

# 基于分类-验证模型的视觉跟踪算法研究

吴敏, 查宇飞, 张园强, 库涛, 李运强, 张胜杰

中国人民解放军空军工程大学航空航天工程学院, 陕西 西安 710038

**摘要** 针对相似度目标跟踪算法主要考虑目标的类内相似, 而忽略不同目标的类间差异的问题, 提出基于分类-验证模型的视觉跟踪算法。该算法通过增加目标的属性(类别)信息, 利用相似度信息与类别信息构建损失函数, 在高维空间学习目标的类内相似和类间差异; 将目标模板与候选目标输入网络模型, 分别通过分类与验证模块实现网络参数更新; 利用训练网络提取目标模板与候选目标的深度嵌入特征, 实现目标跟踪。在 OTB50 和 UAV123 数据库上进行实验, 结果表明, 该算法可以大幅提高跟踪效果, 对相似目标具有较强的稳健性。

**关键词** 机器视觉; 目标跟踪; 分类-验证网络模型; 类内相似; 类间差异

中图分类号 TN919.82

文献标识码 A

doi: 10.3788/AOS201838.0515003

## Visual Tracking Algorithm Based on Classification-Validation Model

Wu Min, Zha Yufei, Zhang Yuanqiang, Ku Tao, Li Yunqiang, Zhang Shengjie

*Aeronautics and Astronautics Engineering College, Air Force Engineering University of PLA,  
Xi'an, Shaanxi 710038, China*

**Abstract** In order to solve the problems that the similarity target tracking algorithm mainly considers the intraclass similarity of targets and ignores the interclass differences of different targets. A visual tracking algorithm based on classification-validation model is proposed, which adds attribute information to the similarity algorithm. The proposed algorithm constructs the loss function with similarity and class information, and learns intraclass similarity and interclass differences in high dimensional space. The classification and verification module is adopted to update network parameters when the target template and candidate target input into the network model. With the trained network, the deep embedding feature of target and candidate target is extracted, thus, the target tracking is achieved. Experiments are carried out on the OTB50 and UAV123 databases. Results show that the proposed algorithm can improve the tracking effect with increased target information, and has strong robustness to the similar targets.

**Key words** machine vision; target tracking; classification-validation model; intraclass similarity; interclass differences

**OCIS codes** 150.1135; 100.4999; 100.4996

## 1 引 言

视觉目标跟踪是计算机视觉领域的一个基本问题, 其通过给定初始帧目标的位置和尺度, 快速准确地寻找目标在后续视频帧中的位置和大小。由于尺度、旋转和形变等内因, 以及相似目标、复杂背景和光照等外因的相互作用, 目标外观会发生一定改变, 因此, 目标跟踪容易失败<sup>[1-2]</sup>。

相似性跟踪算法因其可以直接学习目标之间的

匹配关系, 提高目标的表征能力, 近年来得到学术界的重视<sup>[3-4]</sup>。目前, 相似度跟踪算法采用双路神经网络提取相似度信息, 采用离线训练方法, 利用端对端网络, 通过反向传播训练表示目标的层次卷积特征。在跟踪阶段, 采用回归或者排序算法在搜索区域中寻求最佳匹配。一方面, 相似度算法能同时学习特征和相似性度量, 如采用双路神经网络的用于视觉追踪的双流卷积神经网络(YCNN)算法<sup>[5]</sup>和在 CVPR2016 中性能突出的相似性实例搜

收稿日期: 2017-11-06; 收到修改稿日期: 2017-12-12

基金项目: 国家自然科学基金(61472442, 61773397, 61701524)、陕西省科技新星资助项目(2015kjxx-46)

作者简介: 吴敏(1994—), 男, 硕士研究生, 主要从事计算机视觉、图像处理与分析方面的研究。

E-mail: kj123123213@163.com

导师简介: 查宇飞(1979—), 男, 博士, 副教授, 硕士生导师, 主要从事计算机视觉、模式识别、图像处理方面的研究。

E-mail: 463431261@qq.com(通信联系人)

索跟踪(SINT)算法<sup>[6]</sup>,都通过提取目标的相似度信息来提高跟踪性能;另一方面,其具有深度特征的强大表征能力,在跟踪时具有良好的实时性,如采用全卷积网络(CNN)直接匹配的基于全连接孪生网络的目标跟踪(Siamese-fc)算法<sup>[7]</sup>,在跟踪过程中没有更新,其跟踪速度尤为突出。

相似度学习考虑目标的类内相似,而忽略了不同目标之间的类间差异。本文在此基础上,通过增加目标的属性信息-类别信息,给予不同目标不同的标签,使训练网络识别目标的能力更强;利用相似度与类别信息建立分类-验证损失函数,在高维空间学习目标的类内相似和类间差异;分别利用类别与相似度信息更新网络参数,使训练网络具有更强的识别相似目标和区分不同目标的能力。与相似度跟踪算法相比,本文算法在相似目标视频序列上能够取得令人满意的结果。

## 2 基于相似性学习的跟踪算法

基于相似性学习的目标跟踪算法采用先离线训练再实时跟踪的思路。首先,利用离线训练卷积网络得到匹配函数;然后,进行在线跟踪,即根据匹配函数选择与初始帧标定目标最匹配的候选目标,将其作为跟踪结果。

给出目标模板  $\mathbf{o}$  与候选目标  $\mathbf{u}$  之间的标签,离线训练网络学习相似性,则有

$$L = \sum_j h[m(\mathbf{o}, \mathbf{u}_j), y_j], \quad (1)$$

式中  $L$  为相似性目标跟踪的损失函数,函数  $m(\cdot)$  计算目标模板  $\mathbf{o}$  与第  $j$  个候选目标  $\mathbf{u}_j$  的响应值,对比函数  $h(\cdot)$  与真实标签  $y_j$ ,可以减小损失值,实现网络的训练。

实时跟踪是将初始给定的第一帧目标边界框和下一帧的搜索区域输入已训练的网络,利用提取到的深度特征进行目标匹配,从所给出的候选边界框中检索最匹配的候选框,表达式为

$$\hat{\mathbf{x}}_t = \arg \max_{\mathbf{x}_{j,t}} m(\mathbf{x}_{t=1}, \mathbf{x}_{j,t}), \quad (2)$$

式中  $\mathbf{x}_{t=1}$  为初始给定的第一帧目标框,  $\mathbf{x}_{j,t}$  为第  $t$  帧的候选目标,  $\hat{\mathbf{x}}_t$  为预测的第  $t$  帧的目标位置。函数  $m(\cdot)$  用于计算  $\mathbf{x}_{t=1}$  与  $\mathbf{x}_{j,t}$  的响应值,然后取响应值最大的候选目标作为真实目标。

在当前大多数基于 CNN 的目标跟踪算法普遍较慢的情况下,基于相似性学习的目标跟踪算法利用深度网络提取目标的相似度信息,实现了超实时的帧率。但相似性算法在训练网络参数时,只利用

了目标的相似度信息,其对目标的特征表达能力有限,不足以实现目标的稳健跟踪。

## 3 基于分类-验证模型的跟踪算法

相似性目标跟踪算法利用目标模板与候选目标的相似性信息,在进行网络模型训练时,定义离目标模板近的候选目标为正样本,其余为负样本,训练网络可以学习同一目标的类内相似,但其不考虑不同目标之间的类间差异。增加目标的类别信息可以在高维空间中学习目标的类内相似和类间差异,基于该特点,本文增加目标的属性信息-类别信息,提出基于分类-验证模型的算法,同时利用目标的相似度和类别信息更新网络参数,使所训练网络能够更好地区分目标与背景。图 1 为学习相似度信息和类别信息的示意图。

分类-验证网络模型由分类模型和验证模型组成。其中,分类模型利用目标的类别信息建立分类损失函数,将目标模板及其对应的类别标签输入分类网络,提取目标的深度嵌入式特征,用 soft max 函数对目标进行分类,根据损失值更新网络。通过对不同目标进行分类,使网络学习获取目标的类别信息,从而提高网络区分不同目标的能力,表达式为

$$l_1(f, t) = \sum_{i=1}^K -p_i \ln(\hat{p}_i), \quad (3)$$

式中

$$\hat{p}_i = \text{soft max}[f(\mathbf{x})] = \frac{\exp[\mathbf{w}_i^T f(\mathbf{x})]}{\sum_{j=1}^N \exp[\mathbf{w}_j^T f(\mathbf{x})]}, \quad (4)$$

式中  $l_1$  为分类损失函数,  $f(\mathbf{x})$  为提取出来的深度嵌入式特征,  $N$  为训练数据中目标的总类别数,  $\hat{p}_i$  为预测目标为第  $i$  类的概率,  $p_i$  为目标对应的标签,当  $i=t$  时,  $p_i=1$ , 否则  $p_i=0$ ,  $k$  为同时输入的目标对数目,利用损失值  $l_1$  实现网络参数的更新。

验证模型利用目标的相似度信息建立验证损失函数,通过验证网络提取目标模型和候选目标的深度嵌入式特征,利用 logist 函数实现目标匹配,与其对应的标签进行比较,完成网络参数更新。通过对网络参数的更新,提高目标匹配的响应值,提高正样本与目标模板之间的相似度,减小同一目标之间的类内距离,则有

$$l_2(y, u) = \ln[1 + \exp(-yv)], \quad (5)$$

式中  $l_2$  为验证损失函数,其通过目标模板与候选目标之间的响应值  $v$  和标签  $y$  实现对网络参数的更新。

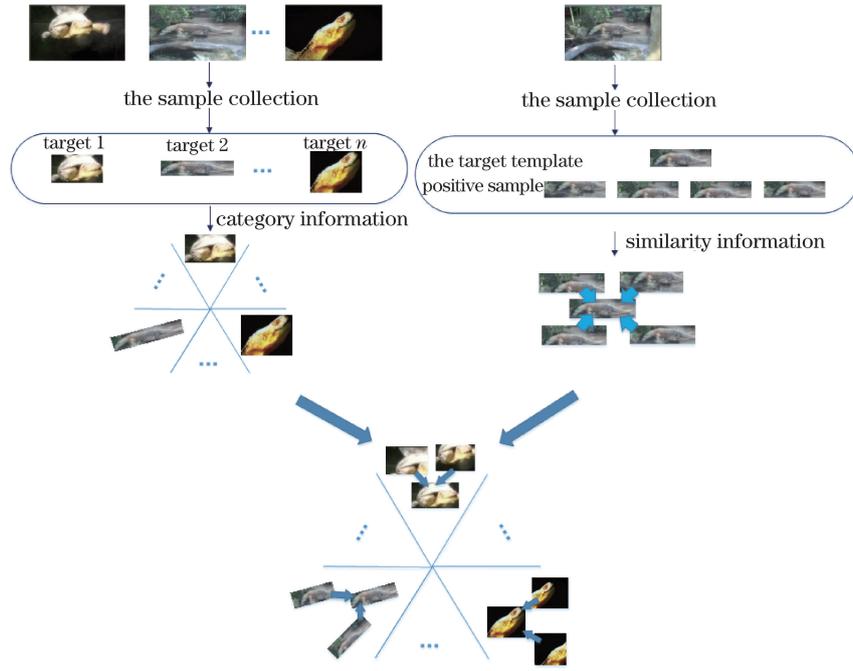


图 1 学习相似度信息和类别信息示意图

Fig. 1 Information of learning similarity and classification

相似度信息方法采用目标之间的匹配程度，其无法解决相似目标的跟踪问题，而利用目标的类别信息，可以提高模型对相似目标的区分能力，实现相似度与类别信息之间的互补。利用目标的相似度和类别信息，建立分类-验证损失函数，学习同一目标之间的类内相似和不同目标之间的类间差异，即有

$$L = \underbrace{\frac{1}{k} \sum_{i=1}^k \frac{1}{|D|} \sum_{u \in D} \ln\{1 + \exp[-y(u)v(u)]\}}_{\text{target validation}} + \underbrace{\lambda \sum_{i=1}^k -p_i \ln(\hat{p}_i)}_{\text{target classification}}, \quad (6)$$

式中  $L$  为分类-验证损失函数，前一部分为目标验证损失函数，学习目标的相似度信息；后一部分为目标分类损失函数，学习目标的类别信息，以权重  $\lambda$  结合两个损失函数，能够使分类-验证损失函数同时学习目标的相似度和类别信息，提高同一目标的相似匹配和不同目标的类间区分能力， $D$  为候选目标的位置， $y(u)$  为中心位置为  $u$  的候选目标标签，将离目标模板中心距离为  $d$  的候选目标作为正样本，以  $+1$  表示；将离目标模板中心距离大于  $d$  的候选目标作为负样本，以  $-1$  表示，则有

$$y[u] = \begin{cases} +1, & \text{if } k \|u - c\| \leq d \\ -1, & \text{otherwise} \end{cases}, \quad (7)$$

式中  $c$  为候选图片中心的位置， $u$  为候选目标中心

位置。

### 3.1 离线训练共享网络

建立分类-验证网络模型实现对网络参数的训练。图 2 所示为深度分类-验证网络模型。该模型主要分为目标分类模块和验证模块，两者共享同一网络。其中，目标验证模块用于提取目标的相似度信息，对应目标验证损失函数；目标分类模块用于提取目标的类别信息，对应目标分类损失函数。

将目标模板  $z$  和候选目标  $x$  输入分类-验证网络模型，利用共享网络分别提取目标模板和候选目标的深度嵌入特征；将目标模板及其所对应的类别代入分类损失函数，得到目标的损失值，进而实现对网络参数的更新；将目标模板、候选目标以及所对应的匹配标签代入验证损失函数，利用得到的损失值对网络参数实现再次更新。

输入目标模板及所对应的类别，通过分类模块对目标进行识别，将其与目标类别进行比较，得到模型的损失值，从而实现网络参数的更新。对(6)式中的目标验证模块进行求导，可得

$$\Delta w_1 = \frac{1}{k} \sum_{i=1}^k \left[ \frac{\exp(z_{n_i})}{\sum_j \exp(z_{n_j})} - 1 \right] z_{n_i}, \quad (8)$$

式中  $\Delta w_1$  为分类损失函数得到的网络参数更新值， $z_{n_i}$  为第  $i$  对样本目标经过  $n$  层卷积后得到的结果。

以目标为中心，从标注的视频数据库中提取目

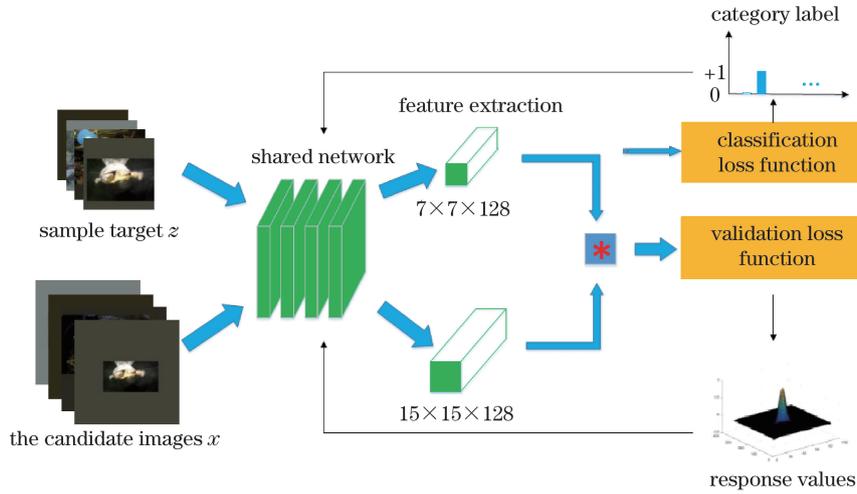


图 2 深度分类-验证网络模型

Fig. 2 Deep classification-validation network model

标模板和候选目标,作为验证模块更新的数据。选取相隔若干帧的两帧图片,每个图片的目标大小都进行无损纵横比归一化。将离目标模板距离小的候选目标作为正样本,将离目标模板距离大的候选目标作为负样本,以验证损失函数得到候选目标的响应值,然后对损失函数求导实现网络的训练。同样对(6)式中的目标分类模块进行求导,可得

$$\Delta w_2 = \frac{1}{k} \sum_{i=1}^k \frac{1}{|D|} \sum_{u \in D} \frac{-y(u) \exp[-y(u)v(u)]}{1 + \exp[-y(u)v(u)]} \cdot z_{n_i} \cdot [w_n^T x_{n_i}(u)], \quad (9)$$

式中  $\Delta w_2$  为验证损失函数得到的网络参数更新值,

$x_{n_i}$  为第  $i$  对候选目标经过  $n$  层卷积后的结果,  $w_n$  为第  $n$  层卷积的卷积核,  $z(u)$  是候选区域中的候选目标。同时有

$$w' = w + \Delta w_1 + \lambda \Delta w_2, \quad (10)$$

式中  $w$  为更新前的共享网络参数,  $w'$  为更新后的共享网络参数。

### 3.2 在线实时跟踪

离线训练过程旨在学习匹配机制,从大量外部视频中学习匹配函数。在线实时跟踪过程是一个利用训练好的网络提取目标特征,进而实现跟踪的过程,如图 3 所示。

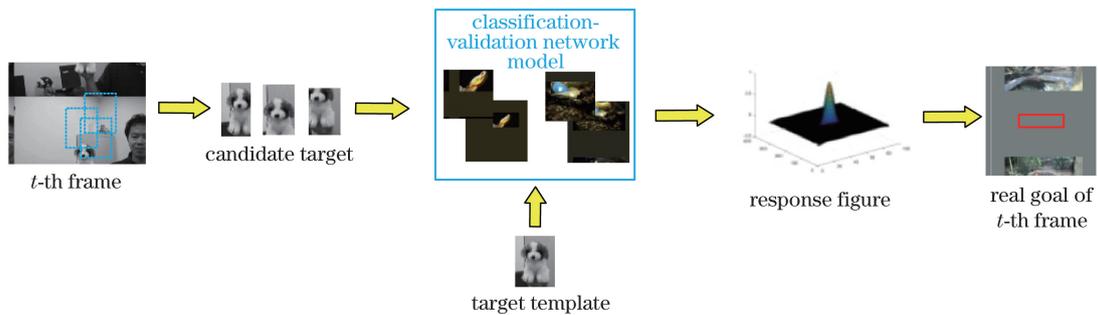


图 3 在线实时跟踪流程

Fig. 3 Online real-time tracking process

选取第一帧的目标作为目标模板  $x$ , 以前一帧目标位置为中心,目标模板大小的若干倍为候选范围,得到候选目标  $z$ ,将目标模板和候选目标带入训练好的共享网络中,分别提取目标模板与候选目标的深度嵌入式特征。

函数  $f(x, z)$  负责比较目标模板  $x$  和候选目标  $z$  的相似度。为了定位目标在新图像里的位置,需要遍历进行比较,选出相似度最大的目标。以目标模板的 0.9、1 和 1.1 倍与候选目标进行卷积,得到 3 个

目标响应矩阵,选取最大值最大的候选目标作为该帧的真实目标,则有

$$z' = \arg \max_{u \in D} \{f[x_i, z(u)]\} \quad i = 1, 2, 3, \quad (11)$$

$$f[x_i, z(u)] = x_i * z(u), \quad (12)$$

式中  $x_1, x_2, x_3$  分别代表样本目标的 0.9、1 和 1.1 倍,  $z'$  为真实目标位置。

## 4 实验结果及分析

在 Matlab 2016a 平台上进行试验。对用于检

测任务的视频序列 ILSVRC2015<sup>[8]</sup> 进行改造,得到专门用于训练孪生网络的数据,测试视频来自于当前目标跟踪领域最常用的 OTB50<sup>[9]</sup> 和 UAV123<sup>[10]</sup>,模板的更新率为 0.1,权重  $\lambda=1$ ,采用先训练后测试的方法<sup>[11]</sup>,即先训练出适合孪生网络的目标跟踪模型,再利用训练好的模型进行目标跟踪。

#### 4.1 算法验证

为了证明本文算法的优越性,选取基于相似程度信息的算法、基于类别信息的算法和本文算法,

表 1 视频序列及其描述

Table 1 Video sequence and its description

Video sequence	Sequence characteristics description	Resolution	Number of frames
Liquor	Similar goals, fast moving	812×575	1741
Subway	Similar goals, occlusion	524×383	175
Football	Similar goals, complicated background	524×383	74
Ironman	Similar goals, complicated background	720×304	166
Matrix	Similar goals, fast moving	800×336	100
Soccer	Similar goals, occlusion	640×360	392

由图 4 可以看出,本文算法在相似目标、背景干扰、遮挡以及快速移动下均能够实现相似目标的稳健跟踪。Liquor、Matrix 序列具有相似目标和快速移动等特点,在 Liquor 序列第 708 帧中,相似度算法、类别算法和本文算法都能对目标进行精确跟踪,但在第 1111 帧中,目标周围出现相似目标时,相似度算法会跟踪到相似目标,而本文算法依然能准确跟踪目标;在 Matrix 序列第 68 帧和第 90 帧中,当目标周围出现相似目标时,本文算法仍能实现有效跟踪。

Subway、Soccer 序列具有相似目标和遮挡等特点,在 Subway 序列第 35 帧中,相似度算法、类别算法和本文算法都能对目标进行精确跟踪,但在第 39 帧中,相似目标对真实目标产生遮挡,相似度算法发生跟踪漂移,而本文算法由于增加了对目标的约束,使提取的特征能够反映真实目标,克服了遮挡的影响,确保了对相似目标的有效跟踪。在 Soccer 序列中,当目标发生遮挡时,本文算法仍能实现较好的跟踪效果。

从图 4 可以看到,Football、Ironman 序列的特点是具有相似目标和复杂背景。相似目标的干扰以及复杂背景使得目标跟踪尤为困难,本文算法通过利用相似度信息与类别信息,提高了模型对目标的判定能力,实现了比相似度算法更好的跟踪效果。

#### 4.1.2 定量分析

为了进一步验证算法的优越性,在 ILSVRC 数据库中选择部分数据,根据训练网络所用的目标信

在 ILSVRC 数据库中选择部分数据对网络模型进行训练,分别从定性与定量两个角度对算法进行验证。

#### 4.1.1 定性分析

本文算法增加了目标的类别信息,能够提升模型对相似目标的判定,同时学习目标的类内相似和类间差异。为了证明其对相似目标的判定能力,从 OTB50 数据库中选取 6 个具有相似目标的视频序列进行验证,各序列的特点如表 1 所示。

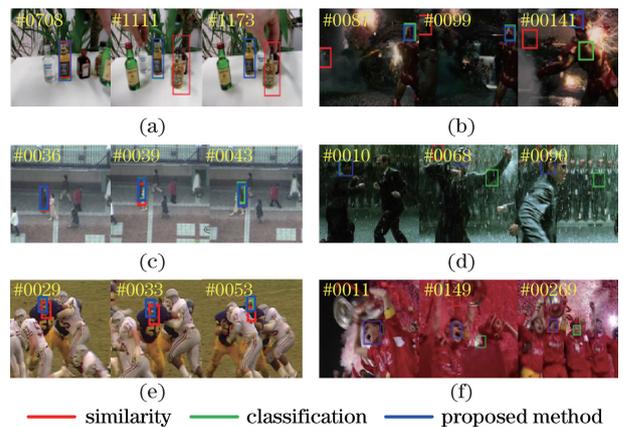


图 4 跟踪结果示意图。(a) Liquor 序列;

(b) subway 序列; (c) football 序列; (d) ironman 序列;

(e) matrix 序列; (f) soccer 序列

Fig. 4 Diagram of the tracking results. (a) Liquor sequence;

(b) subway sequence; (c) football sequence;

(d) ironman sequence; (e) matrix sequence;

(f) soccer sequence

息进行实验。 $\lambda$  为不同信息所占权重,当  $\lambda=0$  时,所训练的网络只利用样本目标的相似度信息;当  $\lambda=\infty$  时,所训练的网络只利用目标的类别信息。

图 5 所示为在 OTB50 数据库上分别用单程评估(OPE)、时间稳健性评估(SRE)和空间稳健性评估(TRE)进行评价得到的结果,将覆盖率为 0.5 时的算法成功率作为排名依据<sup>[12]</sup>,实验结果表明,只利用相似度信息的网络和只利用类别信息的网络都无法取得很好的效果,当同时利用样本的相似度和

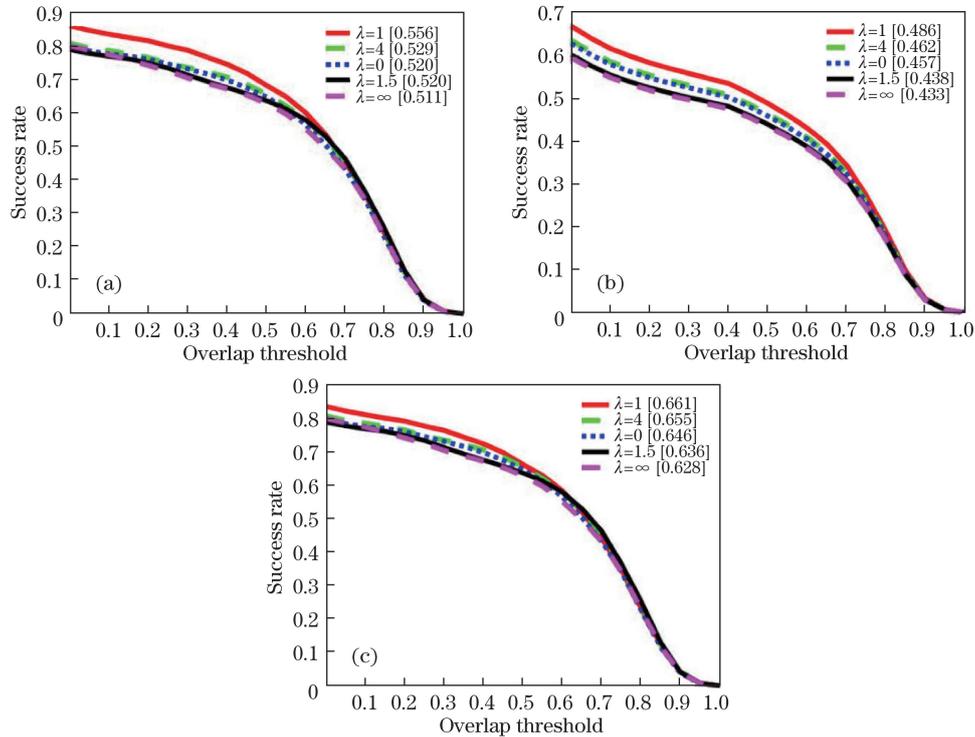


图 5 不同参数的成功率对比。(a) OPE; (b) SRE; (c) TRE

Fig. 5 Comparison of success rate for different parameters. (a) OPE; (b) SRE; (c) TRE

类别信息时,跟踪精确度和成功率都有很大的变化,当 $\lambda=1$ 时,取得最好的跟踪效果;同时利用目标的相似度和类别信息大幅提升了跟踪效果。

#### 4.2 OTB50 数据库评估

OTB50 数据库共包含 50 个视频序列,涉及尺度变化、形变和遮挡等 11 种属性变化<sup>[13-14]</sup>。本文算法首先在该数据库上进行整体性能比较,对比算法有基于连续卷积算子的视觉跟踪(CCOT),SINT, Siamese-fc,基于全卷积网络的视觉跟踪(FCNT),对冲深度追踪(HDT),基于可信赖区域的稳健性视觉跟踪(RPT),基于精确尺度估计的稳健性视觉跟踪(DSST),无训练的卷积网络稳健视觉跟踪(CNT),通过对 51 个序列进行成功率评价,得到平均成功率的指标图,如图 6 所示。

由图 6 可以看出,本文算法的成功率达到了 80.7%,要高于只利用相似度信息的 Siamese-fc 算法和 SINT 算法,说明在相似度信息的基础上增加目标的类别信息,确实提高了目标的表征能力;同时,在跟踪性能上,本文算法和最新的 CCOT 算法相当,但本文算法的跟踪速度高于 CCOT 算法,说明本文算法的整体跟踪性能更优越。

#### 4.3 UAV123 数据库评估

UAV123 数据库是最近的无人机数据库,其包含 123 个视频序列,涉及 12 种属性变化<sup>[15]</sup>。相比

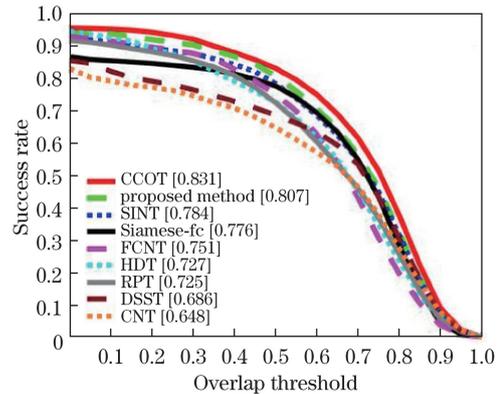


图 6 本文算法在 OTB50 上成功率对比

Fig. 6 Comparison of success rate of proposed algorithm on OTB50

于 OTB2013 数据库,UAV123 数据库主要涉及在低空环境下,由于无人机拍摄的角度与高度的不一致引发的目标形变以及尺度变化的数据,与 OTB50 数据库相比,处理 UAV123 数据库时面临的挑战更大。对比算法有基于熵权最小化的多模型稳健跟踪(MEEM),基于核的结构化输出跟踪(struck), DSST,基于判别相关滤波的跟踪(DCF),基于检测的跟踪(TLD)。

从图 7 可以看出,和 OTB50 数据库相比,各种算法在 UAV123 数据库上的跟踪成功率普遍偏低,这说明在 UAV123 数据库上进行跟踪更具有挑战

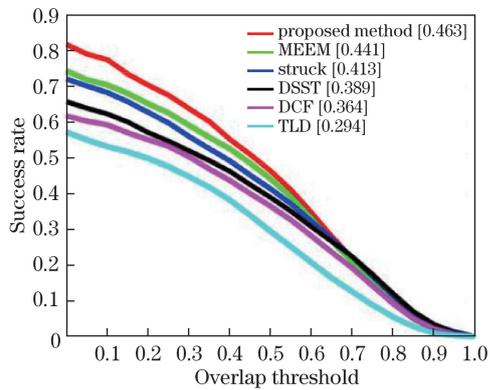


图7 本文算法在 UAV123 上成功率对比

Fig. 7 Comparison of success rate of proposed algorithm on UAV123

性;在该数据库中,本文算法的成功率图的曲线下面积(AUC)为 46.3%,跟踪效果最佳。上述实验结果表明,在无人机视频中,本文算法比其他算法更具有稳健性。

## 5 结 论

结合目标的相似度信息和类别信息,建立了一种分类-验证网络模型,利用相似度和类别信息对共享网络进行训练,使学习到的深度嵌入式特征更好地实现目标表征。本文算法利用目标的相似度信息缩小了类内距离,利用类别信息对不同目标进行了有效区分,即同时学习了同一目标的类内相似和不同目标的类间差异。实验结果表明,本文算法在 OTB50 和 UAV123 数据库上的平均成功率超过了对比算法,实现了目标的精确跟踪。

## 参 考 文 献

- [1] Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.
- [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] Li S S, Zhao G P, Wang J Y. Distractor-aware object tracking based on multi-feature fusion and scale-adaption[J]. Acta Optica Sinica, 2017, 37(5): 0515005.  
李双双, 赵高鹏, 王建宇. 基于特征融合和尺度自适应的干扰感知目标跟踪[J]. 光学学报, 2017, 37(5): 0515005.
- [4] Held D, Thrun S, Savarese S. Learning to track at 100 fps with deep regression networks[C]. European Conference on Computer Vision (ECCV), 2016: 749-765.
- [5] Chen K, Tao W B. Once for all: a two-flow convolutional neural network for visual tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 99: 1-15.
- [6] Tao R, Gavves E, Smeulders A W M. Siamese instance search for tracking[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 1420-1429.
- [7] Bertinetto L, Valmadre J, Henriques J F, et al. Fully-convolutional Siamese networks for object tracking[C]. European Conference on Computer Vision (ECCV), 2016: 850-865.
- [8] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.
- [9] Zeng X, Ouyang W, Yang B, et al. Gated bi-directional CNN for object detection[C]. European Conference on Computer Vision (ECCV), 2016: 354-369.
- [10] Ma C, Huang J B, Yang X, et al. Hierarchical convolutional features for visual tracking[C]. IEEE International Conference on Computer Vision (ICCV), 2015: 3074-3082.
- [11] Leal-Taixe L, Canton-Ferrer C, Schindler K. Learning by tracking: siamese CNN for robust target association[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016: 418-425.
- [12] Wu Y, Lim J, Yang M-H. Online object tracking: a benchmark[C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013: 2411-2418.
- [13] Hong Z, Chen Z, Wang C, et al. Multi-store tracker (muster): a cognitive psychology inspired approach to object tracking[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 749-758.
- [14] Danelljan M, Robinson A, Khan F S, et al. Beyond correlation filters: learning continuous convolution operators for visual tracking[C] // European Conference on Computer Vision (ECCV), 2016: 472-488.
- [15] Wang N Y, Li S Y, Gupta A, et al. Transferring rich feature hierarchies for robust visual tracking[J]. Computer Science, 2015.