

# 基于在线检测和尺度自适应的相关滤波跟踪

王艳川, 黄海, 李邵梅, 高超

国家数字交换系统工程技术研究中心, 河南 郑州 450000

**摘要** 针对相关滤波跟踪在遮挡及目标尺度变化等情况下容易跟踪失败的问题, 提出一种基于在线检测和尺度自适应的相关滤波跟踪算法。相关滤波跟踪器融合方向梯度直方图特征、颜色属性特征和光照不变特征进行目标定位; 通过局部稀疏表示模型的重构残差进行遮挡判别, 如果发生遮挡则进行在线支持向量机检测, 实现目标重定位; 进行由粗至精的尺度估计, 通过尺度预估计和牛顿迭代法得到目标的精确尺度。采用均衡的模型更新策略, 固定更新相关滤波器, 保守更新稀疏表示模型和支持向量机。实验结果表明: 与现有跟踪算法相比, 所提算法能有效降低遮挡、目标尺度变化等复杂因素的干扰, 并在 50 组测试序列上取得较高的距离精度和成功率, 其整体性能优于其他对比算法。

**关键词** 机器视觉; 目标跟踪; 相关滤波; 在线检测; 尺度估计; 模型更新

中图分类号 TP391

文献标识码 A

doi: 10.3788/AOS201838.0215002

## Correlation Filter Tracking Based on Online Detection and Scale-Adaption

Wang Yanchuan, Huang Hai, Li Shaomei, Gao Chao

National Digital Switching System Engineering Technological Research Center, Zhengzhou, Henan 450000, China

**Abstract** In correlation filter tracking, occlusion and object scale change can lead to tracking failure easily. To deal with this problem, a correlation filter tracking algorithm based on online detection and scale-adaption is proposed. The target is initially located through a correlation filter tracker fusing histogram features of oriented gradient, color attribute features and illumination invariant features. The reconstruction residual of local sparse representation model is used for occlusion discrimination. If occlusion occurs, online support vector machine detection will be carried out and target relocating will be realized. Scale estimation from coarse to precise is carried out, and precise scale of target is obtained by scale pre-estimation and Newton iterative method. A balanced model updating strategy is used to update correlation filter regularly and update sparse representation model and support vector machine conservatively. Experimental results show that, compared with existing tracking algorithms, the proposed algorithm can effectively reduce the occlusion, target scale change and other complicated factors, and can gain higher distance precision and success rate on 50 groups of test sequences. The overall performance of the proposed algorithm is better than other contrast algorithms.

**Key words** machine vision; object tracking; correlation filtering; online detection; scale estimation; model update

**OCIS codes** 150.0155; 100.1160; 110.1085

## 1 引 言

目标跟踪是指给出视频中目标的初始状态(如位置和大小等),并在后续序列中对目标的位置状态进行检测标识<sup>[1]</sup>。作为计算机视觉的重要研究课题,目标跟踪广泛应用于智能交通、视频监控和人机交互等领域<sup>[2]</sup>。近年来,研究者在目标跟踪算法的研究方面取得了很大进展,但仍面临着诸多挑战,

例如复杂背景、尺度变化和目标遮挡等。

根据目标的表现建模方式,可将目标跟踪算法模型分为两类<sup>[1]</sup>:产生式模型和判别式模型。产生式模型是对目标进行建模,搜索与之匹配度最高的候选区域进行跟踪;判别式模型基于检测判别机制,通过在线更新分类模型来实现目标的定位跟踪。随着机器学习等领域技术的发展,目前多数学者主要研究判别式跟踪算法<sup>[1]</sup>。其中,相关滤波(CF)跟踪

收稿日期: 2017-08-28; 收到修改稿日期: 2017-09-29

基金项目: 国家自然科学基金(61601513,61379151)、河南省杰出青年基金(144100510001)

作者简介: 王艳川(1987—),男,硕士研究生,主要从事视觉目标跟踪方面的研究。E-mail: 87-chuan@163.com

导师简介: 黄海(1975—),男,博士,研究员,硕士生导师,主要从事计算机视觉方面的研究。E-mail: 13503827672@139.com

模型以其较好的跟踪性能和较高的计算效率,成为当前目标跟踪领域的研究热点之一<sup>[3]</sup>。该模型主要有两个优势:1)通过对样本的循环密集采样,可实现对有限训练数据的扩展;2)采用循环相关和快速傅里叶变换,可有效减小训练和检测的运算量<sup>[3]</sup>。因此,相关滤波跟踪模型具有训练样本稀少且实时性强的特点,更适用于跟踪领域。Bolme等<sup>[4]</sup>提出的最小输出平方误差和相关滤波跟踪(MOSSE)算法将相关滤波思想引入跟踪领域,跟踪速度可达数百帧每秒。Henriques等<sup>[5]</sup>通过训练正则化最小二乘(RLS)分类器进行跟踪判定,之后他们又提取目标的方向梯度直方图(HOG)特征并建立了多通道特征融合模型<sup>[6]</sup>,取得了较好的跟踪效果。Danelljan等<sup>[7]</sup>将颜色属性(CN)信息进行主成分分析(PCA)降维与灰度特征叠加,得到了更加精确、稳健的目标表观模型。Qi等<sup>[8]</sup>充分利用卷积神经网络(CNN)层级所抽取的特征建立多个弱相关滤波跟踪器,并采用自适应对冲思想将多个弱跟踪器整合成强跟踪器。但上述算法<sup>[4-8]</sup>存在两个问题:1)采用固定尺度的滤波模板,不能适应目标尺度变化;2)在目标遮挡后,采用固定的学习机制会导致误差累积,引起跟踪漂移。为解决尺度变化问题,Danelljan等<sup>[9-10]</sup>采用尺度金字塔思想,但该方法存在两个问题:1)该策略以目标原始大小为中心层,向上下层进行权重递减叠加,当目标尺度发生变化时,固定的权重叠加会导

致真实目标尺度响应的减弱;2)各层间具有一定的间隔,由金字塔对应层直接得到目标尺度,会出现估计偏差较大的问题。针对遮挡问题,Ma等<sup>[11]</sup>根据目标回归CF跟踪模型的响应值进行遮挡判断,但该过程易受目标表观变化的干扰。

为增强CF跟踪模型对遮挡及目标尺度变化的稳健性,本文在相关滤波跟踪框架下提出基于在线检测和尺度自适应的相关滤波跟踪算法,主要开展了以下4个方面的工作:1)将HOG特征、CN特征和光照不变特征(IIF)<sup>[12]</sup>进行融合,以增强CF跟踪模型目标表征的精确性;2)提出一种在线检测机制,联合目标的局部稀疏表示模型<sup>[13]</sup>和在线支持向量机(SVM)模型<sup>[12,14]</sup>,实现了目标遮挡的检测与遮挡目标的重定位;3)给出一种“由粗到精”的尺度目标估计方法,解决了尺度金字塔策略中权重配比和尺度估计不精确的问题;4)采用均衡学习策略,固定更新相关滤波系数以适应目标表观变化,根据跟踪结果的可信度保守更新稀疏表示字典和SVM检测器,以保留对目标的长时记忆,从而提高算法的适应性。

## 2 算法描述

所提算法主要分为4个阶段:1)采用相关滤波模型对目标进行跟踪定位;2)针对遮挡问题对目标进行在线检测;3)针对尺度变化问题对目标进行尺度估计;4)对各模块进行更新。算法框架图如图1所示。

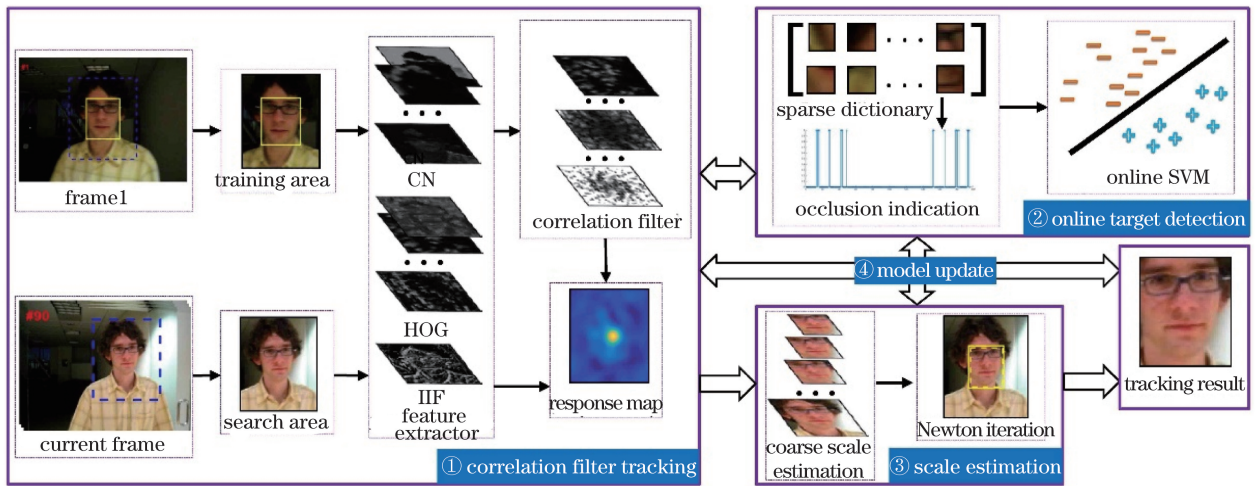


图1 所提算法的框架图

Fig. 1 Framework of proposed algorithm

### 2.1 基于多特征融合的相关滤波跟踪器

在CF跟踪模型中,算法的目标是从训练样本 $\{(\mathbf{x}_i, y_i)\}_i$ 中学习一个多通道的卷积(相关)滤波器。假设第*i*个训练样本为 $\mathbf{x}_i$ ,对应标签为 $y_i$ ,若用函数 $f(\mathbf{x}_i) = \mathbf{w}^T \mathbf{x}_i$ 来表示输入与输出,则分类器

的训练可转化为目标函数的最小化,即 $\min_{\mathbf{w}} \sum_i L[f(\mathbf{w}, \mathbf{x}_i), y_i] + \lambda \|\mathbf{w}\|_2^2$ ,其中 $L(\cdot)$ 为损失函数, $\mathbf{w}$ 为权重向量, $\lambda$ 为正则化参数。 $L(\cdot)$ 的形式为 $L[f(\mathbf{w}, \mathbf{x}_i), y_i] = [y_i - f(\mathbf{w}, \mathbf{x}_i)]^2$ 。

设训练区域  $x$  的尺寸大小为  $[M \ N]$ , 循环样本可表示为  $x_{m,n}$ ,  $(m, n) \in \{0, 1, \dots, M-1\} \times \{0, 1, \dots, N-1\}$ , 训练样本  $x_{m,n}$  对应的标签  $y(m, n) = \exp[-(m-M/2)^2 + (n-N/2)^2 / (2\delta_{\text{label}}^2)]$ , 其中  $\delta_{\text{label}}$  为标准差。通过核函数将样本  $x$  映射至 Hilbert 空间  $\varphi(x)$ , 可实现非线性模型的训练并提升分类性能<sup>[15]</sup>。设核函数  $\mathbf{\kappa}(\mathbf{x}, \mathbf{z}) = \langle \varphi(\mathbf{x}), \varphi(\mathbf{z}) \rangle$ ,  $\mathbf{w}$  可表示为  $\mathbf{w} = \sum_{m,n} a_{m,n} \varphi(x_{m,n})$ , 则  $\mathbf{w}$  可转化为求对偶空间的向量  $\mathbf{a} = (a_{m,n})_{M \times N}$ 。由(1)式推导可得

$$\hat{\mathbf{a}} = (\hat{\kappa}^{xx} + \lambda)^{-1} \hat{\mathbf{y}}, \quad (1)$$

式中  $\hat{\kappa}^{xx}$  表示训练区域  $x$  映射至 Hilbert 空间后的离散傅里叶变换(DFT),  $\hat{\mathbf{y}}$  为对应标签  $\mathbf{y}$  的 DFT。设当前帧目标的检测样本为  $\mathbf{z}$ , 最终得到的响应置信图为

$$f(\mathbf{z}) = \mathcal{F}^{-1}(\hat{\kappa}^{xz} \odot \hat{\mathbf{a}}), \quad (2)$$

式中  $\mathcal{F}^{-1}$  表示傅里叶逆变换,  $\odot$  表示哈达玛积(对应元素相乘)。  $f(\mathbf{z})$  最大值的位置即为目标位置。

较早的相关滤波算法, 如基于检测的循环结构核跟踪算法(CSK)<sup>[5]</sup>, 通常采用灰度特征, 故图像的噪声干扰会对跟踪性能产生较大的影响。核相关滤波的运算基于点乘和范数。根据 DFT 的线性特征, 只需将各特征通道进行求和运算, 对跟踪速度的影响很小<sup>[6]</sup>。假设特征通道数为  $p$ , 描述向量  $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_p]$ ,  $\mathbf{z} = [z_1 \ z_2 \ \dots \ z_p]$ , 由高斯核可得

$$\kappa^{xx} = \exp\left\{-\frac{1}{\delta^2} \left[ \|\mathbf{x}\|_2^2 + \|\mathbf{z}\|_2^2 - 2\mathcal{F}^{-1}\left(\sum_p \hat{\mathbf{x}}^* \odot \hat{\mathbf{z}}\right)\right]\right\}, \quad (3)$$

式中  $*$  表示共轭运算,  $\delta$  为标准差。通过多特征融合, 可获得更为有效的目标特征描述, 以提升跟踪性能<sup>[1]</sup>。基于 HOG 特征计算得到的图像局部区域的梯度强度和方向信息, 反映的是图像各区域块的边缘形状信息; CN 特征将 RGB 空间转换至 CN 空间, 反映的是区域的 11 维主题颜色信息; IIF 将 CIE Lab 空间的亮度通道进行局部非参数变换, 具有对剧烈光照变化的抑制能力。上述三种特征具有对目标描述的互补性, 此处将 HOG 特征、CN 特征和 IIF 融合后进行目标表征, 进而提升算法的跟踪性能。

## 2.2 在线目标检测机制

有效的抗遮挡机制是设计跟踪器时需要考虑的重要因素。对于相关滤波跟踪器, 在目标被遮挡时, 如果持续进行模板更新, 会产生误差累积, 当遮挡逐步消除后, 则很容易误判目标。为此, 提出一种在线

检测机制, 具体方法为: 首先将当前帧相关滤波的跟踪结果输入稀疏模型; 然后根据重构残差来判别是否发生遮挡及遮挡的严重程度, 若遮挡严重, 则根据保守更新的在线 SVM 进行目标再判定; 最后通过计算新判定结果的遮挡指示值进行目标重选定。

稀疏表示跟踪模型是将当前的跟踪结果在基字典子空间进行映射, 得到目标的稀疏表示。局部的稀疏表示跟踪模型能更好地捕获目标的表现变化, 在遮挡序列集上性能更佳<sup>[16]</sup>。采用目标的原始像素进行局部表示, 双线性插值后变为固定目标模板, 再在目标模板上进行重叠掩模得到  $m$  个子块集, 对各子块进行一维向量变换并归一化得到  $\{\mathbf{y}_{\text{sparse},i} \mid i=1, 2, \dots, m\} \in \mathbb{R}^d$ ,  $d$  为子块的维数。稀疏字典  $\mathbf{D} \in \mathbb{R}^{d \times k}$ 。采用  $k$ -means 聚类算法对首帧的子块集进行聚类, 得到  $k$  个聚类中心。当前跟踪结果的子块  $\mathbf{y}_{\text{sparse},i}$  可用字典线性表示为

$$\min_{\mathbf{b}_i} \|\mathbf{y}_{\text{sparse},i} - \mathbf{D}\mathbf{b}_i\|_2^2 + \lambda_{\text{sparse}} \|\mathbf{b}_i\|_1, \text{ s.t. } \mathbf{b}_i \geq 0, \quad (4)$$

式中  $\lambda_{\text{sparse}}$  为调整重构残差和稀疏约束的参数。根据得到的稀疏表示向量  $\mathbf{b}_i$  可计算得到子块在字典集上的重构残差  $\epsilon_i = \|\mathbf{y}_{\text{sparse},i} - \mathbf{D}\mathbf{b}_i\|_2^2$ , 它反映了字典集对当前跟踪结果各子块的表示程度。显然, 重构残差越大, 当前目标子块被遮挡的可能性就越大, 反之越小。设  $\mathbf{o} = (o_1, o_2, \dots, o_m)$  表示遮挡指示向量, 其中的元素可表示为

$$o_i = \begin{cases} 1, & \epsilon_i < \epsilon_0 \\ 0, & \text{other} \end{cases}, \quad (5)$$

式中  $o_i$  表示对应的子块  $\mathbf{y}_{\text{sparse},i}$  是否被遮挡,  $\epsilon_0$  表示子块遮挡判别阈值。通常情况下, 当目标发生小部分的遮挡时, 不会对相关滤波跟踪器的判定结果造成太大影响, 但如果出现大面积(指遮挡面积超过 2/3)或者长时遮挡, 则跟踪器很容易丢失目标。根据遮挡子块的数量来进行遮挡程度的判定, 假设  $O = \sum_i o_i$  表示目标遮挡值,  $O_0$  表示预定的目标遮挡阈值。当  $O \geq O_0$  时, 说明目标发生严重遮挡, 当前相关滤波跟踪结果可信度较低, 需进行目标检测, 反之则无需进行目标检测。

通过训练保守的在线 SVM 模型来实现遮挡目标的再检测。在目标位置附近以目标真实大小的窗口滑动采集样本, 提取阶梯采样特征<sup>[12]</sup>后进行在线 SVM 训练。假设给定的训练数据集  $G = \{(\mathbf{x}_{\text{SVM},j}, \mathbf{y}_{\text{SVM},j}), j=1, 2, \dots, r\}$ ,  $r$  为数据集的大小,  $\mathbf{x}_{\text{SVM},j}$  为训练样本,  $\mathbf{y}_{\text{SVM},j}$  为样本标签。若样本边



框  $\ell_s$  与目标真实边框  $\ell_g$  的重叠率<sup>[16]</sup>  $S = |\ell_s \cap \ell_g| / |\ell_s \cup \ell_g|$ , 设定当  $S > 0.9$  时为正样本,  $S < 0.5$  时为负样本。训练目标函数为

$$\min_{\mathbf{w}_{SVM}, b_{SVM}} \frac{1}{2} \|\mathbf{w}_{SVM}\|_2^2 + C_{SVM} \sum_{i=1}^r L_h[y_{SVM,i}, f(\mathbf{x}_{SVM,i})], \quad (6)$$

式中  $C_{SVM}$  为惩罚参数,  $L_h(\cdot)$  为合页损失,  $\mathbf{w}_{SVM}$  为权重向量。当判定目标出现严重遮挡时, 在上一帧的跟踪位置采集样本, 再将采集到的样本输入 SVM 分类器, 根据输出的最大值来确定新定位的目标。目标重定位的策略如下: 将 SVM 新检测的目标输出至稀疏表示模型进行遮挡检测, 假设相关滤波跟踪结果的目标遮挡值为  $O_{cf}$ , SVM 重检测的目标遮挡值为  $O_{SVM}$ 。若  $O_{SVM} < 0.5O_{cf}$ , 则采用 SVM 重检测的目标位置, 否则保持跟踪结果不变。

### 2.3 由粗至精的目标尺度估计

在跟踪中, 保持跟踪标识框大小不变, 则目标尺度变小会引入背景干扰, 目标尺度变大会导致目标信息缺失。因此, 实现目标尺度自适应是进行稳定跟踪的重要条件。基于尺度金字塔策略<sup>[9]</sup>提出了一种由粗至精的尺度估计方法, 主要改进有: 1) 调整尺度窗函数, 增加了尺度预估计的准确性; 2) 在置信图频域引入牛顿迭代法, 提高了尺度估计的精细程度。

以原始目标为中心, 采集一系列尺度变化的子图, 通过双线性插值将其变为相同尺度, 提取 HOG 特征叠加窗函数后进行尺度相关滤波器学习。假设上一帧的目标大小  $T_{g,sz} = [T_{g,w} \ T_{g,h}]$ , 其中  $T_{g,w}$  为目标的宽度,  $T_{g,h}$  为目标的高度; 尺度金字塔层数为  $Q$ , 目标候选域尺度参数为  $\{\tau^i * T_{g,sz} \mid i = -(Q-1)/2, \dots, (Q-1)/2\}$ ,  $\tau$  为尺度变化因子, 对每个候选域进行尺度变换并提取 HOG 特征后进行一维向量化, 可得  $\mathbf{X}_{sf} = [\mathbf{X}_{sf,1} \ \mathbf{X}_{sf,2} \ \dots \ \mathbf{X}_{sf,Q}]$ ,

与汉宁窗相乘得到训练样本  $\mathbf{X}_{st} = [\mathbf{X}_{st,1} \ \mathbf{X}_{st,2} \ \dots \ \mathbf{X}_{st,Q}]$ , 则

$$\mathbf{X}_{st,i} = \mathbf{X}_{sf,i} * \frac{1}{2} \{1 - \cos[2\pi(i-1)/(Q-1)]\}, \quad (7)$$

式中  $i = 1, 2, \dots, Q$ , 再利用(1)式得到尺度滤波系数  $\hat{\mathbf{a}}_s$ 。检测时, 在新的一帧抽取检测子图及其特征向量  $\mathbf{Z}_{sf} = [\mathbf{Z}_{sf,1} \ \mathbf{Z}_{sf,2} \ \dots \ \mathbf{Z}_{sf,Q}]$ 。如果目标尺度发生变化, 直接对  $\mathbf{Z}_{sf}$  叠加汉宁窗会削弱真实目标尺度的响应, 影响尺度置信图的准确性。给出了一种窗函数调整方法, 以提高真实目标尺度的置信度。将目标进行分块处理, 若  $T_{g,h} > T_{g,w}$ , 则将目标均分为上下两个子图, 反之将目标分为左右两个子图, 而后在各个子图上分别训练两个独立的相关滤波模型。以目标上下划分为例, 设在第  $t$  帧上下子图的定位坐标分别为  $P_{1,t} : (x_{1,t}, y_{1,t})$  和  $P_{2,t} : (x_{2,t}, y_{2,t})$ , 定义  $\varphi_{s,t}$  为窗函数偏移因子, 可表示为

$$\varphi_{s,t} = \begin{cases} +1, & y_{1,t} < y_{1,t-1} \ \& \ y_{2,t} > y_{2,t-1} \\ -1, & y_{1,t} > y_{1,t-1} \ \& \ y_{2,t} < y_{2,t-1} \\ 0, & \text{other} \end{cases} \quad (8)$$

式中 +1 表示尺度变化呈变大趋势, -1 表示尺度变化呈变小趋势, 0 表示尺度不变。由此得到检测样本

$$\mathbf{Z}_{st,i} = \frac{1}{2} \mathbf{Z}_{sf,i} \left[ 1 - \cos\left(2\pi \frac{i-1-\varphi_{s,t}}{Q-1}\right) \right]. \quad (9)$$

将(9)式得到的检测样本代入(2)式可得尺度响应置信图  $f_s(\mathbf{Z}_{st})$ , 尺度预估流程如图 2 所示。利用(9)式可增强真实目标尺度样本的窗口叠加权重, 提高真实目标的响应值。金字塔层数之间具有一定的间隔,  $f_s(\mathbf{Z}_{st})$  为离散值, 因此得到的目标尺度参数较为粗略。下面给出了一种基于牛顿迭代法的尺度精细化估计策略。

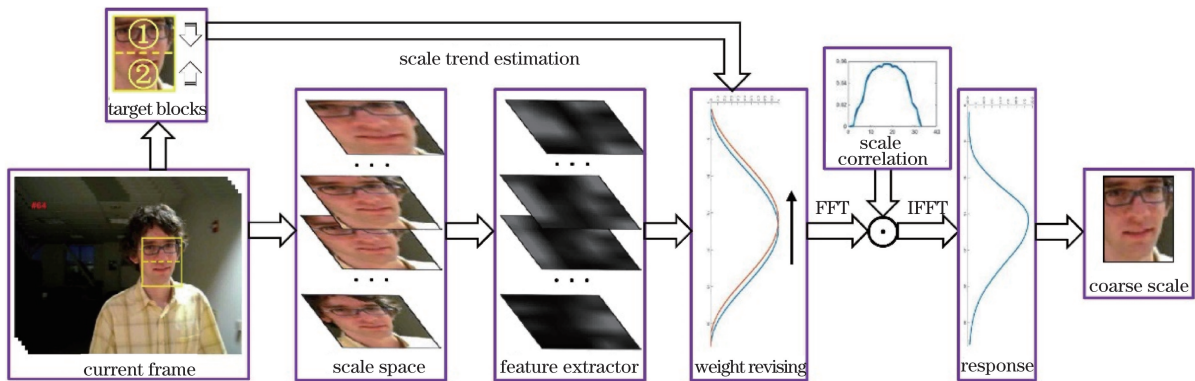


图 2 尺度预估流程图

Fig. 2 Flow chart of scale pre-estimation



若  $\hat{f}_s(\mathbf{Z}_{st}) = [\hat{f}_s(0) \ \hat{f}_s(1) \ \dots \ \hat{f}_s(Q-1)]$ , 对目标候选域进行连续尺度采样,  $\nu \in [0, Q]$  为连续尺度参数,  $f(\nu)$  为连续响应置信图, 则  $f(\nu)$  可通过对  $\hat{f}_s(m)$  进行插值运算得到, 即

$$f(\nu) = \frac{1}{Q} \sum_{m=0}^{Q-1} \hat{f}_s(m) \exp\left(\frac{i2\pi\nu m}{Q}\right). \quad (10)$$

尺度参数  $\nu_{\text{opt}}$  可表示为

$$\nu_{\text{opt}} = \arg \max_{\nu \in [0, Q]} f(\nu). \quad (11)$$

由此可得到当前目标的精确尺度为  $T_{g\_sz} \tau^{\nu_{\text{opt}} - (Q-1)/2}$ , 对(11)式的求解可借助于牛顿迭代法。设由  $f_s(\mathbf{Z}_{st})$  得到的尺度位置  $\nu_0 \in \{0, 1, \dots, Q-1\}$  为迭代起始值, 实验发现仅需数次迭代就可达到收敛状态, 这是由于离散  $f_s(\mathbf{Z}_{st})$  与连续  $f(\nu)$  尺度参数最大值位置距离比较近。

#### 2.4 基于均衡策略的模型更新

在序列中, 视频场景及目标自身是复杂且不断变化的, 如果跟踪器在序列第一帧建立目标的表现模型后不再改变, 显然不能适应目标的表现变化, 最终会导致跟踪失败, 因此用最新的目标表现数据对模型进行更新是很有必要的。如果引入的更新样本有误差, 会使更新后的模型存在偏差, 引起跟踪漂移, 同样会导致跟踪失败。鉴于上述原因, 采用一种均衡模型更新策略: 对相关滤波模型采用持续更新方式, 以适应目标表现变化; 对稀疏表示模型和 SVM 检测器采用保守更新方式, 以保持目标的长时记忆。

符号  $\hat{\mathbf{a}}_t, \hat{\mathbf{x}}_t$  分别表示第  $t$  帧更新后的滤波模板系数和目标模板系数,  $\hat{\mathbf{x}}$  表示当前帧目标的 DFT,  $\eta$  为学习速率, 则相关滤波跟踪模型更新机制可表示为

$$\begin{cases} \hat{\mathbf{a}}_t = \eta (\hat{\mathbf{c}}_t^{\text{rx}} + \lambda)^{-1} \hat{\mathbf{y}} + (1 - \eta) \hat{\mathbf{a}}_{t-1} \\ \hat{\mathbf{x}}_t = \eta \hat{\mathbf{x}} + (1 - \eta) \hat{\mathbf{x}}_{t-1} \end{cases}. \quad (12)$$

尺度金字塔滤波器和目标分块滤波器均采用(12)式的更新机制。由第 2.2 节的分析可知, 目标遮挡值  $O$  代表跟踪结果的可信度, 因此可将  $O$  作为判别进行更新设定。当跟踪结果可信度较高时, 对稀疏表示模型和 SVM 分类器进行更新, 否则保持不变, 如此便确保了目标检测机制的稳定性和记忆性。符号  $\mathbf{Y}_{\text{sparse}, t} \in \mathbb{R}^{d \times m}$  表示第  $t$  帧的目标子块向量集,  $\mathbf{D}_{t-1} \in \mathbb{R}^{d \times k}$  表示第  $t-1$  帧的稀疏字典。如果  $O < O_{\text{update}}$ , 将  $\mathbf{Y}_{\text{sparse}, t}$  与  $\mathbf{D}_{t-1}$  采用  $k$ -means 算法得到更新的字典  $\mathbf{D}_t$ , 否则不予更新。在线 SVM 分类器的更新方法为: 设第  $t$  帧时 SVM 分类器保存的先前帧的训练集  $G_{1,t-1} = \{\zeta_i = (\mathbf{u}_i, \gamma_i, s_i) \mid i = 1, 2, \dots, B\}$ ,  $\mathbf{u}_i, \gamma_i, s_i$  分别代表样本实例、标签和该实

例所代表的支持向量个数,  $G_t = \{(\mathbf{x}_i, y_i) \mid i = 1, 2, \dots, J\}$  表示当前帧提取的数据实例。实现 SVM 在线学习的目标函数可表示为

$$\min_{\mathbf{w}_{\text{SVM}}, b_{\text{SVM}}} \frac{1}{2} \|\mathbf{w}_{\text{SVM}}\|_2^2 + C_{\text{SVM}} \left[ \sum_{i=1}^B \frac{s_i}{N_{\gamma_i}} L_h(\gamma_i, \mathbf{u}_i; \mathbf{w}_{\text{SVM}}) + \sum_{i=1}^J \frac{1}{N_{y_i}} L_h(y_i, \mathbf{x}_i; \mathbf{w}_{\text{SVM}}) \right], \quad (13)$$

式中  $N_+ = \sum_{\gamma_i = +} s_i + \sum_{y_i = +} 1$ ,  $N_- = \sum_{\gamma_i = -} s_i + \sum_{y_i = -} 1$ 。由(13)式可从  $G_t$  得到新的支持向量。设定训练集  $G_{1,t-1}$  的最大容量为  $U$  (正实例容量  $U_+$  和负实例容量  $U_-$ ), 如果新的支持向量与  $G_{1,t-1}$  中的支持向量数量之和不超过最大容量, 则直接加入更新, 否则进行更新:

$$\mathbf{u}' = (s_{i_1} \mathbf{u}_{i_1} + s_{i_2} \mathbf{u}_{i_2}) / (s_{i_1} + s_{i_2}), \quad (14)$$

式中  $s_{i_1}, s_{i_2}$  分别为与训练集某支持向量距离最近的新支持向量个数和训练集该支持向量的个数,  $\mathbf{u}_{i_1}, \mathbf{u}_{i_2}$  分别为与训练集某支持向量距离最近的新支持向量的样本实例和训练集该支持向量的样本实例。通过(14)式可实现距离最小的支持向量的融合, 得到新的训练集  $G_{1,t} = \{\zeta'_i = (\mathbf{u}'_i, \gamma'_i, s'_i) \mid i = 1, 2, \dots, B'\}$ 。如果  $O > O_{\text{update}}$ , 则 SVM 分类器保持原参数不变, 不进行更新。

#### 2.5 算法实现

首先进行各模块初始化, 然后进行位置检测和尺度估计, 最后对各模块进行更新。具体步骤如下:

- 1) 输入。视频序列和首帧的目标位置  $P_0$  和大小  $T_{g\_sz_0}$ 。
- 2) 输出。序列中各帧的目标位置  $P_t$  和大小  $T_{g\_sz_t}$ 。
- 3) 初始化阶段。在训练区域输入相关滤波跟踪模型, 由(1)式得到滤波器系数  $\hat{\mathbf{a}}_1$  和目标模板  $\hat{\mathbf{x}}_1$ ; 将目标输入稀疏表示模型并采用  $k$ -means 聚类算法得到字典  $\mathbf{D}_1$ , 由(5)式得到  $G_1$ ; 分别提取分块及尺度相关滤波器的训练样本, 由(1)式得到各自的滤波器系数和模板。

#### 4) 跟踪阶段

for  $t = 2; F_{\text{frame\_num}}$  ( $F_{\text{frame\_num}}$  为帧数)

// 位置检测

在跟踪位置  $P_{t-1}$  处抽取检测样本  $\mathbf{z}_t$ , 提取特征后由(2)式得到相关滤波跟踪位置  $P_t = P_{\text{cf}, t}$ ;

利用  $P_{\text{cf}, t}$ , 通过(4)式、重构残差和(5)式得到遮挡值  $O_{\text{cf}, t}$ ;

if  $O_{\text{cf}, t} \geq O_0$

根据  $P_{t-1}$  和参数  $\mathbf{w}_{\text{SVM}, t-1}, b_{\text{SVM}, t-1}$  得到 SVM 响应最大值的位置  $P_{\text{SVM}, t}$ ;

根据  $P_{SVM,t}$ , 通过(4)式、重构残差和(5)式得到遮挡值  $O_{SVM,t}$ ;

if  $O_{SVM,t} < 0.5O_{cl,t}$ ,  $P_t = P_{SVM,t}$ ;

end

end

// 尺度估计

根据  $P_t$  和  $T_{g_{sz_{t-1}}}$  抽取样本  $Z_{sf,t}$ , 由分块相关滤波器和(8)式得到  $\varphi_{s,t}$ ;

根据  $Z_{sf,t}$ 、 $\varphi_{s,t}$  和(9)、(2)式, 计算尺度响应图  $f_s(Z_{st,t})$  和目标尺度位置  $\nu_{0,t}$ ;

根据(10)式和  $\nu_{0,t}$  并采用牛顿迭代法可得  $\nu_{opt,t}$ , 从而得到目标精确尺度  $T_{g_{sz_t}}$ ;

// 模型更新

根据(12)式对各相关滤波模块的系数和模板进行更新;

根据  $P_t$  和  $T_{g_{sz_t}}$  计算得到目标遮挡值  $O_t$ ;

if  $O_t > O_{update}$

提取  $Y_{sparse,t}$  后更新  $D_{t-1}$ , 根据(13)、(14)式更新  $G_{1,t-1}$ ;

else

保持原模型参数不变;

end

end

### 3 实验结果及讨论

设计了两个针对性实验和一个综合性实验对算法的有效性进行评估: 1) 对算法的抗遮挡性能进行对比测试; 2) 对算法的尺度适应性进行对比测试; 3) 与现阶段性能较好的跟踪算法进行综合实验对比。实验所用对比算法均采用作者公开的源码设置, 测试序列均来自标准数据集 Benchmark<sup>[16]</sup>, 评价标准采用文献[16]所提的指标: 中心位置误差(CLE)、距离精度(DP)和成功率(SR)。CLE指跟踪结果位置与目标基准位置之间的欧氏距离, 单位为 pixel; DP是 CLE 小于某个阈值(通常为 20 pixel)<sup>[16]</sup>的帧数占序列总帧数的百分比; 当重叠率大于给定阈值(通常为 50%)时, 认为该帧跟踪成功; SR表示跟踪成功帧数与总帧数的比率。

#### 3.1 实验环境及参数

实验环境为: CPU 型号为 3.60 GHz Intel(R) Core(TM) i7, 内存为 16.0 GB, 操作系统为 Windows 7 64 bit, 软件平台为 Matlab R2014b。基于实验测试分析, 并参考文献[9, 12-13], 将具体参数设置如下: 相关滤波跟踪器中, 训练检测区域为目标大小的

2.5 倍, HOG 特征的元胞大小为 4 pixel × 4 pixel, 统计梯度方向数为 9, 正则化参数  $\lambda = 10^{-4}$ ,  $\delta = 0.5$ ,  $\delta_{label} = 0.1$ , 学习速率  $\eta = 0.025$ ; 尺度估计模型中, 尺度层数  $Q = 33$ , 尺度变化因子  $\tau = 1.02$ , 牛顿法迭代次数  $\Gamma = 5$ ; 局部稀疏表示模型中, 目标模板大小为 32 pixel × 32 pixel,  $\lambda_{sparse} = 0.01$ ,  $d = 36$ ,  $k = 50$ ; 子块遮挡判别阈值  $\epsilon_0 = 0.04$ ; 在线 SVM 模型中,  $C_{SVM} = 100$ , 最大容量  $U = 80$ ,  $U_+ = 10$ ,  $U_- = 70$ ,  $O_{update} = 5$ 。

目标遮挡阈值是一个重要参数, 关系到在线检测机制的目标遮挡判别, 对算法的跟踪位置精确度和运算速度影响较大。选用 20 组具有遮挡属性的标准序列 (basketball, carscale, coke, david, david3, dudek, doll, faceoccl, faceocc2, football, freeman4, jogging-1, jogging-2, skating1, lemming, tiger2, walking2, woman, walking, subway) 进行测试, 得到算法在不同目标遮挡阈值下的平均中心位置误差和平均运算速度, 结果如图 3 所示。其中平均运算速度反映的是跟踪算法的效率, 单位为 frame/s。图 3 表明当目标遮挡阈值小于 20 时, 目标再检测的次数

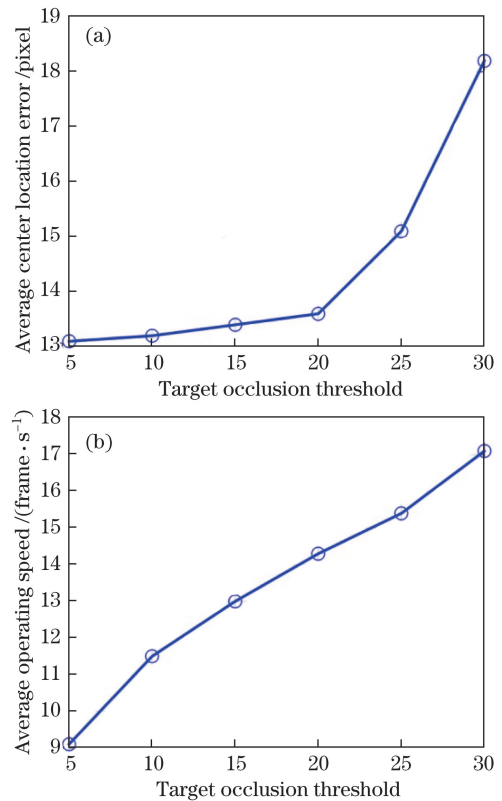


图 3 目标遮挡阈值对跟踪性能的影响。

(a) 平均中心位置误差; (b) 平均运算速度

Fig. 3 Influence of target occlusion threshold on tracking performance. (a) Average center location error; (b) average operating speed

随着该值的减小而增加,导致算法的运算速度大幅降低,而跟踪位置误差减小并不明显,这是因为CF跟踪器自身能够适应小部分遮挡等微弱表观变化。因此不能通过增加检测次数而无限制地减小跟踪位置误差;当目标遮挡阈值大于20时,目标再检测的次数随着该值的增加而减小,虽然跟踪速度有所提升,但在线检测机制不能有效地进行遮挡判别,算法的抗遮挡能力降低,跟踪位置误差明显增大。综合以上分析,并根据图3所示的实验结果,将目标遮挡

阈值设置为20。

### 3.2 遮挡检测的性能测试

为验证所提算法的性能,选择3种具有代表性的抗遮挡算法进行对比,包括基于相关滤波的遮挡算法LCT<sup>[11]</sup>、自适应结构稀疏表示模型ALSA<sup>[17]</sup>和核化结构SVM算法Struck<sup>[18]</sup>,并选取了4组典型的遮挡序列tiger2、lemming、walking2和faceocc2进行测试。图4和图5分别为4种算法针对4组测试序列的跟踪结果图及中心位置误差图。

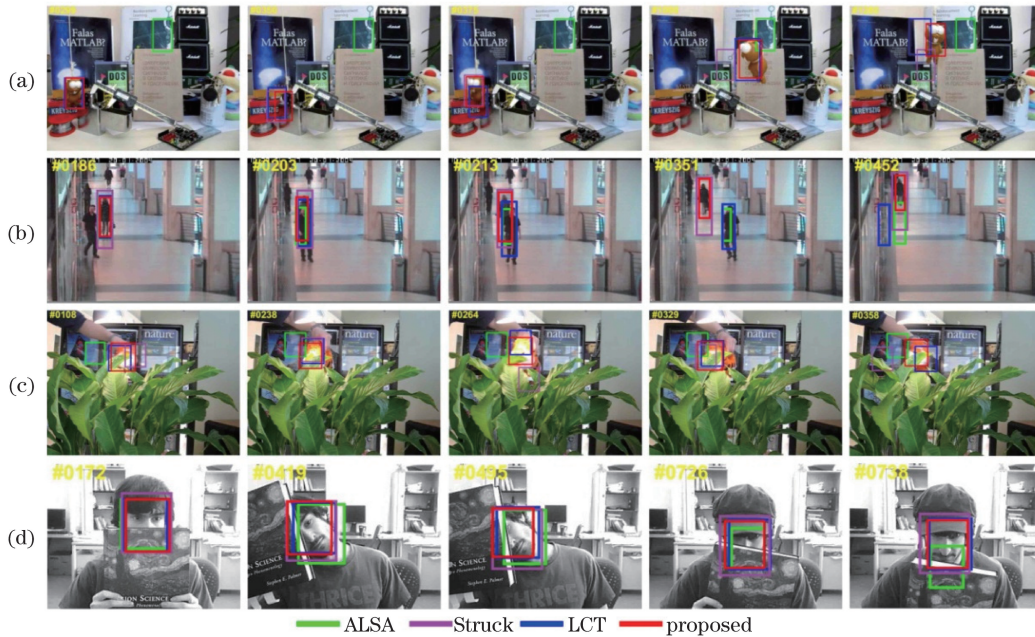


图4 4种算法对4组遮挡序列的跟踪结果。(a) Lemming; (b) walking2; (c) tiger2; (d) faceocc2

Fig. 4 Tracking results of four occluded sequences with four algorithms.

(a) Lemming; (b) walking2; (c) tiger2; (d) faceocc2

图4(a)和图5(a)所示序列Lemming主要存在长时的严重遮挡问题。其中,目标在299~350帧逐渐受到大面积的长时遮挡,并在350~375帧缓慢复出。从图5(a)可以看到,ALSA已经丢失目标,LCT出现较大跟踪漂移,所提算法采用的局部稀疏表示模型能更好地进行遮挡判别,并通过目标长时记忆的在线SVM在第368帧实现了有效的目标检测跟踪。在939~1008帧,目标出现尺度变化并伴有旋转、正面遮挡情况,Struck跟踪失败,LCT定位误差较大并在1280帧丢失目标。在面对长时严重遮挡问题时,所提算法的遮挡检测机制能实现更稳定的持续跟踪。

图4(b)和图5(b)所示序列walking2主要存在短时的相似目标遮挡干扰问题。自186帧起,相似目标逐渐进入目标区域,ALSA在203帧开始定位至错误目标上,LCT则在213帧丢失目标,而所提

算法不受相似目标遮挡的干扰,原因在于其相关滤波器采用互补性特征融合的方法,确保了目标描述更为准确、有效。从整个序列来看,目标逐渐远离镜头,导致尺度变小,Struck虽然实现了对目标的定位,但误差较大且不能应对目标的尺度变化问题,如第351帧和第452帧。由图5(b)可知,所提算法的定位精度优于其他算法。

图4(c)和图5(c)所示序列tiger2主要存在目标反复遮挡的问题。目标在左右来回移动的过程中外观变化较大,并不断地受到树叶的遮挡干扰,如第108,238,264,329,358帧等。所提算法的均衡更新策略能很好地适应该序列的复杂情况,一方面,相关滤波的固定学习机制增强了对目标外观变化的适应性;另一方面,保守更新的在线检测模块确保了对目标的稳定记忆,可对遮挡目标进行再检测,有效地抑制了遮挡问题的干扰。因此,



在整个跟踪过程中,所提算法始终保持着较小的位置误差,而其他算法均有不同程度丢失目标的情况。

图 4(d)和图 5(d)所示序列 faceocc2 主要存在目标部分遮挡的问题。序列在第 172, 419, 495,

726, 738 帧等出现书本遮挡目标面部的情况,ALSA 自 726 帧逐渐丢失目标。由分析可知,对于部分遮挡问题,LCT、Struck 和所提算法均能较好地完成跟踪,且 Struck 和所提算法能更好地减弱部分遮挡的干扰,跟踪精确度更高。

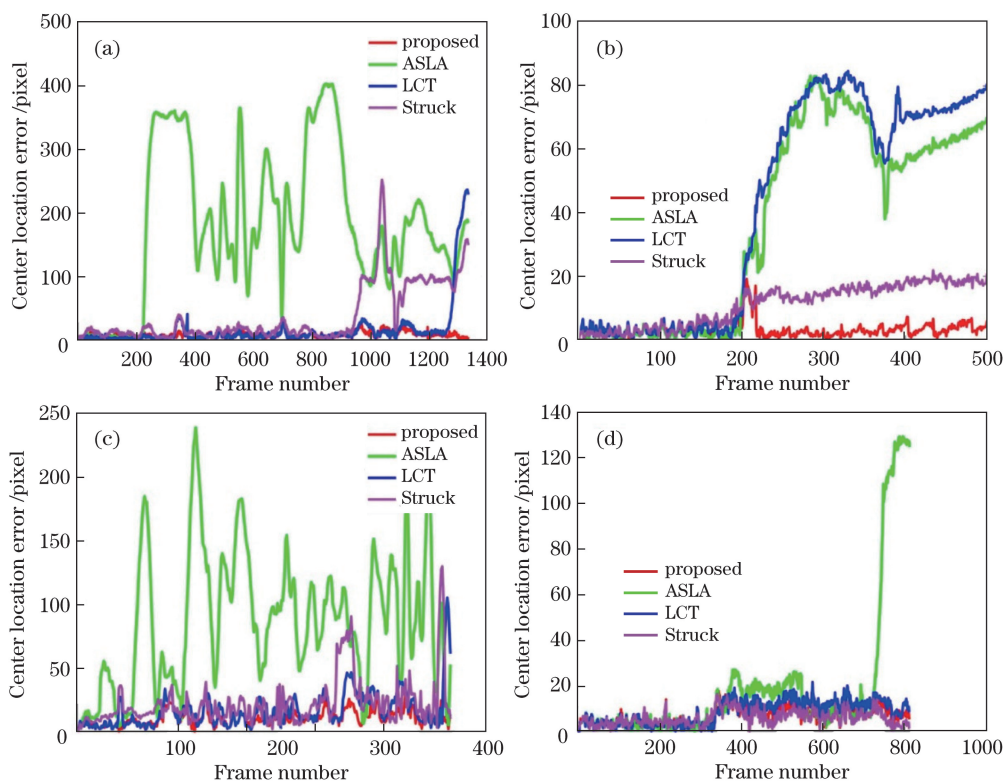


图 5 4 种算法在 4 组遮挡序列的中心位置误差。(a) Lemming; (b) walking2; (c) tiger2; (d) faceocc2

Fig. 5 Center location error of four occluded sequences with four algorithms.

(a) Lemming; (b) walking2; (c) tiger2; (d) faceocc2

### 3.3 尺度自适应性能测试

根据跟踪尺度变化的准确性对所提算法的性能进行验证,针对性地选择 3 种尺度自适应算法进行对比:基于金字塔策略的相关算法 DSST<sup>[9]</sup>、检测跟踪算法 TLD<sup>[19]</sup>和正向稀疏表示模型 L1APG<sup>[20]</sup>,用于测试的 4 组尺度变化序列为 trellis、girl、walking 和 freeman1。跟踪面积<sup>[21]</sup>是指算法跟踪区域的面积,跟踪面积与目标真实面积之间的差值反映了算法的尺度适应性。4 种算法在 4 组尺度变化序列的面积误差结果如图 6 所示,其中横坐标为序列帧数,纵坐标为跟踪面积与真实面积的差值,即面积误差(OAE),该值越接近于 0,说明算法预测的尺度值与真实目标大小越相近,尺度预测的准确度越高,图中的标注数值为面积误差绝对值的平均值,单位为 pixel。

图 6(a)所示序列 trellis 主要存在目标经历光照与尺度变化的问题。L1APG 在跟踪开始后尺度

不断变大,跟踪结果所含背景区域过大,导致定位偏差越来越大并丢失目标;TLD 在第 221 帧以后不能抑制光照及目标尺度变化的双重干扰,预测尺度发生剧烈变化;DSST 预测的目标尺度始终小于目标的真实尺度。由分析可知,相对于其他 3 种算法,所提算法的尺度适应性更好。

图 6(b)所示序列 girl 主要存在目标在旋转中伴有尺度变化的问题。目标在第 81 帧开始旋转,DSST 不能及时地预测目标尺度。随着跟踪的进行,尺度偏差越来越大。在 55~66 帧和 86~111 帧 TLD 丢失目标,在 161~265 帧 TLD 和 L1APG 对目标的尺度预测偏差大,而所提算法则相对较小。

图 6(c)所示序列 walking 主要存在目标远离镜头逐渐变小的问题。该序列的情景较为单一,对算法的尺度适应性检测针对性更强。TLD 不能处理目标的尺度变化,当面积误差逐渐变大时,DSST 与 L1APG 的性能相当。从图中的数值可以看出,所提

算法的面积误差比 DSST 小 39.1 pixel,这是因为所提算法在金字塔的基础上采用了“由粗到精”的策略,尺度估计更加精细。

图 6(d)所示序列 freeman1 主要存在目标尺度反复变大、变小的问题。整个序列目标尺度先变大

后变小,从图中结果来看,所提算法也出现了尺度预测偏差,但明显优于其他算法,且从定位结果可知,L1APG(自第 127 帧起)、DSST(自第 181 帧起)跟踪失败,TLD 在 135~191 帧和第 236 帧至结束丢失目标。

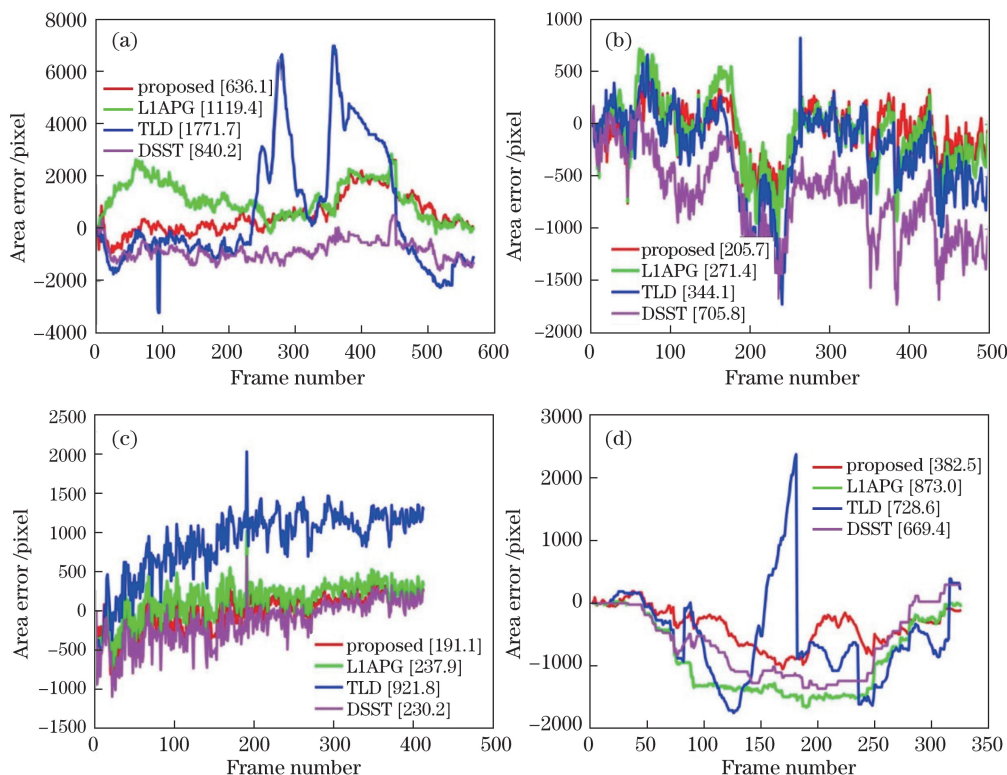


图 6 4 种算法对 4 组尺度变化序列的面积误差结果。(a) Trellis; (b) girl; (c) walking; (d) freeman1

Fig. 6 Area error results of four scale changing sequences with four algorithms.

(a) Trellis; (b) girl; (c) walking; (d) freeman1

### 3.4 与其他稳健跟踪算法的实验比较

为进一步更为全面地评估所提算法的性能,选取了当前主流的算法进行对比,主要包括两类:1)相关滤波算法,KCF<sup>[6]</sup>、SAMF<sup>[10]</sup>、Staple<sup>[22]</sup>和 HDT<sup>[8]</sup>; 2)其他代表性算法,双线性结构 SVM 算法 DLSSVM<sup>[23]</sup>、多专家联合模型 MEEM<sup>[12]</sup>和稀疏协

作模型 SCM<sup>[13]</sup>。为了便于对比,表 1 列出了本文所提算法、KCF、SAMF、Staple 和 HDT 5 种相关滤波跟踪模型的区别。利用标准数据集 Benchmark 提供的 50 个序列进行评测,主要包括复杂背景、尺度变化和目标遮挡等属性,测试方法为对序列首帧进行目标初始化,逐帧检测目标。

表 1 5 种相关滤波跟踪模型比较

Table 1 Comparison of five correlation filter tracking models

Algorithm	Feature	Scale adaption	Occlusion handling
Proposed	HOG,CN,IIF	Coarse-to-precise scale estimation	Online detection module
SAMF	Gray,HOG,CN	Pyramid strategy	No
Staple	Gray,HOG	Pyramid strategy	No
HDT	CNN	No	No
KCF	HOG	No	No

图 7 为 8 种算法在测试集上的整体和主要属性 DP、SR 曲线图,子标题括弧中的数字表示此种属性的视频数量。DP 曲线图的排序标注为 CLE 阈值为

20 pixel 时的距离精度值,SR 曲线图的排序标注为曲线与坐标轴所围成的面积(AUC<sup>[16]</sup>)。从图 7(a)可以看出,在整体 DP 曲线图上,所提算法在 8 种典

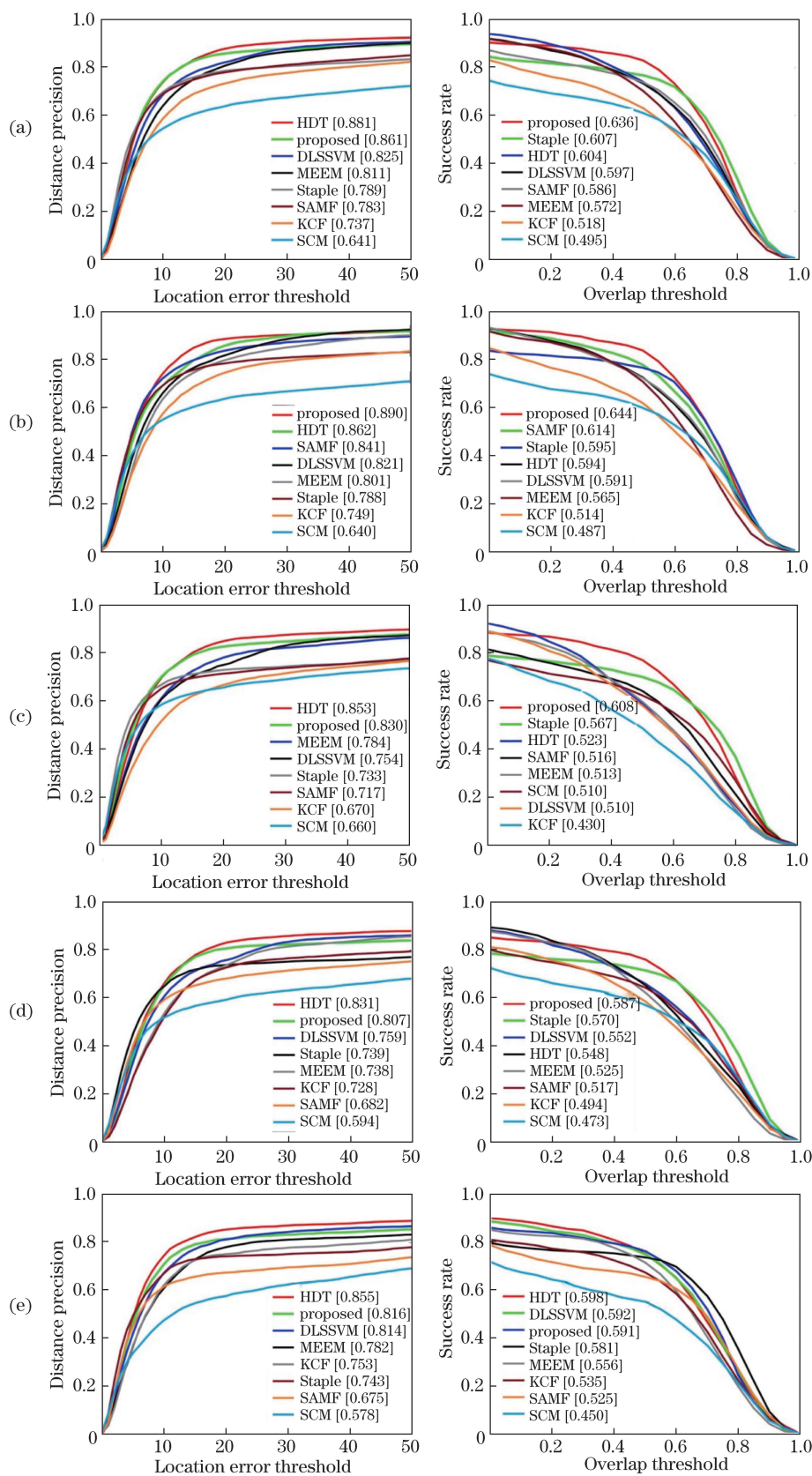


图 7 8 种算法的距离精度和成功率曲线图。(a)整体性能(50);(b)遮挡(29);(c)尺度变化(27);  
(d)光照变化(25);(e)复杂背景(21)

Fig. 7 Curves of distance precision and success rate of eight algorithms. (a) Overall performance (50); (b) occlusion (29);  
(c) scale variation (27); (d) illumination variation (25); (e) complex background (21)



型算法中排名第 2, 仅次于深度学习算法 HDT, 且优于基准相关滤波跟踪 KCF; 相比在整体 DP 图上的排名, HDT、DLSSVM 和 MEEM 在整体 SR 图上的排名均有不同程度的下降, 这是因为这三种算法不能对目标进行尺度估计, 进而降低了算法的覆盖率和成功率, 所提算法、Staple 和 SAMF 则因具备尺度适应模块而在整体 SR 图上的排名有所上升, 分别位列第 1、第 2 和第 5。图 7(b) 中的遮挡子集验证了所提算法的有效性, 可以看出所提算法在距离精度和成功率上均排名第 1, 原因在于所提算法的在线检测模块能更好地保存对目标的记忆性, 在遮挡问题上有强抗干扰力。图 7(c) 为尺度变化子集, 所提算法比 HDT 的 DP 低 2.3%, 比 HDT 的 SR 高 10.1%, 与具备尺度调节功能的 Staple 相比, 所提算法的 DP 和 SR 分别提高了 9.7% 和 4.1%, 因此所提算法的尺度估计策略能更有效地应对目标尺度变化。图 7(d) 为光照变化子集, 与 SAMF 和 Staple 算法相比也有很大幅度的提升, 原因在于所提算法在目标描述上融合了 IIF 通道, 增强了对光照变化的适应能力。相对于其他属性子集, 所提算法在复杂背景子集表现欠佳, 如图 7(e) 所示。这是由于所提算法引入的局部稀疏模块为生成式, 该模块的关键在于对目标进行精确拟合, 如果出现复杂背景(特别是相似背景), 则对判别的准确性影响较大。

标准数据集 Benchmark 中光照变化子集共有

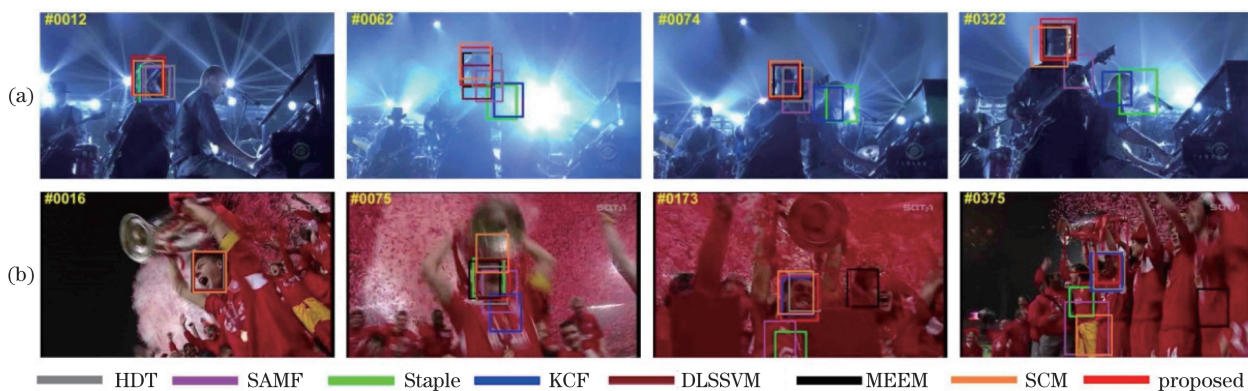


图 8 8 种算法对 2 组典型光照变化序列的跟踪结果。(a) Shaking; (b) soccer

Fig. 8 Tracking results of two typical illumination variant sequences with eight algorithms.

(a) Shaking; (b) soccer

表 2 为 8 种算法针对 50 个测试序列的平均运算速度。从表中可得, 相关滤波跟踪算法因采用的循环相关, 可通过 FFT 来实现高效地检测与训练, 整体上要快于其他算法。HDT 需提取多层 CNN 特征, 本次实验环境均采用 CPU, 运算效率低, 所提

25 组序列, 这里对图 7(d) 进行了整体分析, 为进一步说明所提算法融合光照不变特征方法的有效性, 就算法对光照变化的适应能力进行更为具体的分析。选取两组典型的代表序列 shaking 和 soccer。shaking 序列的目标主要经历背景强光的干扰, soccer 序列的目标主要经历背景阴影的干扰, 图 8 为 8 种算法对两组序列的跟踪结果。图 8(a) 所示为 shaking 序列, 跟踪开始后, HDT、KCF、SAMF 和 Staple 出现了小幅漂移, 如第 12 帧; 第 58 帧后镁光灯亮度开始增强, 目标背景出现剧烈的光照变化, 除了所提算法和 MEEM 外, 其他算法均出现不同程度的跳动, 如第 62 帧; Staple 和 KCF 则在灯光亮度减弱后随灯光定位至镁光灯附近, 跟踪失败, 如第 74 帧; 从整个序列来看, MEEM、DLSSVM、SCM、HDT 和所提算法均完成了跟踪, 但 MEEM 和所提算法的稳定性更好。图 8(b) 所示为 soccer 序列, 在前 31 帧, 目标清晰且亮度较高, 所有算法均能较准确地跟踪, 如第 16 帧; 随后目标在跳跃过程中, 奖杯的上下运动及洒落的烟花对目标跟踪造成了严重的干扰, 如第 75、173 帧等; KCF 在 71~122 帧出现较大偏差, Staple、SAMF、MEEM、DLSSVM 和 SCM 均跟踪失败, 如第 375 帧, HDT 和所提算法实现了对目标的全程跟踪。所提算法能较成功地跟踪上述序列, 原因在于提出的多特征融合方法能减弱光照变化的影响, 实现更准确的目标表征。

算法改进了金字塔的尺度估计策略并增加了目标在线检测模块, 增加了时间开销。需要说明的是, 只有当目标遮挡值超过阈值时才启用 SVM 检测模块, 因此与 SAMF 和 Staple 相比, 跟踪速度有所下降, 但性能有较明显的提升。

表 2 8 种跟踪算法的平均运算速度

Table 2 Average operating speed of eight tracking algorithms

Algorithm	Proposed	HDT	SAMF	Staple	KCF	DLSSVM	MEEM	SCM
Average operating speed / (frame·s <sup>-1</sup> )	15.6	0.8	19.7	18.8	102.1	4.7	6.4	0.4

## 4 结 论

提出了一种基于在线检测和尺度自适应的相关滤波跟踪算法,通过特征融合方法实现了特征间的有效互补,提高了目标表征的准确性。设计了在线检测机制和由粗至精的尺度估计方法,并采用均衡策略进行模型更新,实现了复杂场景下对目标的稳定跟踪。通过对两个针对性实验和一个综合性实验进行验证,结果表明,与其他算法相比,所提算法对光照变化、遮挡和目标尺度变化等场景具有更强的稳健性。

### 参 考 文 献

- [1] Chen Z, Hong Z B, Tao D C. An experimental survey on correlation filter-based tracking[OL]. Computer Science, 2015, 53(6025): 68-83.
- [2] Zhao G P, Shen Y P, Wang J Y. Adaptive feature fusion object tracking based on circulant structure with kernel[J]. Acta Optica Sinica, 2017, 37(8): 0815001.  
赵高鹏, 沈玉鹏, 王建宇. 基于核循环结构的自适应特征融合目标跟踪[J]. 光学学报, 2017, 37(8): 0815001.
- [3] Ma C, Xu Y, Ni B B, *et al.* When correlation filters meet convolutional neural networks for visual tracking[J]. IEEE Signal Processing Letters, 2016, 23(10): 1454-1458.
- [4] Bolme D S, Beveridge J R, Draper B A, *et al.* Visual object tracking using adaptive correlation filters[C]// Proceedings of 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]// Proceedings of 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, Khan F S, Felsberg M, *et al.* Adaptive color attributes for real-time visual tracking[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1090-1097.
- [8] Qi Y, Zhang S, Qin L, *et al.* Hedged deep tracking[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2016: 4303-4311.
- [9] Danelljan M, Hager G, Khan F S, *et al.* Accurate scale estimation for robust visual tracking[C]// Proceedings of British Machine Vision Conference, 2014: 1-11.
- [10] Li Y, Zhu J K. A scale adaptive kernel correlation filter tracker with feature integration[C]// Proceedings of European Conference on Computer Vision, 2014: 254-265.
- [11] Ma C, Yang X K, Zhang C Y, *et al.* Long-term correlation tracking[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015: 5388-5396.
- [12] Zhang J, Ma S, Sclaroff S. MEEM: Robust tracking via multiple experts using entropy minimization[C]// Proceedings of European Conference on Computer Vision, 2014: 188-203.
- [13] Zhong W, Lu H C, Yang M H. Robust object tracking via sparse collaborative appearance model[J]. IEEE Transactions on Image Processing, 2014, 23(5): 2356-2368.
- [14] Wang Z, Vucetic S. Online training on a budget of support vector machines using twin prototypes[J]. Statistical Analysis and Data Mining, 2010, 3(3): 149-169.
- [15] Scholkopf B, Smola A J. Learning with kernels: Support vector machines, regularization, optimization, and beyond[M]. London: MIT Press, 2002: 405-423.
- [16] Wu Y, Lim J, Yang M H. Online object tracking: A benchmark[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2013: 2411-2418.
- [17] Jia X, Lu H C, Yang M H. Visual tracking via adaptive structural local sparse appearance model[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2012: 1822-1829.
- [18] Hare S, Saffari A, Torr P H S. Struck: Structured output tracking with kernels[C]// Proceedings of IEEE International Conference on Computer Vision, 2011: 263-270.
- [19] Kalal Z, Mikolajczyk K, Matas J. Tracking learning detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1409-1422.
- [20] Bao C L, Wu Y, Ling H B, *et al.* Real time robust

- L1 tracker using accelerated proximal gradient approach[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2012: 1830-1837.
- [21] Luo H L, Du F F, Kong F S. Pixel feature-weighted scale-adaptive object tracking algorithm[J]. Journal on Communications, 2015, 36(10): 200-211.  
罗会兰, 杜芳芳, 孔繁胜. 像素点特征加权的尺度自适应跟踪算法[J]. 通信学报, 2015, 36(10): 200-211.
- [22] 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, 2016: 1401-1409.
- [23] Ning J F, Yang J M, Jiang S J, *et al.* Object tracking via dual linear structured SVM and explicit feature map[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2016: 4266-4274.