

# 基于旋转的尺度自适应运动目标跟踪算法

戴煜彤<sup>1</sup>, 陈志国<sup>1\*</sup>, 傅毅<sup>1,2</sup>

<sup>1</sup> 江南大学人工智能与计算机学院, 江苏 无锡 214122;

<sup>2</sup> 无锡环境科学与工程研究中心, 江苏 无锡 214153

**摘要** 针对相关滤波器对平面内旋转和尺度变化等跟踪效果不佳的问题,在 ECO\_HC (efficient convolution operators handcraft)的基础上提出了一种具有旋转特性的尺度自适应相关滤波目标跟踪算法。训练一个旋转和尺度滤波器,采用相位相关算法求得尺度因子和旋转角度;在进行旋转和尺度模型更新时,采取动态自适应更新策略;在位置模型更新阶段,通过融合前序帧的背景信息提高模型的稳定性。实验数据表明,所提方法不仅对平面内旋转和尺度变化有较强的鲁棒性,而且能够达到实时要求。

**关键词** 图像处理; 相关滤波; 平面内旋转; 尺度变化; 模型更新

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.1210019

## Rotation-Based Scale Adaptive Moving Target Tracking Algorithm

Dai Yutong<sup>1</sup>, Chen Zhiguo<sup>1\*</sup>, Fu Yi<sup>1,2</sup>

<sup>1</sup> School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi, Jiangsu 214122, China;

<sup>2</sup> Wu Xi Research Center of Environmental Science and Engineering, Wuxi, Jiangsu 214153, China

**Abstract** Aiming at the problem of poor tracking effect of correlation filters in dealing with in-plane rotation and scale changes, this paper proposed a scale adaptive correlation filter target tracking algorithm with rotation characteristics based on ECO\_HC (efficient convolution operators handcraft). Firstly, we train a scale and rotation filter, and then use the phase correlation algorithm to obtain the scale factor and rotation angle. Secondly, we adopt a dynamic adaptive update strategy for rotation and scale updating. Finally, in the position model update stage, we fuse the background information of front frames to enhance the stability of the template. Experimental data shows that our method is not only robust against in-plane rotation and scale changes, but also can meet real-time requirements.

**Key words** image processing; correlation filter; in-plane rotation; scale change; model update

**OCIS codes** 100.4999; 120.2440; 100.2000

## 1 引言

目标跟踪<sup>[1]</sup>近年来广泛应用于视频监控、无人驾驶、医学图像分析、空中预警等诸多领域,成为计算机视觉的研究热点之一。目标跟踪的主流方法目前正由生成式方法逐渐转向判别式方法,其中基于相关滤波的目标跟踪算法发展尤为迅速<sup>[2]</sup>。

Bolme 等<sup>[3]</sup>提出的最小输出误差平方和(MOSSE)算法将相关滤波引入目标跟踪领域。Henriques 等<sup>[4]</sup>采用了方向梯度直方图(HOG)特征和循环移位采样方式,提出了核相关滤波器(KCF)跟踪算法,大大提高了跟踪算法的精确度。但是,核相关滤波器仅使用了单一的 HOG 特征,不能应对尺度变化的情况。Danelljan 等<sup>[5]</sup>在 MOSSE 算法

收稿日期: 2020-08-10; 修回日期: 2020-09-29; 录用日期: 2020-10-29

基金项目: 江苏省高等学校自然科学研究面上项目(17KJB520039)、江苏省“333 高层次人才培养工程科研项目”(BRA2018147)、2020 年江苏高校“青蓝工程”

\* E-mail: 427533@qq.com

的基础上增加了一个尺度滤波器,提出了 DSST (discriminative scale space tracking) 算法,由两个滤波器单独处理 DSST 算法得到的位置与尺度估计值,DSST 的结果只是尺度最优的,位置却不一定是最优的。循环移位策略的应用使得基于相关滤波的目标跟踪算法受到边缘效应的影响,Danelljan 等<sup>[6]</sup>提出基于空间正则化惩罚系数的 SRDCF (spatially regularized discriminant correlation filter),并取得了一定的效果。之后,Danelljan 等<sup>[7]</sup>提出了 C-COT (continuous convolution operators for visual tracking) 算法,通过插值函数将不同分辨率的特征图插值到连续的空间域上。随后,Danelljan 等<sup>[8]</sup>在 C-COT 的基础上提出了 ECO (efficient convolution operator),该算法分为基于卷积特征和基于手工特征的两种 ECO\_HC 算法。ECO 通过高效卷积操作融合多个特征以获取特征图,从模型的大小、训练集的更新策略以及模型更新的频率三个方面对 C-COT 进行改进,从而有效降低算法的复杂度。C-COT 和 ECO 均沿用了 DSST 的尺度更新策略,但并未对尺度估计进行深入研究。Mukhopadhyay 等<sup>[9]</sup>提出了 Siam-RPN (siamese-region proposal network) 网络架构算法,并在其基础上提出了 Siam-RPN++ 算法<sup>[10]</sup>,这些深度学习的算法着重从特征提取的角度对跟踪算法进行改进。

以上算法均不具备旋转特性,当目标发生旋转时,由于跟踪框无法跟随目标发生旋转,所以提取的特征会产生偏差,跟踪的精确度会有所下降。ECO\_HC 算法采用固定尺度模板,只能在模板范围内选择最优值,若跟踪目标的尺度变化较剧烈,则不能很好地估计尺度因子。因此,本文在 ECO\_HC 算法的基础上,提出具有旋转特性的尺度自适应目标跟踪算法。该方法使得目标框可跟随目标的旋转进行旋转,并且能较好地估计目标的尺度变化。本文的主要工作如下:1) 采用相位相关算法训练旋转和尺度滤波器,计算目标的尺度因子和旋转角度;2) 采用一种动态的旋转和尺度模型更新策略,以适应不同程度的尺度和旋转角度变化;3) 为了弱化由目标框旋转而引入的噪声,在位置模型更新阶段融合前序帧的模型信息,以减少背景剧烈变化导致的模型漂移现象。

## 2 ECO\_HC 跟踪基本原理

相关滤波器<sup>[11-14]</sup>的核心是通过前序帧训练的滤波器判定跟踪目标在下一帧的位置。ECO\_HC 算

法首先在第一帧中选定跟踪目标,获取其在图片帧中的实际位置及其他特征信息以初始化训练跟踪器,并且同时训练一个一维尺度滤波器。后续帧中,根据跟踪器中预测的位置对其周围像素点进行采样,取最高响应值的位置作为跟踪目标的所在位置,然后在尺度滤波器中以目标位置为中心提取 33 个不同尺度下的样本,再对样本进行响应值计算,取最大响应值对应的尺度因子对跟踪器进行更新。相关滤波是根据前序帧训练的模型去预测跟踪目标下一帧的位置,若跟踪目标发生快速形变,如尺度变化或平面内旋转,模型就会学习到错误的信息,导致模型对后续帧的跟踪失败。

### 2.1 相关滤波

相关滤波的核心思想是将设计好的滤波模型  $h$  与输入的图像  $f$  做相关运算,求得响应值  $g$  为

$$g = f \otimes h, \quad (1)$$

式中: $\otimes$ 为叉乘运算。

响应值  $g$  越大,则  $f$  和  $h$  的相关度越高。当输出的响应值取最大值时,该响应值对应的位置即为当前跟踪目标的位置。为了提高计算速度,可将(1)式利用快速傅里叶变换转至频域,即

$$\mathcal{F}(g) = \mathcal{F}(f \otimes h) = \mathcal{F}(f) \odot \mathcal{F}(h)^*, \quad (2)$$

式中: $\mathcal{F}$ 为傅里叶变换; $*$ 为共轭; $\odot$ 为对应元素相乘,(2)式可简化为

$$G = F \odot H^*, \quad (3)$$

$$H^* = \frac{G}{F}, \quad (4)$$

式中: $G$ 、 $F$ 、 $H$  分别表示相关性输出、输入图片的频域表示和滤波器的频域表示; $H^*$  为  $H$  的共轭,为所求的滤波器,其更新公式为

$$H^* = \frac{\sum_i G_i \odot F_i^*}{\sum_i F_i \odot F_i^*}, \quad (5)$$

式中: $i$  表示图片帧序号。

### 2.2 因式分解的卷积操作

ECO\_HC 是在 C-COT 算法的基础上进行改进的,C-COT 通过学习一个含有  $M$  个训练样本的集合  $\{x_j\}_1^M \subset \mathcal{X}$  获得一个卷积滤波器,其中每一个通道  $d$  的特征层  $x_j^{(d)} \in \mathbb{R}^{(N_d)}$  都有独立的分辨率  $N_d$ ,  $x_j$  为第  $j$  个训练样本; $\mathcal{X}$  为样本的采样空间。对于每一个特征通道  $d$ ,通过插值模型  $J_d$  将其转至连续空间域  $t [t \in [0, T]]$ ,即

$$J_d \{x^{(d)}\} (t) = \sum_{n=0}^{N_d-1} x^{(d)} [n] b_d \left( t - \frac{T}{N_d} n \right), \quad (6)$$

式中:  $b_d$  为周期为  $T(T>0)$  的插值内核, 所得插值特征层  $J_d\{x^{(d)}\}$  是一个连续的周期为  $T$  的函数;  $x^{(d)}$  为特征通道;  $n$  为离散空间变量。在 C-COT 中, 使用  $J_d\{x\}$  表示整个插值特征图, 将其与训练的一个连续周期为  $T$  的多通道卷积滤波器  $f = (f^{(1)}, \dots, f^{(D)})$  进行相关运算, 得到的响应分数  $S_f\{x\}$  为

$$S_f\{x\} = f * J\{x\} = \sum_{d=1}^D f^{(d)} * J_d\{x^{(d)}\}, \quad (7)$$

式中:  $x$  为训练样本;  $f^{(d)}$  为特征层  $d$  的滤波器;  $J\{x\}$  为插值样本;  $D$  为特征通道的总数。

ECO\_HC 在特征提取的阶段使用一组起决定性作用, 即响应分数较高的  $C(C < D)$  个滤波器  $f = (f^1, \dots, f^C)$  替代 C-COT 中每个通道单独训练的滤波器  $f = (f^1, \dots, f^D)$ , 则求得的新的检测函数为

$$S_{Pf}\{x\} = \mathbf{P}f * J\{x\} = \sum_{c,d} p_{d,c} f^{(c)} * J_d\{x^{(d)}\} = f * \mathbf{P}^{(T)} J\{x\}, \quad (8)$$

式中:  $f^{(c)}$  为特征层  $c$  的滤波器;  $\mathbf{P}$  为大小为  $D \times C$  的系数矩阵;  $\mathbf{P}^{(T)}$  为线性降维算子;  $\sum_{c,d} p_{d,c} f^{(c)}$  是特征层  $c$  的滤波器的线性组合;  $p_{d,c}$  为滤波器的学习系数。

### 2.3 生成样本空间模型

ECO\_HC 更改了训练集的样本更新策略, 用高斯混合模型 (GMM) 生成不同的分量, 每个分量是一组相似度较高的样本, 分量之间差异较大。GMM 可表示为

$$p(x) = \sum_{l=1}^L \pi_l N(x; \mu_l; \mathbf{I}), \quad (9)$$

式中:  $l$  为模型分量;  $L$  为高斯分量  $N(x; \mu_l; \mathbf{I})$  的数量;  $\pi_l$  为先验权重;  $\mu_l$  为高斯分量的平均值; 单位矩阵  $\mathbf{I}$  为协方差矩阵。起初, 将每一个新样本  $x_j$  都初始化为一个新的分量  $m$ :

$$\pi_m = \gamma, \quad \mu_m = x_j, \quad (10)$$

式中:  $\pi_m, \mu_m$  分别为分量  $m$  的先验权重和高斯分量的均值;  $\gamma$  为学习速率。当分量个数超过设定好的上限  $L$  时, 如果有一组分量的权重  $\pi_l$  小于设定的阈值, 则丢弃该分量; 否则, 将两个最相似的分量  $k$  与  $l$  合并为一个公共分量  $r$ :

$$\begin{cases} \pi_r = \pi_k + \pi_l \\ \mu_r = \frac{\pi_k \mu_k + \pi_l \mu_l}{\pi_k + \pi_l} \end{cases} \quad (11)$$

$|\mu_k - \mu_l|$ , 即两组之间的距离, 是在傅里叶域

使用帕斯瓦尔定理进行计算得到的。

## 3 改进的 ECO\_HC 算法

ECO\_HC 算法采用尺度金字塔模型来更新目标尺度, 跟踪过程中无法很好地应对不同程度的尺度变化, 致使尺度估计产生误差。此外, 现实中目标运动状态也有可能出现平面内的旋转, 而目前大多数目标跟踪算法<sup>[15-20]</sup>的检测框无法随着目标的旋转而旋转, 也会产生较大的误差, 导致跟踪漂移。因此, 本文算法基于相关滤波器 ECO\_HC 提出了一种具有旋转特性的尺度自适应目标跟踪算法, 保留了 ECO\_HC 的位置滤波器, 增加了一个二维的旋转和尺度滤波器, 并对两个滤波器的模型更新策略进行了改进。

### 3.1 旋转和尺度滤波器

ECO\_HC 算法的尺度更新策略是用尺度滤波器建立金字塔模型并在固定的尺度池内选择不同尺度的图像, 用于计算响应值, 取其最大响应值对应的尺度因子。该方法没有考虑目标的旋转以及目标大小剧烈变化对尺度估计产生的影响, 通过相位相关算法计算出前后连续两帧的尺度变化因子和目标的旋转角度。

若第  $i$  帧旋转和尺度滤波器提取的目标特征  $f_i(x, y)$  是由第  $i-1$  帧的目标特征  $f_{i-1}(x, y)$  经过尺度变化、目标旋转以及目标位移后得到, 且尺度因子为  $s$ , 旋转角度为  $\theta_0$ , 位移为  $(\Delta x, \Delta y)$ , 则前后两帧之间的关系为

$$f_i(x, y) = f_{i-1}[s(x \cos \theta_0 + y \sin \theta_0) + \Delta x, s(-x \sin \theta_0 + y \cos \theta_0) + \Delta y]. \quad (12)$$

利用傅里叶变换将 (12) 式转至频域进行计算, 则有

$$|F_i(u, v)| = s^{-2} |F_{i-1}[s^{-1}(u \cos \theta_0 + v \sin \theta_0), s^{-1}(-u \sin \theta_0 + v \cos \theta_0)]|, \quad (13)$$

式中:  $(u, v)$  表示频域中像素点坐标;  $|\cdot|$  为频谱幅度。在频域中, 前后两帧的关系仅与尺度因子和旋转角度有关, 可由旋转和尺度滤波器进行计算得到。为了简化计算, 将 (13) 式由笛卡儿坐标系转至对数极坐标中进行计算, 如图 1 所示, 若极点为  $(x_0, y_0)$ , 极轴为笛卡儿坐标系中的  $x$  轴, 则有

$$\begin{cases} \rho = \ln \left[ \sqrt{(x - x_0)^2 + (y - y_0)^2} \right] \\ \theta = \arctan \left( \frac{y - y_0}{x - x_0} \right) \end{cases}, \quad (14)$$

式中:  $\rho$  为极径;  $\theta$  为极角。

在对数极坐标系中, 前后两帧之间的关系为

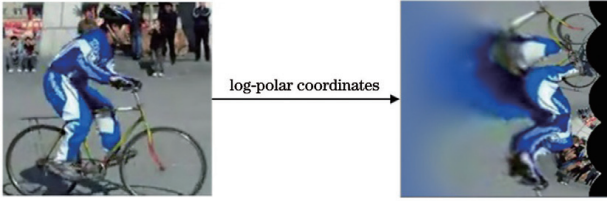


图 1 对数极坐标示意图

Fig. 1 Schematic diagram of log-polar coordinates

$$p_i(\theta, \rho) = p_{i-1}(\theta + \theta_0, \rho + \ln s), \quad (15)$$

式中:  $p_i$  和  $p_{i-1}$  分别表示前后两帧的极坐标。

由(15)式可知,在对数极坐标系中,旋转角度与尺度因子的变化可看作是沿极轴的平移运动。此时应用傅里叶变化位移理论,则(15)式变为

$$F_{p_i}(\omega, \xi) = F_{p_{i-1}}(\omega, \xi) \exp[j2\pi(\omega\theta_0 + \xi \ln s)], \quad (16)$$

式中:  $F_{p_i}$  为  $F_i$  的对数-极坐标变换;  $(\omega, \xi)$  表示像素坐标。

前后两帧的互功率谱为

$$W(\omega, \xi) = \frac{F_{p_{i-1}}(\omega, \xi) F_{p_i}^*(\omega, \xi)}{|F_{p_i}(\omega, \xi) F_{p_{i-1}}^*(\omega, \xi)|} \times \exp[j2\pi(\omega\theta_0 + \xi \ln s)], \quad (17)$$

式中:  $W$  为幅度谱;  $F^*$  为  $F$  的复共轭。对(17)式进行傅里叶逆变换,可得归一化相关函数,从而求得  $\theta_0$  和  $\ln s$  满足

$$\mathcal{F}^{-1}\{W(\omega, \xi)\} = \delta(\theta_0, \ln s), \quad (18)$$

式中:  $\mathcal{F}^{-1}$  为傅里叶逆变换;  $\delta$  为旋转与尺度因子变换函数。

最后,由傅里叶梅林变换算法,对  $\ln s$  取  $e$  为底的指数函数进行转化,则

$$s = \exp(\ln s). \quad (19)$$

至此,通过计算得出前后两帧的旋转角度和尺度因子。根据目标中心点的位置  $(p_{\text{pos},x}, p_{\text{pos},y})$  实现对目标框的尺度和旋转变换:

$$\begin{bmatrix} p'_{\text{pos},x} \\ p'_{\text{pos},y} \\ 1 \end{bmatrix} = \begin{bmatrix} s \cos \theta_0 & s \sin \theta_0 & \Delta x \\ -s \sin \theta_0 & s \cos \theta_0 & \Delta y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{\text{pos},x} \\ p_{\text{pos},y} \\ 1 \end{bmatrix}, \quad (20)$$

式中:  $p'_{\text{pos},x}$  为变换后的中心点横坐标;  $p'_{\text{pos},y}$  为变换后的中心点纵坐标。

### 3.2 旋转和尺度模型的更新策略

本文算法采用了前后两帧的相似度来估计尺度因子和旋转角度。跟踪过程中跟踪目标的尺度和旋转角度的变化程度不一致,即模型参数的变化程度也是变化的,而 ECO\_HC 以及其他相关滤波模型

的尺度更新策略一般采用固定值线性加权策略,所以当相邻帧目标发生较大变化时,其尺度滤波器模型的参数也会有较大的改变,使得尺度滤波器摄入过多的噪声。因此,本文采用变化的更新率来抑制相邻帧之间旋转和尺度模型的参数变化,即模型的更新率与参数变化成反比,该方法对旋转和尺度滤波器的估计值更加准确,更新率  $\gamma$  可表示为

$$\gamma = \frac{M}{\|\alpha_i - \alpha_{i-1}\|_1}, \quad (21)$$

式中:  $\alpha_i$  与  $\alpha_{i-1}$  为相邻两帧的旋转和尺度模型;  $M$  为一个常量,实验中  $M=0.5$ 。根据(21)式,可得旋转与尺度模型的更新方程为

$$\alpha_i = \gamma \alpha_i + (1 - \gamma) \alpha_{i-1}. \quad (22)$$

### 3.3 位置模型更新策略

跟踪框加入旋转策略后,随着视角的改变,背景信息会产生较大的变化,滤波器会获取一些不必要的背景信息,从而因引入噪声而产生误差,造成漂移现象。为了提高模型的稳定性,本文算法在 ECO\_HC 算法的位置模型更新的基础上加强前序帧的背景信息,将第  $i-q$  帧模型与本帧模型进行线性融合:

$$H_{\text{Hf},i} = \lambda_1 H_{\text{Hf},i} + \lambda_2 H_{\text{Hf},i-q}, \quad (23)$$

式中:  $H_{\text{Hf},i}$  与  $H_{\text{Hf},i-q}$  分别为第  $i$  帧与第  $i-q$  帧的位置模型,  $q$  的取值为算法中模型更新的帧率,  $q=5$ ;  $\lambda_1, \lambda_2$  为融合系数,  $\lambda_1=0.9, \lambda_2=0.1$ 。

### 3.4 流程图

算法流程图如图 2 所示。如果当前跟踪帧为第  $i$  帧,则在笛卡儿坐标系中,由第  $i-1$  帧的位置滤波器预测出当前帧中目标的位移变化,在对数极坐标系中,由旋转和尺度滤波器采用相位相关算法并根据第  $i-1$  帧和第  $i$  帧的相似度计算出尺度因子和旋转角度,通过将两个结果相结合预测当前帧的目标框;其次,根据目标的尺度变化程度来控制尺度滤波器模型的更新率;最后,在位置模型更新阶段融合前序帧的模型信息以提高模型的稳定性。

## 4 实验结果及分析

### 4.1 实验环境与参数

本文实验环境的处理器为 Intel(R) i5-8400 2.8 GHz,内存为 8 GB,采用 Matlab2018b 作为编程语言。本文算法与 ECO\_HC 算法的基本参数设置一致,高斯标签带宽  $\sigma$  的值为 0.0625,学习率为 0.009,存储的最大训练样本的数量为 30。为了验证本文算法的有效性,本文基于数据集 OTB-50

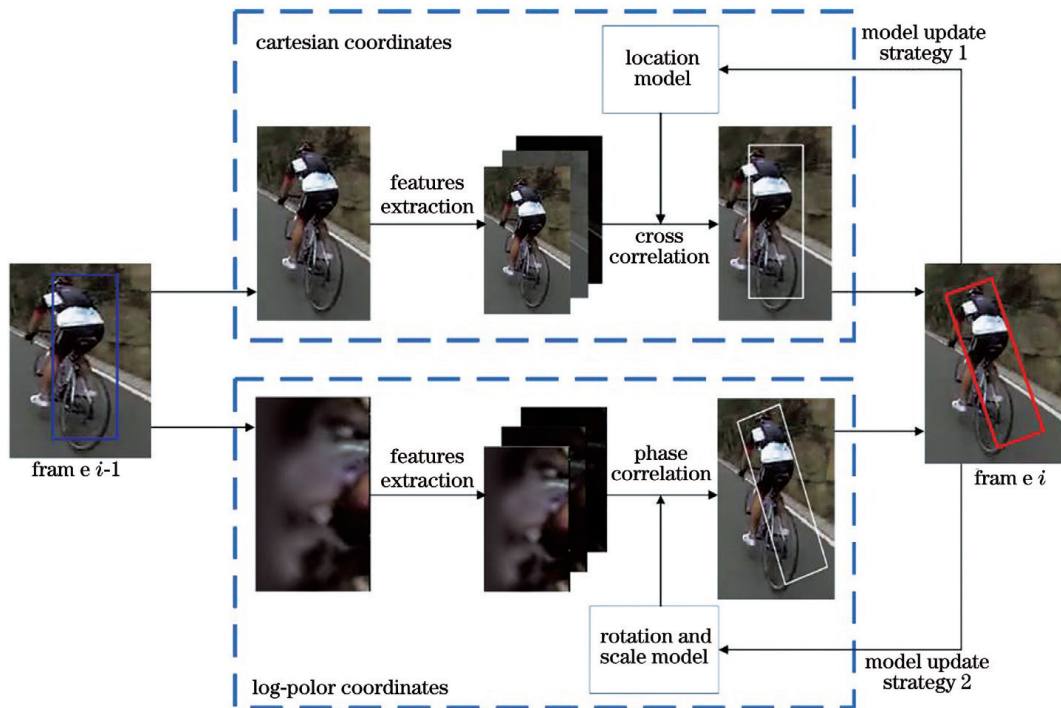


图 2 改进算法的流程图

Fig. 2 Flow chart of improved algorithm

对算法的综合性能进行测评,并在 12 个包含尺度变化和平面内旋转的视频序列上对改进方法进行测试,视频序列如表 1 所示。除了与 ECO\_HC 进行对比之外,本文还选取了 DSST、KCF、LDES、SAMF、

SRDCF、STRCF 进行比较,分别从精确度(DP)、覆盖率(OP)、中心误差(CLE)和速度(FPS)4 个评价指标对算法进行分析。CLE 的值越小,则 DP、FPS 和 OP 的值越大,跟踪效果越优。

表 1 实验中测试视频

Table 1 Test videos in experiment

Video	Number of frames	Main Factor	Video	Number of frames	Main Factor
Bikeshow_ce	361	SV,DEF,FM,IPR,OPR	Human6	792	SV,OCC,DEF,FM,OPR,OV
Singer2	366	IV,SV,DEF,IPR,OPR,BC	Car4	659	IV,SV
Car2	913	IV,SV,MB,FM,BC	Boat_ce1	377	SV,OCC,IPR,OPR
Dog1	1350	SV,IPR,OPR	Doll	1000	IV,SV,OCC,IPR,OPR
Bike_ce2	812	SV,IPR,LR	Carscale	252	OPR,SV,OCC,FM,IPR
Gym	766	SV,DEF,IPR,OPR	Dancer	225	SV,DEF,IPR,OPR

#### 4.2 OTB-50 数据集评价结果

本实验选用公开数据集 OTB-50,此数据集包含 49 个不同的视频序列,覆盖形变、旋转、遮挡和快速运动等多种跟踪场景,用于对跟踪算法的综合性能进行评价。测评结果如图 3 所示,本文算法的精确度为 0.796,综合排名第二,成功率为 0.719,综

合排名第三,与 ECO\_HC 相比,本文算法的精确度和成功率均提高了 0.015。表 2 给出各个算法在 OTB-50 上的运行速度,本文算法的速度相较 ECO\_HC 有所下降,但是基本达到了实时的标准(帧频大于 24 frame/s),较排名第一的算法 STRCF 在速度上有明显的优势。

表 2 算法在 OTB-50 上的运行速度

Table 2 Running speed of algorithms on OTB-50

Algorithm	DSST	ECO_HC	KCF	LDES	SAMF	SRDCF	STRCF	Ours
Running speed (frame·s <sup>-1</sup> )	47.18	50.83	366.29	12.62	20.78	7.16	18.59	23.36

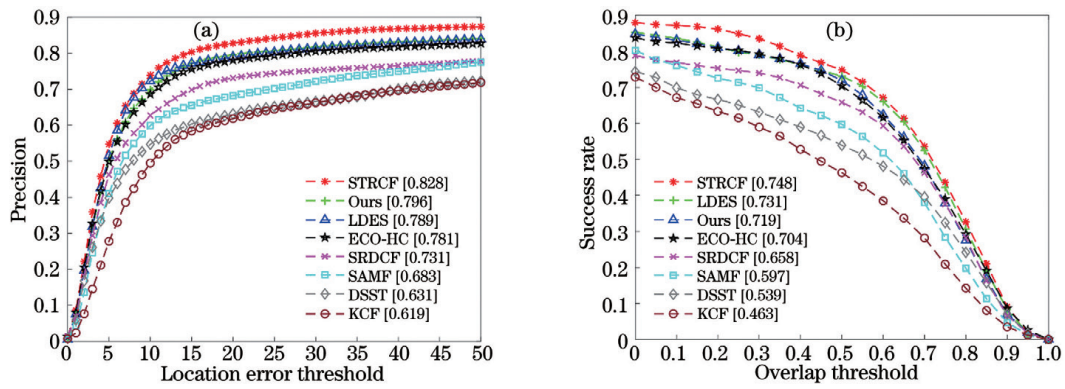


图 3 OTB-50 评价结果。(a)精确度图;(b)成功率图

Fig. 3 Evaluation results of OTB-50. (a) Precision plot; (b) success rate plot

### 4.3 跟踪效果分析

本实验选用了 12 个视频进行跟踪效果分析,部分视频序列的跟踪效果如图 4 所示。从图 4(a)可以看出,在视频序列 Bike\_ce2 上:在第 100 帧中,除了不含尺度变化的 KCF 算法以外,其他算法都能较好地估算目标的尺度;在第 400 帧中,SAMF 和 DSST 发生了漂移,由于目标变化程度较剧烈,ECO\_HC 不能准确地估计尺度,目标框较目标大小偏差较大;在第 800 帧中,KCF、ECO\_HC 和 SRDCF 的尺度估计均出现了问题,从此帧中可以明显看出本文算法相较其他算法的跟踪效果更佳。在视频序列 Bikeshow\_ce 上,DSST 和 SRDCF 都出现了目标丢失的情况,从图 4(b)的三帧中可以看出目标的尺度变化程度不大,但是旋转角度变化很大,

除了 DSST 和 SRDCF,其他几种算法都能够很好地应对尺度变化,本文算法由于估算了目标的旋转角度,对位置的估计也更为准确。图 4(c)中,视频序列 Boat\_ce1 的背景区分度不高,DSST、SAMF 和 KCF 无法准确估计目标尺度,出现漂移,第 100 帧、第 200 帧、第 300 帧中目标的尺度一直在缩小,虽然 ECO\_HC 和 LDES 的跟踪框也一直在变小,但是较目标大小而言,尺度估计值偏差较大。视频序列 Gym 和 Dancer 的尺度变化较小,因此本文算法和其他算法均表现不错,并且本文算法的目标框能准确地随着目标的旋转而旋转。图 4(f)中,跟踪目标受光照因素的干扰,且目标一直在平面内旋转,导致 DSST、ECO\_HC、LDES 一开始就跟踪失败了。

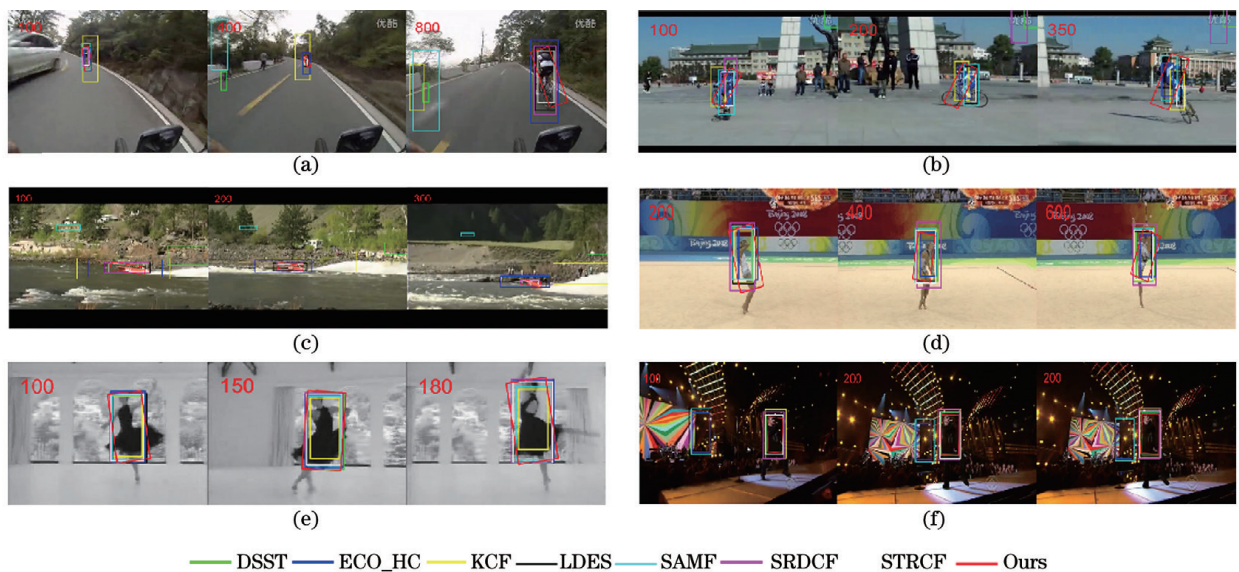


图 4 部分跟踪效果。(a) Bike\_ce2 #100-#400-#800;(b) Bikeshow\_ce #100-#200-#350;(c) Boat\_ce1 #100-#200-#300;(d) Gym #200-#400-#600;(e) Dancer #100-#150-#180;(f) Singer2 #100-#200-#300

Fig. 4 Partial tracking effect. (a) Bike\_ce2 #100-#200-#300; (b) Bikeshow\_ce #100-#200-#300; (c) Boat\_ce1 #100-#200-#300; (d) Gym #200-#400-#600; (e) Dancer #100-#150-#180; (f) Singer2 #100-#200-#300

### 4.4 定量分析

表 3 是本文算法与其他几种算法在 12 个视频序列上的精确度、中心位置误差、覆盖率和速度 4 个评价指标的均值。与 ECO\_HC 比较,本文算法的精确度提高了 0.099,平均中心位置误差值下降了 11.831,覆盖率提高了 0.086,其在精确度、

中心位置误差和覆盖率上均有所提升。在对比算法中,DSST 在部分视频集上出现跟踪丢失的情况,导致整体在 DP 和 CLE 评价上表现不佳,DP 达到 0.653, CLE 达到 47.780。在速度方面,SRDCF、LDES 和 SAMF 都没有达到实时性的要求。

表 3 DP、CLE、OP 和 FPS 比较

Table 3 Comparison of DP, CLE, OP, and FPS

Algorithm	DSST	ECO_HC	KCF	LDES	SAMF	SRDCF	STRCF	Ours
DP	0.653	0.762	0.702	0.753	0.718	0.755	0.846	0.861
CLE	47.780	17.530	15.721	18.191	23.235	12.589	5.349	5.699
OP	0.662	0.637	0.604	0.686	0.593	0.734	0.743	0.723
FPS	39.083	50.126	259.368	17.009	23.767	6.398	26.383	25.244

因此,从 8 种算法在精确度、中心位置误差和覆盖率的综合结果可以看出,本文提出的算法具有一定的优势。实验中绘制了部分视频序列的精确度和中心位置误差评价曲线,从图 5 可以看出,本文算法的精确度较其他算法表现好;从图 6 可以看出,本文算法的中心位置误差明显低于其他算法。图 7 是

12 个视频的平均中心误差、平均重叠率和平均距离精确度的评价曲线,图中方块+虚线为本文算法的结果,可以看出,本文算法在中心位置误差和距离精确度这两个评价指标上优势明显,在重叠率曲线上表现也不错。因此,本文算法在尺度变化和目標旋轉方面有較強的魯棒性。

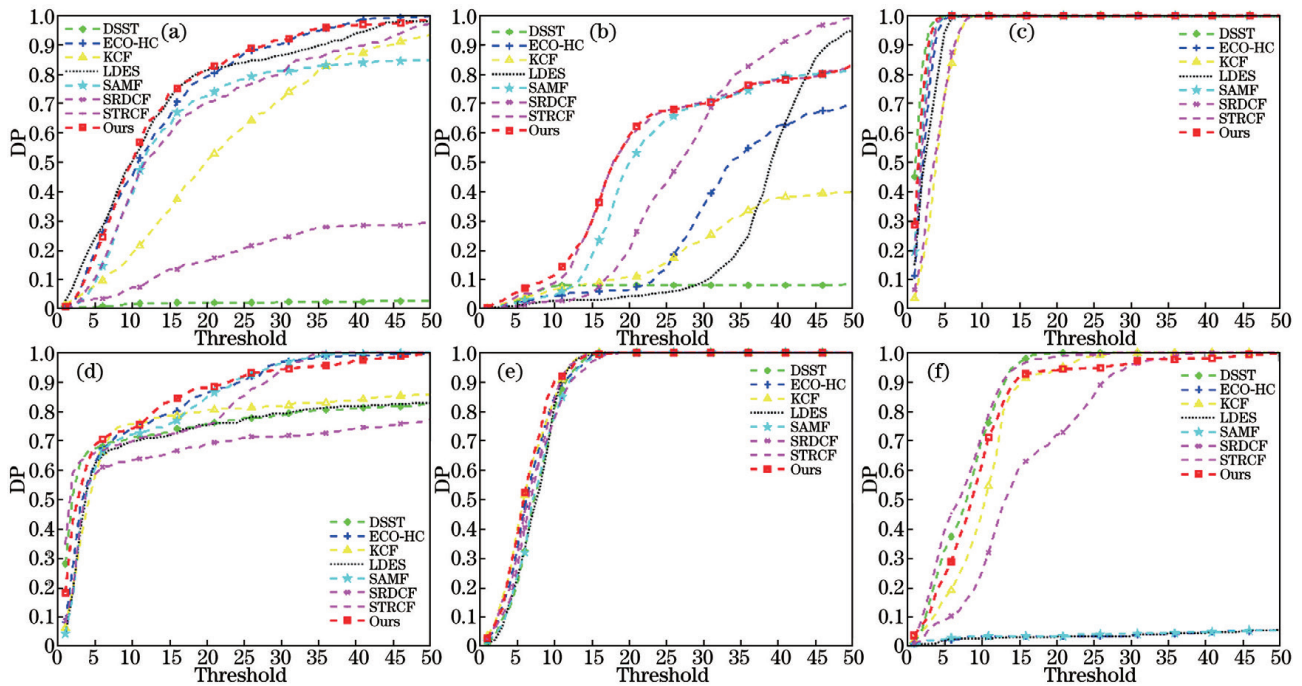


图 5 不同算法在部分视频集上的 DP。(a) Bikeshow\_ce; (b) Boat\_cel; (c) Car2; (d) CarScale; (e) Dancer; (f) Singer2

Fig. 5 DP of different algorithms for partial videos. (a) Bikeshow\_ce; (b) Boat\_cel; (c) Car2; (d) CarScale; (e) Dancer; (f) Singer2

## 5 结 论

针对 ECO\_HC 算法无法应对目标尺度大幅度变化以及平面内旋转的问题,提出一种旋转的尺度

自适应跟踪算法。本文算法在面对尺度的快速变化以及目标旋转时表现优异,采用相位相关算法准确地预测目标的尺度因子和旋转角度,并且动态地更新旋转和尺度滤波器的模型,以适应不同程度的尺

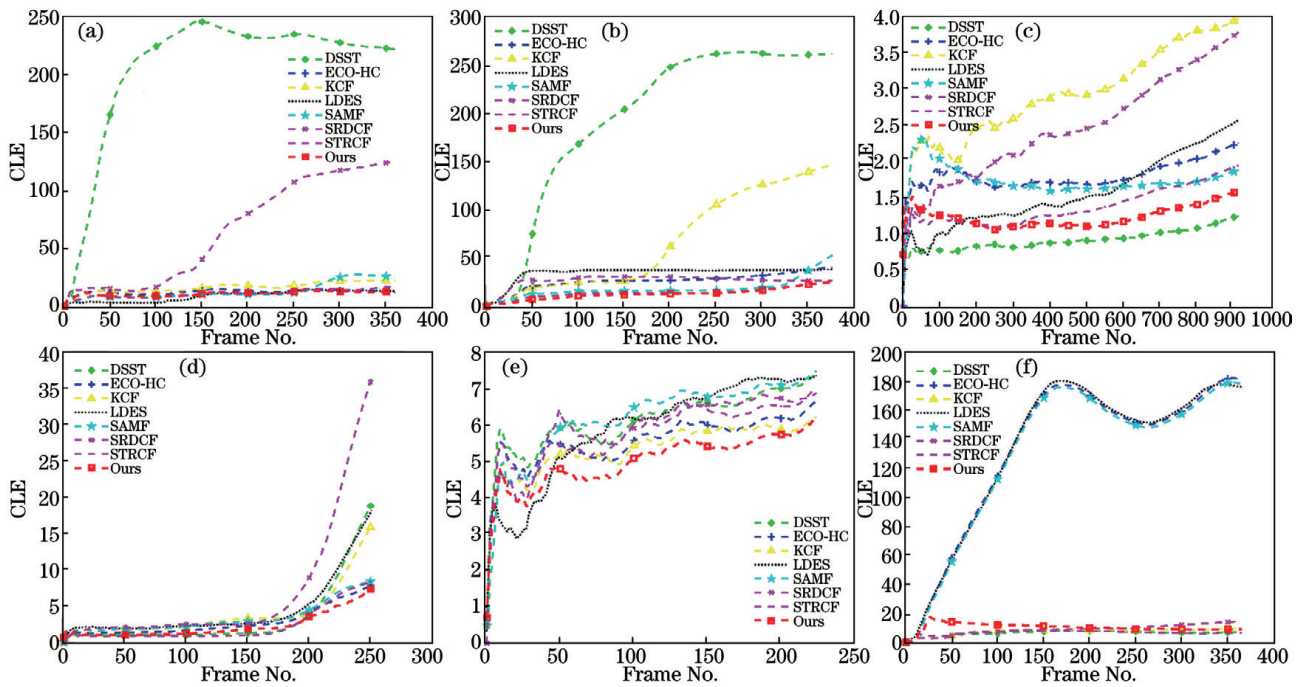


图 6 不同算法在部分视频集上的中心位置误差。(a) Bikeshow\_ce; (b) Boat\_cel; (c) Car2; (d) CarScale; (e) Dancer; (f) Singer2

Fig. 6 Center location error of different algorithms on partial videos. (a) Bikeshow\_ce; (b) Boat\_cel; (c) Car2; (d) CarScale; (e) Dancer; (f) Singer2

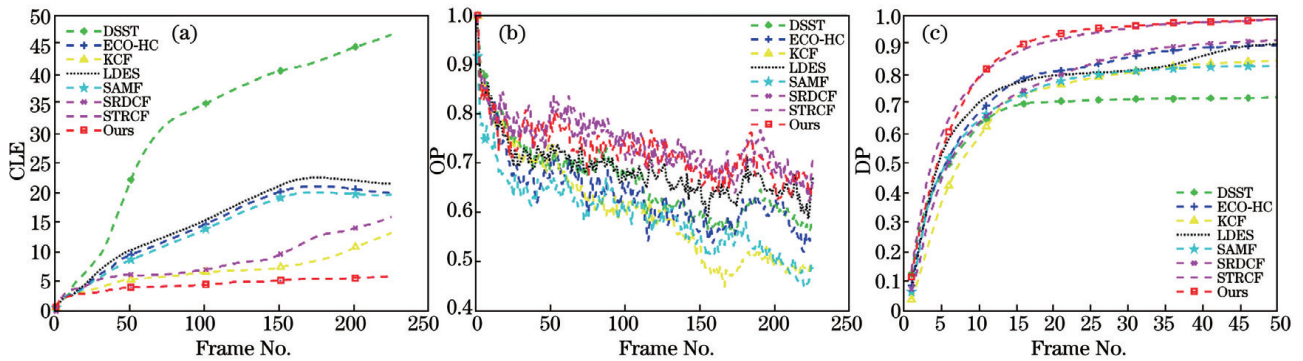


图 7 均值曲线图。(a)中心位置误差;(b)重叠率;(c)距离精确度

Fig. 7 Mean value plots. (a) CLE; (b) OP; (c) DP

度变化、提高跟踪性能。同时,在原先位置滤波器模型的更新基础上,融合前序几帧的背景信息,以加强目标特征、提高模型的稳定性。实验数据表明,本文算法能有效地应对尺度变化和平面内旋转带来的挑战,不仅对平面内旋转和尺度变化有很好的效果,而且能够达到实时要求,具有很好的鲁棒性。本文算法虽然提高了跟踪的精确度,但是速度却受到了较大影响,这个问题将是下一步研究的方向。

参 考 文 献

[1] Wu X J, Xu T Y, Xu W B. Review of target tracking algorithms in video based on correlation filter [J]. Command Information System and Technology,

2017, 8(3): 1-5.

吴小俊, 徐天阳, 须文波. 基于相关滤波的视频目标跟踪算法综述 [J]. 指挥信息系统与技术, 2017, 8(3): 1-5.

[2] Zhu W Q, Liu Y, Bian L, et al. Survey on object tracking method base on generative model [J]. Microprocessors, 2017, 38(1): 41-47.

朱文青, 刘艳, 卞乐, 等. 基于生成式模型的目标跟踪方法综述 [J]. 微处理机, 2017, 38(1): 41-47.

[3] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters [C]// 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York:



- IEEE Press, 2010: 2544-2550.
- [4] 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.
- [5] Danelljan M, Häger G, Khan S F, et al. Accurate scale estimation for robust visual tracking [C] // *Proceedings of the British Machine Vision Conference 2014*, September, 2014, Nottingham. Nottingham: British Machine Vision Association, 2014: 1-11.
- [6] Danelljan M, Hager G, Khan S F, 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.
- [7] Danelljan M, Robinson A, Khan S 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.
- [8] Danelljan M, Bhat G, Khan F S, et al. ECO: efficient convolution operators for tracking[C]//2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6931-6939.
- [9] Mukhopadhyay S, Basu S, Stagg M, et al. Computer implemented system and method for high performance visual tracking: US10757369[P]. 2020-08-25.
- [10] Li B, Wu W, Wang Q, et al. SiamRPN++: evolution of Siamese visual tracking with very deep networks[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: 4277-4286.
- [11] Liu M M, Pei D, Liu J, et al. Filter tracking based on time regularization and background-aware [J]. *Laser & Optoelectronics Progress*, 2019, 56(23): 231503.  
刘明明, 裴东, 刘举, 等. 基于时间正则化及背景感知的滤波器跟踪[J]. *激光与光电子学进展*, 2019, 56(23): 231503.
- [12] Xu J X, Wu X D, Yang K Y. Context-aware correlation filter tracking based on Gaussian output constraint[J]. *Laser & Optoelectronics Progress*, 2020, 57(4): 041508.  
许敬祥, 伍雪冬, 杨开运. 基于高斯输出约束的上下文相关滤波跟踪[J]. *激光与光电子学进展*, 2020, 57(4): 041508.
- [13] Shen Y L, Wu Z D, Zhao R J, et al. Long-term object tracking based on model updating and fast re-detection[J]. *Acta Optica Sinica*, 2020, 40(3): 0315002.  
沈玉玲, 伍忠东, 赵汝进, 等. 基于模型更新与快速重检测的长时目标跟踪[J]. *光学学报*, 2020, 40(3): 0315002.
- [14] Hare S, Golodetz S, Saffari A, et al. Struck: structured output tracking with kernels [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(10): 2096-2109.
- [15] Zhang S L, Wei L, Xing W W, et al. Learning scale-adaptive tight correlation filter for object tracking[J]. *IEEE Transactions on Cybernetics*, 2018, 50(1): 270-283.
- [16] Li Y, Zhu J K. A scale adaptive kernel correlation filter tracker with feature integration[M]//Agapito L, Bronstein M M, Rother C. *Computer vision-ECCV 2014 workshops. Lecture notes in computer science*. Cham: Springer, 2015, 8926: 254-265.
- [17] Kart U, Lukežič A, Kristan M, et al. Object tracking by reconstruction with view-specific discriminative 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: 1339-1348.
- [18] Zhang M D, Wang Q, Xing J L, et al. Visual tracking via spatially aligned correlation filters network[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11207: 484-500.
- [19] Wang N, Song Y B, Ma C, et al. Unsupervised deep tracking[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: 1308-1317.
- [20] Bhat G, Danelljan M, van Gool L, et al. Learning discriminative model prediction for tracking[C]//2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 6181-6190.