

文章编号 1004-924X(2018)08-2100-12

自适应模板更新和目标重定位的相关滤波器跟踪

刘教民, 郭剑威, 师 硕*

(河北工业大学 计算机科学与软件学院, 天津 300401)

摘要:针对核相关滤波器在跟踪中因目标快速运动导致的目标易丢失和部分遮挡问题,本文在多特征尺度自适应核相关滤波器(Scale Adaptive with Multiple Features tracker, SAMF)基础上,提出一种融合自适应模板更新和预测目标位置重定位的核相关跟踪算法。采用联合目标移动速度和特征变化的模板更新机制增大对目标快速运动适应性,根据长时滤波器和短时滤波器协作跟踪提出目标位置修正和重定位模型提升跟踪器应对目标部分遮挡的能力。在 OTB-2015 视频序列集 100 组序列中与序列集提供的算法进行对比,本算法跟踪精度相比 SAMF 提升 2%。在目标发生快速移动时本文算法具有更好的追踪目标能力,目标重定位也很好解决了目标部分遮挡问题。

关键词:视觉跟踪;相关滤波器;自适应模板更新;目标重定位

中图分类号:TP394.1;TH691.9 **文献标识码:**A **doi:**10.3788/OPE.20182608.2100

Correlation filter tracking based on adaptive learning rate and location refiner

LIU Jiao-min, GUO Jian-wei, SHI Shuo*

(School of computer Science & Engineering, Hebei University of Technology, Tianjin 300401, China)

* Corresponding author, E-mail: shishuo@scse.hebut.edu.cn

Abstract: To overcome the problem of loss of target caused by fast motion and the issue of partial occlusion in the tracking of kernel correlation filters, this paper proposed a new kernel correlation tracking algorithm that combines adaptive template updating and the prediction of the relocation of a target, based on the scale adaptive with multiple features tracker (SAMF). A template updating mechanism that combines target velocity and feature changes was proposed to improve the adaptability to fast movement of the target. Based on cooperative tracking of long time and short time filters, a target position correction and relocation model was proposed to improve the ability of the tracker to cope with partial occlusion of the target. In 100 sequences of OTB-2015 video set, the proposed algorithm was compared with the algorithms based on sequence sets and the SAMF algorithm. The tracking accuracy of the proposed algorithm is 2% higher than that of the SAMF algorithm, and the success rate is increased by 1%. The proposed algorithm has better tracking ability for fast moving targets and the target relocation scheme effectively addresses the problem of partial occlusion of the target.

Key words: vision tracking; correlation filter; adaptive learning rate; location refiner

收稿日期:2017-12-06;修订日期:2018-02-20.

基金项目:天津市科技计划资助项目(No. 14RCFGX00846;No. 15ZCZDNC00130;No. 17ZLZDZF00040);河北省自然科学基金资助项目(No. F2015202239)

1 引言

目标跟踪是自动监控、人机交互、无人驾驶等计算机视觉应用领域中最活跃的研究课题之一^[1]。由于光照变化、部分遮挡、目标形变、尺度变化、复杂背景和快速运动等原因,实现快速且有效的跟踪仍面临许多挑战^[2]。

近十年来,相关滤波器(Correlation Filter, CF)的出现,使跟踪效果取得了很大提升,成为近年来视觉跟踪的研究热点^[3]。CF在傅立叶域中计算相关性,对场景中每个感兴趣目标建立高相关响应值的同时对背景区域给出低响应值,再通过计算样本区域与模板的高相关响应值的位置,预测目标出现位置以实现跟踪。Bolme等^[4]在2010年提出了最小误差平方和(Minimum Output Sum of Squared Error, MOSSE)滤波器,首次将相关滤波器应用至跟踪领域,MOSSE计算简单,能快速跟踪目标,但在目标外观变化时不能保证准确跟踪。2012年,Henriques等^[5]在MOSSE基础上提出核循环矩阵跟踪(Circulant Structure Tracking with Kernels, CSK),提升了MOSSE的目标判别能力,但不能应对复杂场景的跟踪。2014年,Danelljan等^[6]在CSK基础上引入颜色名称(Color Names, CN^[7])特征,并将CSK模板更新机制调整为适应多通道特征。CN更好地提升了目标的判别能力,但是对目标遮挡、旋转、出视野等情况跟踪精度不高。Henriques等^[8]在CSK基础上引入多通道HOG特征^[9]得到核相关滤波器(Kernelized Correlation Filters tracking, KCF),KCF是目前基于相关滤波跟踪算法的基础框架,核方法和HOG特征提升了滤波器的判别能力,但对目标旋转、部分遮挡、快速运动等适应性仍需提升。Danelljan等^[10]在CSK的基础上引入特征金字塔和3维相关滤波器,提出了判别尺度空间跟踪器(Discriminative Scale Space Tracker, DSST),DSST一定程度上解决了尺度问题并具有良好的移植性,但仍然无法克服CSK算法的缺陷问题。Li等^[11]将CN和HOG特征融合并用尺度池预测目标尺度变化,提出了多特征尺度自适应跟踪器(Scale Adaptive with

Multiple Features tracker, SAMF), SAMF在KCF基础上提升了目标的判别力,并一定程度上解决了尺度变化问题,具有较好的跟踪效果,但对快速运动目标以及部分遮挡还有待提高。Li等^[12]提出可靠补丁跟踪器(Reliable Patch Trackers, RPT),引入了蒙特卡洛框架预估置信度高的图像补丁的分布,RPT利用部分组件以及生成式跟踪用的概率模型提升了跟踪性能,但时间效率低^[13]。2015年张雷等^[14]提出采用核相关滤波器的自适应尺度跟踪,较KCF提升了目标应对尺度变化的鲁棒性。2016年,王春平等^[15]通过研究HIS空间引入饱和度和梯度特征并且加入尺度自适应来提升KCF的跟踪性能。同年王暉等^[16]提出了特征融合和模板自适应更新相结合的相关滤波跟踪,利用CN和HOG特征互补性设计特征向量,通过响应图的峰值旁瓣比计算模板的更新。

本文在SAMF算法基础上,研究了模板更新机制,提出一种自适应模板更新和目标重定位的相关滤波跟踪器(Adaptive Learning rate and Location Refiner tracker, ALLR)。根据目标的运动速度和特征变化,设计一种自适应模板更新策略,提升了模板更新的可靠性,进而根据跟踪中的响应值提出了目标重定位机制,提高了应对遮挡的能力。

2 SAMF跟踪器

SAMF以KCF跟踪器为基础,采用可用于循环矩阵的高斯核,对多通道图像提取灰度信息、HOG和颜色名称3种特征进行融合,设计尺度池实现尺度自适应,取得了良好的跟踪效果。

2.1 KCF跟踪器

KCF跟踪器的关键是通过增加负样本提高跟踪检测器的辨别能力,同时探索循环矩阵的高效率结构。假设一维数据 $\mathbf{x}=[x_1, x_2, \dots, x_n]$, \mathbf{x} 的循环移位记为 $\mathbf{P}\mathbf{x}=[x_n, x_1, x_2, \dots, x_{n-1}]$ 。所有循环移位视觉样本 $\{\mathbf{P}^u\mathbf{x} \mid u=0 \dots n-1\}$ 形成数据矩阵 $\mathbf{X}=\mathbf{C}(\mathbf{x})$,因矩阵全由循环数据 \mathbf{x} 构成,称为循环矩阵。所有的循环矩阵描述如下^[17]:

$$\mathbf{X}=\mathbf{F}^H \text{diag}(\mathbf{F}\mathbf{x}) \mathbf{F}, \quad (1)$$

其中: \mathbf{F} 是离散傅立叶变换 (Discrete Fourier Transform, DFT) 矩阵, \mathbf{F}^H 是 \mathbf{F} 的埃尔米特转置。应用循环矩阵的分解可简化求线性回归的解, 线性岭回归的目标函数如式(2)所示:

$$\min_{\mathbf{w}} \sum_i^n (f(\mathbf{x}_i) - y_i)^2 + \lambda \|\mathbf{w}\|^2, \quad (2)$$

其中 f 为基础样本的线性组合: $f(\mathbf{x}) = \mathbf{w}^T \mathbf{x}$ 。岭回归具有近似解, $\mathbf{w} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y}$ 。代入循环矩阵后, 得到结果 $\hat{\mathbf{w}}^* = \frac{\hat{\mathbf{x}}^* \odot \hat{\mathbf{y}}}{\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}}^* + \lambda}$, 其中 $\hat{\mathbf{x}} = \mathbf{F} \mathbf{x}$ 表示 \mathbf{x} 的 DFT, $\hat{\mathbf{x}}^*$ 表示 $\hat{\mathbf{x}}$ 的复共轭。在非线性回归情况下, 应用核函数机制 $f(\mathbf{z}) = \sum_{i=1}^n \alpha_i \kappa(\mathbf{z}, \mathbf{x}_i)$ 可引入更强大的分类器, 并且在求解过程中利用循环矩阵的特性可避免矩阵求逆从而提升计算效率^[8]。对偶空间系数 α 可由下式得到:

$$\hat{\alpha}^* = \frac{\hat{\mathbf{y}}}{\hat{\mathbf{k}}^{xx} + \lambda}, \quad (3)$$

其中 \mathbf{k}^{xx} 是核相关性计算。类似线性回归的情况, 对偶系数在傅立叶域计算。在文献[8]中, 引入带循环矩阵的高斯核, 如式(4)所示:

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

可看出 KCF 只需要点积运算和离散傅立叶变换及其逆变换, 时间复杂度为 $O(n \log n)$ 。训练标签 y 是高斯函数, 它从值为 1 的中心目标平滑地衰减到值为 0 的其他移位。0 表示负样本, 需要放大原始目标边界框来包含负样本。SAMF 中采用原始目标框 2.5 倍大小的窗口进行训练。虽然循环移位丢失了原始帧的大量信息, 但分类器获得的密集样本可更精确地匹配模板^[11]。

2.2 多特征融合

提取特征的主要目的是区分对象和背景, SAMF 重要改进是选取鲁棒性高的多通道多个特征。假设多通道的特征融合为一个向量 $\mathbf{x} = [x_1, x_2, \dots, x_c]$, 公式(4)可以重写如下:

$$\mathbf{k}^{xx'} = \exp\left(-\frac{1}{\sigma^2} (\|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2) - 2\mathbf{F}^{-1}\left(\sum_c \hat{\mathbf{x}}_c \odot \hat{\mathbf{x}}_c'^*\right)\right), \quad (5)$$

这样可以使用除了原始灰度像素之外的其他

特征。SAMF 选取三类特征: 灰度信息, GRAY; 梯度直方图特征, HOG^[9]; 颜色名称特征, CN^[7]。其中, HOG 强调图像的梯度变化信息, CN 侧重颜色信息, 灰度表示图像强度, 这些特征相互补充, 其特征组合提升了 SAMF 跟踪的鲁棒性^[11]。

2.3 多尺度核相关滤波器

KCF 无法处理视频中尺度变化, SAMF 提出一种尺度自适应方式, 较好地处理了尺度变化问题。SAMF 中, 固定模板大小为 $\mathbf{s}_T = (s_w, s_h)$, s_w 和 s_h 分别表示样本的宽度和高度, 并定义尺度池 $\mathbf{S} = \{t_1, t_2, \dots, t_k\}$, $\{t_i, i = 1, 2, \dots, k\}$ 表示相对样本的放大和缩小, i 表示尺度池的索引值。假设目标窗口在初始图片区域的大小为 s_i 。在当前帧, 从原始图像空间采集 k 个样本集 $\{t_i \mathbf{s}_i \mid t_i \in \mathbf{S}\}$ 寻找预测目标。SAMF 中采用双线性插值调整所有样本为固定模板 \mathbf{s}_T 的大小, 最终响应值定义为式(6):

$$\arg \max \mathbf{F}^{-1} \hat{f}(\mathbf{z}^i), \quad (6)$$

式中 \mathbf{z}^i 表示调整大小为 $t_i \mathbf{s}_i$ 的样本到模板大小 \mathbf{s}_T 后的值, 这样搜索窗口的多尺度样本均被放缩至相同大小。在增加尺度适应的同时保持滤波器模型的样本和系数维度 s_w 和 s_h 不变, 只需要操作两个数据: 集合 α 和提取的样本数据 $\tilde{\mathbf{x}}_{\text{new}}$ 。集合 α 是在岭回归中利用核技巧 $\mathbf{w} = \sum_i \alpha_i \varphi(\mathbf{x}_i)$ 中组合系数 α_i 的集合, 又因在求解中引入核函数 κ 表示两个样本在高维特征空间的内积 $\kappa(\mathbf{z}, \mathbf{x}_i) = \varphi(\mathbf{z}) \odot \varphi(\mathbf{x}_i)$, 则 $\alpha = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{Y}$, 其中 \mathbf{K} 为保存训练样本间核函数值的核矩阵且 $\mathbf{K}_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{k}^{x_i x_j}$ 。文献[8]证明高斯核函数可在傅里叶域中利用公式(3)快速求解。根据公式(7)所示, 通过 α 以及 $\tilde{\mathbf{x}}_{\text{new}}$ 计算滤波器模板的更新方式:

$$\bar{\mathbf{T}}_{\text{new}} = \theta \mathbf{T}_{\text{new}} + (1 - \theta) \bar{\mathbf{T}}_{\text{old}}, \quad (7)$$

其中: $\bar{\mathbf{T}}_{\text{new}} = [\alpha_{\text{new}}^T, \tilde{\mathbf{x}}_{\text{new}}^T]$ 表示更新的模板, $\bar{\mathbf{T}}_{\text{old}} = [\alpha_{\text{old}}^T, \tilde{\mathbf{x}}_{\text{old}}^T]$ 表示原滤波器模板, θ 表示更新系数。

3 ALLR

本文提出自适应模板更新和目标重定位的相关滤波跟踪器 (Adaptive Learning rate and Location Refiner tracker, ALLR) 是在 SAMF 基础上, 着重研究模型的更新策略, 很好地解决了更新

过程中精确性和稳定性的矛盾,同时采用目标重定位机制,提升了相关滤波器处理遮挡的能力。

3.1 自适应模板更新

相关滤波跟踪需要在视频序列第一帧中输入目标跟踪框,用跟踪框内特征生成滤波器。跟踪过程中目标的任何变化都可导致提取特征的变化,需要及时更新滤波器模型。通常有两种模板更新方式:(1)仅考虑上一帧目标和当前目标的线性关系;(2)考虑前面所有帧和当前目标的线性关系^[13]。相关滤波算法中采用第二种模板更新方式,更新率通常固定在 0.01~0.15^[18]。

实际更新过程中,如模型更新过慢,不能跟上目标特征变化,从而不能满足精确性;盲目增大更新率又会引入误差、噪声等造成模型漂移,而不能满足稳定性。研究发现滤波器模板更新不仅与选取的特征有关,还应根据目标运动速度进行动态更新,才能保证更好的跟踪效果^[13-14,18]。因此 ALLR 从目标速度和特征变化两方面制定更新机制,如图 1 所示。

为自适应速度变化,构造与速度相关的更新率 θ_1 。目标的运动速度 v 定义为当前预测框的中心与上一帧目标框中心像素的欧式距离,速度 v 的变化与目标周围环境及自身外观,如光照、尺度变化等有很大关系。速度 v 较小时,周围环境和自身外观变化都相对较小;速度 v 较大时,周围环境变化剧烈且自身变化相对较大。因此,当速度 v 很小时,尽量保留原来模板的信息, θ_1 随 v 变化的幅度很小;当速度 v 逐渐提升时,需增强对目标周围环境变化以及目标自身外观变化的适应性, θ_1 随 v 变化幅度逐渐增大;当速度 v 大到一定程度时,模板更新率已经很大。更新率 θ_1 与运动速度 v 关系如式(8)所示:

$$\theta_1 = \frac{1}{1 + \left(\frac{6}{1+v}\right)^5}. \quad (8)$$

θ_1 的函数图如图 2 所示,图中目标运动速度 v 大于 5 时认为是快速运动^[15]。目标运动速度 v 在 4~5.21 范围内,取 $\Delta v=0.01$ 时各点斜率大于等于 0.2,最大斜率为 0.21 出现在 $v=4.95$ 处。因此,曲线在速度为 5 附近达到最大斜率,之后随着速度的提升斜率降低,当速度大于 10 时急剧减小。根据目标速度梯度更新模板,更好地适

应了目标的运动变化。与文献[18]相比, θ_1 函数不仅符合更新率的变化规律而且适应的速度范围也更广。

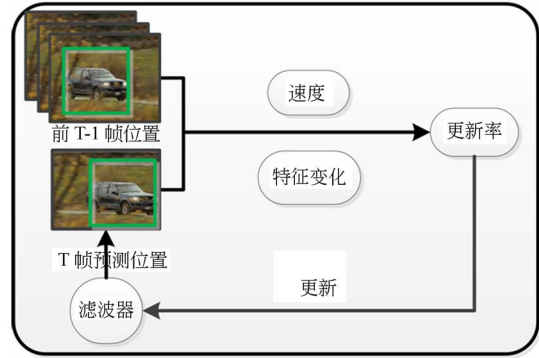


图 1 自适应更新框架

Fig. 1 Framework of adaptive learning rate

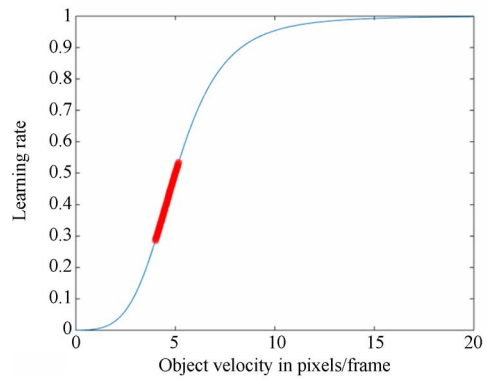


图 2 速度-更新率曲线

Fig. 2 Curve of velocity and learning rate

文中定义的目标速度为两帧间目标框中心像素的距离,实际视频帧中速度会有误差,甚至会出现目标和摄像机相对静止的情况,此时所定义的速度 v 不一定能反应出目标周围环境及自身的变化过程。因此将更新率与提取的目标样本特征变化相关联,定义为更新率 θ_2 。当目标特征变化较小时,认为当前预测位置非常可靠,可适当增大模板更新率;反之,当目标特征变化较大时,认为当前跟踪出现了遮挡,为降低模板受遮挡影响,应减少模板的更新率直至趋于 0。定义特征变化为当前样本特征与模板特征的相关系数 c ,更新率 θ_2 与目标特征变化 c 计算公式如(9)所示:

$$\theta_2 = \frac{1}{2} \frac{e^{5c-\frac{5}{2}} - e^{-5c+\frac{5}{2}}}{e^{5c-\frac{5}{2}} + e^{-5c+\frac{5}{2}}}. \quad (9)$$

更新率 θ_2 的函数图如图 3 所示。当前区域

提取的目标特征与模板特征很接近时,两特征的相关系数趋近为 1,认为是可靠的跟踪并更新模板;反之,若提取的特征与模板特征有很大的差距,那么两特征的相关系数 c 趋近 0,认为跟踪出现了问题且不更新模板。引入目标的相关系数 c 不仅能解决相对静止的速率更新,更能适应目标的特征变化,从而提高应对特征变化的能力。

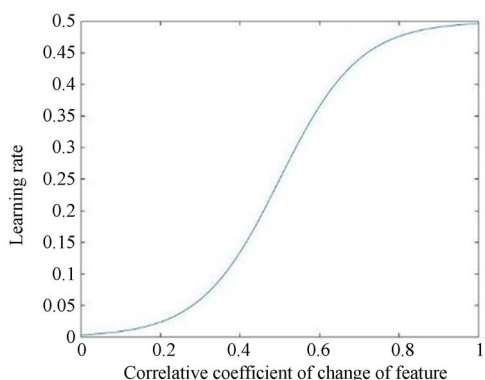


图 3 特征变化-更新率曲线

Fig. 3 Curve of feature change and learning rate

考虑运动速度和特征变化的互补性,本文将 θ_1 与 θ_2 通过线性组合的形式,得到最终的模板更新率:

$$\theta = \alpha \times \theta_1 + \beta \times \theta_2, \quad (10)$$

3.2 目标重定位

文献[19]表明跟踪可仿照人脑记忆模型中的短期记忆和长期记忆设计为短时跟踪器和长时跟踪器,本文据此将跟踪过程分解为短时滤波和长时滤波两个滤波跟踪过程。短时滤波为原跟踪框

架下的滤波器,长时滤波依据 SVM 通过第一帧给定目标位置框内提取的特征建立。长时滤波在跟踪过程中按照设定的阈值进行更新,在尽量保证目标原外观特征的同时增强后续跟踪的适应性。

跟踪过程依靠短时滤波和长时滤波相互合作。在上一帧的预测位置获得搜索窗口,使用短时相关滤波器获取目标的融合特征,并且获得预测区域的响应值,如图 4 所示。

通过设定高阈值 $TH1$ 和低阈值 $TH2$ 来判断重定位组件的操作:

$$Refiner(res) = \begin{cases} \text{更新长时滤波器,} & res > TH1 \\ \text{修正目标位置,} & TH2 < res < TH1 \\ \text{目标重定位,} & res < TH2 \end{cases} \quad (11)$$

式中: res 代表每帧预测的目标位置所对应的响应值。当预测响应值大于阈值 $TH1$ 时,认为本次短时跟踪器得到的位置是精确的,不需要长时滤波参与且更新长时滤波;当响应值在阈值 $TH1$ 和阈值 $TH2$ 之间时,认为目标发生了一定量的偏移,需要长时滤波器来修正短时滤波器的位置且不对长时滤波进行更新;当响应值小于下阈值 $TH2$ 时,认为目标发生了遮挡等重大变化,依靠长时滤波在预测区域搜索窗口得到响应位置并依据长时滤波更新短时滤波,若所得位置响应值依然低于 $TH2$ 则不预测目标位置且不做任何更新动作,保留上一帧位置以及长时滤波,直到后续帧再次符合阈值条件时,预测位置并更新短时滤波。

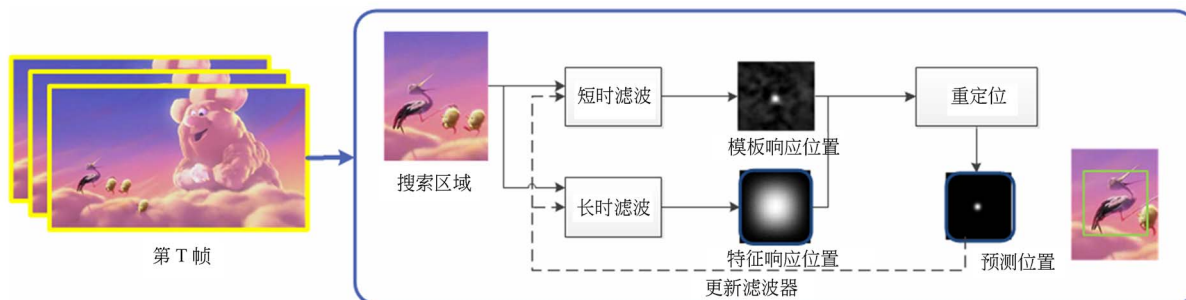


图 4 长时滤波和短时滤波协作示意图

Fig. 4 Cooperation between Long-term filter and Short-term filter

3.3 算法步骤

综上所述,融合动态模版更新率和目标重定位机制,ALLR 算法的基本步骤如下。

算法 ALLR 整体流程

1: frame = 1, 预定目标框样本 $\tilde{\mathbf{x}}_{old}$ 作为模板, 对应的特征矩阵 \mathbf{F}_{old} 、空间系数 \mathbf{a}_{old} , 建立短时滤波器 CF_{Short} , 长时滤波器 CF_{Long}

2: frame > 1 且 t_i 在尺度集 $S = \{t_1, t_2, \dots, t_K\}$ 中

(1) 取大小为 z_i (i 为尺度集的索引值) 图像块, 记作 S_i , 提取特征矩阵 \mathbf{F}_{new}

(2) 用 $\tilde{\mathbf{x}}_{old}$ 、 \mathbf{a}_{old} 计算每个尺度索引 i 的响应值, 并得到最大的响应值 $\hat{f}(z_i)$

3: 取 CF_{Long} 、 CF_{Short} 以及预定义的上阈值 $TH1$, 下阈值 $TH2$ 进行重定位或修正目标位置操作, 最终得到预测位置 \mathbf{P}_{new} 和所对应尺度 $t_i \cdot s_i$

4: 通过 \mathbf{P}_{new} 、 $t_i \cdot s_i$ 得到预测的目标取样模板 $\tilde{\mathbf{x}}_{new}$, 新的空间系数 \mathbf{a}_{new}

5: 通过 \mathbf{P}_{new} 和 \mathbf{P}_{old} 计算 \mathbf{F}_{new} 以及 \mathbf{F}_{old} 计算 θ_2 , 得到新的模板更新率 θ

6: 用 θ 、 $\tilde{\mathbf{x}}_{new}$ 、 \mathbf{a}_{new} 更新模板的 $\tilde{\mathbf{x}}_{old}$ 、 \mathbf{a}_{old}

7: 如果有下一帧跳到步骤 2, 没有则结束跟踪

4 实验

4.1 实验环境

本文采用 Object Tracking Benchmark 2015 数据集^[1]作为实验对象, 并和其中提供的主流跟踪算法 Struck^[20]、TLD^[21]、CXT^[22]、LSK^[23], 以及 KCF^[8]、SAMF^[11]、RPT^[12] 共 7 种算法进行对比。实验在 Intel(R) Core(TM) i7-4790 CPU, 主频 3.6 GHz, 8 GB 内存的电脑上采用 Matlab R2014b 实现。本文算法的运行速度为 6.44 f/s (frames per second)。

4.2 实验系数的确定

为得到最符合目标跟踪的更新率 θ_1 与 θ_2 的权重 α 和 β , 选用平均中心位置误差 (Center Location Error, CLE) 为评价指标。CLE 是算法跟踪结果与手动标定值之间的像素差, 较小的值意味着更精确的结果, CLE 通常选择 20 像素差内的跟踪精度。在速度和特征变化中, 速度为主要影响因素, 但速度比例过大时, 精度又会下降。实验结果如表 1 所示, 当选择更新率阈值 $\alpha=0.01$, $\beta=0.005$ 时, 具有最好的跟踪正确率。

表 1 不同权重下的跟踪正确率

	α	0.005,	0.01,	0.005,	0.05,	0.1,
	β	0.005	0.005	0.01	0.005	0.005
CLE(20)	0.738	0.761	0.742	0.755	0.751	

在最佳更新阈值下实验选择重定位阈值系数, 所选评价指标与更新率相同。实验结果如表 2 所示, 可见当上阈值 $TH1$ 为 0.4, 下阈值 $TH2$ 为 0.15 时, 效果最好, 因此选定重定位阈值为 $TH1=0.4$, $TH2=0.15$ 。

表 2 不同阈值下的跟踪正确率

	$TH1$	0.5,	0.6,	0.4,	0.3,	0.4,	0.4,
	$TH2$	0.15	0.15	0.15	0.15	0.1	0.2
CLE(20)	0.769	0.772	0.782	0.777	0.779	0.780	

4.3 定量分析

采用一次通过性评估 (One-Pass Evaluation, OPE) 选取精度图和成功率图两个评价指标评价本文跟踪算法的性能。精度图 (Precision plot) 显示了预计位置在手动标定给定阈值距离内的帧的百分比, 采用阈值为 20 像素作为每个跟踪器的精度数值。成功率图 (Success plot) 中, 给定边界框和手动标定框 r_a , 重叠值定义为 $s = |r_i \cap r_a| / |r_i \cup r_a|$, 其中 \cap 和 \cup 表示两个区域的交集和并集, $|\cdot|$ 表示该区域的像素数, s 用于确定算法是否成功地追踪一帧中的目标。成功率图评估大于给定阈值 (通常取 0.5) 的成功帧数比率为成功率, 由于阈值在 0 和 1 之间变化, 因此成功率会随之发生变化。为了衡量一个完整序列的性能, 本文通过计算跟踪器对所有序列的平均重叠数 (Average Overlap Score, AOS) 来评价算法。

经过大量实验, 确定更新系数 $\alpha=0.01$, $\beta=0.005$ 和重定位阈值系数 $TH1=0.4$, $TH2=0.15$, 在此系数下本文算法和目前 Struck^[20]、TLD、CXT、LSK 和相关滤波算法 KCF、SAMF、RPT 进行对比。实验结果如图 5, 图 5(a) 成功率图在阈值取 0.5 时的数值显示, ALLR 比 SAMF 高 0.6%; 图 5(b) 精度图, 在 20 像素内的数值显示, ALLR 比 SAMF 高 2%。

OPE 是跟踪算法的常用评价标准, 由于跟踪

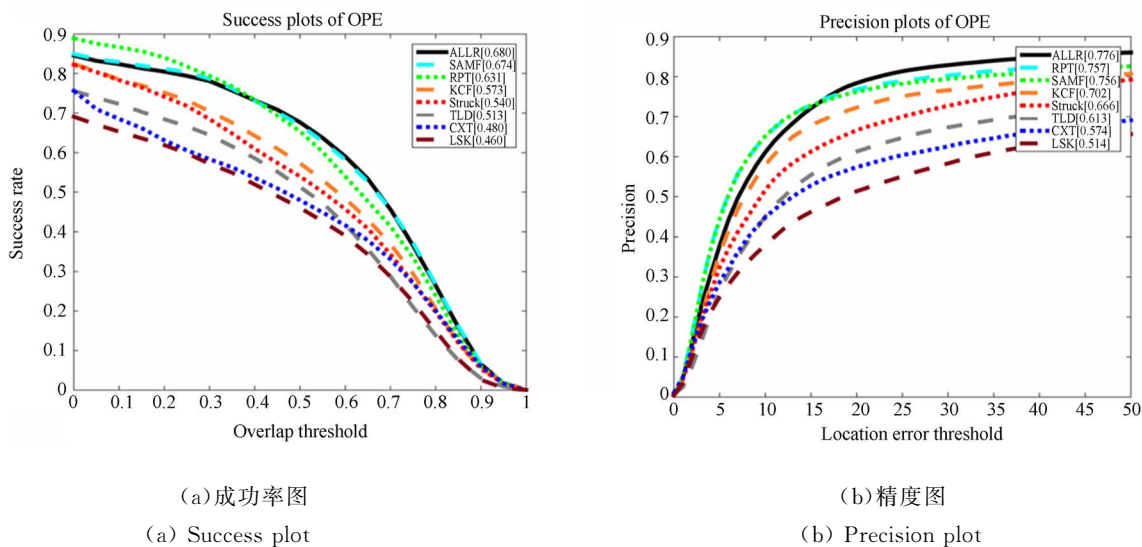


图 5 各跟踪算法 OPE 的成功率图和精度图

Fig. 5 Success plot and precision plot of OPE for the trackers

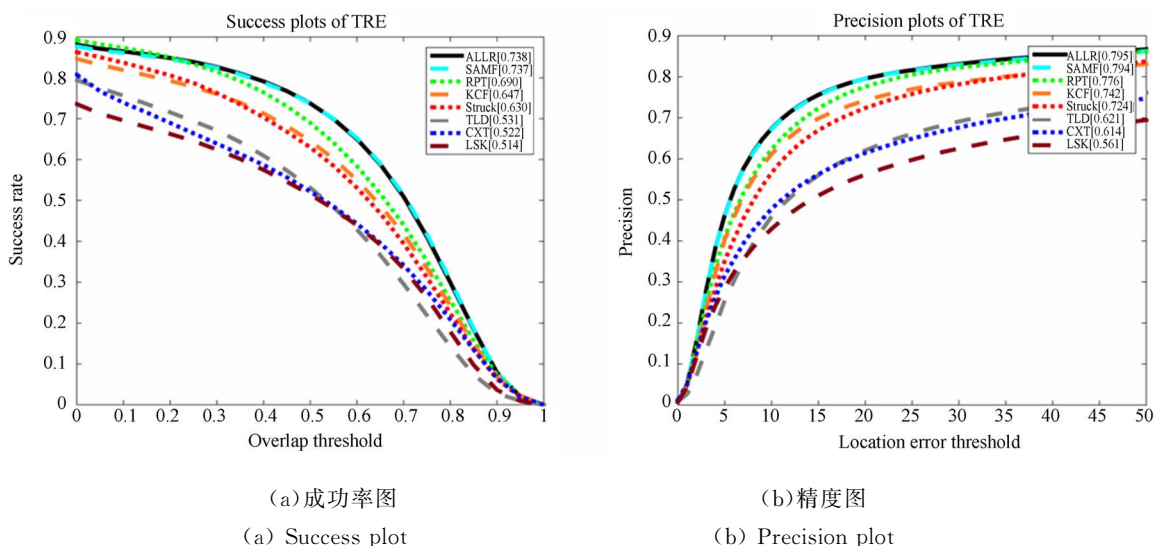


图 6 各跟踪算法 TRE 的成功率图和精度图

Fig. 6 Success plot and precision plot of TRE for the trackers

器对初始化敏感并且在不同起始帧处初始化的表现有所不同,因此引入空间鲁棒性(Spatial Robustness Evaluation, SRE)和时间鲁棒性(Temporal Robustness Evaluation, TRE)两个标准分别分析在不同边界框初始化和不同帧初始化时跟踪器的鲁棒性。在 OTB-2015 上按照数据库的测试要求分别进行了 SRE 和 TRE 的测试,结果如图 6、图 7 所示。TRE 评价中,图 6(a)和图 6(b)显示,ALLR 与 SAMF 基本持平。SRE 评价中,图 7(a)显示 ALLR 在 0.5 处成功率比 SAMF 高 1.6%;图 7(b)显示 ALLR 在 20 像素内精度比

SAMF 高 1.2%。

为进一步测试 ALLR 应对快速运动和遮挡的能力,对 OTB-2015 数据库中具有快速运动属性的 39 个数据集和具有遮挡属性的 49 个数据集进行了精度图分析,如图 8、图 9 所示。在应对快速目标物体时,图 8(a)显示,ALLR 成功率在 0.5 阈值时比 SAMF 高 1.6%;图 8(b)显示,ALLR 在 20 像素内精度比 SAMF 高 1.9%。虽然在快速运动中跟踪性能不及 RPT,但 RPT 算法的时间消耗,数倍于本文算法。RPT 算法分块跟踪,对每个块进行可靠性分析,并通过蒙特卡洛框架

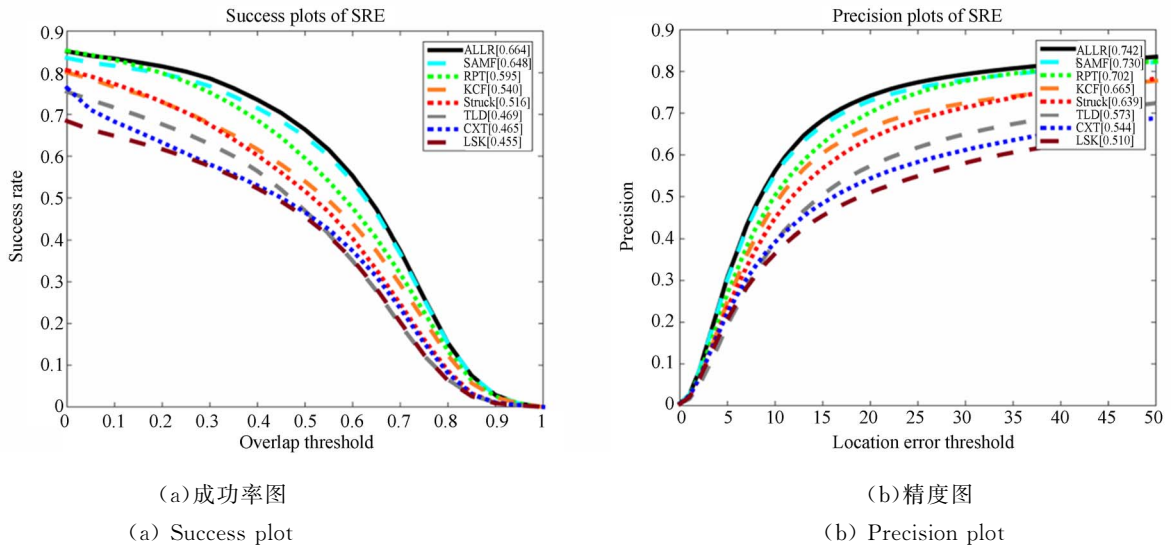


图 7 各跟踪算法 SRE 的成功率图和精度图

Fig. 7 Success plot and precision plot of SRE for the trackers

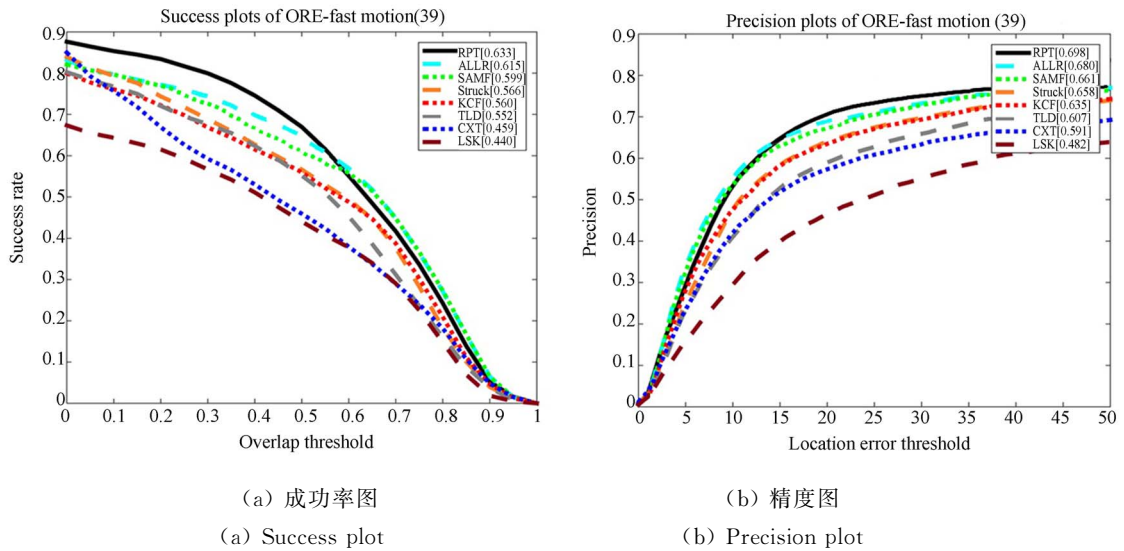


图 8 各跟踪算法在快速运动视频上 OPE 的成功率图和精度图

Fig. 8 Success plot and precision plots of OPE in the videos with fast movement for the trackers

估计分块的分布,十分消耗计算时间。因 RPT 在本文环境下,ALLR 和 RPT 分别是 6.44 f/s 和 1.80 f/s。在应对目标遮挡时,图 9(a)显示,ALLR 成功率在 0.5 阈值时比 SAMF 高 1.7%;图 9(b)显示,ALLR 整体精度在 20 像素内比 SAMF 高 1.9%。

实验结果表明,用 SRE、TRE、OPE 3 种评价指标衡量,本文提出的自适应模板更新和目标重定位机制的 ALLR 算法,在所有数据集及包含快速运动和遮挡情况的数据集中较 SAMF 均有提升。

4.4 定性分析

本文算法是为了增强 SAMF 应对快速运动和遮挡的能力,接下来分别分析两种情况的跟踪效果。由于 Struck、TLD、CXT、LSK 未提供完整的实验结果,这里无法进行对比分析。

4.4.1 快速运动情况

本文算法 ALLR 和 SAMF、KCF、RPT 在序列 Vase 序列上进行快速运动的实验,结果如图 10 所示。第 122 帧时跟踪框定位在盒子正面的图案,第 198 帧时摄像机突然拉远再拉近,第 221 帧跟踪框显示只有本文算法还定位在盒子正面的图案。第

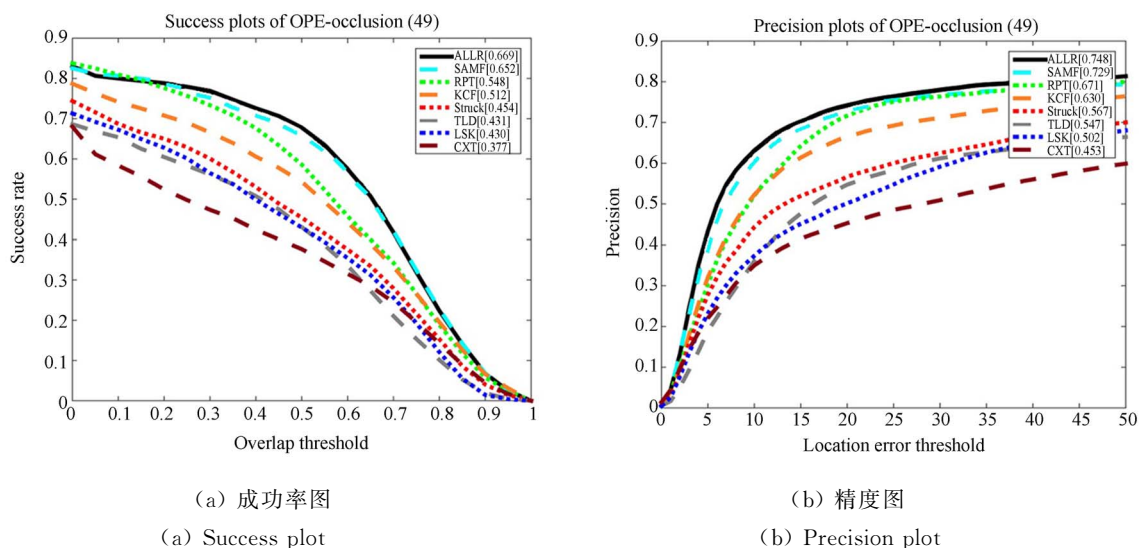


图 9 各跟踪算法在遮挡视频上 OPE 的成功率图和精度图

Fig. 9 Success plot and precision plots of OPE in the videos with occlusion for the trackers

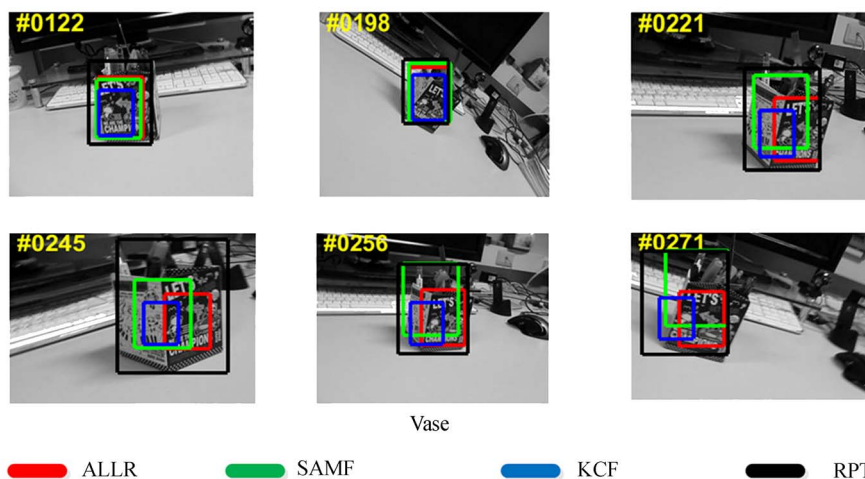


图 10 跟踪算法在快速运动视频集上定性结果

Fig. 10 Qualitative results of the trackers in the video with fast movement

245 帧摄像机再做快速向左拉近且在 256 帧迅速返回原位置,可见本文算法依然定位在正面盒子的图案上。实验表明 ALLR 在上述问题中依然能锁定盒子并且保持盒子正面图案在跟踪框内,没有受到相对速度变化的影响从而产生误差。

4.4.2 遮挡情况

本文算法 ALLR 对比 SAMF、KCF、RPT,进行了多种不同程度的遮挡实验。在序列 Liquor 上应对遮挡的情况如图 11 所示,1 511 帧时所有算法跟踪正常,1518 帧时移动右侧物体发生遮挡,1 528 帧时遮挡结束后,只有本文算法和 RPT 依然能定位到目标瓶。在 1 559 帧时左移目标,

发现 RPT 不能准确跟踪到目标物体。1 741 帧跟踪结束,因目标移动位置经过跟踪失败的 KCF 跟踪框,所以结束时包含 KCF。算法在序列 Girl2 上应对遮挡的情况如图 12 所示。10 帧时,各跟踪器正常跟踪目标,326 帧目标通过台阶,KCF 已丢失目标,RPT 和 SAMF 出现跟踪漂移。405 帧时,SAMF 锁定黑色栏杆,RPT 偏离目标程度变大。至 1 380 帧目标发生遮挡,除本文算法,其他算法已完全偏离目标。1 391 帧完全遮挡,1 405 目标再次出现,ALLR 仍然能够正确跟踪目标。实验结果表明 ALLR 算法,在目标经历不同程度的遮挡之后依然能够准确定位。

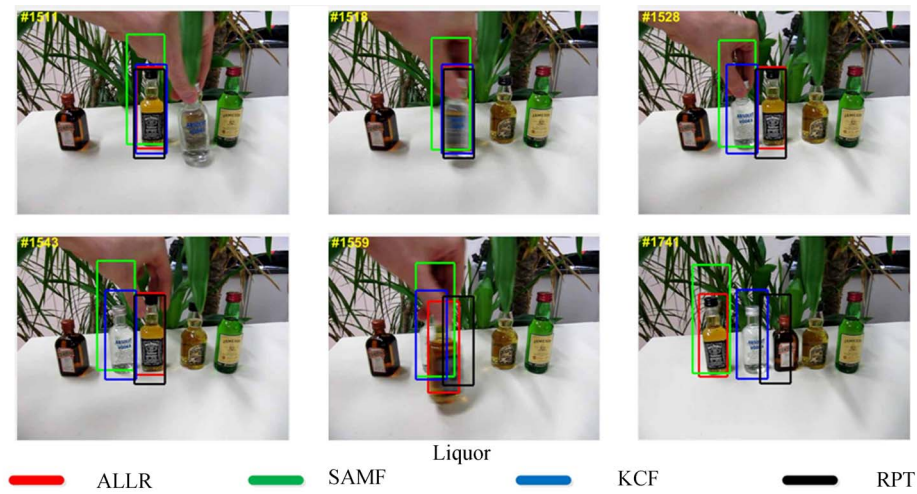


图 11 跟踪算法在遮挡视频集 Liquor 上定性结果

Fig. 11 Qualitative results of the trackers in the Liquor video with occlusion

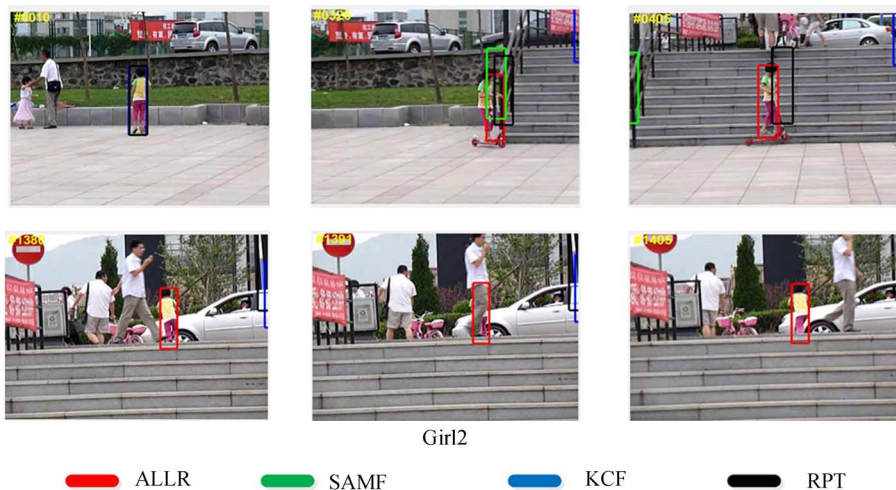


图 12 跟踪算法在严重遮挡视频集 Girl2 上定性结果

Fig. 12 Qualitative results of the trackers in the Girl2 video with heavy occlusion

5 结 论

本文从模板动态更新和目标位置重定位两个方面对 SAMF 进行改进。首先将目标运动速度和目标特征变化两部分作为调节更新率的参数,并通过线性加权融合两个因素。然后,为提升

SAMF 跟踪中应对目标发生遮挡的能力,提出长时滤波和短时滤波相结合的策略。通过对所提算法的详细分析并在 OTB-2015 数据库与多种主流算法进行实验对比,本算法跟踪精度相比 SAMF 提升 2%。本文算法在运动目标快速运动和遮挡的情况下,较其他算法的跟踪性能优。

参考文献:

[1] WU Y, LIM J, YANG M H. Online Object Tracking: A Benchmark[C]. *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Com-

puter Society, 2013:2411-2418.

[2] WU Y, LIM J, YANG M H. Object tracking benchmark[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, 37(9): 1834-48.

- [3] KRISTAN M, PFLUGFELDER R, MATAS J, *et al.*. The visual object tracking VOT2015 challenge results[C]. *IEEE International Conference on Computer Vision Workshop*. IEEE, 2016;564-586.
- [4] BOLME D S, BEVERIDGE J R, DRAPER B A, *et al.*. Visual object tracking using adaptive correlation filters[C]. *Computer Vision and Pattern Recognition*. IEEE, 2010;2544-2550.
- [5] RUI C, MARTINS P, BATISTA J. Exploiting the circulant structure of tracking-by-detection with kernels [C]. *European Conference on Computer Vision*. Springer, Berlin, Heidelberg, 2012;702-715.
- [6] DANELLJAN M, KHAN F S, FELSBURG M, *et al.*. Adaptive color attributes for real-time visual tracking[C]. *Computer Vision and Pattern Recognition*. IEEE, 2014;1090-1097.
- [7] WEIJER J V D, SCHMID C, VERBEEK J, *et al.*. Learning color names for real-world applications [J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2009, 18(7):1512-1523.
- [8] HENRIQUES J F, RUI C, MARTINS P, *et al.*. High-Speed tracking with kernelized correlation filters[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, 37(3):583-596.
- [9] FOIRSYTH D. Object detection with discriminatively trained part-based models[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2014, 47(2):6-7.
- [10] DANELLJAN M, HAGER G, KHAN F S, *et al.*. Discriminative scale space tracking[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2016, PP(99):1-1.
- [11] LI Y, ZHU J. A scale adaptive kernel correlation filter tracker with feature integration[J]. 2014, 8926;254-265.
- [12] LI Y, ZHU J, HOI S C H. Reliable Patch Trackers; Robust visual tracking by exploiting reliable patches[C]. *Computer Vision and Pattern Recognition*. IEEE, 2015;353-361.
- [13] 魏全禄, 老松杨, 白亮. 基于相关滤波器的视觉目标跟踪综述[J]. *计算机科学*, 2016, 43(11):1-5.
- WEI Q L, LAO S Y, BAI L. Visual object tracking based on correlation filters: A Survey [J]. *Computer science*, 2016, 43(11):1-5. (in Chinese)
- [14] 张雷, 王延杰, 孙宏海, 等. 采用核相关滤波器的自适应尺度目标跟踪[J]. *光学 精密工程*, 2016, 24(2):448-459.
- ZHANG L, WANG Y J, SUN H H, *et al.*. Adaptive scale object tracking with kernelized correlation filters [J]. *Opt. Precision Eng.*, 2016, 24(2):448-459. (in Chinese)
- [15] 王春平, 王晔, 刘江义, 等. 基于色度饱和度-角度梯度直方图特征的尺度自适应核相关滤波跟踪[J]. *光学 精密工程*, 2016, 24(9):2293-2301.
- WANG CH P, WANG W, LIU J Y, *et al.*. Scale adaptive kernelized correlation filter tracking based on HHS-OG feature [J]. *Opt. Precision Eng.*, 2016, 24(9):2293-2301. (in Chinese)
- [16] 王晔, 王春平, 李军, 等. 特征融合和模型自适应更新相结合的相关滤波目标跟踪[J]. *光学 精密工程*, 2016, 24(8):2059-2066.
- WANG W, WANG CH P, LI J, *et al.*. Correlation filter tracking based on feature fusing and model adaptive updating [J]. *Opt. Precision Eng.*, 2016, 24(8):2059-2066. (in Chinese)
- [17] GRAY R M. *Toeplitz and circulant matrices: a review*[M]. Now Publishers Inc. 2005.
- [18] ASHA C S, NARASIMHADHAN A V. Adaptive learning rate for visual tracking using correlation filters[J]. *Procedia Computer Science*, 2016, 89: 614-622.
- [19] MA C, YANG X, ZHANG C, *et al.*. Long-term correlation tracking [C]. *Computer Vision and Pattern Recognition*. IEEE, 2015;5388-5396.
- [20] HARE S, SAFFARI A, TORR P H S. Structured output tracking with kernels[C]. *International Conference on Computer Vision*. IEEE Trans Pattern Anal Mach Intell, 2011; 263-270.
- [21] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking learning detection. [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2012, 34(7):1409-22.

- [22] DINH T B, VO N, MEDIONI G. Context tracker: Exploring supporters and distracters in unconstrained environments[C]. *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, 2011:1177-1184.
- [23] LIU B, HUANG J, YANG L, *et al.*. Robust tracking using local sparse appearance model and K-selection[C]. *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, 2011:1313-1320.

作者简介:

刘教民(1958—),男,河南西峡人,博士,教授,博士生导师,1998年于河北工业大学获得博士学位,主要从事计算机智能控制、多媒体技术方面的研究。
E-mail: lmj6667@126.com



师硕(1981—),女,河北易县人,博士,讲师,2006年于东北大学获得硕士学位,2014年于河北工业大学获得博士学位,现为河北工业大学计算机科学与软件学院计算机系教师,主要从事数字图像处理、模式识别等方面的研究。
E-mail: shishuo@scse.hebut.edu.cn