

基于多示例学习和随机蕨丛检测的在线目标跟踪

罗艳^{①③} 项俊^② 严明君^{①③} 侯建华^{*①③}

^①(中南民族大学智能无线通信湖北省重点实验室 武汉 430074)

^②(华中科技大学自动化学院 武汉 430074)

^③(中南民族大学电子信息工程学院 武汉 430074)

摘要: 基于检测的目标跟踪方法目前在计算机视觉领域受到了广泛的关注,这类方法通过训练判别分类器将目标对象从背景中分离出来;分类器的训练是根据当前的跟踪状态从当前帧中提取正负样本来进行,但训练样本的不准确将导致分类器退化产生漂移。该文提出一种能够有效克服目标漂移的跟踪算法,采用检测器和跟踪器相结合的框架,利用中值流算法作为跟踪器,提高跟踪点的可靠性;级联若干个随机蕨弱分类器构成强分类器作为检测器;用在线多示例学习方法更新检测器,提高检测精度;最后将检测器、跟踪器的结果相融合得到最终的目标位置。实验结果表明,与其它方法相比,该方法对目标漂移有更强的鲁棒性。

关键词: 目标跟踪;中值流(MF);随机蕨丛;在线多示例学习(MIL)

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2014)07-1605-07

DOI: 10.3724/SP.J.1146.2013.01358

Online Target Tracking Based on Multiple Instance Learning and Random Ferns Detection

Luo Yan^{①③} Xiang Jun^② Yan Ming-jun^{①③} Hou Jian-hua^{①③}

^①(Hubei Key Laboratory of Intelligent Wireless Communications, South-Central University for Nationalities, Wuhan 430074, China)

^②(College of Automation, Huazhong University of Science and Technology, Wuhan 430074, China)

^③(College of Electronic Information Engineering, South-Central University for Nationalities, Wuhan 430074, China)

Abstract: Recently, a class of tracking techniques called “tracking by detection” receive much attention in computer vision. These methods train a discriminative classifier to separate the object from the background. The classifier bootstraps itself by using the current tracker state to extract positive and negative examples from the current frame. Slight inaccuracies in the tracker lead to incorrectly labeled training examples, which degrade the classifier and cause drift. In this paper, an effective algorithm is proposed to overcome the target drift. It takes the framework of tracking by detection. Median Flow (MF) is used as a tracker to improve the reliability of the tracking point; the detector is constituted with several weak classifiers of random ferns to cascade, and it is updated with online Multiple Instance Learning (MIL). Finally the detector and tracking results are integrated to get the target location. Experiments on a number of challenging video clips show that the proposed method outperforms some state-of-the-art tracking methods, especially for fast motion and drifts.

Key words: Object tracking; Median Flow (MF); Random ferns; Online Multiple Instance Learning (MIL)

1 引言

目标跟踪在近二十年中取得了长足的进展,但受自然场景中光线变化、目标形变、视角变化、遮挡等诸多因素的影响,目标跟踪仍是计算机视觉领

域最具挑战性的研究方向之一^[1]。根据目标外观处理方式的不同,跟踪方法可分为基于生成模型和基于判别模型两大类^[2]。

基于生成模型的算法:用一个模型来描述目标对象,在图像中搜索与之最匹配的区域作为目标位置。2003年Comaniciu等人^[3]提出mean shift跟踪算法,能处理杂乱背景和尺寸问题,但目标不能快速运动;文献[4]为了解决部分遮挡问题,提出了基于分块的跟踪算法,但缺乏模板更新机制,不能处

2013-09-05 收到, 2013-12-24 改回

国家自然科学基金(61141010), 武汉市科技供需对接计划项目(201051824575), 湖北省自然科学基金(2012FFA113)和中南民族大学中央高校基本科研业务费专项资金(CZY13033)资助课题

*通信作者: 侯建华 zil@scuec.edu.cn

理目标外观变化; Ross 等人^[5]构建一个自适应的外观模型,能较好地处理因光照和姿态变化引起的目标形变,但目标被严重遮挡时跟踪失败。一般该类方法要求训练得到一个完善的外观模型,这需要大量的样本,而在应用中,已标记的样本很少,只能在跟踪过程获得,跟踪中获得的样本通常伴随潜在的偏差,会导致跟踪性能下降。

基于判别模型的跟踪算法:将跟踪看作一个二分类问题,通过训练判别分类器将目标从背景中分离。该思想最早由 Collins 等人^[6]提出,将跟踪视为在前景和背景之间的局部二分类问题。该类算法往往与目标检测紧密相关,即判别分类器实际上是起检测器的作用,利用检测器辅助跟踪,因此大部分文献也将基于判别模型的跟踪算法称之为基于检测的跟踪(tracking by detection)^[7]。因该方法能解决目标在长时间遮挡条件下的跟踪问题,受到广泛关注,目前已成为跟踪领域的主流研究方向^[8,9]。

文献[10]采用集成跟踪方法,在参考图像中为每一个像素构建一个特征向量,训练一个自适应的集成分类器将属于目标的像素和属于背景的像素分离;但基于像素的表达使其在处理严重遮挡和杂乱背景时受到制约。在 Oza^[11]在线 Boosting 的基础上,Grabner 等人^[12]提出了在线 AdaBoost 算法并引入到跟踪,该算法有在线学习机制,但是在在线学习中如果出现错误的结果将会产生误差,这种误差随时间累积最终导致跟踪漂移。为了克服漂移问题,Grabner 等人^[13]又提出了一种半监督学习的改进算法,已标注的样本只来自第 1 帧,用后续帧中未标注的样本和已标注的样本来更新分类器;提高了适应性,但若目标快速变化,跟踪性能下降。Babenko 等人^[7]认为漂移是训练样本的标注歧义导致的,提出了在线多示例学习跟踪算法,在当前跟踪状态的周围采集一系列的正样本包用于更新分类器,可以避免训练样本的歧义性;对部分遮挡和克服漂移有很强的鲁棒性,但不能处理大量非刚性的形变。

本文在基于检测的目标跟踪框架下,引入多示例学习机制用以改善检测器精度、克服目标的漂移问题。主要思路如下:利用中值流跟踪器跟踪目标的大致位置;采用随机蕨丛检测器,即级联若干个随机蕨弱分类器构成强分类器;跟踪器和检测器的结果相互验证融合得到最终的目标位置;用在线多示例学习的方法更新检测器,当目标消失重现后由检测器的结果初始化跟踪器。本方法在一些具有挑战性的视频中试验,结果表明,本方法是一个鲁棒的实时目标跟踪算法,在目标快速剧烈运动或者消

失一段时间后重现等复杂情况下有效地克服漂移实现稳定的跟踪。

2 中值流跟踪器

中值流(Median Flow, MF)跟踪算法是由 Kalal 等人^[14]在 2010 年提出,本质是基于点的跟踪,其认为一个有效的跟踪与时间流向无关,即前向跟踪和后向跟踪的结果应该一致。设一个图像序列 $S = \{I_t, I_{t+1}, \dots, I_{t+k}\}$, I_t 为 t 时刻的图像, x_t 为 t 时刻特征点,以 x_t 为初始点,采用 Lucas-Kanade(L-K)光流法^[15]跟踪特征点,得到 t 到 $t+k$ 时刻前向跟踪的轨迹为 $T_f^k = \{x_t, x_{t+1}, \dots, x_{t+k}\}$,则以 x_{t+k} 为初始点,得到相应的 $t+k$ 到 t 时刻的后向跟踪轨迹 $T_b^k = \{x'_t, x'_{t+1}, \dots, x'_{t+k}\}$,其中 $x'_{t+k} = x_{t+k}$,如果前后两次跟踪轨迹相差很大,则说明跟踪失败,反之跟踪有效。两次跟踪误差用欧氏距离来计算,将前向-后向误差定义为

$$\text{FBE}(T_f^k | S) = d(T_f^k, T_b^k) \quad (1)$$

式中 $d(T_f^k, T_b^k) = \|x_t - x'_t\|$ 。

在跟踪时,第 1 帧手工确定要跟踪的目标,用矩形框表示,从框内选择若干个点用于跟踪。用 L-K 光流法跟踪这些点,产生一个稀疏运动流,并估计前向-后向误差 FBE。去除 50% 误差值高的跟踪点,余下 50% 的跟踪点被认为是好的跟踪点,用这些好的点来估计目标的位置,返回一个矩形框。这种跟踪方法就叫做中值流跟踪,相应的跟踪器为中值流跟踪器。中值流的跟踪示意图如图 1 所示。

3 随机蕨丛检测器

检测器的任务是在当前帧中给目标定位。常用的方法分为两类,基于局部图像特征^[16]和滑动窗口^[17]。本文是基于滑动窗口的方法,用各种不同尺寸的滑动窗口扫描输入图像得到图像块(patch),检测器的任务就是区分每一个 patch 否含有感兴趣的

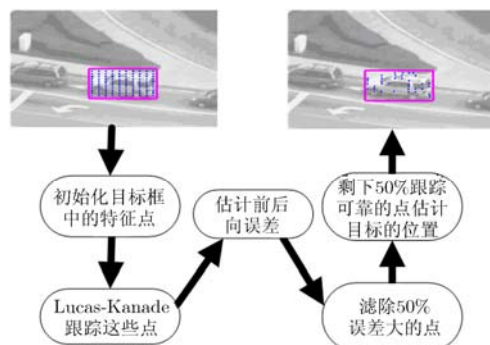


图1 中值流跟踪方法

目标，将含有感兴趣的目标 patch 进行聚簇得到目标的位置。

随机森林(RF)以其特有的优势在计算机视觉中得到了深入的应用。随机森林是 Breiman^[18]首次提出的，由一系列决策树组成，每棵决策树是一个分类器，这些决策树的综合投票得到分类结果；实验证明，RF 的分类准确率高，对异常值和噪声有很好的容忍度，并行快速，且不易出现过拟合。随机蕨丛(random ferns)是 Ozuysal 等人^[19]首先提出的，是 RF 的一种简化，用非层次结构蕨来替换决策树，且验证它有 RF 的特性，且比 RF 的分类效果更好，速度更快，更易于实现。文献[20]把随机蕨丛用于跟踪中，跟踪效果不错；但目标严重遮挡时，跟踪失败。本文采用随机蕨丛作为检测器。

随机蕨丛的结构如图 2 所示，每个蕨有 d 个不同的特征，每个特征输出 0 或者 1，则图像块经过随机蕨编码后的输出有 2^d 种可能。训练时，计算所有训练集的二值特征，并被分成各个蕨，统计每个蕨中这些图像块对应的特征编码 x 属于前景的概率 $P_i(y = 1 | x) = \frac{\#p}{\#p + \#n}$ ， $\#p$ 为特征编码为 x 的已标注为正样本的个数。 $\#n$ 为特征编码为 x 的已标注为负样本的个数。初始化时，概率为 0；随后， $\#p$ 和 $\#n$ 的变化导致编码的条件概率不断更新。当测试 patch 经过随机蕨丛时，每个蕨有一个特征编码，输出与之对应的一个其属于前景的概率，综合每一个蕨对该 patch 的概率进行投票决定该 patch 是否为目标。

假设目标的位置由位置 l^* 处的大小为 $w \times h$ 的图像块来表示，位置 l 处的图像块用特征集 $\mathbf{F} = \{f_1, f_2, \dots, f_n\}$ 来表达，标签集 $\mathbf{C} = \{c_1, c_2, \dots, c_i\}$ ， $c_i \in \{0, 1\}$ ，0 表示背景，1 表示前景。根据 Bayes 准则用后验概率进行分类可得

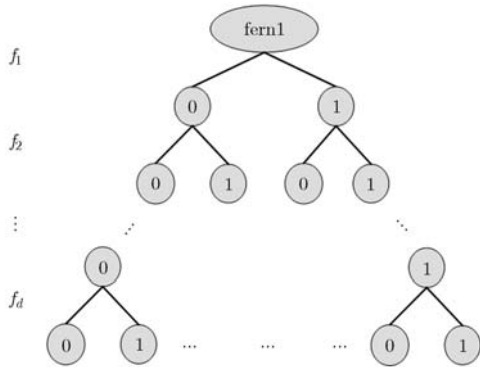


图2 随机蕨丛结构

$$\begin{aligned} c &= \arg \max_{c_i} P(c_i | f_1, f_2, \dots, f_n) \\ &= \arg \max_{c_i} \frac{P(f_1, f_2, \dots, f_n | c_i) P(c_i)}{P(f_1, f_2, \dots, f_n)} \\ &\propto \arg \max_{c_i} P(f_1, f_2, \dots, f_n | c_i) \end{aligned} \quad (2)$$

实际应用中，一般特征数目庞大，直接计算式(2)的联合概率不可行。一种可行的方法是假设各个特征相互独立，但是会忽略特征之间的相关性。折中的办法是把特征集 \mathbf{F} 分为 q 个子集 $\mathbf{F}_j = \{f_1, f_2, \dots, f_d\}$ ($d = n/q$)，每个子集的特征对应一个蕨。蕨内部保持数据的相关性，蕨与蕨之间是相互独立的。那么联合概率变为

$$P(f_1, f_2, \dots, f_n | c_i) = \prod_{j=1}^m P(\mathbf{F}_j | c_i) \quad (3)$$

本文特征的选择采用局部二值特征，类似于 haar-like 特征^[17]。二值函数如下：

$$f_i = \begin{cases} 1, & \Phi_{i,1} > \Phi_{i,2} \\ 0, & \text{其它} \end{cases} \quad (4)$$

其中 $\Phi_{i,1} = \sum_{k=1}^K \alpha_k \varphi_k$ ， φ_k 为一个 Haar 特征， α_k 为权重系数， α_k 和 φ_k 是随机生成的， k 取 [2,4] 内的整数。这些特征用积分图可进行快速计算^[17]。

4 检测器和跟踪器的融合

在视频序列的首帧，用户初始化目标位置，采集正负样本并训练得到随机蕨丛检测器。下一帧中利用该检测器对一定范围内的图像块检测分类，得到目标可能在下一帧的几个位置。首帧在目标框内初始化一些光流点，用中值流跟踪器跟踪这些点，在下一帧也获得目标的一个位置。融合检测器的几个位置和跟踪器的一个位置得到最终的目标位置输出，融合过程如图 3 所示。

5 在线多示例学习更新检测器

为了自适应目标的外观变化，检测器需要在线更新。更新方法如图 4 所示，图 4(a)用当前帧中跟踪位置处的 patch 为正样本，其周围采集部分 patch 为负样本，构成训练集训练分类器；图 4(b)在目标位置附近采集多个正样本，在其周围采集负样本，构成训练集训练；从图 4(a), 4(b)可以看出，训练的正负样本并不是精确的，用这些有误差的样本训练分类器将导致分类性能退化引起漂移。而在线多示例学习算法通过打包的方式，能从一系列样本中挑选最合适的样本用于训练，以避免漂移问题，如图 4(c)所示。

5.1 在线 MIL

传统的学习算法训练一个二值分类器来估计

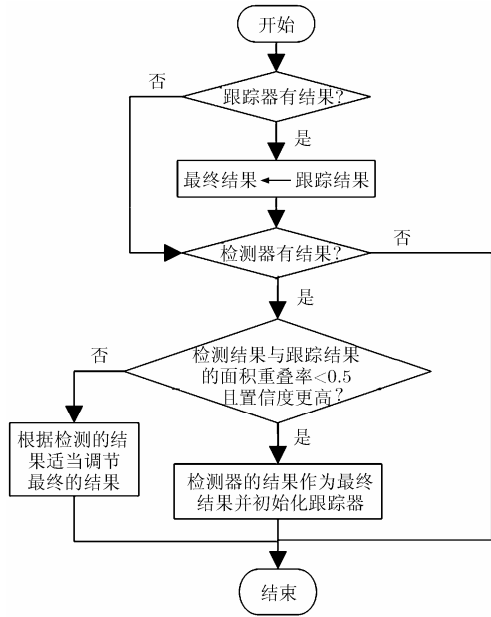


图 3 跟踪器和检测器融合的流程

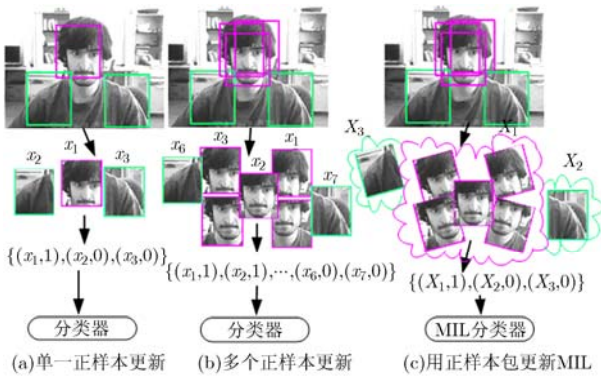


图 4 分类器的更新方法

$p(y|x)$, 要求训练数据的形式为 $\{(x_1, y_1), (x_2, y_2), \dots, (x_s, y_s)\}$, x_i 是一个示例(实验中是一个图像块的特征表达), $y_i \in \{0, 1\}$ 是一个示例标签。在多示例学习中, 假设有 N 个数据包 $\{X_1, X_2, \dots, X_N\}$, 每个包有 s 个示例 $X_i = \{x_{i1}, x_{i2}, \dots, x_{is}\}$ 。 $y_i \in \{0, 1\}$ 是包的标签, $y_{ij} \in \{0, 1\}$ 是示例 x_{ij} 对应的标签。包的标签定义为: $y_i = \max_j y_{ij}, j \in (1, s)$, 则只要包中有一个正例, 该包为正样本包。MIL 跟踪中寻找一个判别分类器 $H(x)$, 返回条件概率 $p(y = 1|x)$ 。 $H(x)$ 是示例分类器, 和示例的条件概率有关。用 NOR 模型(Noisy-Or Model)^[21]建模估算包的概率:

$$p(y_i = 1 | X_i) = 1 - \prod_j (1 - p(y_i = 1 | x_{ij})) \quad (5)$$

其中, 示例的概率表示为

$$p(y_i | x_{ij}) = \sigma(H(x_{ij})) \quad (6)$$

其中 $\sigma(x) = 1/(1 + e^{-x})$ 是 sigmoid 函数, 分类器 $H(x)$

由最大化包的对数似然函数的学习得到, 最大化包的对数似然函数为

$$L = \sum_i (1 - y_i) \lg(y_i \lg p(y_i = 1 | X_i)) + (1 - (y_i = 1 | X_i)) \quad (7)$$

为了适应目标的变化, 分类器 $H(x)$ 是在线自适应更新的。建立一个弱分类器池, 含有 M 个弱分类器, 从中选取 K 个最优的弱分类器构成强分类器 $H(x)$, 选取的原则是最大化包的对数似然函数:

$$h_k = \arg \max_{h \in \{h_1, h_2, \dots, h_M\}} L(H_{k-1} + h) \quad (8)$$

其中 $H_{k-1} = \sum_{j=1}^{k-1} h_j(x)$ 是由前 $k-1$ 个弱分类器级联成的强分类器。这种方法类似于文献[17]中特征的选择方法, 因为弱分类器和特征是一一对应的关系。文献[6]已经证明这种特征选择的方法能有效减少视觉漂移, 而且选取的分类器远少于候选的分类器, 大大提高了分类器的效率。

5.2 检测器的更新

利用融合的最终目标位置, 在目标位置周围采集样本获得训练数据集, 更新随机蕨丛检测器。把每一个 fern 看作一个弱分类器 h , 在线构建一个含有 M 个蕨的弱分类器池。训练数据更新所有的弱分类器 h , 从中挑选最优的 K 个构成随机蕨丛强分类器 ($K \ll M$), 作为检测器。具体的流程如表 1 所示。

表 1 检测器的更新

输入:	训练的数据集 $(X_1, y_1), (X_2, y_2), \dots, (X_N, y_N)$ 有 N 个包, 每个包有 s 个示例
输出:	强分类器 $H(x)$
步骤 1	利用输入数据更新分类器池中的所有随机蕨丛弱分类器 h (即更新特征编码属于前景的条件概率)
步骤 2	初始化 $H(x) = 0$
步骤 3	选取 K 个分类器构成强分类器
	for $k = 1 : K$
	for $m = 1 : M$
	for $i = 1 : N$
	for $j = 1 : s$
	示例的概率 $p_{ij}^m = \sigma(H + h_m(x_{ij}))$
	end
	由式(5)计算包的概率 $p_i^m = 1 - \prod_j (1 - p_{ij}^m)$
	end
	由式(7)计算包的对数似然函数 $L^m = \sum_i (\lg p_i^m)$
	end for
	$m^* = \arg_m \max L^m$
	$h_k(x) \leftarrow h_{m^*}(x)$
	$H(x) = H(x) + h_k(x)$
	end for

更新后的强分类器用于下一帧的检测分类，即每一帧用于分类的检测器均是由从特征池中挑选出来的最优弱分类器构成，增强分类效果。当目标严重遮挡或者消失重现后，跟踪器跟踪失败，此时，用检测器检测到的目标位置重新初始化跟踪器。

6 实验与分析

为了测试本文算法的有效性和稳定性，在常用的经典视频上测试，测试视频的特点如表 2 所示。

表 2 测试视频的特点

视频名称	帧数	主要特点
david	761	光照，视角，姿态，外观
car	945	Camera 运动，完全遮挡，相似目标
Faceocc2	812	遮挡，视角
pedestrian	338	遮挡，相似目标，尺寸，视角
tiger2	365	剧烈运动，旋转，遮挡
panda	980	尺寸，遮挡

对以上测试视频分别采用基于分块的跟踪 (Frag)^[4]、在线多示例学习的 MIL^[7]、基于稀疏表达的 SR^[22]、基于中值流的 MF^[14]和本文算法进行比较。对不同的测试视频所有算法使用相同的参数，其中每个随机蕨丛的特征数 $d = 13$ 是随机生成的，在线候选弱分类器池中 $M = 50$ ，每次选取 $K = 10$ 个弱分类器构成强分类器。

应用两种测量方式度量跟踪的性能。一种度量方式为跟踪的成功率， $sc = \frac{\text{area}(\text{ROI}_T \cap \text{ROI}_G)}{\text{area}(\text{ROI}_T \cup \text{ROI}_G)}$ ，

其中 ROI_T 是跟踪的边界框， ROI_G 是真实的目标框，当 $sc > 0.5$ 时认为跟踪成功。另一种是目标中心位置偏移，即跟踪目标框的中心和真实的目标中心的像素偏移。为了使结果一般化，每个算法对每组视频测试 10 次，将这 10 次的结果平均作为算法的结果。实验结果如表 3 和表 4 所示。

由表 3 可以看出，本文算法在 5 个视频序列中

表 3 不同算法的成功率比较(%)

视频序列	Frag	MIL	SR	MF	本文算法
David	1.0	36.0	20.8	100	100
car	4.0	51.5	61.6	55.8	98.9
Faceocc2	100	100	100	54.4	100
pedestrian	10.0	73.4	19.8	10.0	93.4
tiger2	21.2	54.9	84.6	10.9	76.4
panda	100	100	99.4	10.2	100

表 4 中心位置误差(像素)

视频序列	Frag	MIL	SR	MF	本文算法
David	176.7	41.1	62.6	2.5	<u>2.7</u>
car	184.8	112.3	93.9	<u>83.6</u>	16.4
Faceocc2	<u>15.1</u>	21.9	20.7	106.7	13.7
pedestrian	173.9	<u>78.3</u>	116.5	173.3	40.3
tiger2	67.4	64.8	55.5	77.4	<u>59.3</u>
panda	<u>6.6</u>	6.1	9.6	70.3	7.8
Mean	104.1	<u>54.1</u>	59.8	85.6	23.4

注：字体加粗下划线为最优结果，下划线表示次优结果

的成功率均是最高的。Frag 算法对光照敏感；MF 和 SR 在发生严重遮挡时，很容易跟丢；MIL 算法对全部遮挡、光照敏感，在 david 和 car 中识别率低。

从表 4 也可以看出，本文算法在 3 组视频上均是最优的，另两组为次优，从 6 组视频的均值来看，本文算法的效果优于其它算法。其它算法都有局限性。比如 MF 算法是基于点的跟踪，一旦目标被遮挡，则跟踪丢失。MIL 算法可以很好地应对遮挡问题，但是目标长时间消失重现以后就无法进行正确跟踪。SR 算法基本能正确跟踪，但跟踪的误差较大。综合来看，本文算法对于目标光照变化、视角变化、遮挡、相似目标、消失重现等均具有较好的跟踪效果。从图 5 误差曲线图也可以看出本文算法的中心误差较小(gt 为文献[7]中手工标定的目标真实位置)。

图 6 给出了部分视频跟踪结果，在 david 中，本文算法能很好地处理光照问题；在 car 中，本文算法主要应对相似目标的严重遮挡和消失重现；在 faceocc2 中，本文算法能处理遮挡和视角变化的问题；在 pedestrian 中，本文算法主要处理分辨率低、相似目标、目标较小、阴影等问题。对于以上大小为 320×240 的视频帧，在主频为 2.7 GHz 的普通 PC 机上本文算法的跟踪帧率为 15 f/s 以上，具有较好的实时性。

7 结论

本文采用检测和中值流跟踪相融合的框架，当跟踪丢失时，检测器能重新初始化跟踪器，能解决目标消失重现的问题。检测器通过在线学习机制不断得到提升，从而提升分类器自适应能力，即能够学习目标的诸多外观变化，比如尺寸、光照、视角、部分遮挡；然而在线学习存在的不足是容易出现因检测累计误差导致的漂移问题。多示例学习的本质就是挑选正确的样本用于训练分类器，提高了检测

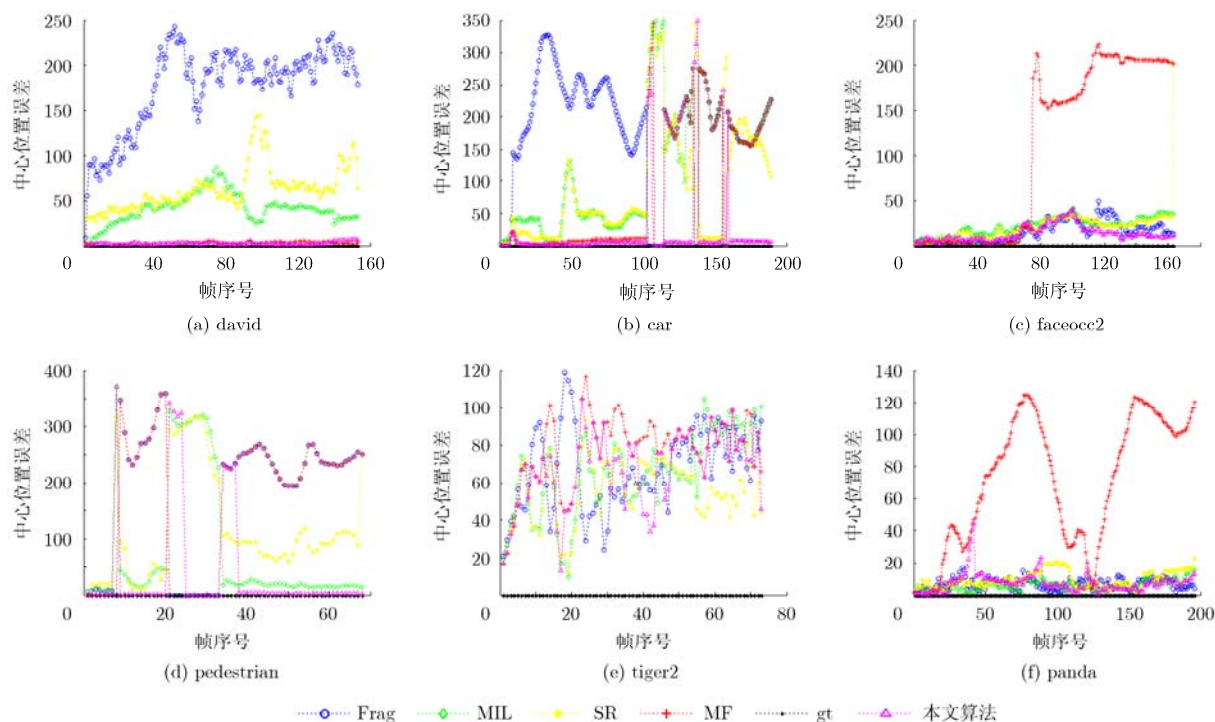


图5 中心位置误差曲线

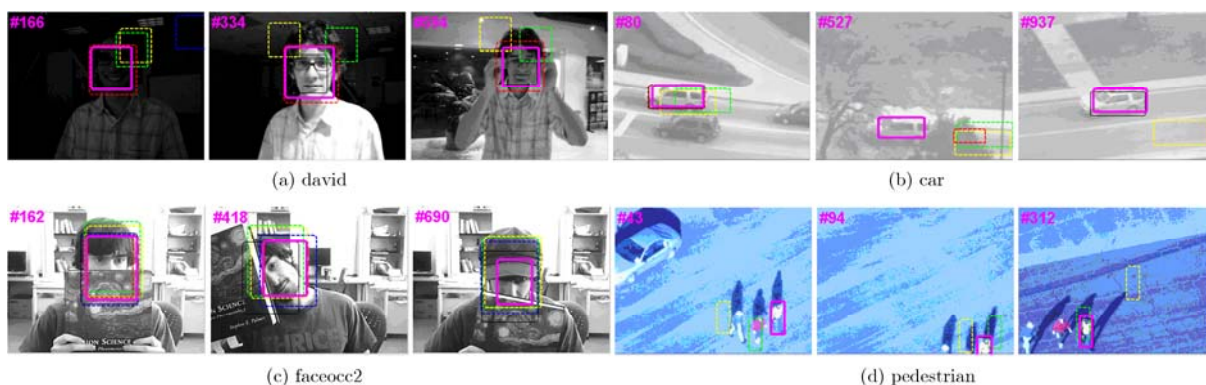


图6 不同跟踪算法的比较

器的精度，对相似目标有一定的识别作用，从而解决了跟踪中相似目标的漂移问题。基于以上理论分析，本文算法能处理跟踪中的多个问题。在测试视频上的实验结果也表明，本文提出的基于多示例学习和随机森林检测的在线目标跟踪算法能够很好地克服目标的漂移，达到实时有效的跟踪效果。

参考文献

- [1] Salti S, Cavallaro A, and Di Stefano L. Adaptive appearance modeling for video tracking: survey and evaluation[J]. *IEEE Transactions on Image Processing*, 2012, 21(10): 4334-4348.
- [2] Zhang Tian-zhu, Ghanem B, Liu Si, et al. Robust visual tracking via structured multi-task sparse learning[J]. *International Journal of Computer Vision*, 2013, 101(2): 367-383.
- [3] Comaniciu D, Ramesh V, and Meer P. Kernel-based object tracking[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(5): 564-575.
- [4] Adam A, Rivlin E, and Shimshoni I. Robust fragments-based tracking using the integral histogram[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, New York, 2006: 798-805.
- [5] Ross D, Lim J, Lin R S, et al. Incremental learning for robust visual tracking[J]. *International Journal of Computer Vision*, 2008, 77(1): 125-141.
- [6] Collins R T and Liu Y. Online selection of discriminative tracking features[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(10): 1631-1643.
- [7] Babenko B, Yang M H, and Sivic B. Robust object

- tracking with online multiple instance learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(8): 1619–1632.
- [8] Izadinia H, Ramakrishna V, Kitani K M, *et al.* Multi-pose multi-target tracking for activity understanding[C]. *IEEE Workshop on the Applications of Computer Vision (WACV)*, Florida, 2013: 385–390.
- [9] Sznitman R, Richa R, Taylor R, *et al.* Unified detection and tracking of instruments during retinal microsurgery[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(5): 1263–1273.
- [10] Avidan S. Ensemble tracking[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(2): 261–271.
- [11] Oza N C. Online ensemble learning[D]. [Ph.D. dissertation], University of California, Berkeley, 2002.
- [12] Grabner H, Grabner M, and Bischof H. Real-time tracking via on-line boosting[C]. *Proceedings British Machine Vision Conference (BMVC)*, Edinburgh, 2006: 47–56.
- [13] Grabner H, Leistner C, and Bischof H. Semi-supervised on-line boosting for robust tracking[C]. *Proceedings European Conference on Computer Vision (ECCV)*, Marseille, 2008: 234–247.
- [14] Kalal Z, Mikolajczyk K, and Matas J. Forward-backward error: automatic detection of tracking failures[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, Istanbul, 2010: 23–26.
- [15] Lucey S, Navarathna R, Ashraf A, *et al.* Fourier Lucas-Kanade algorithm[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(6): 1383–1396.
- [16] Miksik O and Mikolajczyk K. Evaluation of local detectors and descriptors for fast feature matching[C]. *21st International Conference on Pattern Recognition*, Tsukuba, 2012: 2681–2684.
- [17] Viola P and Jones M. Rapid object detection using a boosted cascade of simple features[C]. *IEEE conference on Computer Vision and Pattern Recognition*, Kauai, 2001: 511–518.
- [18] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5–32.
- [19] Ozuysal M, Calonder M, Lepetit V, *et al.* Fast keypoint recognition using random ferns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(3): 448–461.
- [20] Cong Rao, Cong Yao, Xiang Bai, *et al.* Online random ferns for robust visual tracking[C]. *21st International Conference on Pattern Recognition*, Tsukuba, 2012: 1447–1450.
- [21] Chen Si and Li Shaozi. Robust visual tracking with a novel online semi-supervised multiple instance boosting algorithm [C]. *2013 Fourth International Conference on Intelligent Control and Information Processing (ICICIP)*, Beijing, 2013: 426–431.
- [22] Zhang K, Zhang L, and Yang M H. Real-time compressive tracking[C]. *Computer Vision–ECCV*, Florence, 2012: 866–879.
- 罗 艳: 女, 1986 年生, 硕士生, 研究方向为图像处理与模式识别.
- 项 俊: 女, 1984 年生, 博士生, 研究方向为计算机视觉.
- 严明君: 女, 1987 年生, 硕士生, 研究方向为图像处理与模式识别.
- 侯建华: 男, 1964 年生, 博士, 教授, 硕士生导师, 研究方向为图像处理、机器视觉、视频监控.