2

Fig. 6 DP of algorithms. (a) Basketball; (b) carscale; (c) David; (d) girl; (e) shaking; (f) singer2

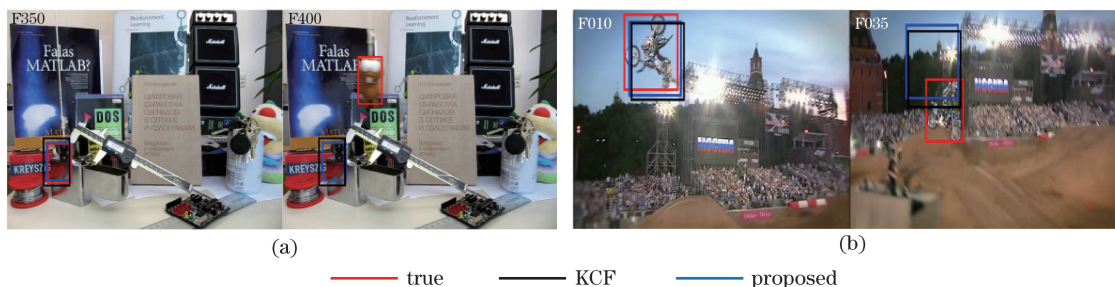


图 7 跟踪漂移现象。(a) Lemming; (b) MotorRolling

Fig. 7 Tracking drift phenomenon. (a) Lemming; (b) MotorRolling

4 结 论

在传统 KCF 跟踪算法的基础上提出了一种基于融合颜色特征的尺度自适应目标跟踪改进方法。引入 Lab 颜色空间下的图像颜色特征, 将其与 HOG 特征融合, 并利用相关滤波分类器计算输出响应, 颜色特征的引入增强了相关滤波分类器对彩色图像序列的跟踪性能; 另外, 通过引入尺度金字塔模型与 Lab 颜色特征直方图对目标尺度进行估计和更新, 有效地解决了目标尺度变化带来的跟踪问题。实验结果表明, 与其他算法相比, 所提算法在复杂场景中具有更强的稳健性和更高的跟踪精度, $76 \text{ frame} \cdot \text{s}^{-1}$ 的平均跟踪速度也能满足实时性的要求; 另外, 当目标出现长时间、大范围的遮挡或高速运动现象时, 跟踪结果不理想, 之后的相关研究可以采取自适应学习策略对算法进行改进。

参 考 文 献

- [1] Wu Y, Lim J, Yang M H. Online object tracking: a benchmark[C]. 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 2013: 2411-2418.
- [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] Zhang X G, Zhang Y, Wang Y N. Covariance tracking based on forgetting factor and Kalman filter[J]. Acta Optica Sinica, 2010, 30(8): 2317-2323.
张旭光, 张云, 王艳宁, 等. 基于遗忘因子与卡尔曼滤波的协方差跟踪[J]. 光学学报, 2010, 30(8): 2317-2323.
- [4] Vojir T, Neskova J, Matas J. Robust scale-adaptive mean-shift for tracking[J]. Pattern Recognition Letters,

- 2014, 49: 250-258.
- [5] Liu W, Zhao W J, Li C. Long-term visual tracking based on spatio-temporal context[J]. *Acta Optica Sinica*, 2016(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.
- [6] Cai Y Z, Yang D D, Mao N, *et al.* Visual tracking algorithm based on adaptive convolutional features[J]. *Acta Optica Sinica*, 2017, 37(3): 0315002.
蔡玉柱, 杨德东, 毛宁, 等. 基于自适应卷积特征的目标跟踪算法[J]. *光学学报*, 2017, 37(3): 0315002.
- [7] Bolme D S, Beveridge J R, Draper B A, *et al.* Visual object tracking using adaptive correlation filters[C]. 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, USA, 2010: 2544-2550.
- [8] Henriques J F, Caseiro R, Martins P, *et al.* Exploiting the circulant structure of tracking-by-detection with kernels[C]//Fitzgibbon A, Lazebnik S, Perona P, *et al.* *Computer Vision-ECCV 2012*. ECCV 2012. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 2012, 7575: 715-720.
- [9] 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.
- [10] 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, Columbus, OH, USA, 2014: 1090-1097.
- [11] Danelljan M, Häger G, Khan F S, *et al.* Accurate scale estimation for robust visual tracking[C]//*Proceedings of the British Machine Vision Conference*, Guildford: BMVA Press, 2014: 1-11.
- [12] Xu L, Wang M, Wen Y. Moving object detection using LAB color space[J]. *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, 2013, 41(s1): 219-222.
许莉, 王敏, 温月. 基于 Lab 颜色空间的运动目标检测[J]. *华中科技大学学报(自然科学版)*, 2013, 41(s1): 219-222.
- [13] Xu Y L, Wang J B, Li Y, *et al.* Scale adaptive correlation tracking combined with color features[J]. *Application Research of Computers*, 2017, 34(3): 945-948.
徐玉龙, 王家宝, 李阳, 等. 融合颜色特征的尺度自适应相关跟踪[J]. *计算机应用研究*, 2017, 34(3): 945-948.
- [14] Li S S, Zhao G P, Wang J Y. Distractor-aware object tracking based on multi-feature fusion and scale-adaption[J]. *Acta Optica Sinica*, 2017, 37(5): 0515005.
李双双, 赵高鹏, 王建宇. 基于特征融合和尺度自适应的干扰感知目标跟踪[J]. *光学学报*, 2017, 37(5): 0515005.
- [15] Everingham M, Gool L V, Williams C K I, *et al.* The pascal visual object classes (voc) challenge[J]. *International Journal of Computer Vision*, 2010, 88(2): 303-338.
- [16] Li Y, Zhu J K, Hoi S C H. Reliable patch trackers: robust visual tracking by exploiting reliable patches[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 2015: 353-361.
- [17] Li Y, Zhu J K. A scale adaptive kernel correlation filter tracker with feature integration[C]//Agapito L, Bronstein M, Rother C. *Computer Vision-ECCV 2014 Workshops*. ECCV 2014. Lecture Notes in Computer Science, Cham: Springer, 2014, 8926: 254-265.
- [18] Gao J, Ling H, Hu W, *et al.* Transfer learning based visual tracking with Gaussian processes regression[C]//Fleet D, Pajdla T, Schiele B, *et al.* *Computer Vision-ECCV 2014*. ECCV 2014. Lecture Notes in Computer Science, Cham: Springer, 2014, 8691: 188-203.