

基于在线半监督 boosting 的协同训练目标跟踪算法

陈思 苏松志 李绍滋* 吕艳萍 曹冬林

(厦门大学信息科学与技术学院 厦门 361005)

(福建省仿脑智能系统重点实验室(厦门大学) 厦门 361005)

摘要: 基于自训练的判别式目标跟踪算法使用分类器的预测结果更新分类器自身, 容易累积分类错误, 从而导致漂移问题。为了克服自训练跟踪算法的不足, 该文提出一种基于在线半监督 boosting 的协同训练目标跟踪算法(简称 Co-SemiBoost), 其采用一种新的在线协同训练框架, 利用未标记样本协同训练两个特征视图中的分类器, 同时结合先验模型和在线分类器迭代预测未标记样本的类标记和权重。该算法能够有效提高分类器的判别能力, 鲁棒地处理遮挡、光照变化等问题, 从而较好地适应目标外观的变化。在若干个视频序列的实验结果表明, 该算法具有良好的跟踪性能。

关键词: 目标跟踪; 在线学习; 半监督学习; 协同训练

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2014)04-0888-08

DOI: 10.3724/SP.J.1146.2013.00826

A Novel Co-training Object Tracking Algorithm Based on Online Semi-supervised Boosting

Chen Si Su Song-zhi Li Shao-zi Lü Yan-ping Cao Dong-lin

(School of Information Science and Technology, Xiamen University, Xiamen 361005, China)

(Fujian Key Laboratory of the Brain-like Intelligent Systems (Xiamen University), Xiamen 361005, China)

Abstract: The self-training based discriminative tracking methods use the classification results to update the classifier itself. However, these methods easily suffer from the drifting issue because the classification errors are accumulated during tracking. To overcome the disadvantages of self-training based tracking methods, a novel co-training tracking algorithm, termed Co-SemiBoost, is proposed based on online semi-supervised boosting. The proposed algorithm employs a new online co-training framework, where unlabeled samples are used to collaboratively train the classifiers respectively built on two feature views. Moreover, the pseudo-labels and weights of unlabeled samples are iteratively predicted by combining the decisions of a prior model and an online classifier. The proposed algorithm can effectively improve the discriminative ability of the classifier, and is robust to occlusions, illumination changes, etc. Thus the algorithm can better adapt to object appearance changes. Experimental results on several challenging video sequences show that the proposed algorithm achieves promising tracking performance.

Key words: Object tracking; Online learning; Semi-supervised learning; Co-training

1 引言

目标跟踪是人工智能和计算机视觉的重要研究课题^[1]。然而, 由于受到场景中各种复杂因素的影响, 如光照、旋转、移动和遮挡等, 目标跟踪已成为一项极具挑战的任务^[2]。目前研究者已经提出诸多目标跟踪算法^[3-12], 大致分为生成式算法^[3-5]和判别式

算法^[6-12]。判别式目标跟踪算法^[6-12]将跟踪视为目标和背景的 2 类分类问题, 受到学者的广泛关注。现有的大多数判别式目标跟踪算法^[6-8], 例如在线 boosting^[6,7]和在线半监督 boosting 算法^[8], 利用分类器的预测结果更新分类器自身, 即称为自训练 (self-training)^[13,14]。此类方法对类别噪声比较敏感, 容易累积分类错误。

为了克服自训练目标跟踪算法的不足, 本文提出一种基于在线半监督 boosting 的协同训练目标跟踪算法 (Co-training based on online Semi-supervised Boosting, 简称 Co-SemiBoost)。本文的

2013-06-07 收到, 2013-09-03 改回

国家自然科学基金(61201359, 61202143), 福建省自然科学基金(2011J01367, 2012J05126)和高等学校博士学科点专项科研基金(20090121110032)资助课题

*通信作者: 李绍滋 szlig@xmu.edu.cn

主要创新点如下：(1)提出了一种新的在线协同训练框架，利用未标记样本协同训练不同视图中的弱分类器，能够在克服目标漂移的同时保持对目标外观的自适应更新；(2)基于半监督技术，结合目标先验模型和在线分类器迭代预测未标记样本的伪类别标记和权重。本文算法能够鲁棒地处理目标遮挡、光照变化、快速运动和复杂背景等问题。若干具有挑战性的视频序列实验结果表明，本文提出的算法具有良好的目标跟踪性能。

2 基于在线半监督 boosting 的协同训练目标跟踪算法

2.1 目标跟踪流程

Co-SemiBoost 算法是一种基于检测的跟踪算法(tracking-by-detection)^[10]。Co-SemiBoost 目标跟踪算法的具体流程如图 1 所示。

跟踪过程中，Co-SemiBoost 算法迭代执行以下两个步骤：(1)分类器更新：在第 t 帧目标位置周围区域内获取若干图像块作为未标记样本，并依次利用在线协同训练框架更新强分类器；(2)目标检测：利用更新后的强分类器预测第 $t+1$ 帧中搜索区域内各个图像块的置信度，置信度最高的图像块被视为新的目标位置。其中，分类器更新是目标跟踪算法的关键步骤。Co-SemiBoost 算法采用一种有效的基于在线半监督 boosting 的协同训练框架来更新强分类器，从而克服自训练跟踪算法的不足。

2.2 在线协同训练

协同训练算法^[15,16]首先利用少量有标记样本训练两个初始分类器，然后在学习过程中，这些分类器挑选若干个置信度高的未标记样本进行标记并用于更新对方分类器，此过程不断迭代，从而有效提高分类性能。理论已证明^[16]，当数据集有两个充分冗余的视图，即两个满足下述条件的特征集：(1)如果训练样本足够，在每个特征集上都足以学到一个

强分类器；(2)在给定类别标记时，每个特征集都条件独立于另一个特征集，那么协同训练算法可以有效地利用未标记样本提升分类器的性能。本文利用 Haar^[17]特征和 LBP^[18]特征分别描述灰度图像，获得两个充分冗余的视图，并在这两个视图上进行协同训练，从而有效提高分类性能。Haar 和 LBP 特征具有不同的目标辨别能力。Haar 特征主要关注局部纹理变化信息(即对目标的姿态、形变等变化比较鲁棒)，而 LBP 特征更多关注边缘变化信息(即对光照等变化比较鲁棒)。因此，对 Haar 特征难以区分的样本，LBP 特征可能拥有较好的区分能力，反之亦然。

Co-SemiBoost 算法所采用的在线协同训练框架如图 2 所示，其中视图 1 和视图 2 分别表示 Haar 和 LBP 特征视图。本文使用了“选择器”^[6]的概念，其定义如下：给定一个 M 个弱分类器的集合，一个选择器根据某一优化准则从中选出一个最优弱分类器。训练一个选择器意味着 M 个弱分类器被更新并从中选出一个最优弱分类器。假设每个特征视图包含 N 个选择器，每个选择器包含 M 个弱分类器，则第 j 个视图中第 n 个选择器 Sel_n^j 包含的 M 个弱分类器记为 $\{h_{n,1}^j, h_{n,2}^j, \dots, h_{n,M}^j\}$ ($j=1, 2; 1 \leq n \leq N$)，其中每个弱分类器对应该特征视图下的一个随机特征。

如图 2 所示，Co-SemiBoost 算法采用在线学习方式^[6]将当前帧目标周围区域的未标记样本依次作为协同训练的输入样本，即每次利用一个未标记样本更新所有选择器。该算法的在线协同训练过程具体描述如下：首先，随机初始化每个特征视图的 N 个选择器。假设第 1 帧的目标位置已知，则从第 1 帧中获取有标记数据 $X^L = \{ \langle \mathbf{x}_1, y_1 \rangle, \dots, \langle \mathbf{x}_{X^L}, y_{X^L} \rangle \}$ ，即目标区域作为正样本，非目标区域作为负样本。初始时，在两个特征视图上分别利用有标记数据 X^L 训练先验分类器 $H_p^1(\mathbf{x})$ 和

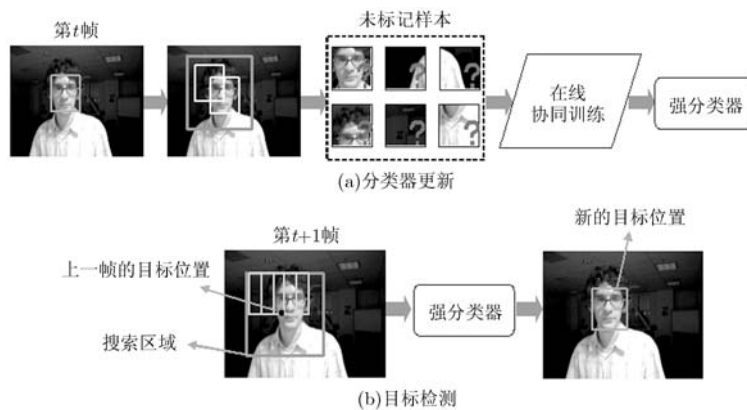


图 1 Co-SemiBoost 目标跟踪算法流程图

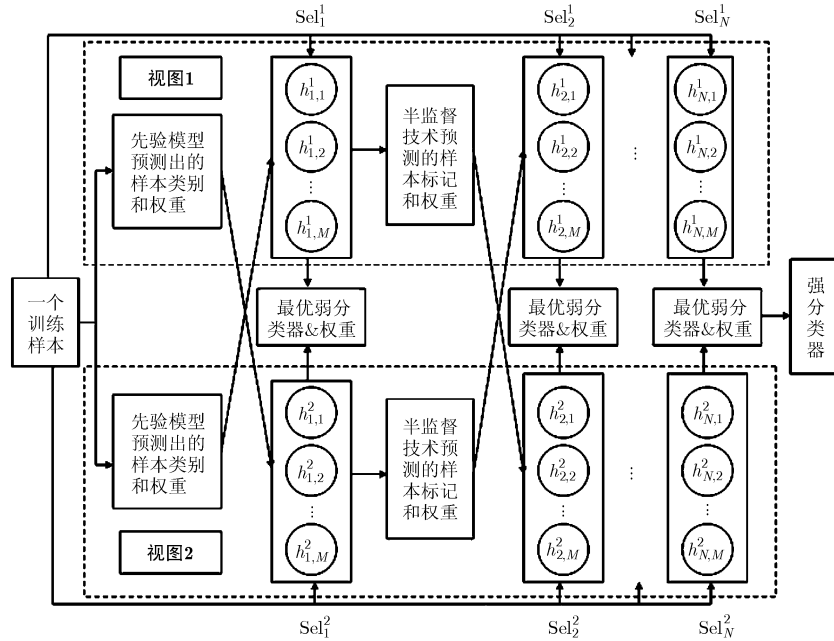


图2 Co-SemiBoost 算法的在线协同训练框架

$H_p^2(\mathbf{x})$ ，这里采用 boosting 方法进行训练。其次，在两个视图中分别利用先验模型预测未标记样本的类别和权重，并用于更新对方视图的选择器 Sel_1^j ($j = 1, 2$) 中的 M 个弱分类器。然后，当选择器 Sel_{n-1}^j ($j = 1, 2; n = 2, 3, \dots, N$) 中的 M 个弱分类器被更新后，利用半监督技术预测该未标记样本的伪类别标记和权重，并用于更新对方视图的下一个选择器 Sel_n^{3-j} 中的 M 个弱分类器，此步骤不断迭代，直到更新完第 N 个选择器为止。同时在上述步骤中，两个视图中的每个选择器依次选出最优弱分类器并计算相应的投票权重。最后将各最优弱分类器及其权重集成为最终的强分类器。

2.3 类别预测与权重估计

基于在线半监督 boosting 算法^[8]，Co-SemiBoost 算法通过衡量未标记样本与有标记样本的相似度来判断未标记样本的类别标记。为了更新第 j 个特征视图 ($j = 1, 2$) 中的每个选择器 Sel_n^j ($n = 1, 2, \dots, N$)，该算法首先利用第 $3-j$ 个特征视图的分类信息来预测未标记样本 \mathbf{x} 属于正类和负类的置信度，具体计算如下：

$$p_n^j(\mathbf{x}) \approx e^{-H_{n-1}^{3-j}(\mathbf{x})} \sum_{\mathbf{x}_i \in X^+} S(\mathbf{x}, \mathbf{x}_i) \approx \frac{e^{-H_{n-1}^{3-j}(\mathbf{x})} e^{H_p^{3-j}(\mathbf{x})}}{e^{H_p^{3-j}(\mathbf{x})} + e^{-H_p^{3-j}(\mathbf{x})}} \quad (1)$$

$$q_n^j(\mathbf{x}) \approx e^{H_{n-1}^{3-j}(\mathbf{x})} \sum_{\mathbf{x}_i \in X^-} S(\mathbf{x}, \mathbf{x}_i) \approx \frac{e^{H_{n-1}^{3-j}(\mathbf{x})} e^{-H_p^{3-j}(\mathbf{x})}}{e^{H_p^{3-j}(\mathbf{x})} + e^{-H_p^{3-j}(\mathbf{x})}} \quad (2)$$

其中 $H_p^{3-j}(\mathbf{x})$ 是第 $3-j$ 个视图的先验分类器； $H_{n-1}^{3-j}(\mathbf{x}) = \sum_{t=1}^{n-1} \alpha_t^{3-j} h_t^{3-j}(\mathbf{x})$ 是根据第 $3-j$ 个视图

中前 $n-1$ 个选择器学习得到的强分类器； X^+ 和 X^- 分别表示有标记的正样本和负样本的集合； $S(\cdot, \cdot)$ 是相似度函数，其利用先验模型 $H_p^{3-j}(\mathbf{x})$ 来估计未标记样本 \mathbf{x} 属于正类或负类的概率。

为了更新第 j 个视图的第 n 个选择器 Sel_n^j ，本文结合式(1)和式(2)得到未标记样本 \mathbf{x} 的伪软类别标记(pseudo-soft-label) $z_n^j(\mathbf{x})$ ，计算如式(3)：

$$z_n^j(\mathbf{x}) = p_n^j(\mathbf{x}) - q_n^j(\mathbf{x}) = \frac{\sinh(H_p^{3-j}(\mathbf{x}) - H_{n-1}^{3-j}(\mathbf{x}))}{\cosh(H_p^{3-j}(\mathbf{x}))} \quad (3)$$

从而，用于更新选择器 Sel_n^j 中的 M 个弱分类器的未标记样本 \mathbf{x} 的伪类别标记 $y_n^j(\mathbf{x})$ 和权重 $\lambda_n^j(\mathbf{x})$ 的计算公式如式(4)和式(5)：

$$y_n^j(\mathbf{x}) = \text{sign}(p_n^j(\mathbf{x}) - q_n^j(\mathbf{x})) \quad (4)$$

$$\lambda_n^j(\mathbf{x}) = |p_n^j(\mathbf{x}) - q_n^j(\mathbf{x})| \quad (5)$$

其中 $\text{sign}(\cdot)$ 是符号函数。权重 $\lambda_n^j(\mathbf{x})$ 也用于计算弱分类器的更新次数^[19]。样本权重越高，则表明对样本的分类置信度越高。当目标被遮挡时，在线分类器 $H_{n-1}^{3-j}(\mathbf{x})$ 会受到遮挡区域的影响，难以对目标的外观进行建模，而通过结合先验模型 $H_p^{3-j}(\mathbf{x})$ ，即保存了第 1 帧中目标的外观特征，可以减少遮挡对在线分类器的影响。

在协同训练过程中，当选择器 Sel_n^j 被更新后，从中选出一个最低错误率的弱分类器(最优弱分类器)，即 $h_n^j = h_{n,m^+}^j$ ($m^+ = \arg \min_m (e_{n,m}^j)$)，其中 $e_{n,m}^j$ 表示弱分类器 $h_{n,m}^j$ 的分类错误率，具体计算如式(6)：

$$e_{n,m}^j = \frac{\lambda_{n,m,w}^j}{\lambda_{n,m,c}^j + \lambda_{n,m,w}^j} \quad (6)$$

其中, 变量 $\lambda_{n,m,c}^j$ (或 $\lambda_{n,m,w}^j$) 表示在跟踪过程中从第 1 个训练样本至当前训练样本为止被弱分类器 $h_{n,m}^j$ 正确分类(或错误分类)的样本的权重累加和。最优弱分类器 $h_n^j = h_{n,m^+}^j$ 对应的投票权重 α_n^j 定义如式(7):

$$\alpha_n^j = \frac{1}{2} \ln \left(\frac{1 - e_n^j}{e_n^j} \right) \quad (7)$$

其中 $e_n^j = e_{n,m^+}^j$ 是最优弱分类器的分类错误率。分类器的错误率越高, 投票权重越低, 反之亦然。最后, 两个视图中每个选择器选出的最优弱分类器 h_n^j ($j = 1, 2; n = 1, 2, \dots, N$) 及其投票权重 α_n^j 集成为最终的强分类器 $H(\mathbf{x}) = \sum_{n=1}^N \sum_{j=1}^2 \alpha_n^j h_n^j$ 。

2.4 不同视图的弱分类器构建

本文使用阈值法^[6]构建 Haar 特征对应的弱分类器。将每个样本上随机位置和大小 Haar 值作为一个特征值 $f_{n,m}^j(\mathbf{x})$, 并构建 $f_{n,m}^j(\mathbf{x})$ 对应的弱分类器 $h_{n,m}^j(\mathbf{x})$:

$$h_{n,m}^j(\mathbf{x}) = p_{n,m}^j \cdot \text{sign}(f_{n,m}^j(\mathbf{x}) - \theta_{n,m}^j) \quad (8)$$

其中 $\text{sign}(\cdot)$ 是符号函数; $\theta_{n,m}^j = |\mu^+ + \mu^-|/2$; $p_{n,m}^j = \text{sign}(\mu^+ - \mu^-)$; μ^+ 和 μ^- 分别为正负有标记样本的均值, 并假设正负样本满足高斯分布。本文利用 Kalman 滤波技术^[6]迭代更新均值 μ^+ 和 μ^- 。

本文使用最近邻方法^[6]构建 LBP 特征对应的弱分类器。对随机图像区域里的每个像素点计算其 4 个邻点的 LBP 值, 并对该图像区域生成 16-bin 的 LBP 直方图特征, 然后构建该特征 $f_{n,m}^j(\mathbf{x})$ 对应的弱分类器:

$$h_{n,m}^j(\mathbf{x}) = \text{sign}(D(f_{n,m}^j(\mathbf{x}), P_{n,m}^j) - D(f_{n,m}^j(\mathbf{x}), N_{n,m}^j)) \quad (9)$$

其中, 距离函数 $D(\cdot, \cdot)$ 采用 Chi square (χ^2) 距离度量, $P_{n,m}^j$ 和 $N_{n,m}^j$ 是正负簇中心。

Co-SemiBoost 算法的伪代码描述如表 1 所示。

3 实验结果与分析

本文实验分为两部分: (1) 对比在线半监督 boosting 算法分别在 Haar 和 LBP 特征视图进行自训练, 以及本文提出的 Co-SemiBoost 算法在 Haar 和 LBP 特征视图上进行协同训练的跟踪性能。(2) 对比 Co-SemiBoost 算法和 3 种新近跟踪算法的性能。3 种跟踪算法分别为在线半监督 boosting 算法 (SemiBoost)^[8], 在线 boosting 算法 (OAB)^[6] 和在线多示例跟踪器 (MIL)^[10]。4 种方法均使用 Haar 和 LBP 两种特征进行训练。本文选择 4 个具有挑战性的公共视频序列^[20]进行实验, 分别为 Coupon Book, Cola Can, Tiger 1 和 Tiger 2。实验使用统一的参数设置。选择器的个数 N 设为 25。对于 Haar 特征和 LBP 特征, 弱分类器的个数 M 均设为 100。搜索区域为上一帧目

表 1 Co-SemiBoost 算法的伪代码

算法: Co-SemiBoost

初始化: 第 1 帧中有标记训练样本; 先验分类器 $H_p^1(\mathbf{x})$ 和 $H_p^2(\mathbf{x})$;

强分类器 $H_0^2(\mathbf{x}) = 0$; 每个弱分类器的分类权值

$\lambda_{n,m,c}^j = \lambda_{n,m,w}^j = 1$ 。

输入: 未标记的训练样本 \mathbf{x}

for $n = 1, 2, \dots, N$

for $j = 1, 2$

(1) 利用式(4)更新 \mathbf{x} 的伪类别标记 $y_n^j(\mathbf{x})$;

(2) 利用式(5)更新 \mathbf{x} 的权重 $\lambda_n^j(\mathbf{x})$;

(3) for $m = 1, 2, \dots, M$

(a) 利用 $y_n^j(\mathbf{x})$ 和 $\lambda_n^j(\mathbf{x})$ 更新 $h_{n,m}^j$;

(b) 如果 $h_{n,m}^j(\mathbf{x})$ 等于 $y_n^j(\mathbf{x})$, 则 $\lambda_{n,m,c}^j = \lambda_{n,m,c}^j + \lambda_n^j(\mathbf{x})$;

否则 $\lambda_{n,m,w}^j = \lambda_{n,m,w}^j + \lambda_n^j(\mathbf{x})$;

(c) 利用式(6)计算 $h_{n,m}^j$ 的分类错误率 $e_{n,m}^j$;

(4) end for

(5) 选出最优弱分类器 $h_n^j = h_{n,m^+}^j$ ($m^+ = \arg \min_m (e_{n,m}^j)$), 其错误率 $e_n^j = e_{n,m^+}^j$;

(6) 利用式(7)计算 h_n^j 的投票权重 α_n^j ;

(7) 生成在线分类器 $H_n^j(\mathbf{x}) = \sum_{t=1}^n \alpha_t^j h_t^j(\mathbf{x})$;

end for

end for

输出: 最终的强分类器 $H(\mathbf{x}) = \sum_{n=1}^N \sum_{j=1}^2 \alpha_n^j h_n^j(\mathbf{x})$ 。

标窗口区域的 2 倍。SemiBoost^[8], OAB^[6] 和 MIL 算法^[10] 的其余参数均使用默认参数设置。所有视频序列仅仅已知第 1 帧中目标位置。实验平台为 Windows XP 系统, 3.01 GHz 处理器和 3.25 GB 内存的个人计算机。

实验记录了视频序列的目标中心位置的平均误差值, 即所有帧的目标中心位置误差值之和与视频帧数的比值。此外, 实验也采用了“目标跟踪精确度图”^[10]来验证跟踪性能。该图记录了随着阈值的变化, 算法跟踪到的目标中心和实际目标中心之间距离小于给定阈值的帧数百分比。本文阈值区间定为 $[0, 100]$ 。本文还给出阈值为 20 时目标跟踪的精确度, 其相当于算法跟踪到的目标窗口和实际目标窗口有至少 50% 覆盖面积的帧数百分比^[10]。此外, 实验还对比了各算法的运行速度, 即在 4 个视频序列上每秒平均运行的帧数。

表 2 列出了对比算法的目标中心位置的平均误差值和运行速度(平均帧数/s), 其中最小平均误差值用粗体字表示。由表 2 可知, 本文的 Co-SemiBoost 算法的平均误差值明显低于 SemiBoost(Haar) 和 SemiBoost(LBP) 算法。由于 SemiBoost(Haar) 和 SemiBoost(LBP) 包含的弱分类器个数是 Co-SemiBoost 算法的一半, 故其运行速度略快于 Co-SemiBoost 算法。但 SemiBoost(Haar) 和 SemiBoost

表2 SemiBoost(Haar), SemiBoost(LBP)与本文算法的目标中心位置平均误差值(像素)和运行速度(平均帧数/s)

视频序列	SemiBoost (Haar)	SemiBoost (LBP)	Co-SemiBoost
Coupon Book	86.16	69.75	7.35
Cola Can	68.91	95.74	21.81
Tiger 1	68.35	53.47	24.22
Tiger 2	76.29	83.11	33.09
运行速度	9.12	7.93	6.78

表3 4种对比算法的目标中心位置平均误差值(像素)和运行速度(平均帧数/s)

视频序列	SemiBoost	OAB	MIL	Co-SemiBoost
Coupon Book	27.40	64.88	22.36	7.35
Cola Can	80.22	39.51	84.74	21.81
Tiger 1	78.14	41.25	82.41	24.22
Tiger 2	63.53	75.30	97.00	33.09
运行速度	6.35	8.91	7.54	6.78

(LBP)容易累积分类错误,其目标位置误差值远大于Co-SemiBoost算法。可见,本文的在线协同训练机制是有效的,克服了自训练的不足。

本实验对比了Co-SemiBoost算法与SemiBoost算法(同时使用Haar和LBP特征视图),OAB算法和MIL算法的跟踪性能。表3给出了SemiBoost, OAB, MIL和Co-SemiBoost算法的平均误差值和运行速度(平均帧数/s),表3表明Co-SemiBoost算法的平均误差值明显低于其它3种对比算法。从运行速度看,当同时使用Haar和LBP特征且弱分类器个数相同时,OAB算法运行速度最快,MIL算法次之,SemiBoost和Co-SemiBoost算法的运行速度相当。由于SemiBoost和Co-SemiBoost算法需要迭代地更新弱分类器,因此增加了时间复杂度。可见,Co-SemiBoost算法具有鲁棒的跟踪性能和良好的跟踪速度。

图3所示为对比算法在阈值区间[0,100]的目标跟踪精确度。从图3可知,Co-SemiBoost算法在较低阈值处具有较高的精确度。表4具体给出了阈值为20

时的目标跟踪精确度,其中最高精确度用粗体字表示。例如Coupon Book视频有约90%视频帧的目标中心位置距离实际中心位置在20个像素以内。然而当阈值为20时,其它3种算法在4个视频中的精确度均明显低于Co-SemiBoost算法。因此,实验结果表明Co-SemiBoost算法采用的协同训练机制是有效可行的,能够自适应目标外观的变化,同时较好地克服目标漂移问题。

图4给出了4个视频中有代表性的目标跟踪效果图。如图4(a)所示,Coupon Book视频在第52帧时

表4 阈值为20时的目标跟踪精确度(%)

视频序列	SemiBoost	OAB	MIL	Co-SemiBoost
Coupon Book	18.18	39.39	27.27	89.39
Cola Can	1.69	25.42	5.08	59.32
Tiger 1	2.82	40.85	1.41	49.30
Tiger 2	1.37	21.92	2.74	41.10

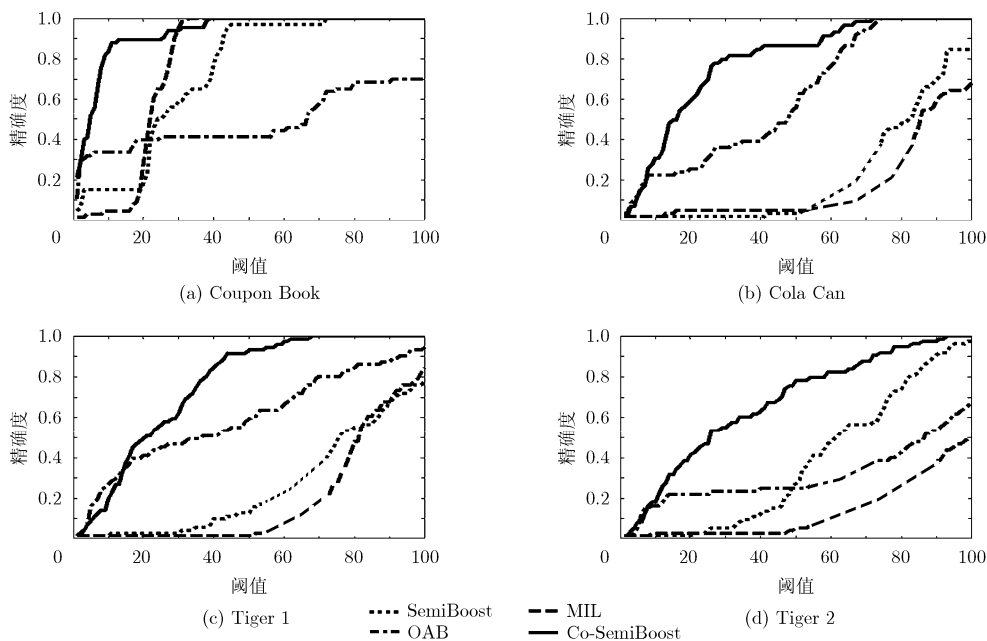


图3 4种对比算法的目标跟踪精确度图

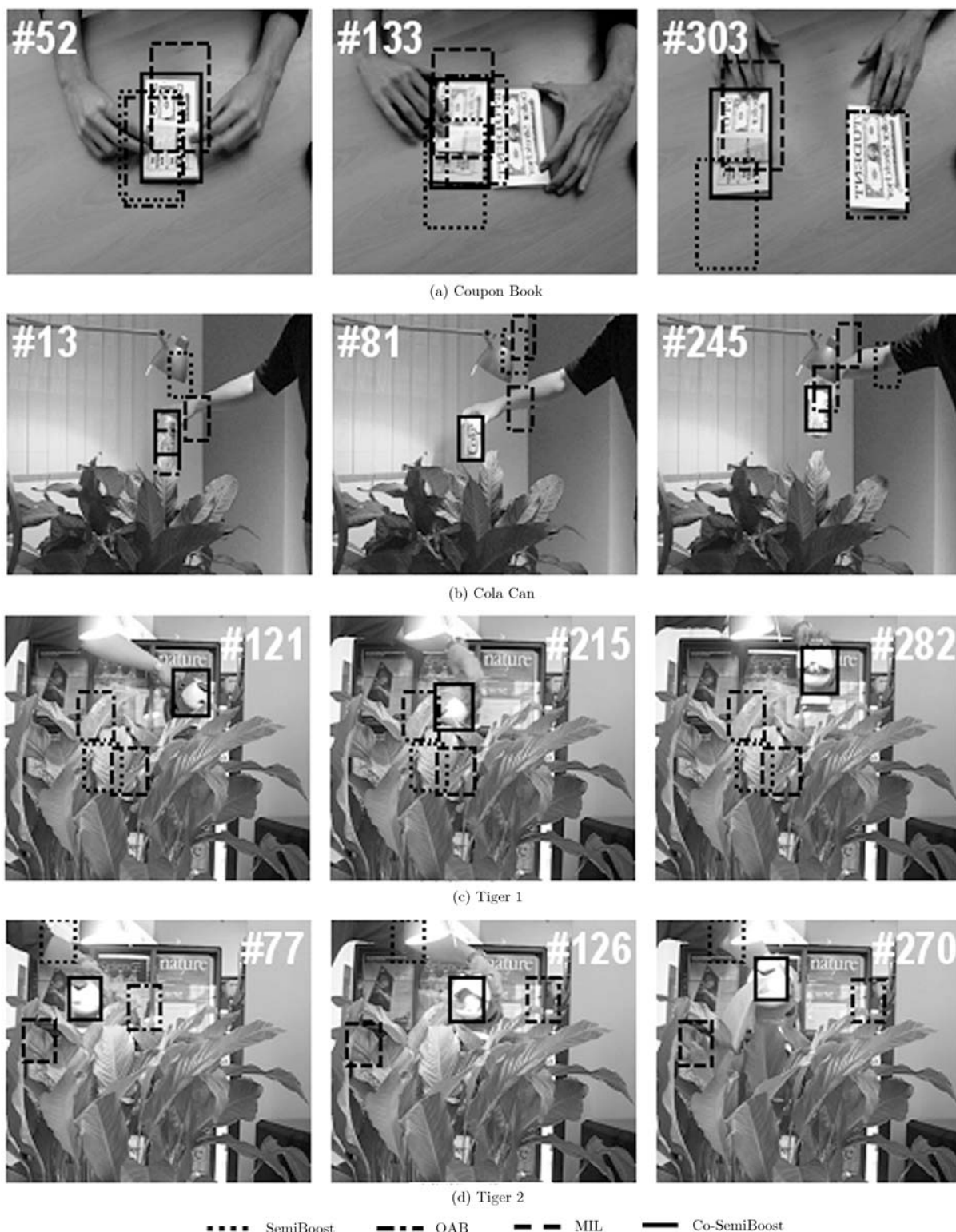


图4 4个视频中4种对比算法的目标跟踪结果示例图

被跟踪的票券被折起，从而目标外观发生变化。在第133帧和第303帧，被跟踪的票券不断受到旁边一张假票券的影响。OAB算法由于没有保存先验模型，发生了严重的目标漂移。SemiBoost算法由于迭代过程中分类错误不断加强，也慢慢地漂离跟踪目

标。而MIL算法由于在学习过程中没有区分正包中不同样本的重要性，因此MIL算法的跟踪性能不稳定。Co-SemiBoost算法由于在协同训练过程中不断提升分类器的判别能力，并且融入了目标先验知识，因此能够较好地地区分相似对象，更稳定地跟踪到目标。

图4(b)给出了Cola Can视频中第13, 第81和第245帧的目标跟踪结果。受到光照、快速移动和旋转的影响, SemiBoost和MIL算法在迭代过程中的分类错误不断累积, 分类器性能下降, 均发生目标漂移。OAB算法跟踪到的目标位置不精确, 导致获取到的正样本不准确。Co-SemiBoost算法能够更准确地跟踪到目标, 因为其利用未标记样本协同训练分类器, 克服了自训练的不足, 提高了分类器的性能。

如图4(c)所示, 在Tiger 1视频中, 玩具老虎在光照变化和复杂背景下被旋转和快速移动。例如第121帧目标被快速移动并且出现部分遮挡。在第215帧时, 目标被快速移动而使图片中目标区域变得模糊, 并且出现较严重的遮挡现象。第282帧老虎面部外观发生明显变化。在这些帧中, Co-SemiBoost算法均能够准确地跟踪到目标, 而其它算法都漂移到了周围的物品。

图4(d)给出了Tiger 2视频在光照条件和复杂背景下第77, 第126和第270帧的跟踪结果, 例如第77帧目标被快速移动出现模糊, 第126帧目标外观出现显著变化, 以及第270帧目标被旋转和部分遮挡。Co-SemiBoost算法能够稳定地跟踪到玩具老虎, 而其它算法都出现漂移问题。

综上所述, Co-SemiBoost算法采用的在线协同训练框架是有效可行的, 明显优于基于自训练的在线boosting和在线半监督boosting算法。并且与新近提出的在线多示例跟踪器相比, Co-SemiBoost算法总体上具有更优的跟踪性能。因此, 在复杂的跟踪环境下, Co-SemiBoost算法既具有自适应能力, 又在一定程度上有效地抑制了目标漂移问题。

4 结束语

本文提出了一种鲁棒的基于在线半监督boosting的协同训练目标跟踪算法(Co-SemiBoost)。该算法采用一种有效的在线协同训练框架, 利用未标记样本协同训练不同视图中的弱分类器, 克服了自训练目标跟踪算法的不足, 从而提高了分类器的判别能力。同时, Co-SemiBoost算法保持了在线半监督boosting算法的优势, 结合目标先验模型和在线分类器来预测未标记样本的伪类别标记和权重, 能够有效处理目标遮挡和光照变化等问题。实验结果表明, Co-SemiBoost算法具有鲁棒的跟踪性能。另外, 由于该算法主要解决单目标的在线跟踪问题, 没有考虑多目标之间的关联性以及多目标运动轨迹估计等问题, 因此不适用于多目标跟踪问题; 如何扩展本文算法用以解决多目标跟踪将是我们重点研究的方向。此外, 由于不同的特征视图可能适合于

不同的跟踪场景, 因而我们将进一步研究在各种场景下不同特征视图的适用性问题。

参考文献

- [1] Yang Han-xuan, Shao Ling, Zheng Feng, *et al.* Recent advances and trends in visual tracking: a review[J]. *Neurocomputing*, 2011, 74(18): 3823-3831.
- [2] Yilmaz A, Javed O, and Shah M. Object tracking: a survey[J]. *ACM Computing Surveys*, 2006, 38(4): 1-45.
- [3] 孙景乐, 唐林波, 赵保军, 等. 基于瑞利分布的粒子滤波跟踪算法[J]. 电子与信息学报, 2013, 35(4): 763-769.
Sun Jing-le, Tang Lin-bo, Zhao Bao-jun, *et al.* A new particle filter tracking algorithm based on Rayleigh distribution[J]. *Journal of Electronics & Information Technology*, 2013, 35(4): 763-769.
- [4] 董文会, 常发亮, 李天平. 融合颜色直方图及 SIFT 特征的自适应分块目标跟踪方法[J]. 电子与信息学报, 2013, 35(4): 770-776.
Dong Wen-hui, Chang Fa-liang, and Li Tian-ping. Adaptive fragments-based target tracking method fusing color histogram and SIFT features[J]. *Journal of Electronics & Information Technology*, 2013, 35(4): 770-776.
- [5] Jia Xu, Lu Hu-chuan, and Yang Ming-hsuan. Visual tracking via adaptive structural local sparse appearance model[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Rhode Island, 2012: 1822-1829.
- [6] Grabner H and Bischof H. On-line boosting and vision[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, 2006: 260-267.
- [7] Grabner H, Grabner M, and Bischof H. Real-time tracking via on-line boosting[C]. Proceedings of British Machine Vision Conference, Edinburgh, 2006: 47-56.
- [8] Grabner H, Leistner C, and Bischof H. Semi-supervised on-line boosting for robust tracking[C]. Proceedings of European Conference on Computer Vision, Marseille, 2008: 234-247.
- [9] Tang Feng, Brennan S, Zhao Qi, *et al.* Co-tracking using semi-supervised support vector machines[C]. Proceedings of the IEEE International Conference on Computer Vision, Brazil, 2007: 1-8.
- [10] Babenko B, Yang Ming-hsuan, 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.
- [11] Zhang Kai-hua, Zhang Lei, and Yang Ming-hsuan. Real-time compressive tracking[C]. Proceedings of European Conference on Computer Vision, Florence, 2012: 866-879.
- [12] Yu Qian, Dinh T B, and Medioni G. Online tracking and reacquisition using co-trained generative and discriminative

- trackers[C]. Proceedings of European Conference on Computer Vision, Marseille, 2008: 678–691.
- [13] Zhu Xiao-jin. Semi-supervised learning literature survey[R]. Computer Sciences TR-1530, USA: University of Wisconsin-Madison, 2007.
- [14] Mallapragada P K, Jin Rong, Jain A K, *et al.* SemiBoost: Boosting for semi-supervised learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(11): 2000–2014.
- [15] Lewandowski M and Orwell J. Co-training on multi-view unlabelled data[C]. Proceedings of the 27th Conference on Image and Vision Computing, New Zealand, 2012: 423–428.
- [16] Blum A and Mitchell T. Combining labeled and unlabeled data with co-training[C]. Proceedings of the Eleventh Annual Conference on Computational Learning Theory, Madison, 1998: 92–100.
- [17] Viola P and Jones M. Rapid object detection using a boosted cascade of simple features[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Kauai, 2001: 511–518.
- [18] Ahonen T, Hadid A, and Pietikainen M. Face description with local binary patterns: application to face recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(12): 2037–2041.
- [19] Oza N. Online ensemble learning[D]. [Ph.D. dissertation], University of California, 2001.
- [20] Babenko B, Yang Ming-hsuan, and Belongie S. Visual tracking with online multiple instance learning[OL]. http://vision.ucsd.edu/~bbabenco/project_miltrack.shtml. 2009.
- 陈思: 女, 1987年生, 博士生, 研究方向为计算机视觉、机器学习、数据挖掘.
- 苏松志: 男, 1982年生, 博士, 讲师, 研究方向为行人检测、行为识别.
- 李绍滋: 男, 1963年生, 教授, 博士生导师, 研究方向为人工智能及其应用、运动目标检测与识别、机器学习、计算机视觉.