

〈图像处理与仿真〉

## 基于上下文感知和尺度自适应的实时目标跟踪

夏爱明, 伍雪冬

(江苏科技大学 电子信息学院, 江苏 镇江 212003)

**摘要:** 针对传统核相关滤波视觉目标跟踪算法在快速运动、背景杂波、运动模糊等情况下跟踪精度低且不能处理尺度变化的问题, 提出了一种基于上下文感知和尺度自适应的实时目标跟踪算法。该算法在核相关滤波算法框架的基础上, 引入了上下文感知和尺度自适应方法, 增加了背景信息且能够处理目标的尺度变化。首先, 利用融合了 fHOG (fusion histogram of oriented gradient)、CN (color names) 和灰度的特征对目标区域进行采样, 训练一个二维位移滤波器, 然后, 在目标区域建立尺度金字塔, 利用 fHOG 对目标区域进行多尺度采样, 训练一个一维尺度滤波器, 最后, 在模型更新阶段改进了更新策略。在标准数据集 OTB-2015 上对 100 组视频序列进行的试验结果表明, 提出的算法比基准算法 (kernel correlation filter, KCF) 精度提高了 13.9%, 成功率提高了 14.2%, 且优于实验中对比的其他跟踪算法。在尺度变化、运动模糊、快速运动等条件下, 提出的算法在准确跟踪的同时, 能够保持较高的速度。

**关键词:** 相关滤波; 特征融合; 上下文感知; 尺度自适应

中图分类号: TP391.41 文献标志码: A 文章编号: 1001-8891(2021)05-0429-08

## Real-time Object Tracking Based on Context Awareness and Scale Adaptation

XIA Aiming, WU Xuedong

(School of Electronics and Information, Jiangsu University of Science and Technology, Zhenjiang 212000, China)

**Abstract:** Because the traditional kernel correlation filter algorithm for visual object tracking has low tracking accuracy under fast motion, background clutter, and motion blurring conditions and cannot deal with scale changes, a real-time object tracking algorithm based on context awareness and scale adaptation is proposed. Based on the kernel correlation filter algorithm framework, context-aware and scale-adaptive methods are introduced to add background information and handle changes in the scale of the target. First, the target region is sampled using the features of the fusion histogram of oriented gradient (fHOG), color names (CN), and gray, and a two-dimensional translation filter is trained. Then, a scale pyramid is established in the target area and multi-scale sampling is performed using fHOG on the target area. Following this, a one-dimensional scale filter is trained. Finally, the update strategy is improved in the model updating stage. The experimental results of 100 sets of video sequences in the standard OTB-2015 dataset show that the proposed algorithm showed an improvement in the accuracy by 13.9% as compared with the benchmark algorithm (kernel correlation filter, KCF), and the success rate improved by 14.2%, which is superior to that of other comparison-tracking algorithms considered in the experiment. Under the conditions of scale change, motion blur, and fast motion, the proposed algorithm can maintain a high speed with accurate tracking.

**Key words:** correlation filters, feature fusion, context-aware, scale-adaptive

收稿日期: 2019-06-08; 修订日期: 2021-03-19.

作者简介: 夏爱明 (1985-), 男, 硕士研究生, 主要研究方向为视觉目标跟踪。E-mail: 362931408@qq.com.

通信作者: 伍雪冬 (1975-), 男, 博士, 教授, 主要研究方向为计算机视觉, 能源预测与环境经济调度等。E-mail: woolcn@163.com.

基金项目: 国家自然科学基金 (61671222)。

## 0 引言

目标跟踪是计算机视觉领域极具挑战性的课题,在服务型机器人、视频监控、人机交互、智能交通等方面有着广泛的应用前景。但在实际的场景中,存在着目标尺度变化、遮挡、姿态变化、光照变化等问题,这些问题影响了目标跟踪的准确性和成功率,因此如何解决这些问题成为视觉目标跟踪方向的研究热点<sup>[1-2]</sup>。

针对在实际场景中存在的问题,研究人员提出了多种算法,但是目前还没有通用的算法可以解决所有问题。总体而言,现有的目标跟踪算法主要分为生成式算法和判别式算法<sup>[3-4]</sup>。

近年来,基于判别式的相关滤波跟踪算法,因其结构简单和速度快而得到了广泛的关注<sup>[5-8]</sup>。Bolme等<sup>[9]</sup>首次将相关滤波引入目标跟踪领域,并提出最小均方误差和(minimum output sum of squared error, MOSSE)滤波器,采用简单的灰度特征即达到了快速高效的跟踪。Henriques等<sup>[10]</sup>提出了经典的核相关滤波算法(kernel correlation filter, KCF),采用方向梯度直方图(histogram of oriented gradient, HOG)<sup>[11]</sup>特征并引入循环采样、核方法和岭回归分类器,算法高速且精度高,成为了后续众多算法的基础,但是该算法没有考虑尺度变化和背景信息,对尺度变化、背景杂波等不够鲁棒。Danelljan等<sup>[12]</sup>提出了自适应颜色属性跟踪算法(color names, CN),利用颜色名<sup>[13]</sup>作为特征进行目标跟踪,取得了不错的效果,但是特征单一,对光照变化不够鲁棒。Danelljan等<sup>[14]</sup>提出了尺度自适应跟踪算法(discriminative scale space tracker, DSST),利用降维后的HOG特征和灰度特征进行位置跟踪,利用HOG特征进行尺度跟踪,实现了对目标尺度的鲁棒跟踪,但是其尺度金字塔尺度较粗,在尺度较细时,跟踪不够鲁棒。Mueller等<sup>[15]</sup>提出了上下文感知相关滤波算法(context-aware correlation filter tracker, CACF),在目标周围提取图像块作为训练样本,增加了训练样本的背景信息,且速度损失较小,但是其更新策略不够鲁棒。

本文针对核相关滤波算法没有考虑尺度变化、背景信息较少的问题,提出了一种改进的基于核相关滤波的实时视觉目标跟踪算法(improved kernel correlation filter, IKCF),该算法引入上下文感知相关滤波来增加背景信息,并改进其更新策略,显著提高其精度和成功率,将HOG特征替换为融合fHOG、CN和灰度的特征,获取更多的特征信息,增强特征的鲁棒性,采用一维尺度金字塔尺度滤波器估计尺度,二维位置滤波器估计位置,并改进尺度金字塔的尺度

粗细,引入了尺度估计且保证算法的实时性。

## 1 基于核相关滤波的改进算法

核相关滤波算法通过对单个样本进行循环移位形成丰富的样本库,利用样本库建立目标表观模型,训练一个正则化的最小二乘(岭回归)分类器,利用核函数在傅里叶域进行分类器训练和响应检测,得到滤波器模板和目标在新一帧的位置,实现了目标的在线训练与跟踪。虽然KCF算法取得了不错的跟踪效果,但是其跟踪准确性和成功率还有进一步提升的空间。

本文针对KCF算法在目标尺度变化、快速运动、背景杂波、运动模糊等情形下出现跟踪失败或跟踪精度不高的问题,对KCF算法在目标特征、背景信息、跟踪策略和更新策略方面进行了改进。

### 1.1 特征融合

在目标跟踪过程中强大的特征可以显著提高跟踪的准确度和检测的性能。

KCF算法在CSK(circulant structure of tracking-by-detection with kernels)算法<sup>[16]</sup>的基础上将灰度特征替换为HOG特征,性能得到了大幅地提升,CN在CSK的基础上用颜色名特征对灰度特征进行了替换,并利用主成分分析技术(principal component analysis, PCA)将11维颜色特征降为2维,取得了很好的效果,可见HOG特征和CN特征对目标表观建模有很强的表达能力。

fHOG<sup>[17]</sup>特征是将36维HOG特征缩减为32维特征而得,包括18维对比度敏感的方向通道、9个对比度不敏感的方向通道、4个纹理通道和一个全零通道,这种改进的HOG特征保持了传统HOG特征判别能力,且缩减了通道维数,从而减少了计算量,加快了跟踪速度,本文去除全零通道,将fHOG缩减为31维。颜色名特征(CN)是人类用来表示世界上颜色的语言标签,包含11个基本的颜色术语:黑色、蓝色、棕色、灰色、绿色、橙色、粉色、紫色、红色、白色和黄色,这种描述比常用的颜色(如RGB、Hue等)具有更好的表达能力。本文没有采用CN算法的降维方法,而仅保留了其中的10维通道。原始的灰度像素特征虽然没有HOG和CN特征表达的信息多,但是其特征表达能力仍不能忽略。

由于核相关函数只需要计算点积和向量范数,因此可以对图像特征应用多个通道。本文中将31维的fHOG、10维的CN和1维的灰度像素共42个通道连接到一个向量中,使得本文提出的算法可以使用更强大的特征,特征的可视化见图1。

## 1.2 环境上下文感知相关滤波

目标跟踪通常在第一帧人工给定一个矩形框，根据给定的矩形框进行跟踪。基于核相关滤波的算法一般都是在给定矩形框的基础上放大 1~2 倍建立搜索区域，但是这样做会丢失很多背景信息，而目标周围的背景对于跟踪的性能有很大影响，如相似物体的背景杂波等。文献[15]提出了在训练阶段向滤波器添加上下文环境信息的环境上下文相关滤波跟踪算法，增加更多的背景信息，采样区域如图 2 所示。

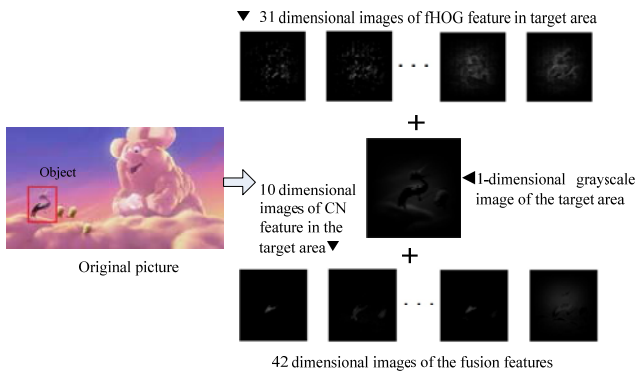


图 1 融合特征的可视化

Fig.1 Visualization of the fusion features



(a) 传统采样区域

(b) 上下文采样区域

(a) Traditional sampling area (b) Context sampling area

图 2 采样区域

Fig.2 Sampling area

在每一帧中，对目标样本  $x$  周围的  $k$  个上下文图像块  $a_j(j=1,2,\dots,k)$  进行采样，它们对应的循环矩阵为  $A_j(j=1,2,\dots,k)$ ，这些上下文图像块可以视为硬性负样本。它们以各种干扰物和不同背景的形式包含目标周围的背景。这样可以学习一个包含上下文样本的滤波器，使得滤波器对目标产生较高响应，对上下文背景样本响应接近零。包含环境上下文图像块的最小化目标函数为：

$$\min_w \sum_i^m (y_i - f(x_i))^2 + \lambda \|w\|^2 + \lambda_1 \sum_{j=1}^k \|A_j w\|^2 \quad (1)$$

式中： $w$  为滤波器； $[y_i - f(x_i)]^2$  为一个损失函数； $\lambda$  和

$\lambda_1$  为防止过拟合的正则化参数； $\lambda \|w\|^2 + \lambda_1 \sum_{j=1}^k \|A_j w\|^2$  为正则化项。

上式在单通道特征原始域中的解为：

$$\hat{w} = \frac{\hat{x} \odot \hat{y}}{\hat{x}^* \odot \hat{x} + \lambda + \lambda_1 \sum_{j=1}^k \hat{a}_j^* \odot \hat{a}_j} \quad (2)$$

式中： $\wedge$  表示变量的快速傅里叶变换形式； $*$  表示变量的复共轭； $\odot$  表示元素之间的点积； $x$  为搜索区域的图像块； $y$  为目标的期望输出。

由于采用的原始域，训练的滤波器为  $\hat{w}$ ，因此只要更新一个系数即可，采用线性插值更新滤波器：

$$\hat{w}^{t+1} = (1 - \theta) \hat{w}^{t-1} + \theta w^t \quad (3)$$

式中： $t$  表示图像的第  $t$  帧； $\theta$  为学习率。

## 1.3 跟踪过程

### 1.3.1 位移相关滤波

基于 HOG 特征的相关滤波跟踪器（比如 KCF，DSST）通过提取灰度图像的方向梯度图对特征进行表达，对运动模糊和光照变化鲁棒性高，但对颜色跟踪效果不鲁棒，CN 特征侧重于颜色信息，对颜色鲁棒性高，对运动模糊和光照变化效果不佳，两者具有互补作用，提出的算法将两者结合并加上灰度特征，在 KCF 框架的基础上引入环境上下文感知，用于 2 维位移相关滤波训练和检测。将表示特征的多个通道连接到一个向量中，连接后的特征向量为  $x_c = [x_1, x_2, \dots, x_{42}]$ ，则公式(2)为：

$$\hat{w} = \frac{\hat{x}_c \odot \hat{y}}{\hat{x}_c^* \odot \hat{x}_c + \lambda + \lambda_1 \sum_{j=1}^k \hat{a}_j^* \odot \hat{a}_j} = \frac{A}{B} \quad (4)$$

式中： $A$  和  $B$  分别为位移滤波器模型的分子和分母。

响应的检测公式为：

$$\hat{R} = \sum_{i=1}^c \hat{z}_i \odot \hat{w} \quad (5)$$

为了训练滤波器，提取包含目标的图像块的特征图，利用公式(4)进行训练，得到一个位移滤波器。在下一帧，提取预测目标位置的特征图  $z$  来估计新帧中的位置，利用公式(5)计算相关得分  $\hat{R}$ ，响应最大的位置为目标的新位置。

### 1.3.2 尺度相关滤波

KCF 跟踪算法采用固定的目标搜索区域，没有考虑目标尺度的变化，跟踪过程中目标尺度变化时，容易导致跟踪失败。为了保证跟踪的成功率，需要对目标的尺度变化进行估计。目前尺度估计算法的一种常见思路是将尺度估计并入位移跟踪器，构成三维滤波

器(如 SAMF, scale adaptive with multiple features tracker<sup>[17]</sup>), 而滤波器维数越高需要的计算成本越高, 这将降低跟踪器的速度。理想情况下, 精确的尺度估计方法应该是鲁棒的, 同时计算效率高。文献[14]提出了一种快速尺度估计方法, 通过分别独立学习一个二维位移滤波器和一个一维尺度估计滤波器, 相比融合尺度和位移的三维滤波器拥有更低的计算成本。此方法可以独立选择特征, 而尺度估计加入 CN 特征对于尺度滤波器跟踪意义不大, 还可能降低跟踪的效果, 因此, 提出的算法在这种方法的基础上采用 31 维的 fHOG 特征进行尺度跟踪, 以估计目标图像块的尺度。

### 1) 一维尺度金字塔构建

设  $P$  和  $R$  分别表示当前帧中目标的宽和高,  $S$  表示尺度滤波器中的尺度数, 其值取为 33。以位移滤波器预测的目标位置为中心, 构建尺度金字塔图像块  $J_n$ , 表达式如下:

$$J_n = b^n P \times b^n R, n \in \left\{ -\frac{S-1}{2}, \dots, \frac{S-1}{2} \right\} \quad (6)$$

式中:  $b$  表示各个特征层之间的尺度因子, 在实验中观察发现, 两帧图像间尺度变化不会特别大, 故本文采用细尺度, 取值为 1.002, 通过式(6)可以得到 33 个以目标尺度为中心的尺度金字塔, 如图 3。

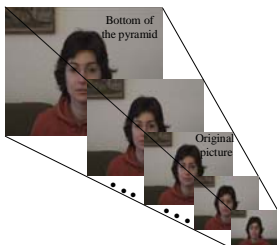


图3 尺度金字塔

Fig.3 Scale pyramid

### 2) 一维尺度估计

在视觉跟踪场景中, 两帧之间的尺度差通常小于平移。因此, 首先应用位移滤波器预测新一帧中的目标位置, 然后利用 31 维 fHOG 特征, 在该位置上提取  $S$  个不同尺度的候选样本, 将这些样本伸缩到和尺度滤波器尺寸一样大小, 采用一维高斯函数  $y_s$  作为尺度滤波器输出:

$$y_s = \exp \left[ -\frac{(i-17)^2}{\sigma_s^2} \right], i \in \{1, 2, \dots, 33\} \quad (7)$$

利用下式计算尺度相关滤波器响应值  $R_s$ , 以最大响应尺度为当前目标尺度。

$$\hat{R}_s = \frac{\sum_{i=1}^d \hat{x}_s^* \odot \hat{y}_s \odot \hat{z}_s}{\hat{x}_s \odot \hat{x}_s^* + \lambda_s} = \frac{C \odot \hat{z}_s}{D + \lambda_s} \quad (8)$$

式中:  $x_s$  为上一帧的特征图像块;  $z_s$  为新目标位置的图像块。 $C$  和  $D$  分别为尺度模型的分子部分和分母部分。

## 1.4 更新策略

在 KCF 中利用简单线性插值作为模型更新策略被证明不是最优的<sup>[12]</sup>, 因为在更新模型时只考虑当前帧, 而且之前帧在模型中的贡献指数衰减过快。为了提高 KCF 的鲁棒性, 本算法采用了 CN 中提出的更新策略, 同时考虑了当前帧和之前所有帧中的目标图像块, 对分子分母分别更新, 公式(4)和(8)的更新方式如下:

$$A^{t+1} = (1 - \eta_1) A^{t-1} + \eta_1 A^t \quad (9)$$

$$B^{t+1} = (1 - \eta_1) B^{t-1} + \eta_1 B^t \quad (10)$$

$$C^{t+1} = (1 - \eta_2) C^{t-1} + \eta_2 C^t \quad (11)$$

$$D^{t+1} = (1 - \eta_2) D^{t-1} + \eta_2 D^t \quad (12)$$

式中:  $\eta_1$ 、 $\eta_2$  分别为位移滤波器和尺度滤波器的学习率。

## 2 算法流程

基于上述策略, 文中算法的跟踪步骤如下。

输入:

当前帧图像  $I_t$ ;

先前目标位置  $p_{t-1}$  和尺度  $s_{t-1}$ ;

位移模型  $A_{t-1}$ 、 $B_{t-1}$  和尺度模型  $C_{t-1}$ 、 $D_{t-1}$ 。

输出:

当前帧目标位置  $p_t$  和尺度  $s_t$ ;

更新位移模型  $A_t$ 、 $B_t$  和尺度模型  $C_t$ 、 $D_t$

位移估计:

1) 在当前帧图像  $I_t$  中位置  $p_{t-1}$  和尺度  $s_{t-1}$  处提取位移样本  $z_{trans}$ ;

2) 在公式(4)中利用  $A_{t-1}$ 、 $B_{t-1}$  计算  $\hat{w}_{t-1}$ ;

3) 在公式(5)中利用  $w_{t-1}$  和  $p_{t-1}$  计算位移响应  $\hat{R}$ ;

4) 将响应  $\hat{R}$  最大处设为目标的新位置  $p_t$ 。

尺度估计:

5) 在当前帧图像  $I_t$  中位置  $p_t$  和尺度  $s_{t-1}$  处提取尺度样本  $z_{scale}$ ;

6) 在公式(8)中利用  $C_{t-1}$ 、 $D_{t-1}$  和  $z_{scale}$  计算尺度响应  $\hat{R}_s$ ;

7) 将响应  $\hat{R}_s$  最大处设为目标的新尺度  $s_t$ 。

模型更新

8) 在当前帧图像  $I_t$  中位置  $p_t$  和尺度  $s_t$  处提取位移样本  $\hat{x}_c$  和尺度样本  $\hat{x}_s$ ;

9) 利用公式(9)、(10)更新位移模型  $A_t$ 、 $B_t$ ;

10) 利用公式(11)、(12)更新尺度模型  $C_t$ 、 $D_t$ 。当输入下一帧图像  $I_{t+1}$  时，重复以上步骤，算法的整体跟踪框架如图4所示。

### 3 实验验证与分析

#### 3.1 特征和参数设置

所有训练集视频的参数设为固定，以防止过拟合。正则化参数设置为  $\lambda=0.0001$ ， $\lambda_1=0.01$ 。

将位移滤波器期望的相关输出的标准偏差设置为目标大小的1/10。滤波器大小设置为初始目标大小的3倍。

位移滤波器的二维高斯标签的带宽  $\sigma=0.4$ ，尺度滤波器的一维高斯标签带宽  $\sigma_1=0.25$ 。位移和尺度滤波器的学习率分别设置为  $\eta_1=0.005$ ， $\eta_2=0.03$ 。

位移滤波器中使用的 fHOG 特征单元尺寸采用  $4 \times 4$ ，尺度滤波器使用的 fHOG 特征单元尺寸采用  $1 \times 1$ 。尺度参数设置参照第1.3.2节。

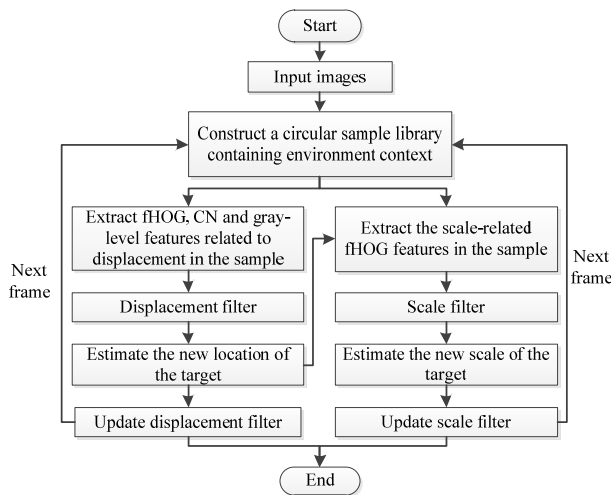


图4 本文算法流程图

Fig.4 The flow chart of the proposed algorithm

#### 3.2 实验设置

本文的所有代码在 MATLAB R2016a 上编译实现。在 CPU 为 Intel G630, 主频 2.70 GHz, 内存为 4 GB 的台式计算机上运行。

##### 3.2.1 数据集

Wu 等<sup>[18]</sup>建立了一套评估目标跟踪算法的视频测试数据集 OTB-2015，其中包含 100 个视频，涵盖了光照变化、颜色变化、快速移动、遮挡、尺度变化等一系列视频目标跟踪的过程中常见的问题，具有较高的代表性，已成为目标跟踪领域的通用测试集，提出的算法在此数据集中的全部 100 个视频序列中运行。

##### 3.2.2 评估方法

为了评估跟踪算法的性能，实验中采用 OTB-2015 测试数据集中两种通用评估方法即精度图 (precision

plots) 和成功率图 (success plots) 作为定量评价标准，评估方式采用一次跟踪评估 (one-pass evaluation, OPE)，即所有的视频序列只跟踪一次。

在精度图中，根据某范围内的阈值来绘制平均距离精度 (distance precision, DP)，使用的指标是中心位置误差 (center location error, CLE)，CLE 定义为算法预测的目标框中心坐标与视频序列中人工标注的框中心坐标之间的平均欧式距离。DP 为 CLE 小于某一阈值的帧数占总帧数的比值，计算公式为  $V_{DP} = m/n$ ， $m$  为算法预测的位置与目标的真实位置小于一定阈值时的帧数， $n$  为跟踪视频的总帧数，实验采用标准阈值 (20 像素)。

成功率图的评估标准是成功率 (Success Rate, SR)，即跟踪目标框与人工标注目标框之间的重叠率超过一定阈值的帧数与跟踪视频的总帧数的比值，用重叠率  $S_R$  表示，计算公式为：

$$S_R = \text{area}(R_T \cap R_G) / \text{area}(R_T \cup R_G) \quad (13)$$

式中： $R_T$  为算法预测目标框； $R_G$  为人工标注目标框， $\text{area}(R_T \cap R_G)$  表示区域重叠面积， $\text{area}(R_T \cup R_G)$  表示区域合并面积。实验中采用标准阈值 0.5，大于该值认为跟踪成功。

为了评估算法的实时性，采用平均跟踪速度进行评估，平均跟踪速度定义为视频帧数除以算法跟踪图像总耗时，单位为 fps (帧/每秒)，文中以文献[19-20]中定义的 15 fps 作为实时性的判别标准。

#### 3.3 与先进算法的对比与分析

实验中选择了 OTB-2015 中提供的可用于运行 100 组视频的 7 种算法，分别为 CT (compressive tracking)、IVT (incremental learning for visual tracking)、DFT (distribution fields for tracking)、ASLA (adaptive structural local sparse appearance model)、L1APG (L1 tracker using accelerated proximal gradient approach)、ORIA (online robust image alignment)、CSK (circulant structure of tracking-by-detection with kernels)，新加入了近年来比较先进的算法 3 种，分别为 KCF、CN 和 DSST。IKCF 为本文提出的算法。为了保证测试的客观性，所有的对比算法都采用其原始代码和参数，并在相同的测试环境中进行。

##### 3.3.1 不同算法的特点

表 1 列出了最新 6 种算法的特点，从表 1 中可看出，本文算法采用了更强的特征来表征目标，并且加入了尺度估计。

##### 3.3.2 不同算法帧率

表 2 列出了全部 11 种算法在不同视频上的跟踪速

度,表中第1行为视频序列,最后1列为运行速度,单位fps,表中加粗字体表示其对应的算法达到了实时性的要求。从表2中可以看出,L1APG算法的跟踪速度仅约0.6fps,远不能达到实时要求,而本文提出的算法可以达到26fps,与DSST相近。

表1 6种最新算法的特点

Table 1 The features of the 6 latest algorithms

Algorithm	Feature	Scale
CSK	gray	No
KCF	HOG	No
DSST	Displacement: HOG, gray; Scale: HOG	Yes
CN	gray, CN	No
ASLA	-	Yes
IKCF	Displacement: gray, HOG, CN; Scale: HOG	Yes

表2 11种算法运行速度对比

Table 2 Comparison of 11 algorithms' running speeds

Algorithm						fps
	CarDark	Car4	David2	Sylvester	Trellis	Average speed
CT	86.93	73.14	54.85	55.39	63.84	66.83
IVT	15.70	16.65	15.20	16.58	17.18	16.26
DFT	13.67	4.07	10.69	7.56	6.38	8.47
ASLA	1.70	1.21	1.56	1.45	1.65	1.51
L1APG	0.77	0.43	0.56	0.57	0.53	0.57
ORIA	22.25	8.19	16.61	4.26	5.61	11.38
CSK	488.46	52.75	482.92	112.83	67.21	240.83
KCF	331.10	22.44	291.70	91.07	42.26	155.72
DSST	54.63	5.32	53.38	14.94	6.96	27.05
CN	155.09	30.38	258.27	81.25	16.42	108.28
IKCF	39.23	9.47	46.56	24.85	9.24	25.87

### 3.3.3 定量分析

将11种算法在标准OTB-2015视频数据集上运行,得到排名前10的10种算法精度曲线图和成功率曲线图图5。图5为全部100组视频的整体精度和成功率图。从图中可看出,文中提出的算法在整体跟踪精度和成功率上均是最优的。在精度图中,提出的算法精度为0.793,比排名第二的基准算法(KCF)提高了13.9%。在成功率图中,本文算法成功率比基准算法(KCF)提高了14.2%,比排名第二的DSST提高了2.6%。

实验中还绘制了不同环境下的精度图和成功率图,从中选出了快速运动、背景杂波、运动模糊、尺度变化4种环境下的精度图和成功率图,列在图6中。从图中可以看出,在这4种跟踪环境下:①提出的算法在精度和成功率上均比基准算法(KCF)有较大提高;②提出的算法的精度均排在第一;③在跟踪成功率方面,除了尺度变化环境下的成功率略低于DSST,其余跟踪环境下的成功率均排在第一。

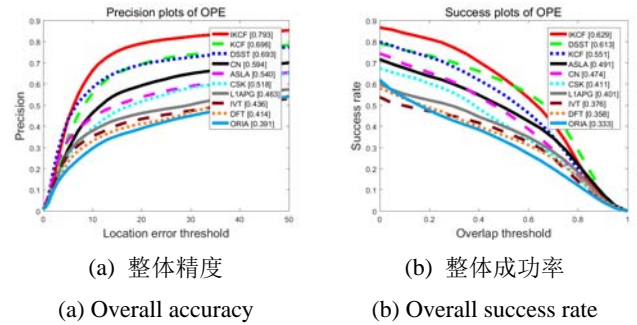


图5 100个视频序列的精度图和成功率图

Fig.5 Accuracy diagram and success rate diagram of 100 video sequences

### 3.3.4 定性分析

为了更直观地体现各个算法的跟踪效果,图7列出了11种对比算法中排名前4位的算法在OTB-2015数据集中部分具有代表性的视频序列的跟踪效果截图。图中视频序列从上到下依次为jogging-1, singer2和shaking。3组视频均存在快速运动、遮挡、运动模糊、尺度变化等多种干扰条件,具有很好的代表性。

1) jogging-1 视频中第76帧左右遇到灯杆的遮挡,在第84帧看到只有本文算法跟踪到目标,其他算法均跟踪失败。

2) singer2 视频中第11帧4种算法都能很好地跟踪,由于目标的快速运动、姿态变化、背景杂波和尺度变化,在15帧左右CN算法跟踪发生了偏离,本文算法和其它算法能够一直跟踪到第366帧,并且跟踪效果最好。

3) shaking 视频序列中由于目标的光照变化、姿态变化、尺度变化和平面内翻转,本文算法在第15帧左右发生了跟踪偏离,但在21帧左右又能重新跟踪到目标,并且跟踪效果较好。

从3组视频序列可以看出,本文算法在保持较高精度的同时,还可以很好地处理目标的尺度变化和周围的环境,说明算法中加入的尺度自适应、上下文感知和更新策略发挥了良好的作用。

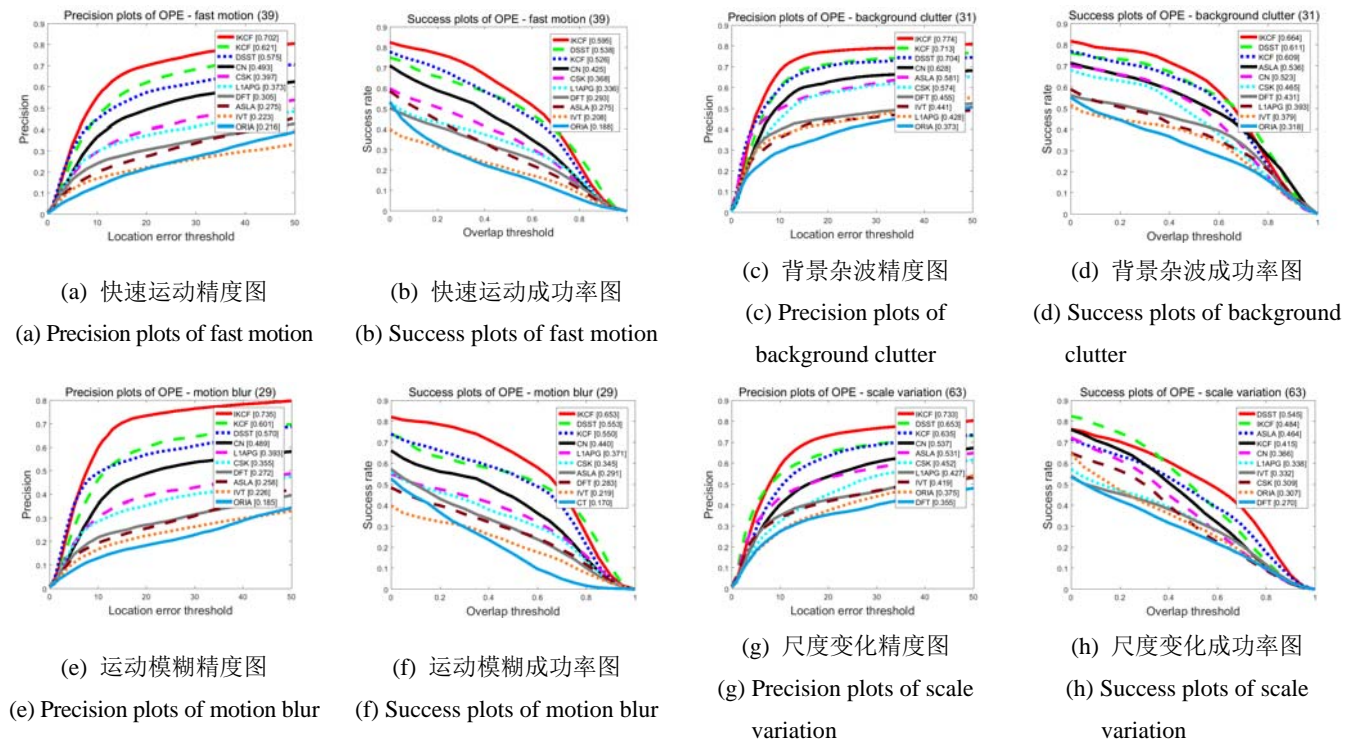


图 6 不同跟踪环境曲线图

Fig.6 The graphs of the different tracking environments

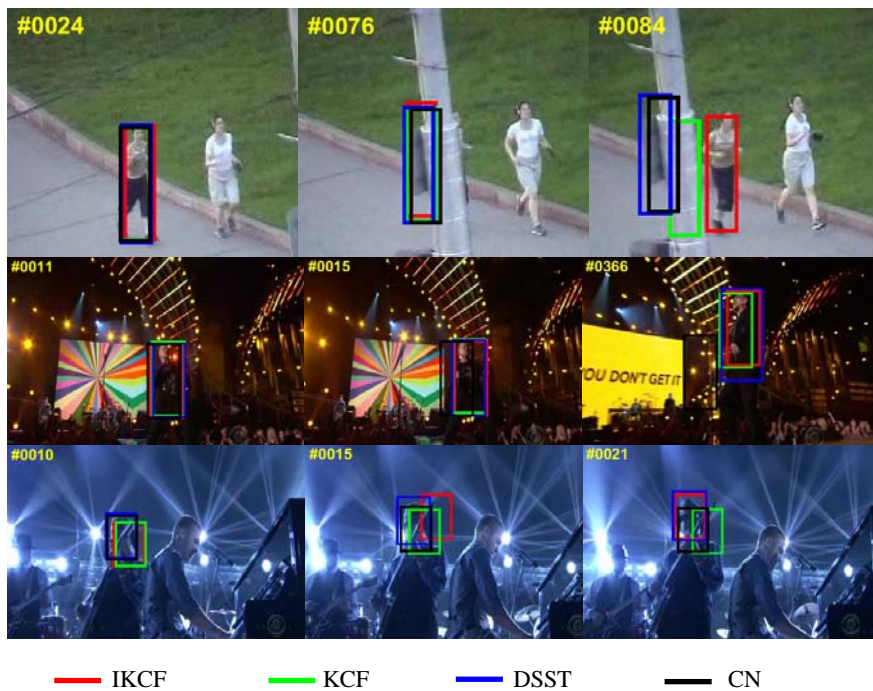


图 7 四种最优算法可视化效果图

Fig.7 Four optimal algorithms visualization

#### 4 结束语

针对 KCF 算法存在背景信息较少和不能很好处理目标尺度变化的问题，在其基础上提出了一种改进算法，增加背景信息的同时，还可以处理目标的尺度

变化。利用融合的多特征训练一个二维位移滤波器估计目标位置，利用 fHOG 特征训练一个一维尺度滤波器估计目标尺度，在模型更新阶段改进了更新策略。在标准数据集上进行的实验表明：提出的算法对基准 KCF 算法有较大改善，与其他当前一些主流跟踪算法

相比整体综合性能较好,验证了提出的算法的可行性和有效性。另外,位移滤波器中使用的3种特征只是简单的矢量链接,没有考虑到不同视频场合下不同特征的主导作用,同时跟踪的成功率还有待进一步提高。

### 参考文献:

- [1] Yilmaz A. Object tracking: a survey[J]. *Acm Computing Surveys*, 2006, **38**(4): 1-45.
- [2] Smeulders A W M, Chu D M, Cucchiara R, et al. Visual tracking: an experimental survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **36**(7): 1442-1468.
- [3] 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):1-5,18.  
WEI Quanlu, LAO Songyang, BAI Liang. Visual Object Tracking Based on Correlation Filters: A Survey[J]. *Computer Science*, 2016, **43**(11): 1-5, 18.
- [4] 孟磊,杨旭.目标跟踪算法综述[J].*自动化学报*,2019,**45**(7): 1244-1260.  
MENG Lu, YANG Xu. A Survey of Object Tracking Algorithms[J]. *Acta Automatica Sinica*, 2019, **45**(7): 1244-1260.
- [5] 徐康,龙敏.增强尺度估计的特征压缩跟踪算法[J].*红外技术*,2018,**40**(12): 1176-1181.  
XU Kang, LONG Min. Feature Compression Tracking Algorithm with Enhanced Scale Estimation[J]. *Infrared Technology*, 2018, **40**(12): 1176-1181.
- [6] 刘教民,郭剑威,师硕.自适应模板更新和目标重定位的相关滤波器跟踪[J].*光学精密工程*,2018,**26**(8): 2100-2111.  
LIU Jiaomin, GUO Jianwei, SHI Shuo. Correlation filter tracking based on adaptive learning rate and location refiner[J]. *Optics and Precision Engineering*, 2018, **26**(8): 2100-2111.
- [7] 叶瑞哲,许卓斌.基于自适应搜索的空时上下文目标跟踪算法[J].*微电子学与计算机*,2018,**35**(6):88-91.  
YE Ruizhe, XU Zhuobin. Spatio-Temporal Context-based Object Tracking Algorithm in Video Surveillance[J]. *Microelectronics & Computer*, 2018, **35**(6): 88-91.
- [8] 胡昭华,钮梦宇,邵晓雯,等.多层深度特征的目标跟踪算法研究[J].*现代电子技术*,2019,**42**(1): 51-56.  
HU Zhaohua, NIU Mengyu, SHAO Xiaowen, et al. Research on object tracking method based on multi-level deep feature[J]. *Modern Electronics Technique*, 2019, **42**(1): 51-56.
- [9] 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 IEEE*, 2010: 2544-2550.
- [10] 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.
- [11] Felzenszwalb P F, Girshick R B, McAllester D, et al. Object detection with discriminatively trained part-based models[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(9): 1627-1645.
- [12] Danelljan M, Shahbaz Khan F, Felsberg M, et al. Adaptive color attributes for real-time visual tracking[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014: 1090-1097.
- [13] Van De Weijer J, Schmid C, Verbeek J, et al. Learning color names for real-world applications[J]. *IEEE Transactions on Image Processing*, 2009, **18**(7): 1512-1523.
- [14] Danelljan M, Häger G, Khan F, et al. Accurate scale estimation for robust visual tracking[C]//*British Machine Vision Conference*, 2014: 1-11.
- [15] Mueller M, Smith N, Ghanem B. Context-aware correlation filter tracking[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 1396-1404.
- [16] Henriques J F, Caseiro R, Martins P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[C]//*European Conference on Computer Vision*, 2012: 702-715.
- [17] LI Y, ZHU J. A scale adaptive kernel correlation filter tracker with feature integration[C]//*European Conference on Computer Vision*, 2014: 254-265.
- [18] WU Y, LIM J, YANG M H. Object Tracking Benchmark[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(9): 1834-1848.
- [19] YUN S, CHOI J, YOO Y, et al. Action-decision networks for visual tracking with deep reinforcement learning[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017: 2711-2720.
- [20] LI Y, FU C, HUANG Z, et al. Keyfilter-aware real-time uav object tracking[C]//2020 *IEEE International Conference on Robotics and Automation (ICRA)*, 2020: 193-199.