

## 基于特征自适应选择的鲁棒跟踪算法

吴成东, 齐苑辰, 陈东岳

(东北大学 信息科学与工程学院, 沈阳 110004)

**摘要:** 为了解决真实场景下视频目标的跟踪问题, 提出一种基于特征自适应选择的鲁棒跟踪算法. 首先, 针对在线AdaBoost算法特征池特征鲁棒性差的问题, 提出一种基于颜色与金字塔梯度方向直方图特征相结合的特征池构造方式; 然后, 针对分类器在更新过程中容易受到错误样本影响的问题, 对每帧跟踪结果增加遮挡检测环节以避免漂移现象的发生. 大量的对比实验表明, 在真实场景下所提出的方法具有更好的效果.

**关键词:** 目标跟踪; 在线学习; 自适应更新; 遮挡检测

**中图分类号:** TP391

**文献标志码:** A

## Robust object tracking with adaptive feature selection

WU Cheng-dong, QI Yuan-chen, CHEN Dong-yue

(College of Information Science and Engineering, Northeastern University, Shenyang 110004, China. Correspondent: QI Yuan-chen, E-mail: qiyuanchen649@163.com)

**Abstract:** In order to solve the tracking problem of video sequences in real-world scenarios, a robust tracking algorithm based on adaptive feature selection is proposed. Firstly, for the problem that the candidate features of the online AdaBoost algorithm are not robust, a construction mode of the candidate feature pool is proposed, which combines color and pyramid gradient orientation histogram features. Then, for the problem that classifiers are vulnerable to the influence of improper samples during the update, a process of occlusion detection is added at each frame after obtaining the tracking result to avoid the phenomena of drift. Lots of comparison experiments show that the proposed algorithm tracks the object accurately and reliably in realistic videos.

**Key words:** object tracking; online learning; adaptive updating; occlusion detection

### 0 引言

目标跟踪是计算机视觉领域的一个重要研究方向, 在智能场景监控、人机交互、机器人导航等领域都有着广泛的应用. 尽管目前已取得了一些进展, 但是对于真实场景下的视频目标跟踪问题仍然具有挑战性, 这主要是因为真实场景中目标会发生更多不受控的表观和运动变化.

目前, 主流的基于视频信息的目标跟踪算法包括3个部分: 观测模型、运动模型和搜索策略. 其中: 观测模型用来描述候选目标与参考模板之间的相似性, 欧氏距离和巴氏距离<sup>[1]</sup>等都可以作为度量标准; 运动模型用于利用新的一帧图像预测目标可能的状态, 可以通过卡尔曼滤波<sup>[2]</sup>或粒子滤波<sup>[3-4]</sup>等实现; 搜索策略用来在当前时刻寻找目标最可能的状态, Mean Shift<sup>[1]</sup>或者 Sliding Window<sup>[5]</sup>等算法都可以达到此目

的. 因为本文更关注于对目标观测模型的构建与更新, 所以更着重笔墨于观测模型.

观测模型的构建分为产生式和判别式两类. 产生式类算法通过学习目标的表观模型来寻找与之最接近的候选目标, 从而确定目标的状态. Mei等<sup>[6]</sup>把压缩感知理论应用于粒子滤波跟踪算法, 将跟踪问题看作是稀疏近似问题, 通过对目标模板和测试模板的线性组合来估计目标的状态, 从而达到跟踪的目的. 但是, 这类算法由于忽略了目标周围的背景信息, 无法准确地将目标从背景中分离出来.

判别式类算法将跟踪问题当作分类问题来处理, 目的是设计出能够有效将目标和背景区分开的最优分类器, 从而通过在指定邻域内进行检测与识别以实现目标的跟踪. Collins等<sup>[7]</sup>通过在线选择对当前场景区分性较高的特征来确定目标的状态; Grabner

收稿日期: 2013-11-05; 修回日期: 2014-05-09.

基金项目: 国家自然科学基金项目(61273078, 61005032); 中央高校基本科研业务费专项基金项目(N1106040065032).

作者简介: 吴成东(1960—), 男, 教授, 博士生导师, 从事图像智能处理、机器人控制等研究; 齐苑辰(1983—), 女, 博士生, 从事图像处理、目标检测、目标跟踪的研究.

等<sup>[5]</sup>设计了一种可以在线更新的 Boosting 分类器,并将其应用于跟踪领域.但上述方法在分类器更新过程中使用的样本都是由每个时刻获得的跟踪结果确定的,因此一旦结果不准确,将直接影响到分类器的性能,从而降低算法的准确性. Boris 等<sup>[8]</sup>虽然针对这一问题设计了基于“样本集打包”的分类器学习方法,但由于其将辨别样本歧义性的任务完全丢给多实例学习算法本身,一方面复杂度有所增加,另一方面也增加了错误跟踪的可能性.

基于以上分析,本文在贝叶斯框架下提出一种鲁棒的基于特征自适应选择的跟踪算法,并通过实验验证了本文算法在真实场景下实现目标跟踪的有效性.

## 1 贝叶斯跟踪方法

贝叶斯方法将跟踪问题转化为非线性非高斯的滤波问题,通过一组在状态空间传播的随机样本对概率密度函数进行近似表达,有

$$p(X_t|Y_{1:t}) \propto p(Y_t|X_t) \int p(X_t|X_{t-1})p(X_{t-1}|Y_{1:t-1})dX_{t-1}. \quad (1)$$

其中:  $X_t = (X_t^c, X_t^s)$  为目标在  $t$  时刻的状态,  $X_t^c$  为目标中心的坐标,  $X_t^s$  用来描述目标的矩形框的大小;  $Y_{1:t}$  为累加到  $t$  时刻的观测序列.

为了获得对目标的最优估计,在当前时刻生成的随机样本中搜索可以使  $p(X_t|Y_{1:t})$  达到最大的样本,并将其作为最终的结果,如下式所示:

$$\hat{X}_t = \arg \max_{X_t^{(l)}} p(X_t^{(l)}|Y_{1:t}), \quad \forall l = 1, 2, \dots, N, \quad (2)$$

其中  $X_t^{(l)}$  为第  $l$  个候选样本的状态.

## 2 基于特征自适应选择的贝叶斯跟踪算法

### 2.1 特征池的构造

在复杂动态场景下,由于受到相机运动以及背景干扰等因素的影响,目标更易发生剧烈的变化,单独使用一种特征很难适应这种变化.如何构造候选特征集从而使其对目标的描述更加准确与全面,以及如何随着目标的运动和环境的自适应选择更具有鉴别性的特征描述目标是本文必须解决的两个关键问题.前者可以通过构造鲁棒的特征池来解决,后者可以通过在线 AdaBoost (OAB) 算法得以解决.

#### 1) 颜色特征.

颜色直方图特征具有计算简单、处理速度快的特点,而且对旋转和缩放等都不敏感,其表达式为

$$\hat{q}_u = C_q \sum_{i=1}^n K(\|X_0 - X_i^*\|^2) \delta[b(X_i^*) - u]. \quad (3)$$

其中:  $X_0$  为目标中心;  $\{X_i^*\}_{i=1, \dots, n}$  为中心在  $X_0$  的目标区域内每个像素点的坐标;  $b(X_i^*)$  为当前像素点  $X_i^*$  的特征值属于该特征的第几个统计区间;  $\delta$  为克罗内

克函数;  $K$  为核函数的剖面函数,对距中心近的像素赋予一个较大的权值,而距中心远的像素赋予较小的权值;  $C_q$  是一个归一化系数,确保单个直方图内的分量之和为 1.

颜色特征是目标跟踪算法中最常用到的特征之一,但由于其隐藏了图像的空间信息且容易受到光照等外界环境因素的影响,仅仅使用颜色特征无法确保基于特征的分类器的有效性.

#### 2) 金字塔梯度方向直方图特征.

Ann 等<sup>[9]</sup>提出了一种金字塔梯度方向直方图 (PHOG) 的方法.该方法首先对在图像上提取的边缘轮廓图分层,每一层都在上一层的基础上分成更小的块;然后分别对各层上的各小块提取梯度方向直方图;最后将这些直方图合并,便构成了金字塔梯度方向直方图.与梯度方向直方图 (HOG)<sup>[10]</sup>相比,该算法具有计算简单、尺度信息丰富的优点.文献 [14] 共选择 3 层组合构成 PHOG,每层被分割成  $2^{2 \times l}$  个小块,其中  $l = 0, 1, 2$ .

通过上述分析可知,层数越高,被分割出的小块的数目就越多,所提取的直方图对图像边缘的刻画也越清晰.鉴于此,本文提出一种新颖的特征池构造方式,在 6 种颜色特征的基础上,按照  $l = 0, l = 1, l = 2, l = 3$  的分层顺序加入 4 种梯度特征,从而可以更加细致地描述目标的形状,即便在颜色信息失效时依然能够精准地将目标从背景中区分出来.由于包含了颜色和梯度两类特征,本文构造的特征池的鲁棒性更好.此外,梯度特征使用了多尺度金字塔方式进行提取,能够兼顾不同尺度的特征,增强算法对具有尺度变化的目标的跟踪稳定性.

### 2.2 运动模型

运动模型用于描述目标在相邻时刻的状态的变化规律,在对目标运动规律没有先验知识的情况下,可假设其服从正态分布,即

$$p(X_t|X_{t-1}) = G(X_t - X_{t-1}, \Psi). \quad (4)$$

其中:  $\Psi$  为对角协方差矩阵,  $X_{t-1}$  为  $t-1$  时刻的目标状态.

### 2.3 观测模型

观测模型用来度量候选样本与参考目标之间的相似性,与传统的基于目标模型构建观测模型的方法不同,本文充分利用了背景信息.将 Boosting 分类器的结果通过双曲函数  $\sigma$  转化为观测模型,如下式所示:

$$p(Y_t|X_t) \propto \sigma(H(X)) = \frac{1}{1 + \exp^{-H(X)}}, \quad (5)$$

$$H(X) = \sum_{n=1}^N \alpha_n h_n(X), \quad (6)$$

其中  $H(X)$  表示由  $N$  个弱分类器  $h_n(X)$  线性组合而成的强分类器, 每个弱分类器对应一个特征. 分类器的性能直接取决于特征的鉴别能力, 如果候选特征对目标的鉴别能力越强, 则分类器的分类性能越好. 因此, 构建一个鲁棒的候选特征池显得尤为重要. 在本文算法中要求每个弱分类器返回一个表征相似性的实数而非某个类别标签, 即

$$h_n(X) = DD(f_n(X), E(f_n^+)) - DD(f_n(X), E(f_n^-)). \quad (7)$$

其中:  $DD$  为两个特征直方图间的扩散距离<sup>[11]</sup>,  $E(f_n^+)$ 、 $E(f_n^-)$  分别为正样本和负样本取第  $n$  个特征时的直方图均值,  $f_n(X)$  表示对  $X$  所对应的图像区域提取的第  $n$  个特征.

上述分类器可按如下过程训练获得: 当有新样本送入时, 首先并行地对特征池里的所有分类器进行更新, 即在线调整式 (7) 中的参数  $E(f_n^+)$  或  $E(f_n^-)$ ; 然后依次从特征池中选出误差  $e$  最小的几个特征, 令  $e_n$  表示第  $n$  个特征的错误率, 其值越小, 表示其对应的特征表征目标的能力越好, 如下式所示:

$$\hat{n} = \arg \min_{n=1,2,\dots,N} e_n, \quad (8)$$

$$e_n = \frac{\lambda_n^{\text{wrong}}}{\lambda_n^{\text{wrong}} + \lambda_n^{\text{correct}}}. \quad (9)$$

其中:  $\hat{n}$  为选出的第  $n$  个弱分类器;  $\lambda_n^{\text{wrong}}$  为在第  $n$  个特征上类别判断错误的样本数量;  $\lambda_n^{\text{correct}}$  为在第  $n$  个特征上类别判断正确的样本数量. 具体而言, 在分类器训练过程中如果当前样本是正样本, 则首先利用候选特征池里的特征分别对这个已知类别标签的样本进行判断: 如果正确, 则该特征对应的  $\lambda_n^{\text{correct}}$  相应地增加; 反之则  $\lambda_n^{\text{wrong}}$  增加. 随着样本依次进来, 每个特征所对应的这两个参数的值也在不断发生变化. 显然, 如果在第  $n$  个弱分类器下求得的  $\lambda_n^{\text{wrong}}$  越小, 则表示该分类器性能越好, 式 (9) 的值就会越小. 所以, 使式 (9) 的取值达到最小的弱分类器将被依次选出以构造强分类器.

## 2.4 候选样本的生成

本文采用 Metropolis Hastings (M-H) 抽样算法<sup>[12]</sup> 从当前帧图像中获取候选目标样本. 首先以上一时刻目标状态  $X_{t-1}$  为均值, 利用式 (4) 随机产生一个新的状态  $X'_t$ , 将其作为 MH 抽样算法的初始状态.

以新生成的随机状态  $X'_t$  为均值, 同样利用式 (4) 随机生成另一个新状态  $X_t^*$ . 该新状态  $X_t^*$  是否能够被接收, 可通过接受概率  $\eta$  进行判断,  $\eta$  可由下式计算得到:

$$\eta = \min \left[ 1, \frac{p(Y_t|X_t^*)Q(X'_t|X_t^*)}{p(Y_t|X'_t)Q(X_t^*|X'_t)} \right]. \quad (10)$$

生成服从标准分布的随机数  $u$ , 如果  $u \leq \eta$ , 则令  $X'_t = X_t^*$ , 否则,  $X'_t$  保持不变. 重复上面的步骤直到候选样本数量达到上限, 算法结束.

## 2.5 遮挡检测以及样本集在线更新策略

每帧图像跟踪完成后, 都要在线调整样本集, 当样本集足够丰富时再对分类器进行更新. 为避免漂移现象, 本文采用一种半遗忘式的样本集更新策略. 半遗忘式是指在更新过程中使用的样本包括两部分: 一部分来自视频的第 1 帧, 即“不遗忘的部分”; 另一部分来自视频的最近 4 帧, 即“遗忘的部分”. 这样做的目的是在获取目标最新信息的同时保留目标的原始信息, 在确保跟踪结果准确的前提下能够在很大程度上缓解漂移现象. 在本文实验中分类器每 2 帧更新一次, 每帧选取 5 个样本, 1 个正样本, 4 个负样本. 正样本为上一帧目标状态所确定的跟踪框内的图像区域; 负样本分别为以正样本跟踪框的 4 个顶点为中心且大小与正样本相同的矩形图像区域. 但是, 在实际跟踪过程中由于受到遮挡等因素的影响, 无法保证每次用来更新分类器的样本都是准确的, 因此为了避免这种干扰, 很有必要对每帧图像获得的跟踪结果增加遮挡检测环节, 如下式所示:

$$\bar{p}(y = 1|X_t^+) - p(y = 1|X_t^*) > \theta. \quad (11)$$

其中:  $p(y = 1|X_t^*)$  表示当前帧的跟踪结果是目标的概率,  $p(y = 1|X_t^+)$  表示样本集中的正样本是目标的平均概率, 均可通过式 (5) 得到; 阈值  $\theta$  通常取 0.2 ~ 0.3 之间. 如果存在遮挡现象, 则当前时刻的状态将由过去一帧的结果所取代, 并且停止更新样本集.

## 3 实验结果及分析

为了验证本文算法的有效性, 分别在“singer”、“basketball”、“bolt”和“OneLeaveShopReenter2cor”4 段视频上进行测试, 这些视频分别来自于音乐会、篮球比赛、田径比赛和走廊监控. 然后分别与 OAB<sup>[5]</sup>、MIL<sup>[8]</sup>、L1<sup>[6]</sup> 3 种算法进行比较.

### 3.1 定性分析

实验在 Matlab R2011b 环境下完成, 目标初始位置由手动给定, 特征池候选特征数目设置为 10, 每帧 4 帧分类器自动更新一次, 每次选择 4 个特征来构建强分类器, 单帧内目标候选样本个数设置为 20.

**实验 1** 在视频“singer”中, 随着视角从近景到远景的切换, 相机获取的图像序列中目标发生了尺度较大的缩放现象, 且背景受到周围剧烈的灯光变化影响. 图 1 分别给出了本文算法与其他 3 种算法的跟踪结果. 可以看出: OAB 算法和 MIL 算法由于没考虑到目标的尺度变化问题, 当相机镜头缩放倍数发生变化时, 跟踪框无法随着目标大小的变化而变化; 而 L1 算

法对光照变化敏感,在第228帧由于受到背景的光照由明到暗的变化影响,误将背景判断为目标.通过比较这4种算法的结果可以发现,相比于其他算法,本文算法由于采用了能够提供多尺度信息的PHOG特征和M-H抽样算法,可以更好地处理目标的尺度缩放和光照变化情况.

**实验2** 图2给出了4种算法在视频“basketball”上的跟踪结果.在该视频中,目标的姿态变化程度较大,且存在遮挡和背景干扰现象,这些都会引起跟踪算法的失效.在第10~第23帧时,目标受到周围队友的遮挡影响,致使L1算法失效,从第19帧开始跟踪框完全漂移到背景上;而对于MIL算法,当目标在第270~第286帧受到周围跑动的同队队友影响时,从第286帧开始跟踪框漂移到同队队友身上;OAB算法也由于受到对方白色球衣队员的干扰,在第478~第493帧时跟踪框开始漂移,第540帧又漂移到非目标的其他绿色球衣队员身上.与其他算法相比,本文算法构造了性能更佳的候选特征池,从而获得了更强的鉴别能力,当目标周围存在背景干扰时仍然能够精准地跟踪目标.

**实验3** 图3给出了4种算法在“bolt”视频上的跟踪结果.该视频中目标受周围背景的干扰严重,由于受到目标身后的跑道标志的影响,从第11帧开始

MIL算法发生了漂移;同样受此影响从第18帧开始,L1算法的结果漂移到4号跑道运动员身上,第55帧开始漂移到绿色草坪上,直到视频结束;OAB算法由于受到周围其他运动员以及背景广告栏的干扰,第133帧发生了漂移现象,跟踪框挪到了周围运动员身上,且此后都没有再跟踪到目标,直到第248帧目标完全脱离出跟踪视野,算法停止工作.与上述3种算法相比,本文算法在候选特征池中采用了4种PHOG特征,使得算法能够更为细致地捕获目标与干扰对象之间的差异,从而取得了更精准的结果.

**实验4** 图4给出了4种算法在“OneLeaveShop-Reenter2cor”视频上获得的结果.在该视频中目标受到其他物体较长时间的干扰和遮挡,其中待跟踪的目标为拎包的女士,作为干扰物的那名男士从第174帧开始出现在画面中,直到257帧以与女士相反的方向行走,此后转身与女士同方向行走.在这个过程中L1算法从第205帧开始失效,跟踪窗口落到男士身上且直到视频结束没有发生变化;而MIL和OAB算法分别从第208和第206帧开始跟踪窗口发生了漂移,且从第369和第372帧开始因为男士重新走向女士身边使得窗口又漂移到女士身上;相比于其他3种算法,本文算法由于增加了遮挡检测环节,在样本集中减少了错误样本出现的概率,从而提高了算法精度.

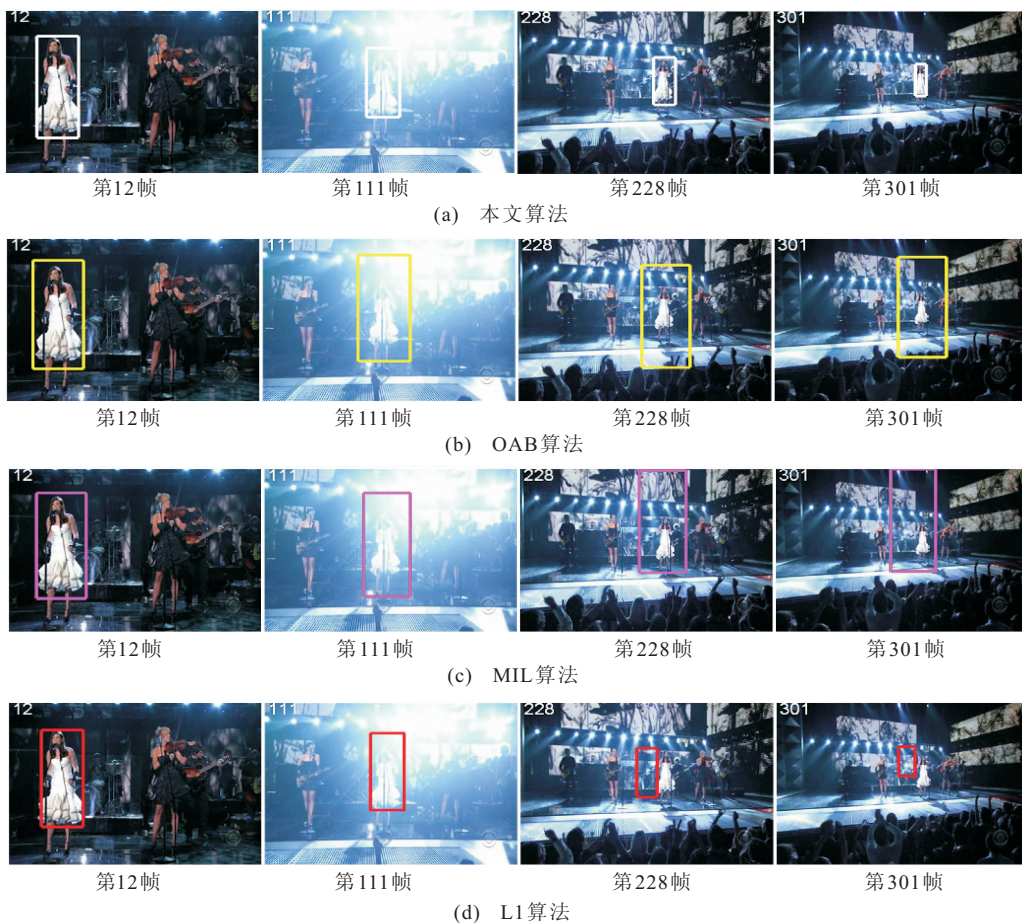


图1 “singer”视频跟踪结果(数据来自文献[12])

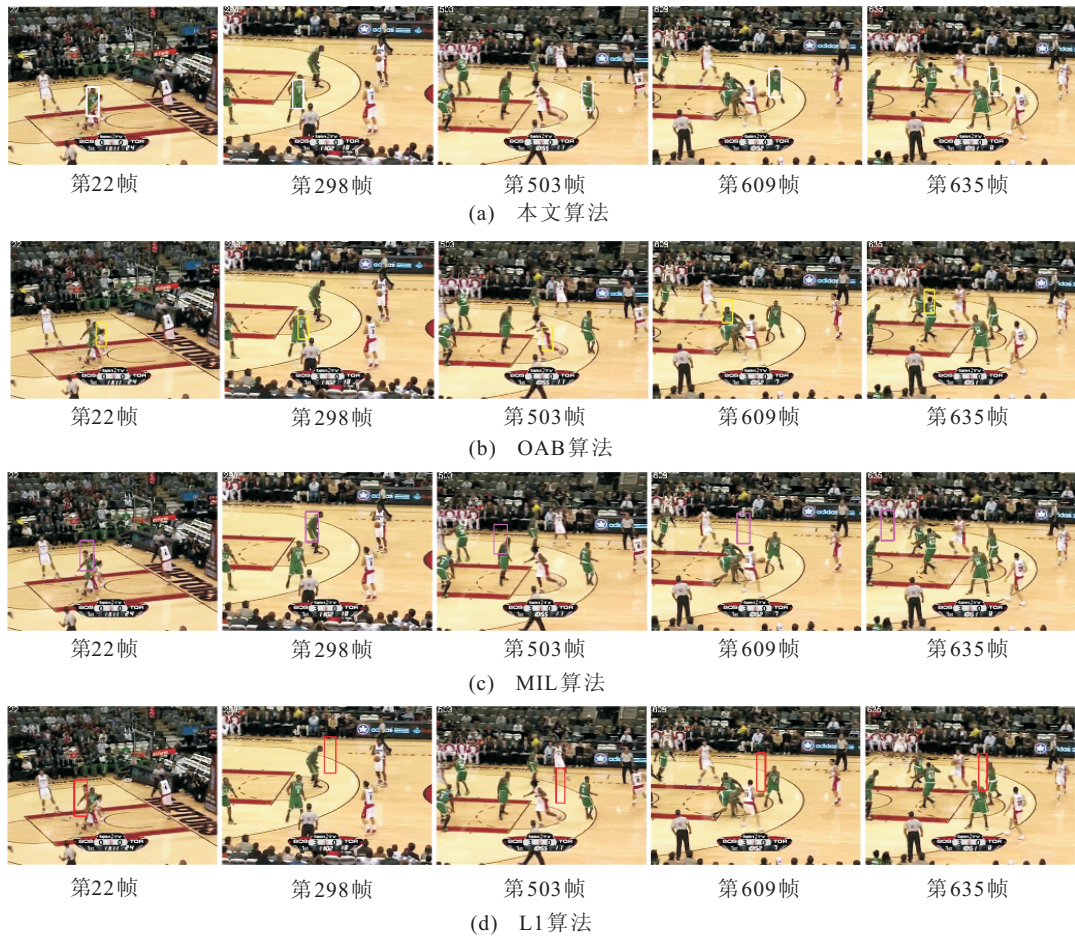


图 2 “basketball” 视频跟踪结果 (数据来自文献 [12])

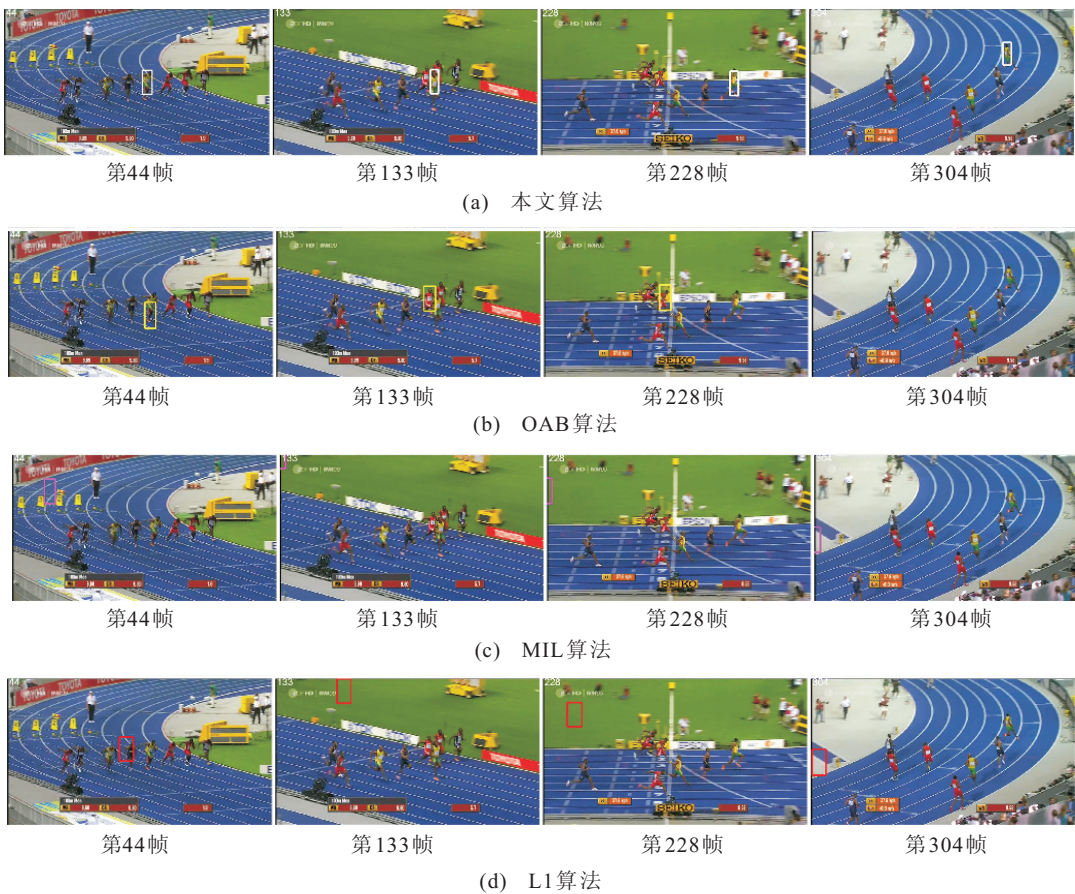


图 3 “bolt” 视频跟踪结果 (数据来自文献 [13])



图 4 “OneLeaveShopRecenter2cor” 视频跟踪结果 (数据来自 PETS2001 数据库)

3.2 定量分析

4 种算法在 4 段视频上的定量对比分析如图 5 所示, 横轴表示图像帧数, 纵轴表示算法求出的目标位置相比于真实位置的距离, 以像素为单位. 表 1 分别给出了 4 种算法的平均定位误差 (取整后的), 同样以

像素为单位, 每个数字代表当前算法在对应视频下定位的平均误差. 为了保证实验数据的可靠性, 每个算法都在 4 段视频下运行 5 次, 并将求出的平均值作为最后的定位误差. 通过对图 5 和表 1 的分析可以看出, 相比于其他算法, 本文算法取得了较为精确的结果.

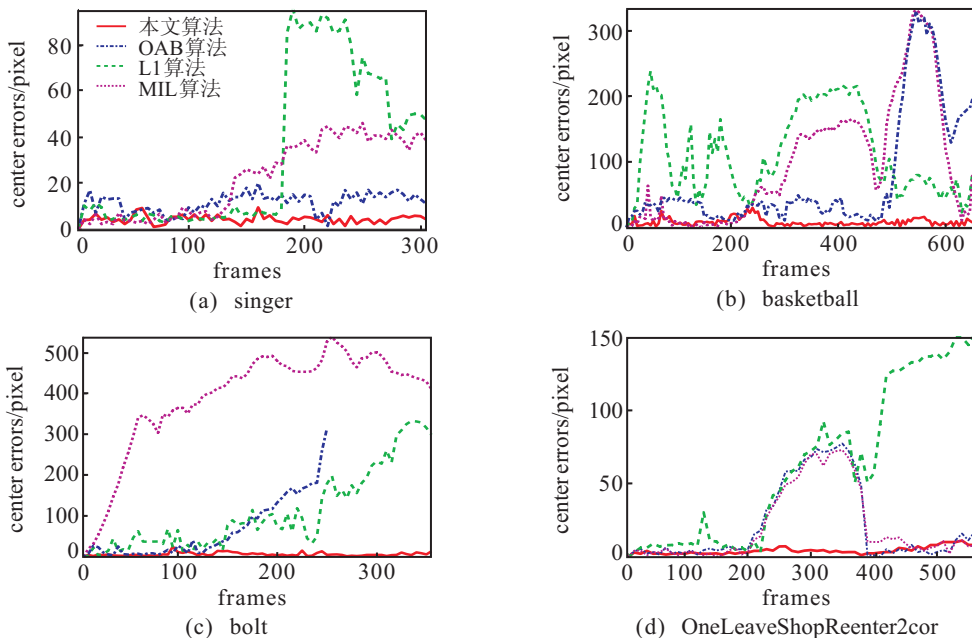


图 5 本文算法与 OAB 算法、MIL 算法、L1 算法跟踪结果的定量对比

表 1 跟踪结果的定量对比

序列	算 法			
	OAB	MIL	L1	本文
basketball	74	96	112	8
bolt	66	390	109	6
singer	12	22	33	4
OneLeaveReenter2cor	21	20	62	4

### 3.3 算法的复杂度分析

一个跟踪算法优劣的评价指标包括精度和效率两个方面, 精度已在第 3.1 和第 3.2 节进行了分析. 在效率方面, 因为不同的程序运行环境对算法的运算时间有着较为明显的影响, 而算法复杂度可以对执行算法所需要的计算工作量给予定量描述, 是衡量程序执行效率的一个非常准确且通用的方法, 所以在此主要对算法复杂度进行分析. 由于 OAB 算法与本文算法的关联度最高, 下面将对这两种算法进行对比分析.

本文算法的复杂度主要包括 3 个方面: 1) 候选样本的生成, 可以表示为  $O(ni)$ . 其中:  $n = 20$  为随机样本的数量,  $i = 10$  为迭代次数. 2) 样本置信度的计算, 可以表示为  $O(nblk)$ . 其中:  $b = 4$  为基分类器的个数,  $l = 4$  为梯度特征的最高层数,  $k = 8$  表示层数为 0 的金字塔梯度直方图的维度. 3) 分类器的更新, 可记为  $O(mq(u + bu))$ . 其中:  $m = 16$  为用来更新分类器的样本数,  $q = 3$  为迭代数,  $u = 10$  为弱分类器的总数. 由于增加的遮挡检测环节对算法的复杂度基本上没有影响, 在此不予计算. 于是本文算法的整体复杂度可以表示为  $O(t(ni + nblk + mqu(1 + b)))$ , 其中  $t$  为测试视频的帧数. 而 OAB 算法的整体复杂度为  $t(w_1h_1(1 + b) + mqu(1 + b))$ , 其中  $w_1$  和  $h_1$  分别表示搜索区域的宽和高, 一般为上一时刻用来描述目标的矩形框大小的二倍. 因此目标所占面积越大复杂度越高, 反之亦然. 而本文算法的复杂度不随着目标大小的改变而改变. 通过对比可以发现, 两者的区别主要在样本的生成和置信度的计算部分, 所以本文只对这两部分的复杂度进行计算. 以目标面积最小的 bolt 视频为例, 本文和 OAB 算法的复杂度分别为  $O(2760)$ 、 $O(30000)$ , 通过这个计算结果可以发现, 本文算法在提高精度的同时降低了复杂度.

## 4 结 论

本文提出了一种应用于单摄像机下的单目标跟踪算法. 该算法在贝叶斯框架下设计了一个基于特征自适应选择机制的观测模型, 通过 M-H 抽样算法获取候选样本, 并分别计算每个样本属于目标的概率, 进而选取概率最大的样本作为当前时刻目标的状态. 为了提高算法的性能, 采用了基于多 PHOG 特征的候选特征池, 使得算法更利于捕获目标与背景间的细微差异, 并在跟踪结束后通过遮挡检测进一步抑制漂移现象的发生. 最后通过大量对比实验验证了本文

算法的有效性.

### 参考文献(References)

- [1] Comaniciu D, Member V, Meer P. Kernel-based object tracking[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-575.
- [2] Wang Dong, Lu Hu Chuan, Yang Ming Hsuan. Least soft-threshold squares tracking[C]. Proc of the IEEE Computer Society Conf on Computer Vision and Pattern Recognition. Oregon: IEEE Press, 2013: 2371-2378.
- [3] Kwon J, Lee K. Tracking by sampling trackers[C]. Proc of the IEEE Int Conf on Computer Vision. Barcelona: IEEE Press, 2011: 1195-1202.
- [4] Yilmaz A, Javed O, Shah M. Object tracking: A survey[J]. ACM Computing Surveys, 2006, 38(4): 1-45.
- [5] Grabner H, Grabner M, Bischof H. Real-time tracking via on-line Boosting[C]. Proc of the British Machine Vision Conf. Edinburgh: Elsevier Press, 2006: 47-56.
- [6] Mei Xue, Ling Hai Bin. Robust visual tracking and vehicle classification via sparse representation[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2011, 33(11): 2259-2272.
- [7] Collins R, Liu Y, Leordeanu M. Online selection of discriminative tracking features[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1631-1643.
- [8] Babenko B, Yang M H, Belongie S. Robust object tracking with online multiple instance learning[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2011, 33(8): 1619-1632.
- [9] Anna B, Andrew Z, Xavier M. Representing shape with a spatial pyramid kernel[C]. Proc of the 6th ACM Int Conf on Image and Video Retrieval. Girona: Elsevier Press, 2007: 401-408.
- [10] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]. Proc of the IEEE Computer Society Conf on Computer Vision and Pattern Recognition. San Diego: IEEE Press, 2005: 886-893.
- [11] Ling H, Okada K. Diffusion distance for histogram comparison[C]. Proc of the IEEE Computer Society Conf on Computer Vision and Pattern Recognition. New York: IEEE Press, 2006: 246-253.
- [12] Kwon J, Lee K. Visual tracking decomposition[C]. Proc of the IEEE Computer Society Conf on Computer Vision and Pattern Recognition. San Francisco: IEEE Press, 2010: 1269-1272.
- [13] Wang S, Lu H, Yang F, et al. Superpixel tracking[C]. Proc of the IEEE Int Conf on Computer Vision. Barcelona: IEEE Press, 2011: 1323-1330.

(责任编辑: 李君玲)