

# 高置信度的尺度自适应核相关跟踪方法

李福进, 刘慧慧\*, 任红格, 史涛

华北理工大学电气工程学院, 河北 唐山 063210

**摘要** 相关滤波器由于近年来在实时准确地跟踪视频信号方面展示出了优异的性能, 为此引起了很多学者的关注。但在应对复杂场景外观变化较大的情况下, 滤波器的跟踪效果易不稳定, 估计尺度变化能力也有待改进, 模型更新过程易引入不可信样本, 不利于精确跟踪目标。为此, 采用颜色直方图模型和相关滤波判别模型相结合的方式对跟踪目标进行精确定位, 再从预测位置处提取尺度样本训练核尺度相关器。针对模型更新过程中可能引入干扰信息的问题, 选择当响应置信度达到一定阈值时, 对模型进行更新。在跟踪数据集(OTB-2015)中所提方法表现良好, 成功率为 0.694, 精确度为 0.794。

**关键词** 机器视觉; 目标跟踪; 相关滤波; 尺度估计

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.0815004

## Scale Adaptive Kernel Correlation Tracking Method with High Confidence

Li Fujin, Liu Huihui\*, Ren Hongge, Shi Tao

College of Electrical Engineering, North China University of Science and Technology, Tangshan, Hebei 063210, China

**Abstract** Correlation filters have demonstrated excellent performance in real-time and accurate tracking of video signals in recent years, which have attracted the attention of many scholars. However, when the appearance of complex scenes changes greatly, the tracking effect of the filter is easy to be unstable, and the ability to estimate scale changes needs improvement. The model update process easily introduces untrustworthy samples, which is not conducive to track the target accurately. Therefore, a combination of a color histogram model and correlation filtering-discrimination model is used to track objects accurately, and then scale samples are extracted from the predicted position to train the kernel scale correlator. Considering the problem of interference information getting introduced in the model update process, the model is updated when the response confidence attains a certain threshold. The proposed algorithm in the tracking data set (OTB-2015) performed well with a success rate and a precision score of 0.694 and 0.794, respectively.

**Key words** machine vision; target tracking; correlation filter; scale estimation

**OCIS codes** 150.1135; 100.4999; 100.3008

## 1 引言

智能视频监控技术在提高风险预警能力、减轻人力负担、保障公共安全以及便捷交通管理和客户服务等方面具有重要的应用价值<sup>[1]</sup>。目标跟踪作为提高视频监控智能化水平的关键研究方法之一, 近

年来取得了很大进展<sup>[2-4]</sup>, 但是在尺度变换、快速运动和背景杂乱等复杂的视频场景中, 跟踪效果不理想, 目标容易发生漂移。

近年来, 目标跟踪数据集和基准<sup>[5-6]</sup>的提出为评价跟踪性能提供了一种比较统一的标准。相关滤波(CF)跟踪方法利用循环矩阵在频域中的对角性质,

收稿日期: 2020-07-30; 修回日期: 2020-09-06; 录用日期: 2020-09-17

基金项目: 河北省自然科学基金(F2018209289)

\* E-mail: 719614681@qq.com

简化了滤波器的训练过程,使其在数据集上展示出了明显的速度优势。引入核技巧和梯度方向直方图(HOG)的核相关滤波器(KCF)<sup>[7]</sup>在保持高帧率的同时,跟踪精度得到进一步提升,其中将采用线性核的方法记为对偶相关滤波器(DCF)。之后,学者们从提高跟踪目标的尺度自适应<sup>[8-9]</sup>、抗遮挡能力<sup>[10]</sup>和样本质量<sup>[11-12]</sup>,以及改进样本标签<sup>[13]</sup>和特征表征力<sup>[14-15]</sup>等角度提出了各种提高跟踪性能的方式。

Li 等<sup>[16]</sup>提出了尺度自适应的多特征融合跟踪(SAMF)方案,该方案使用了结合颜色名(CN)和 HOG 特征的 CF 跟踪器对目标进行多尺度检测,可以有效提高了跟踪精度。SAMF 方案在许多算法中均有应用<sup>[17-18]</sup>。Possegger 等<sup>[19]</sup>结合颜色直方图提出了干扰感知(DAT)模型,可以有效抑制了相似对象的干扰。Wang 等<sup>[20]</sup>提出了基于循环特征的大边缘目标跟踪(LMCF)算法,该算法采取了多峰检测的方式并使用平均峰值相关能量(APCE)自适应地更新模板,从而提高定位精度。由于核互相关器(KCC)不受核函数类型的限制,所以可以将 CF 只能预测平移变换扩展到能够预测旋转和尺度等任意仿射变换<sup>[21]</sup>。Ma 等<sup>[22]</sup>使用了两种判别回归模型来预测位置,当处于低置信度状态时对目标进行检测,可以显著提高跟踪精度。也有研究人员将 CF 跟踪算法与特征建模能力强的深度学习方法结合,通过提取深度特征<sup>[23]</sup>或使用神经网络模拟 CF 实现端到端的训练方式<sup>[24-25]</sup>,可以明显提高了跟踪精度,但其运行速度较采用手工特征的 CF 方法仍有待提高。

尽管 CF 跟踪器已经取得较好的发展,但其在复杂的视频场景中仍然存在图像特征不足以准确反映目标外观变化、尺度变化需通过平移滤波器来得到,以及缺少对响应置信度进行判断等问题,致使模型中易引入不可信样本,不利于精确地跟踪

目标。

针对上述问题,本文采用颜色直方图对象-背景模型与核尺度相关器判别模型结合的方法,提高 CF 算法应对形变的能力,从而提高跟踪的稳定性,实现对跟踪目标的精确定位。接着从最可信的图像帧中训练核尺度相关器,用于对尺度进行预测。针对模型更新过程中可能引入噪声而造成模型损坏的问题,本文预定义阈值,当置信度满足一定阈值时,对模型进行更新。在 100 个序列中的跟踪结果表明,该方法在精度方面较 KCC 方法和 SAMF 方法分别高出 0.039 和 0.043,成功率分别提高 0.129 和 0.020。该方法在 CPU 上可以实现实时跟踪,在尺度变化、形变、光照、旋转、视野外和相似背景干扰等复杂的视频场景中能够有效跟踪目标,有利于智能视频监控技术的发展。

## 2 方法设计

使用颜色直方图来表征目标的对象模型已在相对复杂的视频场景中取得了较好的结果,颜色特征的优越性在 DAT 跟踪器中已展现出来。受到 DAT<sup>[19]</sup>的启发,使用基于颜色直方图的贝叶斯分类器对位置进行估计,同时考虑到颜色直方图特征容易在光照发生较大变化的情况下产生跟踪漂移,因此再次提取对光照具有鲁棒性的 HOG 特征来表征目标,运用实时性能优异的 CF 方法与直方图方法相结合的方式定位目标。

### 2.1 颜色直方图对象-背景模型

训练阶段的搜索区域  $Q \in \{M, \epsilon\}$ , 其中  $M$  为前景目标区域,  $\epsilon$  为背景区域。输入图片  $I$ , 前景和背景直方图模型可分别表示为  $H_M^I$  和  $H_\epsilon^I$ 。对于像素  $x$ , 根据贝叶斯规则, 使用颜色直方图来近似计算  $I$  中的  $x$  属于对象区域  $O$  的概率, 表达式为

$$P(x \in O \mid M, \epsilon, b_x) = \begin{cases} \frac{H_M^I(b_x)}{H_M^I(b_x) + H_\epsilon^I(b_x)}, & \text{if } I(x) \in I(M \cup \epsilon), \\ 0.5, & \text{otherwise} \end{cases}, \quad (1)$$

式中:  $b_x$  表示  $x$  在区域  $Q$  中的颜色直方图  $H$  上所属颜色区间对应的直方条。对于不可见像素的情况, 将  $P$  值设为 0.5。(1)式求得的对象似然概率可以通过查表的方式来高效计算, 得到一个能够将目标从背景中区分出来的像素级颜色概率图。

### 2.2 多通道特征核相关器判别模型

KCC<sup>[21]</sup>为 CF 方法提供了统一的解, 打破了

KCF 核函数类型的限制, 可以用于预测平移、旋转和尺度等自定义的仿射变换。受到 KCC 的启发, 使用多通道核相关器模型对平移及尺度进行预测, 将用于估计平移和尺度的核相关器分别记为 KTC 和 KSC。

测试样本及训练样本分别为  $z \in \mathbb{R}^n$  和  $x \in \mathbb{R}^n$ , 其中  $n$  为维度, 将样本映射到高维特征

空间  $\boldsymbol{\varphi}(\boldsymbol{x})$  中, 则滤波器  $f$  的核化互相关响应函数为

$$\hat{\boldsymbol{g}} = \mathcal{F}[\mathbf{k}_x(\boldsymbol{z}) \otimes f] = \hat{\mathbf{k}}_x(\boldsymbol{z}) \odot \hat{f}^*, \quad (2)$$

式中:  $\boldsymbol{g}$  表示回归目标;  $\mathcal{F}(\cdot)$  表示快速傅里叶变换函数;  $\otimes$  表示相关运算符;  $\odot$  表示点积运算;  $\hat{\cdot}$  和  $*$  分别表示快速傅里叶变换符号和共轭符号;  $\mathbf{k}_x(\boldsymbol{z}) = [k(\boldsymbol{z}, \boldsymbol{x}_1) \cdots k(\boldsymbol{z}, \boldsymbol{x}_m)]^T$ , 其中  $k(\cdot)$  为核函数, 定义  $k(\boldsymbol{u}, \boldsymbol{u}') = \boldsymbol{\varphi}^T(\boldsymbol{u})\boldsymbol{\varphi}(\boldsymbol{u}')$ ,  $m$  为样本数量,  $\boldsymbol{x}_i \in \mathcal{T}(\boldsymbol{x}) \in \mathbb{R}^n$  为基于样本  $\boldsymbol{x} \in \mathbb{R}^n$  的向量, 其中  $\mathcal{T}(\cdot)$  为变换函数,  $\mathcal{T}_{\text{Trans}}(\cdot)$  和  $\mathcal{T}_{\text{Scale}}(\cdot)$  分别为平移和尺度变换函数,  $i \in \{1, \dots, m\}$ 。  $f$  和  $\boldsymbol{g}$  的长度与  $m$  有关, 即  $f \in \mathbb{R}^m$  和  $\boldsymbol{g} \in \mathbb{R}^m$ 。

对于一个图像信号的  $D$  维通道特征图, 在每帧图像中提取  $\boldsymbol{x} \in \mathbb{R}^n$ 。  $\boldsymbol{g}$  选用高斯函数。学习一个最优的多通道特征相关器  $f$ , 则目标函数应满足

$$\min_{f_c} \left\| \sum_{c=1}^D \mathbf{k}_{x_c}(\boldsymbol{x}_c) \otimes f_c - \boldsymbol{g} \right\|^2 + \lambda_{\text{templ}} \left\| \sum_{c=1}^D f_c \right\|^2, \quad (3)$$

式中:  $\boldsymbol{x}_c$  表示第  $c$  维通道的特征样本;  $\lambda_{\text{templ}}$  表示正则化参数。在频域中, (3) 式可以表示为

$$\min_{f_c^*} \left\| \sum_{c=1}^D \hat{\mathbf{k}}_{x_c}(\boldsymbol{x}_c) \odot \hat{f}_c^* - \hat{\boldsymbol{g}} \right\|^2 + \lambda_{\text{templ}} \left\| \sum_{c=1}^D \hat{f}_c^* \right\|^2. \quad (4)$$

对 (4) 式进行求解, 得到对应于第  $c$  维通道特征样本的相关器参数, 表达式为

$$\hat{f}_c^* = \frac{\hat{\boldsymbol{g}} \odot \hat{\mathbf{k}}_{x_c}^*(\boldsymbol{x}_c)}{\sum_{j=1}^D \hat{\mathbf{k}}_{x_j}(\boldsymbol{x}_j) \odot \hat{\mathbf{k}}_{x_j}^*(\boldsymbol{x}_j) + \lambda_{\text{templ}} \boldsymbol{I}}. \quad (5)$$

对于信号  $\boldsymbol{z}$ , 平移变换数为  $q$ 。由于在跟踪问题中, 采用线性核具有计算简单且跟踪效果较好, 因此所提算法采用线性核  $k(\boldsymbol{z}, \boldsymbol{x}_o) = \boldsymbol{z}^T \boldsymbol{x}_o$  进行简化求解, 其中  $o$  取  $o = 1, \dots, q$ , 则

$$\begin{aligned} \mathbf{k}_x(\boldsymbol{z}) &= [k(\boldsymbol{z}, \boldsymbol{x}_1) \cdots k(\boldsymbol{z}, \boldsymbol{x}_n)]^T \\ &= [\boldsymbol{z}^T \boldsymbol{x}_1 \cdots \boldsymbol{z}^T \boldsymbol{x}_n]^T \\ &= \boldsymbol{z} \otimes \boldsymbol{x} \\ &= \mathcal{F}^{-1}(\hat{\boldsymbol{z}} \odot \hat{\boldsymbol{x}}^*), \end{aligned} \quad (6)$$

式中:  $\mathcal{F}^{-1}(\cdot)$  表示逆傅里叶变换。由于 (6) 式只存在元素方向的运算, 因此可以快速地训练一个多通道的线性分类器。

### 2.3 高置信度更新策略

为了适应目标的外观变化, 在图像的第  $t$  帧中, 采用学习率  $\eta_{\text{color}}$  以线性插值的方式来更新颜色对

象-背景模型, 表达式为

$$H_M^I(b_x)_{t, \text{model}} = (1 - \eta_{\text{color}}) H_M^I(b_x)_{t-1, \text{model}} + \eta_{\text{color}} H_M^I(b_x)_t, \quad (7)$$

$$H_\varepsilon^I(b_x)_{t, \text{model}} = (1 - \eta_{\text{color}}) H_\varepsilon^I(b_x)_{t-1, \text{model}} + \eta_{\text{color}} H_\varepsilon^I(b_x)_t. \quad (8)$$

相关器模板参数  $\hat{f}_c^*$  可以简写为  $\hat{f}_c^* = \hat{\boldsymbol{r}}_c / (\hat{\boldsymbol{d}} + \lambda_{\text{templ}} \boldsymbol{I})$ , 其中  $\hat{\boldsymbol{r}}_c = \hat{\boldsymbol{g}} \odot \hat{\mathbf{k}}_{x_c}^*(\boldsymbol{x}_c)$ ,  $\hat{\boldsymbol{d}} = \sum_{j=1}^D \hat{\mathbf{k}}_{x_j}(\boldsymbol{x}_j) \odot \hat{\mathbf{k}}_{x_j}^*(\boldsymbol{x}_j)$ , 则对当前  $t$  帧的模型参数使用 CF 模型的学习率  $\eta$  进行在线更新, 表达式为

$$\hat{\boldsymbol{r}}_{c,t, \text{model}} = (1 - \eta) \hat{\boldsymbol{r}}_{c,t-1, \text{model}} + \eta \hat{\boldsymbol{r}}_{c,t}, \quad (9)$$

$$\hat{\boldsymbol{d}}_{t, \text{model}} = (1 - \eta) \hat{\boldsymbol{d}}_{t-1, \text{model}} + \eta \hat{\boldsymbol{d}}_t. \quad (10)$$

为了提高遮挡处理的能力, 以及避免更新过程中目标被损坏的问题, 采用 LMCF 方法<sup>[20]</sup> 对判断置信度的方法进行改进, 使用 APCE 来衡量检测目标的置信度, 表达式为

$$A_{\text{APCE}, t} = \frac{|S_{\max, t} - S_{\min, t}|^2}{\text{mean} \left[ \sum_{w, h} (S_{w, h, t} - S_{\min, t})^2 \right]}, \quad (11)$$

式中:  $S_{\max, t} = \max(S_t)$ ,  $S_{\min, t} = \min(S_t)$ , 其中  $S_t$  表示第  $t$  帧图像的响应;  $w \in (0, \dots, W)$  表示行,  $W$  为图像块的宽度;  $h \in (0, \dots, H)$  表示列,  $H$  为图像块的高度。将采用两种置信度阈值判断跟踪可靠性的方法进行对比 (详见第 3 节), 发现在满足 APCE 与最大响应均大于更新阈值  $\theta$ , 则对模型参数进行更新, 此时效果最佳。

### 2.4 检测阶段

在位置检测阶段, 计算颜色概率图中所有大小等于预测尺度的矩形框内像素概率的和, 其可以在概率图上查找四次积分图而得到。设矩形框区域的长宽分别为  $u$  和  $v$ , 则该矩形框中心位置的响应图为

$$S_{\text{color}}(x) = [J(p_u, p_v) + J(p_u + u, p_v + v) - J(p_u + u, p_v) - J(p_u, p_v + v)], \quad (12)$$

式中:  $J$  表示积分图, 在位置  $(p_u, p_v)$  处的积分图  $J(p_u, p_v) = \sum_{p_{u1} \leq p_u, p_{v1} \leq p_v} V(p_{u1}, p_{v1})$ , 其为输入图像左上角的所有像素值  $V(p_{u1}, p_{v1})$  的和,  $(p_u, p_v)$  为矩形框左上角的位置。

在前一帧图像的目标位置处提取  $\boldsymbol{z}$ , 则在当前  $t$  帧的图像中其平移相关响应可表示为

$$S_{t, \text{templ}} = \mathcal{F}^{-1} \left[ \sum_{c=1}^D \hat{\mathbf{k}}_{t, x_c}(\boldsymbol{z}_c) \odot \hat{f}_{t-1, c, \text{model}}^* \right]. \quad (13)$$

最终的平移响应得分采用融合因子  $\gamma$  来综合

HOG 特征和颜色统计特征的响应得分,得到当前  $t$  帧图像的平移综合响应分数为

$$S_{t,p} = (1 - \gamma)S_{t,tpl} + \gamma S_{t,color}, \quad (14)$$

式中:  $S_{t,color}$  表示颜色统计特征响应得分。由 (14) 可知,响应的最大峰值位置即为预测的目标位置。

当进行尺度预测时,  $\mathbf{x}_i$  可以通过尺度变换来得到。由于尺度空间不具有周期性,因此存在边界效应,此时可以采用向核向量补零的方式来缓解,即

$$\mathbf{\kappa}_x(\mathbf{z}) = [k(\mathbf{z}, \mathbf{x}_1) \cdots k(\mathbf{z}, \mathbf{x}_m) 0 \cdots 0]^T. \quad (15)$$

核向量的长度增加一倍,即  $\mathbf{\kappa}_x(\mathbf{z}) \in \mathbb{R}^{2m}$  和  $\mathbf{f} \in \mathbb{R}^{2m}$ 。补零是非必要的,即使没有补零,核尺度相关器的尺度估计性能仍然是优于一般算法的<sup>[21]</sup>。

尺度检测阶段,尺度测试样本  $\mathbf{z}_{scale}$  在当前  $t$  帧图像的尺度响应  $S_{t,scale}$  与 (13) 式求解方式相同,最大尺度相关响应即为预测的目标尺度。

### 2.5 算法执行

跟踪方法的流程如图 1 所示,步骤如下,其中  $S_p$  为位置响应得分,  $S_{scale}$  为尺度响应得分。

首先初始化目标的位置和尺度;接着从当前帧的目标框中分别提取 HOG 特征和颜色特征,并分别用于训练 KTC 和直方图模型以计算两种模型的响应分数,使用 (14) 式求得平移综合响应分数,其最大值位置处即为预测的下一帧目标位置;从预测位置处提取 HOG 特征来训练 KSC,并使用 (13) 式求解尺度相关响应,其最大响应值的尺度即为目标尺度;然后采用高置信度的方式更新模型;最后输出预测位置和尺度,对下一帧图像的模型进行训练和更新,直到完成所有序列的跟踪。

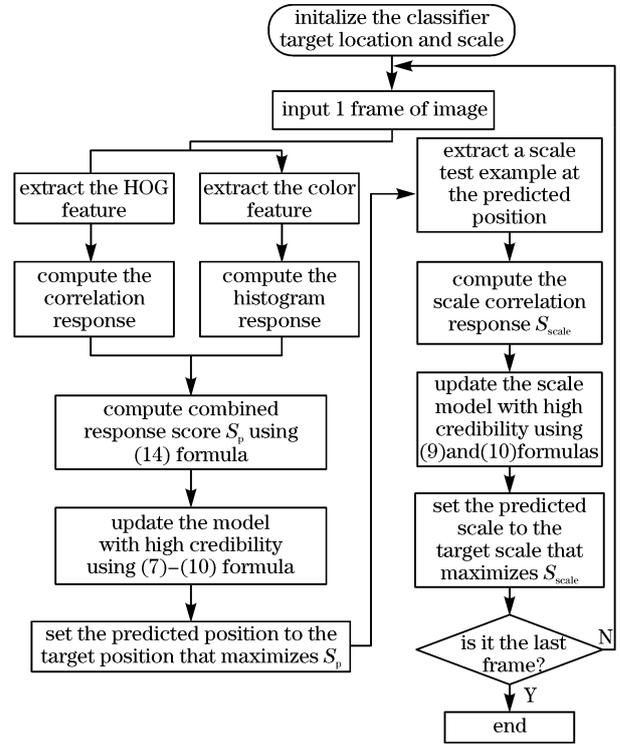


图 1 跟踪方法的流程图

Fig. 1 Flowchart of tracing method

### 3 分析与讨论

将所提方法在 MATLAB R2014a 软件中进行仿真模拟,实验环境为 Intel (R) Core (TM) i5-4200U CPU@1.60 GHz 2.3 GHz,根据在线目标跟踪基准(OTB-2015)对跟踪性能进行评估。

采用不同的置信度阈值来判断所提方法跟踪的可靠性,并与其他方法进行对比。受到文献[22]的启发,当方法 1(ours 1)在 APCE 与最大响应均大于  $\theta$  值时,则认为跟踪响应是可信的,当方法 2(ours 2)在 APCE 与最大响应分别以一定的正确率  $\alpha_1$  和  $\alpha_2$  大于历史均值时,判定跟踪结果可信<sup>[20]</sup>。判别置信度的条件如表 1 所示,其中  $\theta$ 、 $\alpha_1$  和  $\alpha_2$  的设置参考

表 1 跟踪置信度的阈值设置

Table 1 Threshold setting for tracking confidence

Name	Translation confidence criterion	Scale confidence criterion	Value
Ours 1	$\begin{cases} A_{APCE,p,t} > \theta \\ \max [S_{p,t}] > \theta \end{cases}$	$\begin{cases} A_{APCE,scale,t} > \theta \\ \max [S_{scale,t}] > \theta \end{cases}$	$\theta = 0.5$
Ours 2	$\begin{cases} A_{APCE,p,t} > \alpha_1 \frac{1}{t-1} \sum_{i=1}^{t-1} A_{APCE,p,t}^{(i)} \\ \max [S_{p,t}] > \alpha_2 \frac{1}{t-1} \sum_{i=1}^{t-1} \max [S_{p,t}^{(i)}] \end{cases}$	$\begin{cases} A_{APCE,scale,t} > \alpha_1 \frac{1}{t-1} \sum_{i=1}^{t-1} A_{APCE,scale,t}^{(i)} \\ \max [S_{scale,t}] > \alpha_2 \frac{1}{t-1} \sum_{i=1}^{t-1} \max [S_{scale,t}^{(i)}] \end{cases}$	$\begin{aligned} \alpha_1 &= 0.45 \\ \alpha_2 &= 0.70 \end{aligned}$

文献[20,22],  $A_{APCE,p}$  为预测平移位置的平均峰值相关能量,  $A_{APCE, scale}$  为预测尺度的平均峰值相关能量。

模型的参数设置如表 2 所示。

表 2 实验参数的设置

Table 2 Setting of experimental parameters

Parameter	Value
CF (HOG)model learning rate	0.01
Background and foreground colour models learning rate	0.04
Scale learning rate	0.008
Interpolation factor	0.3

将所提方法与其他 4 种性能优异的方法 (DCF<sup>[7]</sup>、KCF<sup>[7]</sup>、SAMF<sup>[16]</sup> 和 KCC<sup>[21]</sup>) 进行对比, 100 个视频序列中的一次性评估精确曲线和成功率曲线分别如图 2 和图 3 所示。不同跟踪方法在 100 个

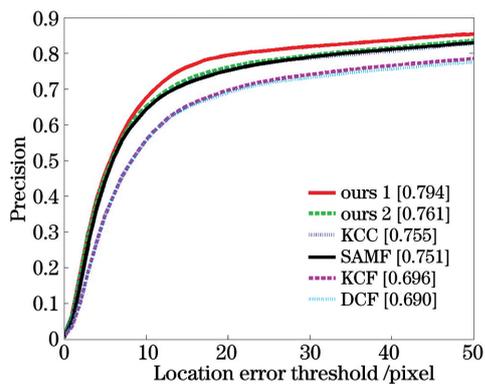


图 2 不同方法在 100 个序列中的精度曲线

Fig. 2 Precision curves of different methods in 100 sequences

表 3 不同跟踪器的性能比较

Table 3 Performance comparison of different trackers

Parameter	Ours 1	Ours 2	SAMF	KCC	KCF	DCF
Average precision	0.794	0.761	0.751	0.755	0.696	0.690
Average success rate	0.694	0.673	0.674	0.565	0.551	0.548
Average frame rate/(frame · s <sup>-1</sup> )	15.340	16.350	5.476	16.680	72.382	105.650

基准数据集使用 11 种属性描述了跟踪过程中的不同挑战性情况。表 4 和表 5 分别为 11 个属性下的平均距离精度和平均重叠成功率, 其中 11 个属性分别为尺度变化(SV)、快速运动(FM)、背景杂乱(BC)、运动模糊(MB)、光照变化(IV)、视线外(OV)、平面内旋转(IPR)、形变(DEF)、低分辨率(LR)、遮挡(OCC)和平面外旋转(OPR), 对比结果排名第一的数值加粗, 排名第二的数值使用下划线来标记, 括号中的数

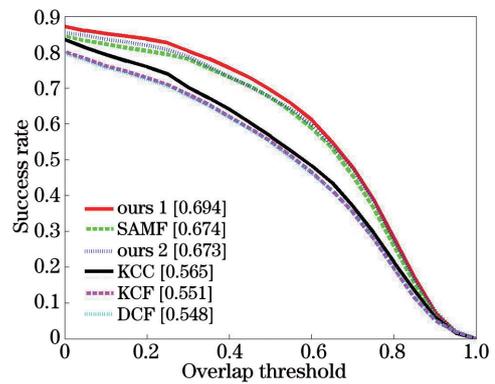


图 3 不同方法在 100 个序列中的成功率曲线

Fig. 3 Success rate curves of different methods in 100 sequences

视频序列中的平均精度、成功率及帧速率, 如表 3 所示。从图 2、图 3 和表 3 可以看到, ours 1 方法和 ours 2 方法较其他方法在成功率方面提高更为明显, ours 1 的跟踪效果较 ours 2 更优; 在位置误差阈值为 20 pixel 的情况下, ours 1 的平均距离精度 (DP) 为 0.794; 在重叠阈值范围为 [0, 1] 的情况下, ours 1 的平均重叠成功率 (OS) 为 0.694, 该方法的精确率和成功率较 DCF 方法提高了 0.104 和 0.146, 充分展示出融入颜色直方图及 KSC 估计尺度方法的跟踪优势, 以及使用置信度阈值判别响应可靠性的优越性; 在精度方面, 采用置信度高的模型进行更新, 其精度较 KCC 方法高出 0.039, 成功率比 SAMF 方法提高 0.020; 该方法在保持高精度和高成功率的情况下, 在 CPU 上的帧速率达到 15.34 frame/s, 虽较实时性最优的 DCF 方法有所差距, 但较融入 CN 特征的 SAMF 方法实时性能较优。

值表示满足每一种属性的图像数。从表 4 和表 5 可以看到, 使用高置信度更新策略且融入多种特征的 KCC 方法, 能够提高跟踪器在具有挑战性的 BC、FM、SV 和 DEF 等属性下的稳定性; 所提方法 ours 1 和 ours 2 在 BC、SV 和 IPR 属性下的成功率及距离精度优势很明显, ours 1 在 MB 和 LR 情况下的精度, 以及在 LR、OCC、MB 和 OPR 情况下的成功率, 虽较排名第一的 SAMF 方法有所降低, 但所差不多。

表 4 不同跟踪挑战下的平均精度

Table 4 Average precision facing different tracking challenges

Attribute	Ours 1	Ours 2	SAMF	KCC	KCF	DCF
SV(63)	<b>0.745</b>	<u>0.715</u>	0.701	0.701	0.635	0.628
FM(39)	<b>0.714</b>	0.680	0.654	<u>0.711</u>	0.621	0.603
BC(31)	<b>0.832</b>	<u>0.814</u>	0.689	0.808	0.713	0.686
MB(29)	<u>0.700</u>	0.691	0.655	<b>0.713</b>	0.601	0.576
IV(37)	<b>0.809</b>	0.790	0.708	<u>0.808</u>	0.724	0.698
OV(14)	<b>0.682</b>	0.610	0.628	<u>0.652</u>	0.501	0.487
IPR(51)	<b>0.802</b>	<u>0.782</u>	0.721	0.750	0.701	0.686
DEF(43)	<b>0.756</b>	0.693	0.680	<u>0.751</u>	0.619	0.623
LR(9)	<u>0.606</u>	0.518	<b>0.685</b>	0.494	0.560	0.564
OCC (48)	<b>0.724</b>	0.685	<u>0.722</u>	0.679	0.632	0.611
OPR (63)	<b>0.754</b>	0.715	<u>0.739</u>	0.716	0.677	0.665

表 5 不同跟踪挑战下的平均成功率

Table 5 Average success rate facing different tracking challenges

Attribute	Ours 1	Ours 2	SAMF	KCC	KCF	DCF
SV(63)	<b>0.606</b>	<u>0.589</u>	0.584	0.495	0.415	0.416
FM(39)	<b>0.640</b>	<u>0.630</u>	0.595	0.599	0.526	0.521
BC(31)	<b>0.783</b>	<u>0.769</u>	0.639	0.625	0.609	0.597
MB(29)	0.640	<b>0.653</b>	0.641	<u>0.651</u>	0.550	0.542
IV(37)	<b>0.707</b>	<b>0.707</b>	<u>0.640</u>	0.599	0.550	0.534
OV(14)	<b>0.603</b>	<u>0.566</u>	0.551	0.516	0.457	0.450
IPR(51)	<b>0.696</b>	<u>0.680</u>	0.641	0.548	0.553	0.544
DEF(43)	<b>0.651</b>	<u>0.608</u>	0.606	0.606	0.503	0.505
LR(9)	<u>0.418</u>	0.382	<b>0.515</b>	0.244	0.295	0.301
OCC (48)	<u>0.653</u>	0.630	<b>0.664</b>	0.530	0.512	0.500
OPR (63)	<u>0.654</u>	0.625	<b>0.660</b>	0.530	0.527	0.519

为了更直观地反映所提方法的跟踪效果,定性评估 6 种方法在三个具有挑战性的序列中的表现,

结果如图 4 所示。

从图 4 可以看到,对于图 4(a),在背景杂乱和

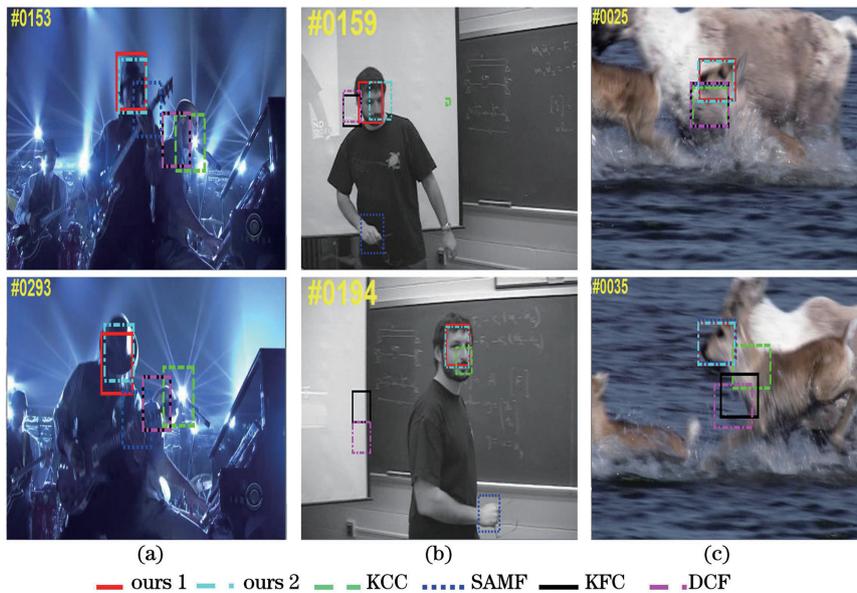


图 4 6 种跟踪器在不同序列中的定性比较结果。(a) shaking 序列; (b) freeman1 序列; (c) deer 序列

Fig. 4 Qualitative comparison results of six trackers in different sequences. (a) shaking sequence; (b) freeman1 sequence; (c) deer sequence

光照较弱的情况下,当出现遮挡或旋转时,所提方法能够准确跟踪目标并能够适应目标的尺度变化,其他方法均出现跟踪失败的情况;对于图 4(b),当其他跟踪器难以应对频繁的旋转变换时,该方法可以很好处理;对于图 4(c),在第 25 frame 的图像中,当有相似的背景干扰时,除了所提方法外,其他方法均出现跟踪漂移的现象,在第 35 frame 的图像中,当处理运动模糊的情况时,所提方法与 SAMF 方法均能准确跟踪目标,而且所提方法的总体跟踪性能表现更佳。

## 4 结 论

为了使图像特征能够准确反映目标的外观变化,采用颜色直方图对象-背景模型与相关滤波判别模型结合的方式来定位目标。针对尺度变化需通过计算平移滤波响应求得的问题,使用尺度变换样本来直接训练核尺度相关器。针对未判断响应置信度就对模型进行更新可能会引入不可信样本,从而造成模型损坏的问题,采用 APCE 和最大响应判据对置信度高的模型进行更新。在 100 个序列中的跟踪结果表明,所提方法在 CPU 上实现实时跟踪,在尺度变化、形变、光照、旋转和相似背景干扰等复杂的视频场景中,能够以较高的精度和成功率跟踪目标,对提高智能视频监控系统的监控效率具有重要价值。

## 参 考 文 献

- [1] Wu G X. Research on key technology of object tracking in intelligent video surveillance [D]. Nanjing: Nanjing University of Science and Technology, 2017.  
吴国星. 智能视频监控下的目标跟踪关键技术研究 [D]. 南京: 南京理工大学, 2017.
- [2] Lu H C, Li P X, Wang D. Visual object tracking: a survey [J]. Pattern Recognition and Artificial Intelligence, 2018, 31(1): 61-76.  
卢湖川, 李佩霞, 王栋. 目标跟踪算法综述 [J]. 模式识别与人工智能, 2018, 31(1): 61-76.
- [3] Mao Z C, Chen H D. Long-term object tracking algorithm based on kernelized correlation filter [J]. Laser & Optoelectronics Progress, 2019, 56(1): 010702.  
茅正冲, 陈海东. 基于核相关滤波的长期目标跟踪算法 [J]. 激光与光电子学进展, 2019, 56(1): 010702.
- [4] Zhang H L, Hu S Q, Yang G S. Video object tracking based on appearance models learning [J]. Journal of Computer Research and Development, 2015, 52(1): 177-190.  
张焕龙, 胡士强, 杨国胜. 基于外观模型学习的视频目标跟踪方法综述 [J]. 计算机研究与发展, 2015, 52(1): 177-190.
- [5] Wu Y, Lim J, Yang M H. Object tracking benchmark [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834-1848.
- [6] 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.
- [7] 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.
- [8] Liu X Y, Wang Y M, Ma W N. Scale-adaptive correlation filter tracking algorithm based on FHOG and LBP features [J]. Laser & Optoelectronics Progress, 2020, 57(4): 041512.  
刘晓悦, 王云明, 马伟宁. 融合 FHOG 和 LBP 特征的尺度自适应相关滤波跟踪算法 [J]. 激光与光电子学进展, 2020, 57(4): 041512.
- [9] Liao X F, Hou Z Q, Yu W S, et al. A scale adapted tracking algorithm based on kernelized correlation [J]. Acta Optica Sinica, 2018, 38(7): 0715002.  
廖秀峰, 侯志强, 余旺盛, 等. 基于核相关的尺度自适应视觉跟踪 [J]. 光学学报, 2018, 38(7): 0715002.
- [10] Wang K Y, Chen Z G, Fu Y. Target anti-occlusion algorithm based on correlation filter [J]. Laser & Optoelectronics Progress, 2019, 56(3): 030401.  
王凯宇, 陈志国, 傅毅. 基于相关滤波器的目标抗遮挡算法 [J]. 激光与光电子学进展, 2019, 56(3): 030401.
- [11] Mueller M, Smith N, Ghanem B. Context-aware correlation filter tracking [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1387-1395.
- [12] Wan X, Zhang C H, Zhang L, et al. Visual tracking via weighted global context-aware correlation filter [J]. Computer Engineering and Applications, 2018, 54(16): 1-12, 67.  
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, 67.
- [13] Bibi A, Mueller M, Ghanem B. Target response adaptation for correlation filter tracking [M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham:

- Springer, 2016, 9910: 419-433.
- [14] Bertinetto L, Valmadre J, Golodetz S, et al. Staple: complementary learners for real-time tracking [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1401-1409.
- [15] Yin K, Li J L, Li L, et al. Adaptive feature update object-tracking algorithm in complex situations [J]. *Acta Optica Sinica*, 2019, 39(11): 1115002.  
尹宽, 李均利, 李丽, 等. 复杂情况下自适应特征更新目标跟踪算法 [J]. *光学学报*, 2019, 39(11): 1115002.
- [16] Li Y, Zhu J K. A scale adaptive kernel correlation filter tracker with feature integration [M] // Agapito L, Bronstein M, Rother C. *Computer vision-ECCV 2014 workshops. Lecture notes in computer science*. Cham: Springer, 2015, 8926: 254-265.
- [17] Danelljan M, Häger G, Khan F S, et al. Learning spatially regularized correlation filters for visual tracking [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 4310-4318.
- [18] Danelljan M, Robinson A, Shahbaz Khan F, et al. Beyond correlation filters: learning continuous convolution operators for visual tracking [M] // Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9909: 472-488.
- [19] Possegger H, Mauthner T, Bischof H. In defense of color-based model-free tracking [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 2113-2120.
- [20] Wang M M, Liu Y, Huang Z Y. Large margin object tracking with circulant feature maps [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 4800-4808.
- [21] Wang C, Zhang L, Xie L, et al. Kernel cross-correlator [EB/OL]. (2018-02-26) [2020-07-29]. <https://arxiv.org/abs/1709.05936>.
- [22] Ma C, Yang X K, Zhang C Y, et al. Long-term correlation tracking [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 5388-5396.
- [23] Dai K N, Wang D, Lu H C, et al. Visual tracking via adaptive spatially-regularized correlation filters [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 4665-4674.
- [24] Valmadre J, Bertinetto L, Henriques J, et al. End-to-end representation learning for correlation filter based tracking [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 5000-5008.
- [25] Guo Q, Feng W, Zhou C, et al. Learning dynamic Siamese network for visual object tracking [C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 1781-1789.