

基于增强群跟踪器和深度学习的目标跟踪

程 帅^① 曹永刚^{①②} 孙俊喜^{*③} 赵立荣^{①②} 刘广文^① 韩广良^②

^①(长春理工大学电子信息工程学院 长春 130022)

^②(中国科学院长春光学精密机械与物理研究所 长春 130000)

^③(东北师范大学计算机科学与信息技术学院 长春 130117)

摘 要: 为解决基于外观模型和传统机器学习目标跟踪易出现目标漂移甚至跟踪失败的问题, 该文提出以跟踪-学习-检测(TLD)算法为框架, 基于增强群跟踪器(FoT)和深度学习的目标跟踪算法。FoT 实现目标的预测与跟踪, 增添基于时空上下文级联预测器提高预测局部跟踪器的成功率, 快速随机采样一致性算法评估全局运动模型, 提高目标跟踪的精确度。深度去噪自编码器和支持向量机分类器构建深度检测器, 结合全局多尺度扫描窗口搜索策略检测可能的目标。加权 P-N 学习对样本加权处理, 提高分类器的分类精确度。与其它跟踪算法相比较, 在复杂环境下, 不同图片序列实验结果表明, 该算法在遮挡、相似背景等条件下具有更高的准确度和鲁棒性。

关键词: 计算机视觉; 群跟踪器; 跟踪-学习-检测; 深度学习; 支持向量机; 深度检测器

中图分类号: TP391.4

文献标识码: A

文章编号: 1009-5896(2015)07-1646-08

DOI: 10.11999/JEIT141362

Target Tracking Based on Enhanced Flock of Tracker and Deep Learning

Cheng Shuai^① Cao Yong-gang^{①②} Sun Jun-xi^③
Zhao Li-rong^{①②} Liu Guang-wen^① Han Guang-liang^②

^①(School of Electronic Information Engineering, Changchun University of Science and Technology, Changchun 130022, China)

^②(Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130000, China)

^③(School of Computer Science and Information Technology, Northeast Normal University, Changchun 130117, China)

Abstract: To solve the problem that the tracking algorithm often leads to drift and failure based on the appearance model and traditional machine learning, a tracking algorithm is proposed based on the enhanced Flock of Tracker (FoT) and deep learning under the Tracking-Learning-Detection (TLD) framework. The target is predicted and tracked by the FoT, the cascaded predictor is added to improve the precision of the local tracker based on the spatio-temporal context, and the global motion model is evaluated by the speed-up random sample consensus algorithm to improve the accuracy. A deep detector is composed of the stacked denoising autoencoder and Support Vector Machine (SVM), combines with a multi-scale scanning window with global search strategy to detect the possible targets. Each sample is weighted by the weighted P-N learning to improve the precision of the deep detector. Compared with the state-of-the-art trackers, according to the results of experiments on variant challenging image sequences in the complex environment, the proposed algorithm has more accuracy and better robust, especially for the occlusions, the background clutter and so on.

Key words: Computer vision; Flock of Tracker (FoT); Tracking-Learning-Detection (TLD); Deep learning; Support Vector Machine (SVM); Deep detector

1 引言

目标跟踪在各个领域得到广泛应用, 例如运动分析、行为识别等, 同时也面临巨大挑战: 遮挡、相似背景、光照变化、外观变化等^[1]因素都能导致目

标偏移甚至跟踪失败。基于外观模型的跟踪算法^[2-5]通过尺度不变特征转换、方向梯度直方图等特征对目标进行表达, 这些特征不能反映目标本质信息, 跟踪过程中常出现误匹配问题。选用复杂的外观模型, 计算成本很高。

外观模型与传统的机器学习算法^[6-9]相结合, 将跟踪视为二值分类问题, 有效地利用背景信息, 以区分目标与背景, 虽改进了跟踪效果, 但也面临部分难题, 没有足够的训练数据来学习分类模型,

2014-10-29 收到, 2015-03-23 改回, 2015-06-01 网络优先出版
国家自然科学基金(61172111)和吉林省科技厅项目(20090512,
20100312)资助课题

*通信作者: 孙俊喜 juxi_sun@126.com

对相似目标分辨能力不强, 常出现错误分类问题。

深度学习^[10]是目前研究热点问题, 通过构建深层非线性网络结构可从少数样本集中学习图像信息的本质特征, 最终提高分类器分类的准确性。

群跟踪器(Flock of Tracker, FoT)^[11]将局部跟踪器与全局运动模型相结合, 可处理遮挡、非刚性目标局部变化的问题, 单元格群跟踪器(Cell FoT)^[12]选择最佳局部跟踪器进行跟踪, 解决目标漂移问题, 使目标跟踪更具鲁棒性。

为解决基于外观模型和传统机器学习目标跟踪算法的问题, 本文提出以跟踪-学习-检测(Tracking-Learning-Detection, TLD)^[13]算法为框架, 基于增强FoT和深度学习的目标跟踪算法。在原有FoT的基础上, 利用局部跟踪器的时空上下文^[14], 增添级联预测器, 提高预测局部跟踪器的成功率。快速随机采样一致性(RANdom SAMple Consensus, RANSAC)^[15]算法评估跟踪器的全局运动模型, 减少迭代次数, 降低模型评估的失败率。深度检测器由深度去噪自编码器(Stacked Denoising AutoEncoder, SDAE)^[16]、支持向量机(Support Vector Machine, SVM)构成, 实现图像信息的有效表达, 利用无监督特征学习和迁移学习解决训练样本不足问题, 结合全局多尺度扫描窗口搜索策略, 实现目标检测。加权P-N学习考虑样本权重, 提高分类器分类准确度。K均值(K-means)聚类算法对在线模板集聚类, 形成二值树, 减少模板匹配数量, 降低计算复杂度。

2 增强 FoT

在原有FoT跟踪器基础上, 增强FoT利用局部跟踪器时空上下文信息增添时空上下文预测器, 并与NCC(Normalized Cross Correlation)预测器^[13]构成级联预测器, 对局部跟踪器进行预测。快速RANSAC算法利用跟踪成功的局部跟踪器评估全局运动模型, 预测下一帧中目标位置信息, 图1为增强FoT框图。

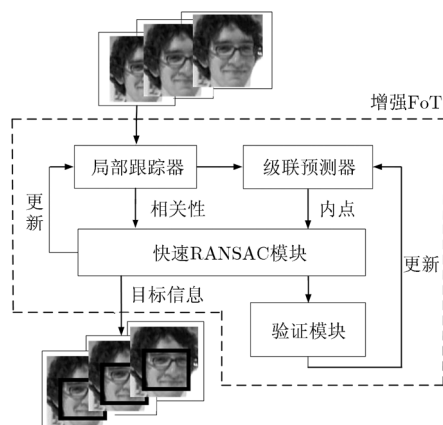


图1 增强 FoT 框图

2.1 级联预测器

级联预测器由NCC预测器、时空上下文预测器构成, 3种预测器间采用级联关系, 局部跟踪器只有符合3种预测器的条件才认定为跟踪成功, 否则跟踪失败。时间上下文预测器利用时间上下文信息构建马尔科夫链模型, 根据当前时刻的信息预测下一时刻局部跟踪器正确跟踪的概率, 判断局部跟踪器是否跟踪成功, 具体实现参考文献[11]。空间上下文预测器利用空间上下文信息假定短时间内相邻局部跟踪器之间具有运动一致性, 利用这一特性通过相邻跟踪器预测局部跟踪器是否跟踪成功, 图2为空间上下文预测器预测过程。

如图2所示, x 为跟踪器 i 的参考点, 每一对相关点 $i, j \in 1, 2, 3, 4$ 得出相似性矩阵变换 T_{ij} , 通过得到的变换矩阵 T_{ij} 对 x 变换, 映射误差为 $\varepsilon_{ij} = \|\hat{x}_{ij} - x'\|^2$, 符合条件 $\varepsilon_{ij} < \varepsilon_N$ 的点 \hat{x}_{ij} , 即以 x' 为圆心, ε_{ij} 为半径的区域内 \hat{x}_{ij} 点的个数达到设定的阈值时, 认定 x 跟踪成功。

2.2 快速 RANSAC

快速RANSAC算法能够从包含大量外点的数据集中快速、准确地估计出最优参数模型, 并使用贝叶斯算法更新内点集合的概率, 减少迭代次数, 降低模型失败的概率。因此, 根据级联预测器得到局部跟踪器集合, 采用快速RANSAC算法评估目标最优全局运动模型, 提高跟踪器对目标位置的预测精度, 解决目标漂移问题, 具体实现参考文献[15]。

3 深度检测器

深度检测器由3个部分构成: (1)滑动窗口, 利用多尺度滑动窗口全局扫描输入图像; (2)SDAE编码器, 即深度去噪自编码器的编码器部分, 提取多尺度扫描窗口内图像的特征, 对图像信息进行有效表达; (3)线性SVM分类器^[17]二值分类提取的特征向量, 预测窗口图像为目标或背景。滑动窗口策略参见文献[13]。

3.1 SDAE

SDAE从施加噪声的训练集中学习、重构原始

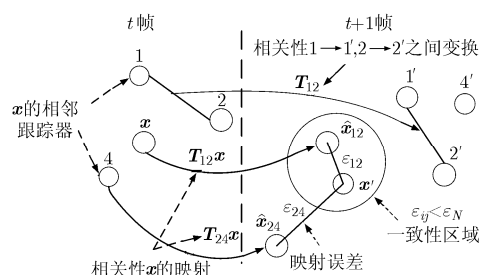


图2 空间上下文预测器预测过程

数据,通过优化重构误差提高深度网络对噪声的鲁棒性。采用无监督特征学习和逐层贪婪算法^[18]预训练多个自编码器,预训练过程如图3(a)。每个编码器隐藏层的网络单元都为输入层网络单元的一半,直到隐藏层网络单元数减到256为止。预训练完成后,展开多个自编码器得到SDAE,并添加超完备滤波层(2560),SDAE网络结构如图3(b)所示。为获得更有效的网络参数,反向传播原始数据与重构数据之间误差微调整个网络,微调后网络结构如图3(c)。由图3(b)、图3(c)的输出结果可看出,微调后得到的重构数据更接近原始数据,实现了网络参数优化。

3.2 SVM 分类器

在二值分类问题中,用线性SVM分类器取代传统的sigmoid分类器,提高深度检测器对相似目标的辨识能力,解决目标漂移或跟踪丢失问题,分类与微调过程如下。

假定训练样本为 (x_n, y_n) , $n = 1, 2, \dots, N$, $x_n \in \mathcal{R}^D$, $y_n \in \{-1, +1\}$, 无约束条件的SVM优化问题为

$$\min_{\mathbf{w}} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{n=1}^N \max(1 - \mathbf{w}^T x_n y_n, 0)^2 \quad (1)$$

式中, C 为规则参数, 式(1)为标准hinge loss函数的二阶范数优化问题(L2-SVM)。

预测数据 x 的类标记公式为

$$\arg \max_y (\mathbf{w}^T x y) \quad (2)$$

为了进一步优化深度分类神经网络,利用经典的反向传播算法,通过SVM分类器微调整个深度网络。令等式(3)为目标函数 $l(\mathbf{w})$,倒数第2层的激活值 h 取代输入数据 x ,优化公式为

$$\frac{\partial l(\mathbf{w})}{\partial h_n} = -2C y_n \mathbf{w} (\max(1 - \mathbf{w}^T h_n y_n, 0)) \quad (3)$$

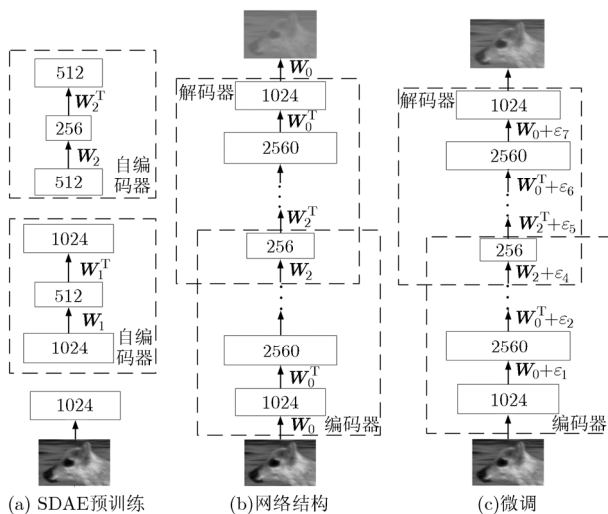


图3 SDAE构建图

4 加权 P-N 学习

在线 P-N 学习算法^[9]在标记样本过程中常出现错误标记问题,用错误标记的样本训练分类器,很大程度上降低分类器性能,导致目标偏移甚至跟踪失败。为解决上述问题,在分类过程中,对训练集合中每个样本赋予两类权值:正样本权值 W^+ 、负样本权值 W^- 。正、负样本权值分别由两部分构成:

$$\left. \begin{aligned} W^+ &= W_b^+ + W_c^+ \\ W^- &= W_b^- + W_c^- \end{aligned} \right\} \quad (4)$$

式中, W_b 为自扩散(bootstrapping)分类过程赋予的权值, W_c 为 SVM 分类过程赋予的权值。+表示正样本, -表示负样本。

在 bootstrapping 分类过程中, W_b^+ 和 W_b^- 计算公式为

$$\left. \begin{aligned} W_b^+ &= f^+ / (f^+ + f^-) \\ W_b^- &= f^- / (f^+ + f^-) \end{aligned} \right\} \quad (5)$$

其中, f^+ , f^- 分别为样本被分类为正样本、负样本的次数。

在分类过程中, SVM 分类器二值分类编码器提取的特征向量 \mathbf{x} , 每个图像块被分类为正样本的后验概率为 $P(y = 1 | \mathbf{x})$, 则分类权值计算公式为

$$\left. \begin{aligned} W_c^+ &= P(y = 1 | \mathbf{x}) \\ W_c^- &= 1 - P(y = 1 | \mathbf{x}) \end{aligned} \right\} \quad (6)$$

样本被最后被分类为正样本或负样本的计算公式为

$$\left. \begin{aligned} W^+ / W^- &\geq 1, \text{ 正样本} \\ W^+ / W^- &< 1, \text{ 负样本} \end{aligned} \right\} \quad (7)$$

5 算法实现

图4为算法流程图。

离线训练阶段,将 32×32 自然图像集^[19]进行归一化处理,并用1024维特征向量表示,每一维对应一个像素。利用预处理的数据,通过无监督特征学习预训练SDAE。预训练完成后,根据输入数据和重构数据之间的误差,通过经典反向传播算法微调整个网络。迁移学习将预训练得到SDAE的编码器转化到在线跟踪过程中。

为加快第1层预训练过程,将大小 32×32 的自然图像分为5个大小为 16×16 的图像块,位置分别在左上角、右上角、左下角、右下角及中心,然后训练这5个有512个隐藏单元的自编码器,根据5个编码器的权值初始化第1层的自编码器。

在线跟踪初始化阶段,首先在第1帧图像中手动选取待跟踪的目标,并对目标图像扭曲、旋转、缩放建立完整的目标全视角数据库^[20],以此作为正

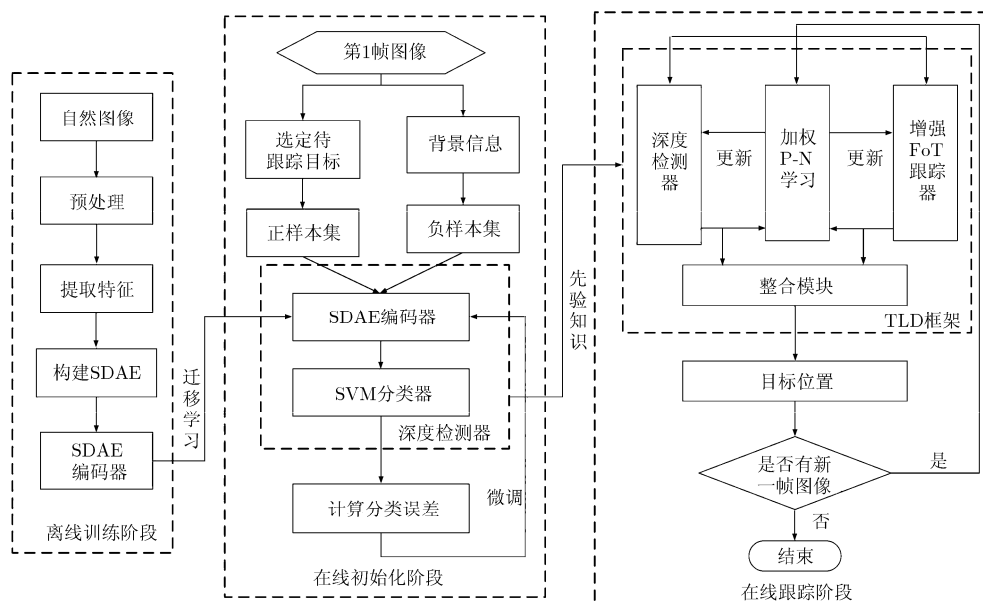


图4 算法流程图

样本，这对目标旋转、缩放、外观局部变化等都具有很好的鲁棒性，选取目标周围的背景区域作为负样本，通过正负样本集有监督训练深度检测器。正负样本集通过 SDAE 编码器提取样本特征，利用提取特征集和类标记训练 SVM 分类器。

在线跟踪阶段，增强 FoT 跟踪器和深度检测器并行处理每一帧图像，增强 FoT 跟踪器预测当前帧中目标所在的位置，深度检测器检测一个或多个可能的目标位置，整合模块整合检测结果和跟踪结果，得到当前帧是否存在目标、目标位置及当前帧跟踪轨迹是否有效等信息。加权 P-N 学习更新跟踪器和检测器。

TLD 中模板更新策略虽然适应目标外观变化及光照等外界环境影响，但是使在线模板数量不断增加，从而增加匹配过程中的计算复杂度。利用 K-means 算法，将在线模板分为两个子集合，构成二值树，计算复杂度从原来的 $O(n)$ 简化为 $O(\lg n)$ ，从而减少了时间复杂度。

6 实验结果与分析

第1部分为 TLD 对比实验，包括跟踪器、检测器及在线学习性能对比。用序列 David^[13], Jumping^[13], Pedestrian1^[13], Pedestrian2^[13], Car^[13] 对跟踪器进行性能比较，其中 T_{FB+NCC} 为 TLD 的跟踪器， $T_{cascade}$ 为有级联预测器的跟踪器， $T_{cascade+RANSAC}$ 为有级联预测器和快速 RANSAC 的跟踪器。通过成功跟踪目标帧数对短时跟踪器进行性能评估，跟踪结果与真实位置之间的重叠率大于 0.5，认定为跟踪成功，表 1 为 3 种跟踪器成功跟踪

帧数对比结果。

从表 1 可得， $T_{cascade+RANSAC}$ 成功跟踪目标的帧数多于其他两种跟踪器，因为级联预测器过滤掉没有正确跟踪局部跟踪器，即外点，减少外点对全局跟踪模型的干扰，同时快速 RANSAC 有效评估全局跟踪模型，有效解决目标漂移问题，实现更鲁棒的目标跟踪。

检测器性能比较，TLD 算法中检测器是基于 2 bit BP 特征和级联分类器，本文算法中检测器是基于深度学习和 SVM 分类器，两者均采用滑动窗口策略，图 5，图 6 为跟踪结果。

如图 5 所示，在 Freeman1^[1], Pktest01^[21] 序列中，TLD 算法(黑色虚线)均出现误匹配问题，将背景或相似目标作为跟踪结果，本文算法(黑色实线)可实现正确的跟踪。由于 TLD 算法仅使用像素级特征(2 bit BP)进行分类与模板匹配，将匹配度最高的作为跟踪结果，在相似背景的干扰下，极易产生错误匹配问题，而本文算法利用深度学习对目标进行有效表达，减少分类误差，解决误匹配问题，对场景中相似目标具有更高的辨识能力。如图 6 所示，在

表 1 目标成功跟踪帧数

图片序列	目标存在帧数	T_{FB+NCC}	$T_{cascade}$	$T_{cascade+RANSAC}$
David	761	761	761	761
Jumping	313	40	81	124
Pedestrian1	140	49	136	140
Pedestrian2	266	100	235	263
Car	860	570	580	640

Woman^[1], Subway^[1]序列中, TLD(黑色虚线)首先是出现漂移问题, 之后在遮挡、相似目标的干扰下, 跟踪失败。本文算法(黑色实线)在样本数不足时, 可实现鲁棒的跟踪。这是因为当训练样本数量不足时, TLD 算法训练得到级联分类器分类能力差, 导致目标偏移甚至跟踪失败。而本文算法可利用自然图像数据预训练深度网络, 再利用迁移学习将预训练知识应用到在线跟踪中, 克服训练样本不足问题。

在线学习机制对比实验, 经 David^[13], Carchase^[13], Panda^[13]序列测试, 图 7 为加权 P-N 学习与 P-N 学习跟踪结果对比图, 每组图片中第 1 列为目标的真实位置, 第 2 列 P-N 学习跟踪结果, 第 3 列为加权 P-N 学习跟踪结果, 可以看出 P-N 学习在跟踪过程中产生不同程度的目标漂移问题, 加权 P-N 学习考虑样本权重, 提高分类器的精确度, 解决目标漂移问题。

第 2 部分为与目前流行的 BSBT^[22], coGD^[23], CXT^[24]算法进行定性对比实验, 图 8 为 4 种跟踪算法在 Carchase^[13], Panda^[13]序列的跟踪结果。

在 Carchase 序列中跟踪目标为运动的小车, 小车在运动过程中常受到遮挡、相似背景及尺度变化等复杂环境的影响。在 238 帧中由于相似目标的干扰, CXT 和 coGD 算法出现目标漂移, BSBT 算法常出现错误跟踪, 本文算法可实现正确的跟踪。在 682, 1290 帧, CXT 和 coGD 算法均已跟踪失败, BSBT 算法同样出现错误跟踪问题, 本文算法依然具有稳定跟踪。这是由于 coGD, CXT, BSBT 算法

均采用像素级特征, 不能对图像进行有效的表达, 所以在相似背景条件下常出现目标漂移甚至错误跟踪, 本文算法利用深度网络提取图像本质特征, 即使有相似目标的干扰, 依然可以实现正确跟踪。

在 Panda 序列中大熊猫在行走时姿态不断发生改变, BSBT, CXT, coGD 算法在跟踪前期就出现漂移现象, 之后在 472 帧, CXT 和 coGD 算法出现目标漂移, BSBT 已跟踪失败, 但本文算法依然能够实现稳定跟踪。在 1472 帧目标重新出现在视场中, 本文算法能够重新准确的定位到目标, BSBT 算法出现目标漂移现象, CXT 和 coGD 算法均跟踪失败。由于目标在行走中姿态不断变化, 导致 BSBT, CXT, coGD 算法在线学习过程中训练样本不足, 同时样本得不到及时更新, 得到的分类器分类精度不高, 最终导致目标漂移甚至跟踪失败。由于跟踪过程中可用训练样本数量很少, 本文算法可利用自然图像集预训练深度检测器, 优化网络参数, 将预训练得到的深度检测器应用到在线跟踪过程中, 解决训练样本不足导致的分类精度不高问题, 实现鲁棒的跟踪。

第 3 部分为与 MIL^[3], OAB^[6], SBT^[7]跟踪算法定量对比实验, 这些算法参数均使用默认值, 具体实现参考相应文献。使用准确率、召回率判定算法优越性, 为了使结果更加一般化, 每个算法在每组序列上均进行 20 次测试, 取其平均值作为最后的跟踪结果, 跟踪结果准确率如表 2 所示。

准确率越高表明算法跟踪精度越高, 由表 2 可

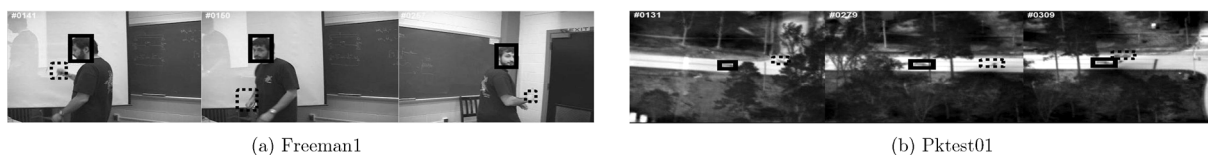


图 5 误匹配跟踪结果

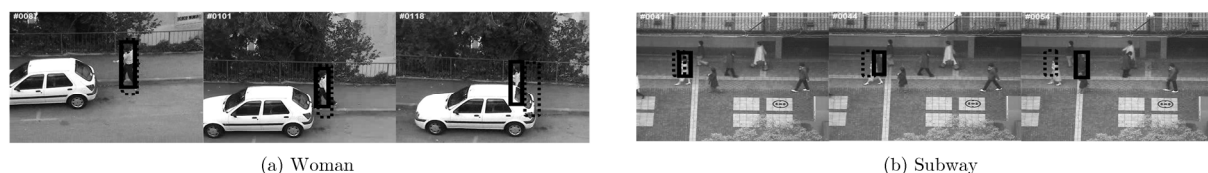


图 6 训练样本不足跟踪结果



图 7 加权 P-N 学习与 P-N 学习跟踪结果

得，在 David 等^[13]10 组图片序列中，本文算法的准确率有 6 组结果为最好，3 组为次好结果，表明本文算法在复杂环境下如遮挡、相似背景、剧烈运动等具有更高跟踪精度。召回率越高表明算法能够正

确检测及跟踪目标的帧数越多，鲁棒性越好。表 3 为跟踪结果的召回率，由表 3 可得，本文算法的召回率均高于其他算法，表明本文算法在目标可见时，实现成功跟踪的比率更大，鲁棒性更好。

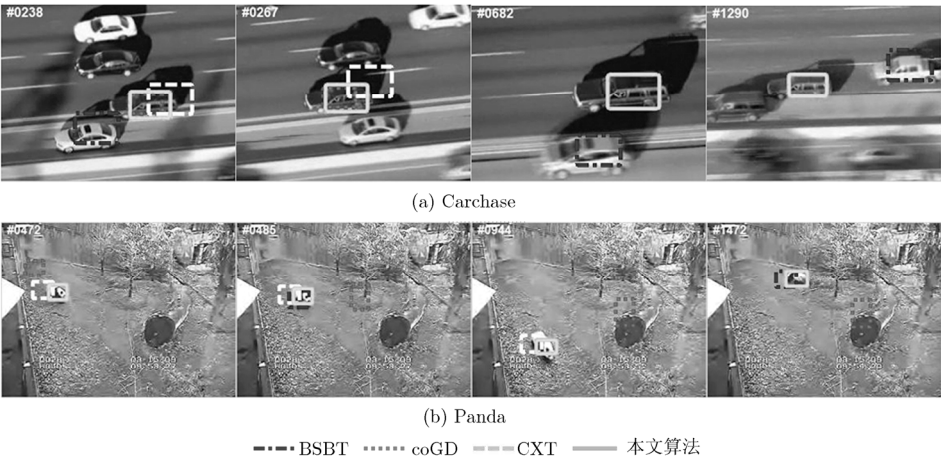


图 8 跟踪结果对比图

表 2 平均准确率

图片序列	OAB	SBT	MIL	FDT
David	0.01	<u>0.27</u>	0.06	1.00
Jumping	<u>0.41</u>	0.14	0.37	0.99
Pedestrian1	0.36	0.20	<u>0.42</u>	1.00
Pedestrian2	<u>0.74</u>	0.55	0.10	0.88
Pedestrian3	1.00	0.41	0.49	<u>0.95</u>
Car	0.89	1.00	0.11	<u>0.93</u>
Motocross	<u>0.13</u>	0.01	0.02	0.71
Volkswagen	0.04	0.01	<u>0.26</u>	0.53
Carchase	<u>0.73</u>	0.79	0.49	0.53
Panda	0.28	0.58	0.14	<u>0.32</u>

注：表中黑体为最好结果，下划线为次好结果

表 3 平均召回率

图片序列	OAB	SBT	MIL	本文算法
David	0.01	<u>0.17</u>	0.06	0.56
Jumping	0.03	0.05	<u>0.20</u>	0.93
Pedestrian1	0.06	0.14	<u>0.32</u>	1.00
Pedestrian2	0.12	<u>0.46</u>	0.11	0.92
Pedestrian3	0.33	0.31	<u>0.58</u>	0.96
Car	0.57	<u>0.67</u>	0.12	0.92
Motocross	0.02	<u>0.03</u>	0.01	0.58
Volkswagen	0.03	0.02	<u>0.04</u>	0.64
Carchase	0.05	<u>0.07</u>	0.06	0.40
Panda	0.27	0.26	<u>0.35</u>	0.86

注：表中黑体为最好结果，下划线为次好结果

7 结束语

本文提出以 TLD 为框架, 基于增强 FoT 和深度学习的目标跟踪算法, 基于时空上下文的级联预测器和快速 RANSAC 算法提高跟踪器的稳定性。深度学习与线性 SVM 构建深度检测器, 克服了基于外观模型和传统机器学习目标跟踪算法的缺点, 将深度学习与计算机视觉(SVM)相结合应用到目标跟踪领域。在复杂环境下与其他跟踪算法进行比较, 实验结果表明, 本算法在复杂环境下具有更高的准确性和鲁棒性, 具有极高的实用价值。下一步工作利用核 SVM、随机深林等分类器取代线性 SVM 分类器, 进一步提高分类精确度。

参 考 文 献

- [1] Wu Y, Lim J, and Yang M H. Online object tracking: A benchmark[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Portland, USA, 2013: 2411-2418.
- [2] Ross D A, Lim J, Lin R S, *et al.*. Incremental learning for robust visual tracking[J]. *International Journal of Computer Vision*, 2008, 77(3): 125-141.
- [3] Babenko B, Yang M H, and Belongie S. Robust object tracking with online multiple instance learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(8): 1619-1632.
- [4] 陈东成, 朱明, 高文, 等. 在线加权多示例学习实时目标跟踪[J]. *光学精密工程*, 2014, 22(6): 1661-1667.
Chen Dong-cheng, Zhu Ming, Gao Wen, *et al.*. Real-time object tracking via online weighted multiple instance learning [J]. *Optics and Precision Engineering*, 2014, 22(6): 1661-1667.
- [5] He S F, Yang Q X, Rynson L, *et al.*. Visual Tracking via Locality Sensitive Histograms[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Portland, USA, 2013: 2427-2434.
- [6] Grabner H, Grabner M, and Bischof H. Real-time tracking via online boosting[C]. Proceedings of British Machine Vision Conference, Edinburgh, UK, 2006: 47-56.
- [7] Grabner H, Leistner C, and Bischof H. Semi-supervised on-line boosting for robust tracking[C]. Proceedings of European Conference on Computer Vision, Berlin, Germany, 2008: 234-247.
- [8] 颜佳, 吴敏渊. 遮挡环境下采用在线 Boosting 的目标跟踪[J]. *光学精密工程*, 2012, 20(2): 439-446.
Yan Jia and Wu Ming-yuan. On-line boosting based target tracking under occlusion[J]. *Optics and Precision Engineering*, 2012, 20(2): 439-446.
- [9] Kalal Z, Matas J, and Mikolajczyk K. P-N learning: bootstrapping binary classifiers by structural constraints[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, New York, USA, 2010: 49-56.
- [10] 郑胤, 陈权崎, 章毓晋. 深度学习及其在目标和行为识别中的新进展[J]. *中国图像图形学报*, 2014, 19(2): 175-184.
Zheng Ying, Chen Quan-qi, and Zhang Yu-jin. Deep learning and its new progress in object and behavior recognition[J]. *Journal of Image and Graphic*, 2014, 19(2): 175-184.
- [11] Tomas V and Jiri M. Robustifying the flock of trackers[C]. Proceedings of Computer Vision Winter Workshop, Graz, Austria, 2011: 91-97.
- [12] 周鑫, 钱秋朦, 叶永强, 等. 改进后的 TLD 视频目标跟踪方法[J]. *中国图象图形学报*, 2013, 18(9): 1115-1123.
Zhou Xin, Qian Qiu-meng, Ye Yong-qiang, *et al.*. Improved TLD visual target tracking algorithm[J]. *Journal of Image and Graphic*, 2013, 18(9): 1115-1123.
- [13] Kalal Z, Mikolajczyk K, and Matas J. Tracking-learning-detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(7): 1409-1422.
- [14] Zhang K, Zhang L, Liu Q, *et al.*. Fast visual tracking via dense spatio-temporal context learning[C]. Proceedings of European Conference on Computer Vision, Zurich, Switzerland, 2014: 127-141.
- [15] Botterill T, Mills S, and Green R D. New conditional sampling strategies for speeded-up RANSAC[C]. Proceedings of British Machine Vision Conference, London, UK, 2009: 1-11.
- [16] Vincent P, Larochelle H, Lajoie I, *et al.*. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion[J]. *Journal of Machine Learning Research*, 2010, 11(2): 3371-3408.
- [17] Tang Yi-chuan. Deep learning using linear support vector machines[C]. Proceedings of International Conference on Machine Learning: Challenges in Representational Learning Workshop, Atlanta, USA, 2013: 266-272.
- [18] Hinton G E and Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504-507.
- [19] Torralba A, Fergus R, and Freeman W T. 80 million tiny images: a large data set for nonparametric object and scene recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(11): 1958-1970.
- [20] 高文, 汤洋, 朱明. 复杂背景下目标检测的级联分类器算法研究[J]. *物理学报*, 2014, 63(9): 094204.
Gao Wen, Tang Yang, and Zhu Ming. Study on the cascade classifier in target detection under complex background[J]. *Acta Physica Sinica*, 2014, 63(9): 094204.
- [21] Collins R T, Zhou X H, and Teh S K. An open source tracking test bed and evaluation web site[C]. Proceedings of IEEE International Workshop on Performance Evaluation of

- Tracking and Surveillance, Breckenridge, USA, 2005: 17–24.
- [22] Stalder S, Grabner H, and Van G L. Beyond semi-supervised tracking: tracking should be as simple as detection, but not simpler than recognition[C]. Proceedings of IEEE Conference on Computer Vision Workshops, Kyoto, Japan, 2009: 1409–1416.
- [23] Dinh T B, Vo N, and Medion G. Context tracker: exploring supporters and distracters in unconstrained environments[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Providence, USA, 2011: 1177–1184.
- [24] Qian Yu, Thang B D, and Gerard M. Online tracking and reacquisition using co-trained generative and discriminative trackers[C]. Proceedings of European Conference on Computer Vision, Marseille, France, 2008: 678–691.
- 程 帅：男，1987 年生，博士生，研究方向为图像处理、目标跟踪、深度学习。
- 曹永刚：男，1972 年生，博士生，研究员，研究方向为光电测控设备总体及时统技术。
- 孙俊喜：男，1971 年生，博士，教授，研究方向为模式识别与智能系统、目标的检测与跟踪、嵌入式车牌识别系统、医学图像处理与分析。
- 赵立荣：女，1971 年生，博士，研究员，研究方向为视频判读、数据处理等。
- 刘广文：男，1971 年生，博士，副教授，研究方向为智能信息处理。
- 韩广良：男，1968 年生，博士，研究员，研究方向为实时视频处理、视频目标识别和跟踪、计算机视觉。