

# 基于特征融合和尺度自适应的干扰感知目标跟踪

李双双, 赵高鹏, 王建宇

南京理工大学自动化学院, 江苏 南京 210094

**摘要** 针对复杂场景下单一颜色特征稳健性差、存在类目标干扰及目标尺度变化的问题, 提出了一种基于特征融合和尺度自适应的干扰感知目标跟踪方法。首先, 综合目标、邻域背景、类似干扰区域的三原色 (RGB) 特征和改进的方向梯度直方图 (HOG) 特征计算得到干扰感知目标模型; 在搜索区域内逐像素点计算目标概率图, 然后进行密集采样得到候选目标, 利用目标概率图的概率值与距离值进行加权, 同时定位目标和类似干扰, 并更新目标模型; 采用 RGB 直方图建立尺度模型, 从当前帧图像上截取不同尺度的图像块并计算其 RGB 直方图, 通过与尺度模型比较, 获得最优尺度估计并更新尺度模型。实验结果表明, 提出的方法对复杂场景下的类目标干扰、局部遮挡、尺度变化等均具有很好的适应性, 同时距离精度、重叠精度等指标优于对比算法。

**关键词** 机器视觉; 目标跟踪; 特征融合; 尺度自适应; 干扰感知

**中图分类号** TP391 **文献标识码** A

**doi:** 10.3788/AOS201737.0515005

## Distractor-Aware Object Tracking Based on Multi-Feature Fusion and Scale-Adaption

Li Shuangshuang, Zhao Gaopeng, Wang Jianyu

*School of Automation, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China*

**Abstract** Aiming at the tracking drift problem caused by the RGB feature, similar appearance and scale change in complex scenes, an improved method of distractor-aware object tracking based on multi-feature fusion and scale-adaption is proposed. Firstly, the distractor-aware object models are established base on the RGB feature and the modified (histogram of oriented gradient) HOG feature, which are extracted from object, surrounding background, and distractors. Secondly, the candidates are extracted by dense sampling in likelihood maps, which are obtained by calculating every pixel in the search region. The locations of the target and the distractor are obtained by vote score and distance score, also the model updating method is given. The RGB feature is extracted to establish a model scale, and the multi-scale feature pyramid method is used to get templates at different scales. The optimal scale is obtained by comparison between the model scale and the template scales. The experimental results indicate that the proposed algorithm can well adapt to environmental variation including distractors, partially blocking and scale variation and outperforms the compared tracking methods in terms of the distance precision and overlap precision.

**Key words** machine vision; object tracking; multi-feature fusion; scale adaption; distractor awareness

**OCIS codes** 150.0155; 100.1160; 110.1085

## 1 引 言

视觉目标跟踪是智能视频监控、交通监测、人机交互以及机器人导航等领域的重要技术之一, 其本质是从视频图像序列中估计出感兴趣目标的位置、尺度以及运行轨迹<sup>[1-2]</sup>。依据模型设计机制的不同, 目标跟踪方法大致可以分为两类: 产生式方法和判别式方法<sup>[1-2]</sup>。产生式方法提取目标自身的外观信息, 然后建立目标模型, 搜索与目标模型相似度最高的目标候选区域对目标位置进行更新。其模型的建立可以分为基于核

**收稿日期:** 2016-12-21; **收到修改稿日期:** 2017-01-14

**基金项目:** 江苏省博士后科研自助计划(1601181B)

**作者简介:** 李双双(1990—), 女, 硕士研究生, 主要从事视觉跟踪和运动分析方面的研究。E-mail: 798806912@qq.com

**导师简介:** 赵高鹏(1983—), 男, 博士, 讲师, 主要从事机器视觉方面的研究。

E-mail: zhaogaopeng@njust.edu.cn (通信联系人)

的方法<sup>[3-4]</sup>、基于子空间的方法<sup>[5]</sup>以及基于稀疏表示的方法<sup>[6-7]</sup>。判别式方法不仅利用了目标自身的外观信息,还利用了目标周围的背景信息,从背景中区分并估计出感兴趣的目标区域。其跟踪方法可以分为基于在线 Boosting 的方法<sup>[8]</sup>、基于支持向量机的方法<sup>[9-10]</sup>、基于随机学习的方法<sup>[11]</sup>以及基于判别分析的方法<sup>[12-13]</sup>。

视觉目标跟踪目前已有大量的研究成果,但在复杂场景下仍存在具有挑战性的问题:1)跟踪过程中出现的类目标干扰、光照变化、部分遮挡等外在因素;2)目标自身运动导致的尺度变化等内在因素。通常类目标干扰与待跟踪目标具有相似的特征,如颜色相同或相近,容易导致跟踪漂移。针对该问题,文献[14]提出了一种干扰感知跟踪(DAT)方法。该方法利用三原色(RGB)特征建立干扰感知目标模型,抑制类目标干扰,然后进行密集采样获得候选目标,利用目标概率图的概率值与距离值加权定位目标,得到跟踪结果。但仅由单一特征描述的目标模型,在一些复杂场景下跟踪效果较差。

在前述研究的基础上,本文提出了一种基于特征融合和尺度自适应的干扰感知目标跟踪算法。综合 RGB 颜色特征与改进的方向梯度直方图(HOG)特征,建立目标模型,逐点计算得到目标概率图,从而同步定位目标和类似干扰,抑制了类目标干扰、光照变化等导致的目标跟踪漂移。采用特征金字塔对目标尺度进行估计和更新,抑制了目标尺度变化导致的目标跟踪丢失。

## 2 干扰感知目标跟踪方法

文献[14]提出的干扰感知目标模型  $P$ ,综合目标、干扰、背景区域的 RGB 颜色直方图信息进行计算,其定义为

$$P(x \in \vartheta | b_x) = \lambda_p P_d(x \in \vartheta | O, D, b_x) + (1 - \lambda_p) P_s(x \in \vartheta | O, S, b_x), \quad (1)$$

$$P_s(x \in \vartheta | O, S, b_x) = \begin{cases} \frac{H_O^l(b_x)}{H_O^l(b_x) + H_S^l(b_x)}, & I(x) \in I(O \cup S), \\ 0.5, & \text{otherwise} \end{cases}, \quad (2)$$

$$P_d(x \in \vartheta | O, D, b_x) = \begin{cases} \frac{H_O^l(b_x)}{H_O^l(b_x) + H_D^l(b_x)}, & I(x) \in I(O \cup D), \\ 0.5, & \text{otherwise} \end{cases}, \quad (3)$$

式中  $P_s$  和  $P_d$  分别为目标-背景模型和目标-干扰模型,即在目标-背景区域和目标-干扰区域像素  $x$  属于目标  $\vartheta$  的概率值; $\lambda_p=0.5$  为权重参数, $b_x$  为图像  $I(x)$  的 RGB 颜色直方图的第  $b$  个量化通道, $O$  为给定的矩形目标区域, $S$  为给定的背景区域(为目标矩形区域的 1.9 倍), $D$  为干扰区域, $H_n^l(b_x)$  为区域  $\Omega \in I$  内计算的非归一化颜色直方图  $H$  的第  $b$  个量化通道。

在跟踪过程中更新目标模型,即

$$P_{1:t}(x \in \vartheta | b_x) = \eta P(x \in \vartheta | b_x) + (1 - \eta) P_{1:t-1}(x \in \vartheta | b_x), \quad (4)$$

式中  $\eta=0.1$  为学习率。

在第  $t$  帧时,在搜索区域(第  $t-1$  帧目标区域的 2.9 倍)密集采样,获得第  $t$  帧的候选区域,根据目标模型得到的概率值和距离值计算候选目标得分,即

$$O_t^* = \arg \max_{O_{t,i}} [s_v(O_{t,i}) s_d(O_{t,i})], \quad (5)$$

式中  $O_{t,i}$  表示第  $t$  帧的第  $i$  个候选目标,概率值  $s_v(O_{t,i}) = \sum_{x \in O_{t,i}} P_{1:t}(x \in \vartheta | b_x)$ ,距离值  $s_d(O_{t,i}) =$

$\sum_{x \in O_{t,i}} \exp\left(-\frac{\|x - c_{t-1}\|^2}{2\sigma^2}\right)$ 。对得分进行排序,得到第  $t$  帧的目标,同时设置阈值  $T$ ,满足  $s_v(O_{t,i}) \geq$

$T s_v(O_t^*)$  的候选目标被认为是当前帧的干扰目标。

DAT 跟踪器只利用目标的 RGB 颜色特征,该特征对旋转、姿态改变、部分遮挡等都具有很好的抗干扰性,但对光照变化、类目标干扰、目标尺度变化敏感。图 1 所示为在这些因素影响下 DAT 跟踪器的跟踪结果,其中黄色矩形框代表真实目标区域,蓝色框代表 DAT 跟踪器跟踪到的目标区域,左上角黄色数字代表视频序列的帧数。当光线发生变化时,DAT 跟踪器会发生跟踪漂移,如图 1(a)、(d)所示;当目标背景变化较快、产生背景干扰或遮挡时,DAT 跟踪器跟踪失败,如图 1(b)、(e)所示;当目标存在较大的尺度变化时,



图 1 在(a)(d)光照变化、(b)(e)背景干扰和(c)(f)尺度变化影响下的 DAT 跟踪结果

Fig. 1 Results of DAT method with effect of (a)(d) illumination variation, (b)(e) background distractor and (c)(f) scale variation

容易导致 DAT 跟踪器跟踪丢失或跟踪不精确,如图 1(c)、(f)所示。

### 3 本文方法

DAT 跟踪器只利用了目标的 RGB 颜色特征,容易受光照变化的影响,而目标的 HOG 特征对光照变化不敏感,因此融合了 RGB 特征与改进的 HOG 特征,由多特征融合的干扰感知跟踪模型完成精确定位。由于 DAT 跟踪器对目标尺度变化适应能力差,因而采用特征金字塔方法完成尺度估计和更新。算法流程如图 2 所示。

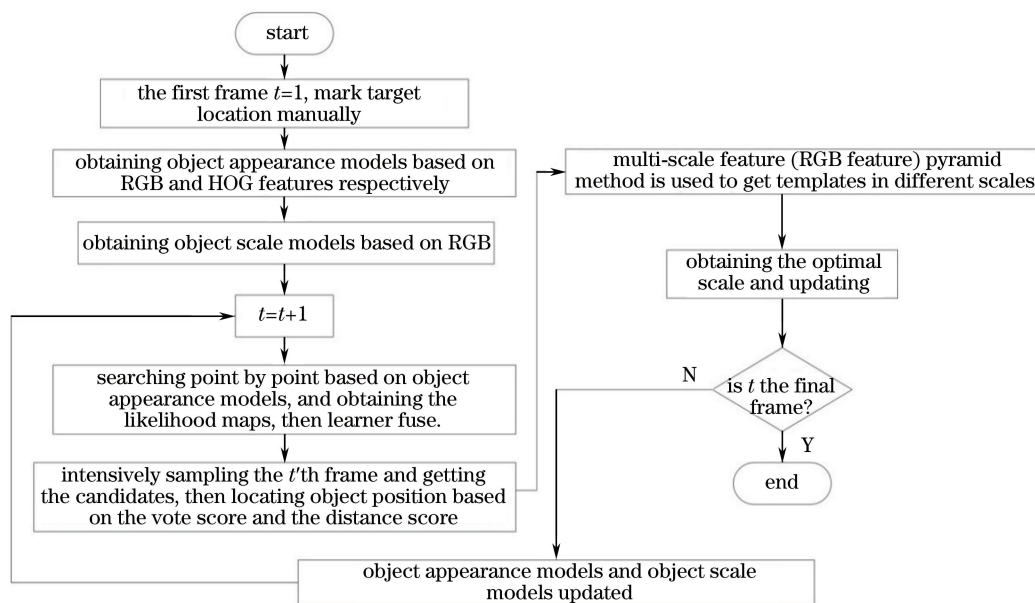


图 2 算法流程图

Fig. 2 Flowchart of the proposed algorithm

#### 3.1 特征表示

颜色直方图描述的是不同色彩在整幅图像中所占的比例,常用的颜色空间是 RGB 颜色空间。首先,将颜色空间划分为若干个小的颜色区间,每个小区间作为直方图的一个 bin;然后,通过计算颜色落在每个小区间内的像素数量就可以得到颜色直方图。

HOG 特征是通过计算和统计图像局部区域的梯度方向直方图来构成特征<sup>[15]</sup>。在给定的细胞单元和块下,不同大小的图像得到的特征维度不同,因此无法直接利用其建立干扰感知目标模型。本文设计了一种简化的 HOG 特征计算方法,首先计算图像各像素的梯度幅值与方向,将梯度方向 360°分为 9 个区间;然后将

梯度幅值归一化后分成 9 个梯度区间,对所有像素的梯度在各个方向区间和幅值区间进行直方图统计。图 3 为一个示例,其中图 3(a)的蓝色框表示目标区域,图 3(b)为该目标区域的改进 HOG 特征。

改进的 HOG 特征提取算法实现过程如下:

1)将源图像进行灰度化处理;

2)计算每个像素的梯度幅值和梯度方向;

3)对梯度幅值进行量化对比,减小梯度值过大或过小对直方图带来的影响;

4)将梯度方向量化为 9 个 bin,在每个方向上将梯度幅值量化为 9 个 bin,从而得到 HOG 特征,特征维数为  $9 \times 9$ 。



图 3 (a)目标区域和(b)对应的改进 HOG 特征

Fig. 3 (a) Object region and (b) corresponding modified HOG features

### 3.2 特征融合的干扰感知目标模型

根据 RGB 颜色信息,分别计算目标区域、背景区域和干扰区域的颜色直方图,从而得到基于 RGB 特征的目标-背景模型与目标-干扰模型,即

$$P_{\text{rgb}_s}(x \in \vartheta | O, S, b_x) = \begin{cases} \frac{H_o^l(b_x)}{H_o^l(b_x) + H_s^l(b_x)}, & I(x) \in I(O \cup S), \\ 0.5, & \text{otherwise} \end{cases}, \quad (6)$$

$$P_{\text{rgb}_d}(x \in \vartheta | O, D, b_x) = \begin{cases} \frac{H_o^l(b_x)}{H_o^l(b_x) + H_d^l(b_x)}, & I(x) \in I(O \cup D), \\ 0.5, & \text{otherwise} \end{cases}. \quad (7)$$

提取改进的 HOG 特征,分别计算目标区域、背景区域和干扰区域的 HOG 特征,得到基于改进 HOG 特征的目标-背景模型与目标-干扰模型,即

$$P_{\text{hog}_s}(x \in \vartheta | O, S, b_x) = \begin{cases} \frac{h_o^l(b_x)}{h_o^l(b_x) + h_s^l(b_x)}, & I(x) \in I(O \cup S), \\ 0.5, & \text{otherwise} \end{cases}, \quad (8)$$

$$P_{\text{hog}_d}(x \in \vartheta | O, D, b_x) = \begin{cases} \frac{h_o^l(b_x)}{h_o^l(b_x) + h_d^l(b_x)}, & I(x) \in I(O \cup D), \\ 0.5, & \text{otherwise} \end{cases}, \quad (9)$$

式中  $h_o^l(b_x)$  为区域  $\Omega \in I$  内计算的改进 HOG 特征  $h$  的第  $b$  个量化通道。

根据(6)~(9)式,得到提出的干扰感知目标模型  $P_{\text{modified}}$ ,采用集合表示为

$$P_{\text{modified}} = \{P_{\text{rgb}_s}, P_{\text{rgb}_d}, P_{\text{hog}_s}, P_{\text{hog}_d}\}. \quad (10)$$

### 3.3 目标定位与模型更新

第  $t$  帧时,根据目标模型  $P_{1:t-1}$ ,得到每个像素点对应的目标概率值,在搜索区域内逐个像素点查找,计算目标概率图,分别可得到 RGB 特征的目标-背景概率图  $L_{\text{rgb}_s}$ 、RGB 特征的目标-干扰概率图  $L_{\text{rgb}_d}$ 、改进 HOG 特征的目标-背景概率图  $L_{\text{hog}_s}$ 、改进 HOG 特征的目标-干扰概率图  $L_{\text{hog}_d}$ ,对 4 幅概率图进行线性融合得到最终的目标概率图为

$$L = \alpha L_{\text{rgb}_s} + \alpha L_{\text{rgb}_d} + \alpha L_{\text{hog}_s} + \alpha L_{\text{hog}_d}, \quad (11)$$



式中  $\alpha$  为概率图融合系数,取值为 0.25。

图 4 为一帧图像目标定位的示例。图 4 中所有图片的尺寸与搜索区域相同,即目标区域的 2.9 倍。图 4(a)为原始图像帧,红色框代表目标区域,绿色框代表背景区域,尺寸为目标区域的 1.9 倍,计算得到相应的 RGB 特征的目标-背景概率图与改进 HOG 特征的目标-背景概率图,分别如图 4(b)和图 4(c)所示。图 4(d)为原始图像帧,黄色框代表干扰区域,与目标区域大小相同,计算得到相应的 RGB 特征的目标-干扰概率图与改进 HOG 特征的目标-干扰概率图,分别如图 4(e)和图 4(f)所示。图 4(h)为融合后的目标概率图,图 4(g)为获得的目标区域。从融合后的图 4(h)可以看出,目标区域更加明显,而干扰区域得到抑制,有利于得到准确的定位结果。

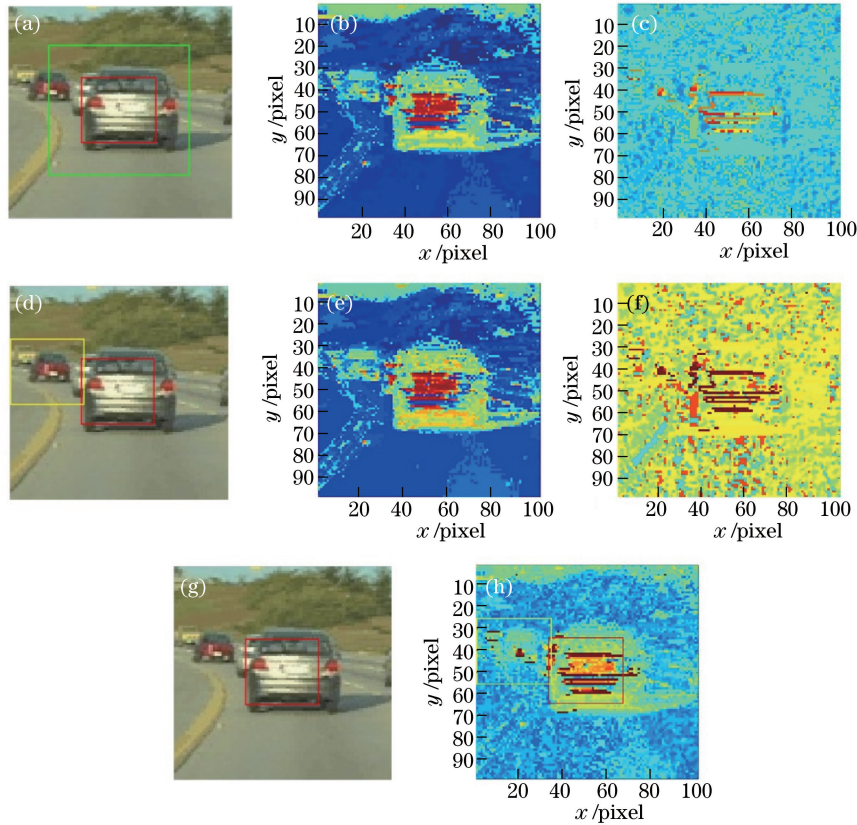


图 4 目标定位示例。(a)目标-背景搜索图;(b)目标-背景 RGB 概率图;(c)目标-背景 HOG 概率图;  
(d)目标-干扰搜索图;(e)目标-干扰 RGB 概率图;(f)目标-干扰 HOG 概率图;(g)目标;(h)目标概率图

Fig. 4 Example of object tracking. (a) Search image of object-surrounding; (b) RGB likelihood map of object-surrounding;  
(c) HOG likelihood map of object-surrounding; (d) search image of object-distractor; (e) RGB likelihood map of  
object-distractor; (f) HOG likelihood map of object-distractor; (g) object; (h) likelihood map of object

获得目标概率图后,在搜索区域内密集采样,即以目标大小的矩形框为采样窗口,以采样窗口 90% 的重叠率在搜索区域内逐行逐列扫描。根据(5)式中的概率值和距离值计算方法逐个对采样区域计算得分,获得每个采样区域的概率值和距离值,然后对得分进行排序,得分最高的候选目标即为当前帧的定位结果,而满足  $s_v(O_{t,i}) \geq Ts_v(O_t^*)$  的采样区域认为是当前帧中存在的干扰目标。

在长时间的跟踪过程中,在得到第  $t$  帧的定位结果后,计算当前帧的目标模型,并采用线性加权来更新目标模型,即

$$P_{1:t}(x \in \mathcal{D} | b_x) = \eta P(x \in \mathcal{D} | b_x) + (1 - \eta_u) P_{1:t-1}(x \in \mathcal{D} | b_x), \quad (12)$$

式中  $\eta_u$  为目标模型更新学习率。

### 3.4 尺度估计与更新

目标的尺度估计是目标跟踪的关键部分,采用特征金字塔对目标尺度进行估计。首先,由第 1 帧图像获得目标位置与尺度,初始化目标尺度模板即 RGB 特征模板  $H_{\text{model}|1}$  和尺寸模板  $S_{\text{model}|1}$ ,同理第  $t - 1$  帧的目

标尺度模板可表示为  $H_{\text{model}|1:t-1}$  和  $S_{\text{model}|1:t-1}$ 。在跟踪到第  $t$  帧图像目标中心位置  $c_{\text{pos}}$  后,以最小尺度  $s_{\text{min}} = 2 \times [(S_{\text{model}|1:t-1}/2)/15]$  为基础,分别得出 33 个不同尺度( $s = i \times s_{\text{min}}$ )的大小。在第  $t$  帧时,以目标中心位置为中心,按照计算得到的不同尺度来截取图像,然后将不同尺度的目标图像通过尺度变换,变成与  $S_{\text{model}|1:t-1}$  大小一样的图,并提取其 RGB 直方图  $H_i$ 。以特征模板  $H_{\text{model}|1:t-1}$  为基础,获得特征差值  $f(i) = |H_i - H_{\text{model}|1:t-1}|$ ,并得到最小值  $f_{\text{min}}$ ,以其尺寸大小作为最优尺寸  $S_{\text{model}}$ ,以其 RGB 特征直方图作为最优特征  $H_{\text{model}}$ 。

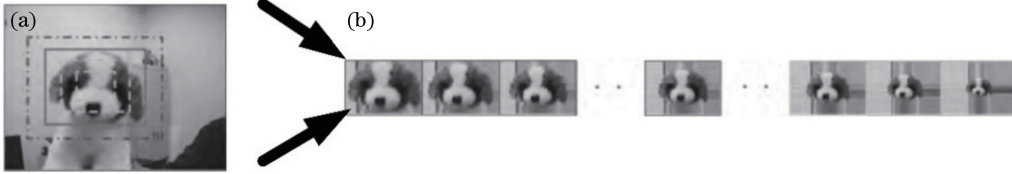


图 5 多尺度特征金字塔。(a)目标图像;(b)不同尺度下的图像

Fig. 5 Multi-scale feature pyramid. (a) Object; (b) screenshots based on different scales

图 5 所示为一个样例,在图 5(a)中黑色实线为尺寸模板  $S_{\text{model}|1:t-1}$  大小,白色虚线为下采样,黑色虚线为上采样。采样得到的图像经过尺寸变换得到与尺寸模板  $S_{\text{model}|1:t-1}$  大小一样的图像,如图 5(b)所示。

当获得的最优尺寸  $S_{\text{model}}$  与模板尺寸  $S_{\text{model}|1:t-1}$  之间的 RGB 直方图特征差异大于阈值  $\lambda_s$  时,即

$$|H_{\text{model}} - H(i_{S_{\text{model}|1:t-1}})| \geq \lambda_s, \quad (13)$$

式中  $i_{S_{\text{model}|1:t-1}}$  代表尺度大小与  $S_{\text{model}|1:t-1}$  值相等的索引,采用线性加权法来更新第  $t$  帧的目标尺度模板,即

$$\begin{cases} H_{\text{model}|1:t} = \epsilon H_{\text{model}} + (1 - \epsilon) H_{\text{model}|1:t-1} \\ S_{\text{model}|1:t} = (1 - \epsilon) S_{\text{model}} + \epsilon S_{\text{model}|1:t-1} \end{cases}, \quad (14)$$

式中  $\epsilon$  为尺度模型更新学习率,反之不更新。

## 4 实验结果与分析

### 4.1 实验环境与参数设置

将提出的跟踪方法与 DAT 跟踪器<sup>[14]</sup>、核循环结构(CSK)跟踪器<sup>[16]</sup>、压缩感知跟踪器(CT)<sup>[17]</sup>、时空上下文(STC)跟踪器<sup>[18]</sup>、多示例学习(MIL)跟踪器<sup>[19]</sup>进行了定性和定量比较。选用公开的标准测试数据集 Benchmark<sup>[20]</sup>,测试视频序列依次为 Boy、RedTeam、Girl2、KiteSurf、Man 和 Car24,其特点如表 1 所示。

表 1 实验视频

Table 1 Experimental videos

Video	Frame number	Characteristic
Boy	602	Scalevariation, motion blur, fast motion
RedTeam	1000	Scalevariation, occlusion, low resolution
Girl2	500	Scalevariation, occlusion, deformation-non-rigid object deformation, motion blur
KiteSurf	84	Illuminationvariation, occlusion
Man	134	Illuminationvariation
Car24	1600	Illuminationvariation, scale variation, background Clutters

本文方法采用 Matlab 实现,进行比较的跟踪器均采用作者公开的代码。本文方法参数设置:候选目标判定阈值  $T=0.5$ ,目标模型更新学习率  $\eta_a=0.1$ ,尺度模型更新学习率  $\epsilon=0.2$ ,尺度更新阈值  $\lambda_s=0.2$ 。

### 4.2 定性评估

图 6 所示为不同算法的跟踪结果(左上角黄色数字代表帧序号)。

1) 快速运动、旋转:图 6(a) Boy 序列中目标运动速度较快且摄像平台的剧烈运动导致运动模糊的出现,压缩感知跟踪器因目标的快速运动在第 265 帧发生漂移,第 462 帧目标发生旋转,CSK 跟踪器与 STC 跟踪器跟踪目标丢失,虽然 DAT 跟踪器、压缩感知跟踪器能全程跟踪目标,但尺度并没有很好的更新,本文算法的全程稳健跟踪充分验证了运动估计模块的优越性,避免了由目标快速运动带来的跟踪漂移。图 6(d)

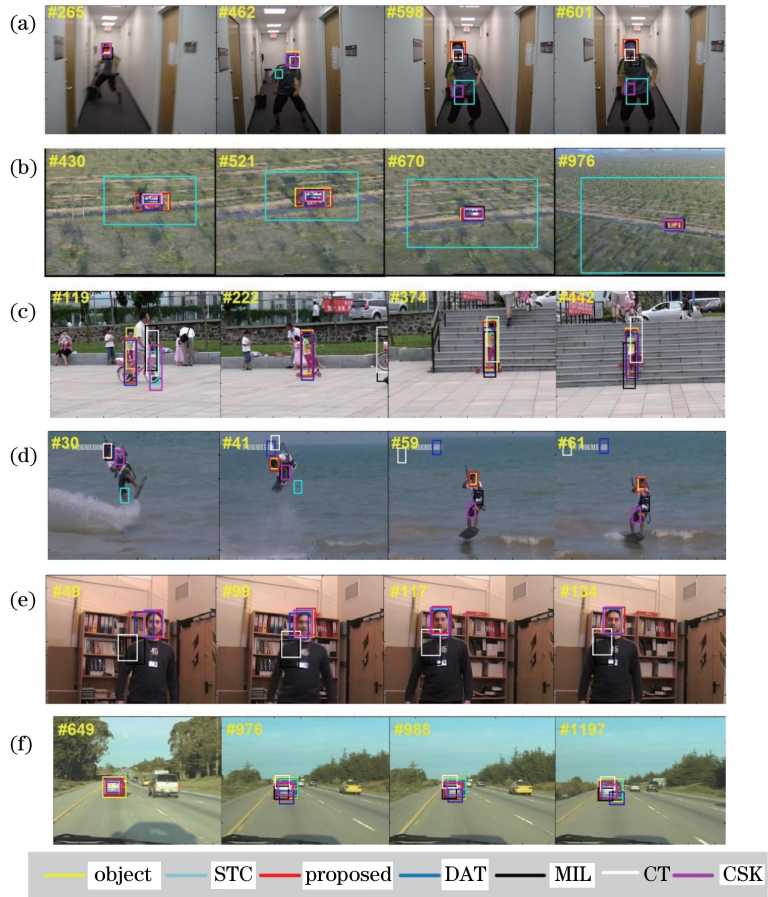


图 6 不同算法的跟踪结果。(a) Boy; (b) RedTeam; (c) Girl2; (d) KiteSurf; (e) Man; (f) Car24

Fig. 6 Tracking results of different algorithms. (a) Boy; (b) Red Team; (c) Girl2; (d) KiteSurf; (e) Man; (f) Car24

KiteSurf 序列在第 30 帧发生旋转,STC、CSK、MIL、压缩感知跟踪器发生跟踪漂移,在第 41 帧时 DAT 跟踪器跟踪目标丢失,而本文算法全程稳健跟踪。

2) 尺度变化:图 6(b) RedTeam 视频序列中尺度改变明显。在第 337 帧时,目标尺度增大,STC 跟踪器尺度更新错误,在第 435 帧随着目标的持续增大,只有本文算法能够准确更新目标尺度,因此本文算法对尺度变化较大的目标具有很好的适应性。

3) 全局和局部遮挡:图 6(c) Girl2 序列显示了本文算法的抗遮挡能力。在第 119 帧时发生全局遮挡,只有本文算法和 DAT 算法能较好地跟踪目标,STC、CT、MIL、CSK 跟踪器因全局遮挡而发生跟踪漂移。

4) 背景干扰和光照变化:图 6(e) Man 序列显示了本文算法抗背景干扰的能力。在第 49 帧光照变强,且在背景干扰的情况下,MIL、压缩感知跟踪器跟踪目标丢失,DAT 跟踪器跟踪目标发生漂移,本文算法与 STC、CSK 跟踪器能很好地全程跟踪目标。

5) 类目标干扰:图 6(f) Car24 序列显示了本文算法抗类目标干扰的能力。视频帧有类目标干扰的情况下,本文算法能够稳定地跟踪目标,在第 976 帧后,由于光照变化,DAT 跟踪器、压缩感知跟踪器跟踪位置发生漂移,本文算法与 STC、CSK 跟踪器能够全程跟踪目标。

### 4.3 定量评估

采用文献中常见的两个评价标准<sup>[21]</sup>对算法进行定量评估。

1) 距离精度。 $P_D = N_{\text{Num}(d < \xi)} / N_{\text{Num}(\text{all})}$ ,其中  $\xi$  为距离阈值,  $d = \sqrt{(x_t - x_c)^2 + (y_t - y_c)^2}$ ,  $(x_t, y_t)$  是跟踪结果的目标位置,  $(x_c, y_c)$  是目标真实位置,取值范围为  $0 \sim 50$ ,  $N_{\text{Num}(\text{all})}$  为数据集总帧数,  $N_{\text{Num}(d < \xi)}$  为距离小于距离阈值的帧数。

2) 重叠精度。 $M = (\gamma_t \cap \gamma_a) / (\gamma_t \cup \gamma_a)$  代表每一帧跟踪结果得分,其中  $\gamma_t$  为跟踪结果区域,  $\gamma_a$  为目标真实区域,  $M > 0.5$  则认为跟踪成功。



当距离精度阈值越小而中心位置误差越大,且重叠精度阈值一定跟踪成功率越高时,表明该跟踪器的跟踪精度越高,稳健性越好。距离精度曲线如图 7 所示,重叠精度如表 2 所示(阈值为 0.5),表中加粗代表最优,下划线代表次优。

从图 7 和表 2 可以看出,本文算法在精确性和稳健性方面优于其他 5 种算法。综合定性评估和定量评估的结果,本文方法对快速运动与目标模糊、背景干扰、光照变化、局部或全局遮挡、尺度变换均具有很好的适应性。

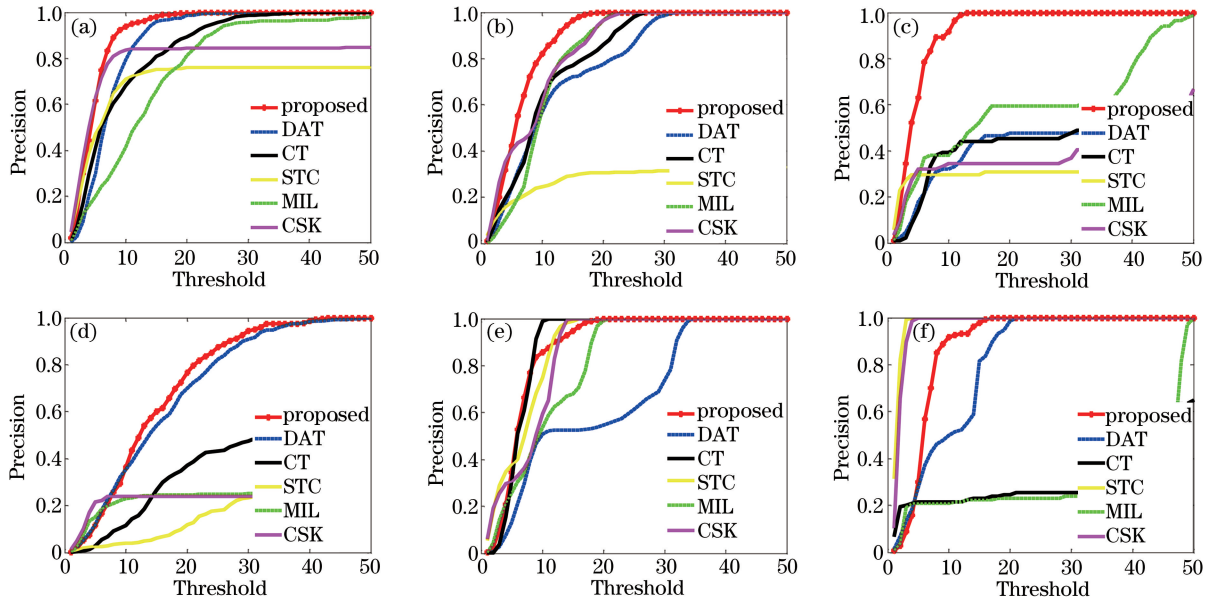


图 7 距离精度曲线图。(a) Boy; (b) RedTeam; (c) Girl2; (d) KiteSurf; (e) Man; (f) Car24  
Fig. 7 Curves of distance precision. (a) Boy; (b) RedTeam; (c) Girl2; (d) KiteSurf; (e) Man; (f) Car24

表 2 重叠精度对比

Table 2 Comparison of overlap precision

Sequence	Proposed	DAT	STC	MIL	CSK	CT
Boy	<b>0.982</b>	0.963	0.844	0.289	0.809	0.535
RedTeam	<b>0.875</b>	0.429	0.350	0.178	<u>0.588</u>	0.367
Girl2	<b>0.822</b>	<u>0.798</u>	0.229	0.288	0.281	0.299
KiteSurf	<b>0.500</b>	0.238	0.333	0.321	<u>0.393</u>	0.190
Man	<u>0.657</u>	0.440	<b>1.000</b>	0.209	<b>1.000</b>	0.216
Car24	<u>0.401</u>	0.226	<b>0.568</b>	0.286	0.292	0.235

#### 4.4 算法的复杂度分析

算法的运行环境:计算机配置 CPU 是 2.30 GHz 的英特尔 i3 处理器,4.00 GB 内存,64 位 Win7 操作系统,Matlab R2014a。以 Boy 视频序列为例,分辨率为 480 pixel×640 pixel,目标大小为 35 pixel×42 pixel。表 3 所示为本文算法与 DAT 算法处理每帧图像的平均消耗时间,其中本文算法的平均消耗时间是在程序未经优化的情况下得出的。由表 3 可知,相比于 DAT 算法,本文算法复杂度相对较高,约为 14.3 frame/s。本文方法利用多种特征,相比于单一特征的跟踪方法而言,计算耗时较大;在目标定位时采用穷举搜索的方式产生候选目标,也会导致计算量较大,因此有待进一步优化。

表 3 复杂度分析

Table 3 Complexity analysis

Sequence	Tracker	Average time per frame /ms
Boy	Proposed	70
	DAT	42



## 5 结 论

提出了一种特征融合和尺度自适应的干扰感知目标跟踪方法。目标模型考虑目标、邻域背景和干扰区域信息,通过采用 RGB 和 HOG 特征融合的方式提高目标特征的维度,并利用特征金字塔方法对目标尺度进行估计和更新,实现了复杂场景下目标的稳定跟踪。采用复杂场景的测试数据集进行测试,并与其他典型算法进行对比分析,实验结果表明,本文算法对类目标干扰、光照变化、局部遮挡、尺度变换等均具有较高的稳健性。进一步研究重点为算法优化的问题,以提高跟踪方法的实时性。

## 参 考 文 献

- [1] Zhang Huanlong, Hu Shiqiang, Yang Guosheng. 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.
- [2] Yin Hongpeng, Chen Bo, Chai Yi, *et al.* Vision-based object detection and tracking: A review [J]. Acta Automatica Sinica, 2016, 42(10): 1466-1489.  
尹宏鹏, 陈波, 柴毅, 等. 基于视觉的目标检测与跟踪综述 [J]. 自动化学报, 2016, 42(10): 1466-1489.
- [3] Zhang Lei, Wang Yanjie, Sun Honghai, *et al.* Adaptive scale object tracking with kernelized correlation filters [J]. Optics and Precision Engineering, 2016, 24(2): 448-459.  
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.
- [4] Yang Xin, Fei Shumin, Li Gang, *et al.* Improved mean shift tracking algorithm based on complicated feature fusion [J]. Control and Decision, 2014, 29(7): 1297-1300.  
杨欣, 费树岷, 李刚, 等. 基于复杂特征融合的改进 mean shift 目标跟踪 [J]. 控制与决策, 2014, 29(7): 1297-1300.
- [5] Sun Rui, Huang Jingru, Ding Wenxiu. A real-time object tracking algorithm based on subspace learning [J]. Opto-Electronic Engineering, 2015, 42(2): 52-58.  
孙锐, 黄静茹, 丁文秀. 一种基于子空间学习的实时目标跟踪算法 [J]. 光电工程, 2015, 42(2): 52-58.
- [6] Liu Wenzhuo, Yuan Guanglin, Xue Mogen. Robust fast visual tracking based on two-stage representation [J]. Acta Optica Sinica, 2016, 36(12): 1215001.  
刘文琢, 袁广林, 薛模根. 基于两阶段稀疏表示的稳健快速视觉跟踪 [J]. 光学学报, 2016, 36(12): 1215001.
- [7] Zhang Xudong, Chen Zhonghai, Hu Liangmei, *et al.* Object tracking method based on sparse representation of joint template [J]. Control and Decision, 2015, 30(9): 1696-1700.  
张旭东, 陈仲海, 胡良梅, 等. 基于联合模板稀疏表示的目标跟踪方法 [J]. 控制与决策, 2015, 30(9): 1696-1700.
- [8] Wang Yawen, Chen Hongchang, Li Shaomei, *et al.* Online boosting tracking algorithm combined with occlusion sensing [J]. Journal on Communications, 2016, 37(9): 2016181.  
王亚文, 陈鸿昶, 李邵梅, 等. 融合遮挡感知的在线 Boosting 跟踪算法 [J]. 通信学报, 2016, 37(9): 2016181.
- [9] Hu zhaohua, Xu Yuwei, Zhao Xiaolei, *et al.* Multi-feature selection tracking based on support vector machine [J]. Journal of Applied Sciences, 2015, 33(5): 502-517.  
胡昭华, 徐玉伟, 赵孝磊, 等. 基于支持向量机的多特征选择目标跟踪 [J]. 应用科学学报, 2015, 33(5): 502-517.
- [10] Liu Kai, Dai Pingyang, Jiang Xiaolian, *et al.* Object tracking algorithm based on ranking support vector machine fused with multiple features [J]. Computer Engineering, 2014, 40(11): 42-45.  
刘楷, 戴平阳, 江晓莲, 等. 基于排序支持向量机的多特征融合目标跟踪算法 [J]. 计算机工程, 2014, 40(11): 42-45.
- [11] Guo Pengyu, Su Ang, Zhang Hongliang, *et al.* Online mixture of random Naive Bayes tracker combined texture with shape feature [J]. Acta Optica Sinica, 2015, 35(3): 0315002.  
郭鹏宇, 苏昂, 张红良, 等. 结合纹理和形状特征的在线混合随机朴素贝叶斯视觉跟踪器 [J]. 光学学报, 2015, 35(3): 0315002.
- [12] Qian Cheng, Xu Shuchang, Zhang Yin, *et al.* Object tracking method based on incrementally updating linear discriminant subspace [J]. Pattern Recognition and Artificial Intelligence, 2010, 23(2): 250-255.  
钱诚, 徐舒畅, 张引, 等. 一种基于线性判别空间增量更新的跟踪方法 [J]. 模式识别与人工智能, 2010, 23(2): 250-255.
- [13] Zhang Ji, Wang Hongyuan. A novel tracking method based on incremental semi-supervised discriminant analysis [J].

Journal of Nanjing University (Natural Sciences), 2012, 48(4): 397-404.

张 继, 王洪元. 一种基于增量半监督判别分析的跟踪方法[J]. 南京大学学报(自然科学), 2012, 48(4): 397-404.

- [14] Possegger H, Mauthner T, Bischof H. In defense of color-based model-free tracking [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 2113-2120.
- [15] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 1: 886-893.
- [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] Zhang K, Zhang L, Yang M H. Real-time compressive tracking[C]. European Conference on Computer Vision, 2012: 864-877.
- [18] Zhang K, Zhang L, Liu Q, *et al.* Fast visual tracking via dense spatio-temporal context learning [C]. European Conference on Computer Vision, 2014: 127-141.
- [19] Babenko B, Yang M H, Belongie S. Robust object tracking with online multiple instance learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(8): 1619-1632.
- [20] Wu Y, Lim J, Yang M H. Online object tracking: a benchmark[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013: 2411-2418.
- [21] Liu Xing, Chen Ying. Target tracking based on adaptive fusion of multi-feature[J]. Opto-Electronic Engineering, 2016, 43(3): 58-65.
- 刘 行, 陈 莹. 自适应多特征融合目标跟踪[J]. 光电工程, 2016, 43(3): 58-65.