

基于多尺度自适应权重的稀疏表示目标跟踪算法

程中建 周双娥 李康

湖北大学计算机与信息工程学院 武汉 430062

(2838588360@qq.com)

摘要 目标跟踪是计算机视觉中的一个重要研究领域,在交通导航、自动驾驶、机器人技术等众多方面有着广泛应用。基于局部稀疏表示的生成式模型算法 ASLA 的速度快、跟踪准确性高,但是在复杂跟踪环境下,例如目标局部遮挡、目标外观剧烈变化等,往往会丢失目标。文中分析原算法跟踪原理得到了产生目标跟踪丢失的原因。基于 ASLA 算法,提出了 3 点改进方法:1)适应跟踪目标区域大小,采用多尺度分块方式,获取互补的目标局部信息;2)在 ASLA 特征池化过程中根据分块重构误差建模分块自适应权重,以区分不同分块中包含的判别信息,且在多尺度池化特征中引入不同尺度下的目标遮挡信息作为权重;3)在模板更新时,利用最近帧跟踪结果的稀疏表示权重,使更新模板更相似最近跟踪结果,提高了算法的鲁棒性。实验结果表明,该算法在复杂跟踪环境下相比 ASLA 等具有更高的跟踪准确度,能够实时、准确地跟踪到目标。

关键词: 目标跟踪;多尺度自适应权重;稀疏表示;ASLA;生成式模型

中图分类号 TP391.41

Sparse Representation Target Tracking Algorithm Based on Multi-scale Adaptive Weight

CHENG Zhong-jian, ZHOU Shuang-e and LI Kang

School of Computer Science and Information Engineering, Hubei University, Wuhan 430062, China

Abstract Target tracking is an important research field in computer vision. It is widely used in many aspects such as traffic navigation, autonomous driving and robotics. The generative model algorithm ASLA based on local sparse representation is fast and has high tracking accuracy, but it often loses its target in the face of complex tracking environment, such as target partial occlusion and dramatic change of target appearance. This paper analyzes the tracking principle of the original algorithm to get the cause of target tracking loss. Based on the ASLA algorithm, a three-point improvement method is proposed. 1) Adaptive tracking of the target area size using multi-scale blocking method to obtain complementary target local information. 2) In the feature pooling process of ASLA, block adaptive weight is modeled based on block reconstruction error to distinguish the discrimination information contained in different blocks, and introducing target occlusion information at different scales as weights in multi-scale pooling features. 3) When the template is updated, the weight of the latest tracking results in subspace sparse representation is enhanced to make the updated template more similar to the recent tracking results, and improve the robustness of the algorithm. Experimental results show that the algorithm has higher tracking accuracy than algorithms such as ASLA in complex tracking environment, and can track the target in real time and accurately.

Keywords Target tracking, Multi-scale adaptive weight, Sparse representation, ASLA, Generative

1 引言

目标跟踪是计算机视觉的重要研究领域之一,在交通导航、虚拟现实、视频监控、机器人技术等领域有着广泛的应用。近年来,针对单目标跟踪算法的研究取得了极大的进展^[1-5]。但由于受到目标局部遮挡、光照强度变化、目标外观运动模糊等众多因素的干扰,开发出高效的跟踪算法仍然具有挑战性。

按照外观模型的不同,现有的目标跟踪算法可以分为两类:判别式模型和生成式模型。判别式算法模型^[6-9]将目标跟踪视为二值分类问题,通过使用采样得到的正样本和负本来训练分类器以分离目标与背景,然后选择具有最大分类响应的待选样本作为跟踪结果。Grabner 等^[6]提出了一种基

于在线增强的判别特征更新的目标跟踪算法,提高了判别模型的跟踪准确度;Jiang 等^[8]提出了一种基于软特征(Soft Feature, SF)的目标前趋预测跟踪方法,通过利用目标前景区域的灰度特征具有可微分的同频率变化这一显著特点,实现了对目标的长时间跟踪;Xue 等^[9]通过建立一种在线鲁棒判别式字典学习模型,结合粒子滤波框架实现了目标跟踪。判别式模型在跟踪外观无剧烈变化的环境中往往会丢失目标。

生成式算法模型^[10-15]基于训练样本学习到目标的外观模型,并在候选样本中选择具有最小重构误差的样本作为跟踪结果。Gao 等^[10]提出了一种用于并发单标签图像分类和标注的组稀疏编码框架;Hu 等^[11]提出了一种多特征联合稀疏表示方法目标跟踪模型;Jepson 等^[12]提出了基于最大期望

基金项目:国家自然科学基金面上项目(61370002)

This work was supported by the National Natural Science Foundation of China (61370002).

通信作者:周双娥(zhouse@hubu.edu.cn)

法的高斯混合模型算法;Mei等^[13]通过建立目标整体外观模型,求解 l_1 最小化问题来实现目标跟踪;Ren等^[14]将局部稀疏表示组合成用于目标检测的稀疏编码直方图,获得了鲁棒的跟踪效果;Tong等^[15]以TLD^[16]框架为基础,通过动态调整图片分辨率,来改进模板匹配方法,实现了一种快速目标跟踪算法。

近年来,基于生成式模型的算法ASLA^[17]因为其跟踪速度快、准确性高,而受到了科研工作者的关注。ASLA将分块思想引入到目标稀疏表示中,将各分块形成的池化特征向量作为候选样本的相似度,结合目标局部特征和整体外观信息提高了算法的跟踪准确度。在模板更新方面,ASLA结合增量子空间学习和稀疏表示方法来适应目标外观变化,在减少遮挡目标模板的影响的同时,减小了目标偏移的可能性。然而,ASLA算法在跟踪过程中仅采用固定大小的分块,未能考虑目标外观具有的不同结构信息,不能准确获取目标局部信息;且在目标区域的池化表征中应该考虑不同分块对于判别特征的重要性,基于重要性赋予不同分块自适应权重;另外,ASLA在模板更新过程中同样忽略了新加入的模板集应与最新的跟踪结果相似这一影响。

基于上述研究,为解决文献[17]中目标局部信息缺失等问题,本文从多尺度目标分块角度出发研究了基于稀疏表示

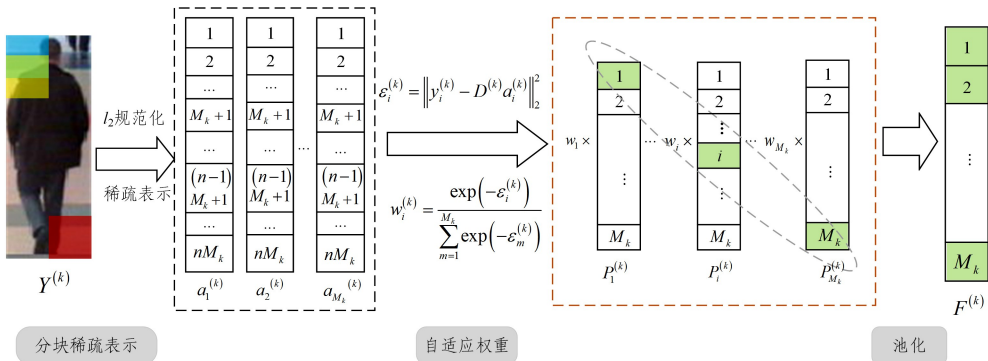


图2 基于权重的局部稀疏外观模型的流程图

Fig. 2 Flow chart of local sparse appearance model based on weights

2.1 基于自适应权重的局部稀疏外观模型

由稀疏表示理论可知,一个目标分块向量可由目标模板及自适应分块稀疏系数表示。预先采集得到大量目标模板 $T = \{T_i | i=1:n\}$,将模板中的目标区域按照不同尺度进行分块将其并作为重构字典,其中尺度个数为 K ,得到 $D^{(k)} = \{d_i^{(k)} | i=1:n \times M_k\} \in R^{d_k \times (n \times M_k)}$,其中 k 表示目标区域的第 k 种分块尺度; $d_i^{(k)} \in R^{d_k}$ 表示 k 尺度分块下的第 i 个分块的图片像素值列向量, d_k 为其维度; n 表示目标模板的个数; M_k 表示 k 尺度分块下的目标区域分块个数。对 $D^{(k)}$ 中的每个列向量进行归一化。由于特定尺度下的各分块对应目标区域的某一固定部分,因此结合各分块信息能够表示出目标的整体结构,在提取不同模板共同特征的字典下,复杂的待选样本分块可被较好地表示,且在结合不同尺寸提取互补目标信息的作用下,提高了准确性。

给定一个待选样本,将其按照不同尺度分块得到 $Y^{(k)} = \{y_i^{(k)} | i=1:M_k\} \in R^{d_k \times M_k}$,其中 $y_i^{(k)} \in R^{d_k}$ 表示待选样本在尺度 k 下的第 i 个分块。则 $y_i^{(k)}$ 可由 $D^{(k)}$ 线性稀疏表示:

$$\min \|y_i^{(k)} - D^{(k)} a_i^{(k)}\|_2 + \lambda_1 \|a_i^{(k)}\|_1, a_i^{(k)} \geq 0 \quad (1)$$

的目标跟踪算法,提出了一种在粒子滤波框架下基于多尺度自适应权重的稀疏表示目标跟踪算法。

2 算法描述

如图1、图2所示,对于每一个候选样本,按照多尺度处理得到不同分块。对于某一分块尺度 $ScaleK$,由预先采集到的模板集合 $D^{(k)}$ 进行稀疏表示,计算重构误差并赋以权重,得到尺度 k 的池化特征向量。根据不同尺度下的待选样本遮挡的严重程度赋权值,得到待选样本与实际样本的相似度,最终选取相似度最大的待选样本作为该帧的跟踪结果。

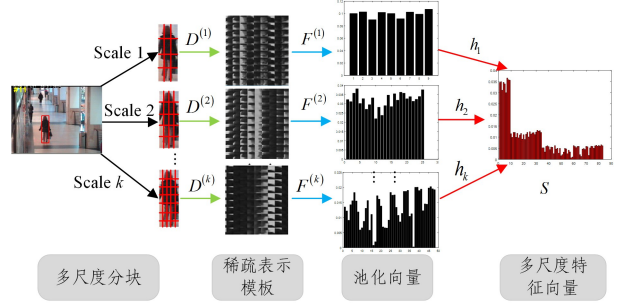
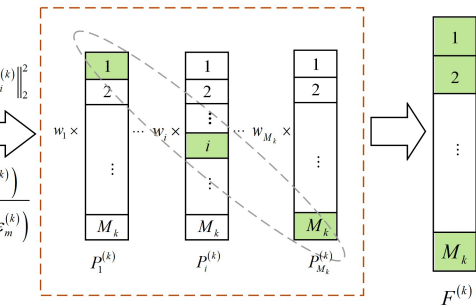


图1 基于多尺度自适应权重的稀疏表示跟踪算法流程图

Fig. 1 Flow chart of a sparse representation tracking algorithm based on multi-scale adaptive weights



得到系数矩阵 $A^{(k)} = [a_1^{(k)}, a_2^{(k)}, \dots, a_{M_k}^{(k)}] \in R^{(n \times M_k) \times M_k}$ 为 k 尺度下的各个分块对应的稀疏表达系数。为了准确表示稀疏表示结果,如图2所示,将分块 k 的稀疏系数分成 n 个部分,即可得到 $(a_i^{(k)})^T = [(a_{i1}^{(k)})^T, (a_{i2}^{(k)})^T, \dots, (a_{im}^{(k)})^T]$,其中 $a_{im}^{(k)} \in R^{M_k \times 1}$,累加向量 $a_{im}^{(k)}$ 得到分块 i 的稀疏表示结果。

$$P_i^{(k)} = \sum_{m=1}^n a_{im}^{(k)}, i=1, 2, \dots, M_k \quad (2)$$

由于利用模板 T 稀疏表示具有外观变化的模板时,不同分块提取的目标特征信息的多少不同,应当赋以不同的特征提取权重,而在ASLA中采用固定权重不能够对不同分块的影响进行区分。因此,本研究利用每个分块的重构误差来有效地表示各分块的权重。

$$\epsilon_i^{(k)} = \|y_i^{(k)} - D^{(k)} a_i^{(k)}\|_2^2 \quad (3)$$

$$w_i^{(k)} = \exp(-\epsilon_i^{(k)}) / \sum_{m=1}^{M_k} \exp(-\epsilon_m^{(k)}) \quad (4)$$

其中, $\epsilon_i^{(k)}$ 表示尺度 k 的第 i 个分块的重构误差,显然对于频繁出现的分块,其具有较小的重构误差,这些分块在表示目标时具有较高权重的特征性。因此,规范化 $P_i^{(k)}$ 为:

$$(P_i^{(k)})' = w_i^{(k)} \times P_i^{(k)}, i=1, 2, \dots, M_k \quad (5)$$

由文献[17]可知,矩阵 $\mathbf{P}^{(k)} = [(P_1^{(k)})', (P_2^{(k)})', \dots, (P_{M_k}^{(k)})'] \in R^{M_k \times M_k}$ 中对角元素形成的向量能够作为最后的池化特征来表示待选样本,即:

$$F^{(k)} = \text{diag}(P^{(k)}) \quad (6)$$

其中, $F^{(k)}$ 表示待选样本在尺度 k 下提取局部和整体特征信息后的池化特征。

2.2 基于多尺度的目标区域分块

目标区域在固定尺度分块稀疏表示后能够较好地表示待选样本的特征,然而由于存在分块尺度的限制,获取的局部结构信息往往不够充分,为了精确获取目标的局部信息,本文采用多尺度的目标区域分块方式,不同分块尺度获取的局部信息相互补充,能较好地表示目标。即:

$$S = [(F^{(1)})^T, (F^{(2)})^T, \dots, (F^{(K)})^T]^T \quad (7)$$

其中, $F^{(k)}$ ($k=1, 2, \dots, K$) 表示在第 k 个分块尺度下的池化特征向量, $S \in R^{\sum_{k=1}^K M_k}$ 表示待选样本的多尺度特征向量。

在跟踪过程中,受到目标外观的变化,不同尺度分块得到的池化特征向量会受到改变,为了适应由于目标外观变换对于不同尺度下得到的池化表征向量的影响,为不同尺度分块方式赋予自适应权重。对于已经计算得到第 k 种尺度下的各分块重构误差 $\epsilon_i^{(k)}$, 设定阈值 ϕ , 则

$$o_i^{(k)} = \begin{cases} 0, & \epsilon_i^{(k)} \leq \phi \\ 1, & \epsilon_i^{(k)} > \phi \end{cases} \quad (8)$$

当重构误差值大于阈值时, k 尺度下 k 的分块 i 被遮挡, 此时 $o_i^{(k)}=1$ 。若尺度 k 下的被遮挡的分块越多, 则其对应的第 k 尺度下的池化特征向量的权重 h_k 越小, 即:

$$G_k = \frac{\sum_{i=1}^{M_k} o_i^{(k)}}{M_k} \quad (9)$$

$$h_k = \frac{\exp(-G_k)}{\sum_{k=1}^K \exp(-G_k)} \quad (10)$$

重写式(7), 计算得到 K 个尺度下的目标特征向量为:

$$S = [h_1(F^{(1)})^T, h_2(F^{(2)})^T, \dots, h_K(F^{(K)})^T]^T \quad (11)$$

其中, $\sum_{k=1}^K h_k = 1$, 在式中引入权重 h_k 能够区分在不同尺度下目标模板与真实跟踪结果的相似性, 且考虑了尺度对相似性的影响。

2.3 模板更新

作为基于生成式模型的跟踪算法中的重要部分, 有效的模板更新方式对跟踪具有重要意义。受到外界光照的变化, 运动模型等因素造成了不可避免的目标外观变换, 固定的模板将会丢失目标。然而, 如果模板集更新过快, 模板产生的误差将会累积并最终导致跟踪器发生漂移, 若模型集更新得过慢, 将会无法更新目标外观变化信息。本文通过改进文献[17]中的模板更新方式, 来提高模板更新的准确率, 以完成模板更新。

文献[17]结合子空间学习与稀疏表示的优点, 改善目标模板集以适应目标外观变化, 并减弱遮挡模板带来的影响。其运用一个递增的概率序列赋以模板集中模板更新的概率。即:

$$L_p = \left\{ 0, \frac{1}{2^{n-1}-1}, \frac{3}{2^{n-1}-1}, \dots, 1 \right\} \quad (12)$$

由满足 $[0, 1]$ 分布的随机变量 r 来判定模板集中的模板是否被更新。该方法在减慢旧模板更新速度的同时加快了新添加模板的更新速度, 减缓了由于模板更新产生的漂移问题。

通过旧模板集合 $T_{\text{old}} = \{T_i | i=1:n\}$ 的跟踪结果可以得到新的模板集合 $T_{\text{new}} = \{T_i | i=n+1:m\}$, $T_i \in R^d$ 为图片的像素值向量。然而, ASLA 算法没有考虑替换模板应该更加相似于最近的模板, 即应与新的跟踪模板集合 T_{new} 更加相似。通过增加模板 T_{new} 提高其奇异值分解的分量, 从而得到新的观测向量。

$$T_{\text{obs}} = [T_{\text{old}}, (\mu n) T_{\text{new}}] \quad (13)$$

其中, μ 为一个常数值, 将跟踪到的新的模板集合重复多次, 与旧模板集合共同形成新的模板矩阵 T_{obs} 。由文献[18]可知, 更新模板 z 可由模板集合 T_{obs} 奇异值分解和平凡模板线性表示, 即:

$$z = Uq + e = [U \quad I] \begin{bmatrix} q \\ e \end{bmatrix} \quad (14)$$

其中, z 为观测向量, U 为 T_{obs} 分解后的基向量, q 为其系数; e 为 z 中存在遮挡或损失的像素点。由于噪声和遮挡产生的误差具有随机性, 将问题转化为 l_1 最小化问题。即:

$$\min \|z - Hc\|_2 + \lambda_2 \|c\|_1 \quad (15)$$

其中, $H = [U \quad I]$, $c = [q \quad e]^T$; λ_2 为参数, 该方法在避免更新模板时由遮挡或噪声对模板集合产生影响的同时, 使得新加入的模板更加相似于最近的跟踪结果, 很好地解决了目标外观变化丢失目标的情况。

算法 1 模板更新算法

输入: 跟踪目标向量 z , 旧模板集合 T_{old} , 新跟踪模板集合 T_{new} , 参数

λ_2

输出: 旧模板集合 T_{old}

步骤 1 由式(13)得到分解矩阵 T_{obs}

步骤 2 $T_{\text{obs}} \xrightarrow{\text{SVD}} U \Sigma V^T$, 得到基向量 U

步骤 3 产生一系列升序排列数字, 并归一化到 $[0, 1]$ 范围内

步骤 4 随机产生满足 $0-1$ 均匀分布的数字, 移除对 T 模板集中的样本

步骤 5 由式(15)计算得到 q, e

步骤 6 将 $z' = Uq$ 模板添加到 T_{old} 中

3 实验结果与分析

3.1 实验环境与数据准备

本文算法在 Matlab2016a 平台下实现, 操作系统为 Win10, CPU 为 2.4GHz Intel i5 处理器, 内存为 8GB。算法中采用的分块尺度大小为 $11 \times 11, 9 \times 9, 7 \times 7$ (经验参数), 尺度个数 $K=3$; 采用已标记的跟踪序列前 10 帧为目标模板, 即 $n=10$ (经验参数); $\mu=0.85$; $\lambda_1=0.01$; $\lambda_2=0.001$; 粒子个数为 600。

为了衡量算法的跟踪效果, 验证跟踪的有效性, 将算法与 ASLA 等跟踪原理相似的算法进行比较, 对比的算法有 CT^[19], MIL^[20], L1APG^[21], ASLA^[17], 从而得到算法在 6 个具有挑战性视频序列中的跟踪结果 (Basketball, Boy, David2, David3, Fish, Jumping), 视频序列中包含跟踪过程中可能遇到的目标形变、部分遮挡、光照强度变化、运动模糊等干扰因素。跟踪序列数据与对比算法的实验数据由文献[22]得到。

3.2 定量分析

本实验采用 2 个指标评价算法的跟踪效果。1) 重合度判别法, 根据跟踪结果的目标框与人工标记的目标框重合度成功率^[22] 评价方法的表现结果, 其中定义重合度为 $score =$

$\frac{area(R \cap G)}{area(R \cup G)}$, R 为算法跟踪结果区域; G 为人工标定的目标区域, \cap 和 \cup 分别表示区域的交集与并集, 函数 $area(\cdot)$ 返回区域像素点个数。2) 中心距离误差判别法, 人工标记目标框中心与实际跟踪结果中心的距离, 单位为像素。在传统的评价方法中, 若某帧跟踪结果 $score$ 值大于设定阈值, 则视为该帧跟踪正确, 一般取阈值为 0.5; 在中心距离误差判别法中, 若某帧中心距离的误差小于给定阈值, 则认为该帧跟踪成功, 一般取阈值为 20 个像素。

图 3 给出了重合度成功率曲线图, 图 4 给出了中心距离误差成功率图。从图 3、图 4 可以看出, 本文算法在处理目标局部遮挡 (如跟踪序列 David3)、光照强度剧烈变化 (如 Fish)、运动模糊 (如 Jumping、Boy)、目标外观剧烈变化和复杂背景 (如 David2 和 Basketball) 时都具有更高的跟踪精确度。相比原算法, 在测试的 6 个视频序列中本文算法均有较大提高。由图 3、图 4 可知, 本文算法得到了良好的跟踪效果, 下文结合图 3, 根据所选测试序列特点和跟踪结果进行具体分析。

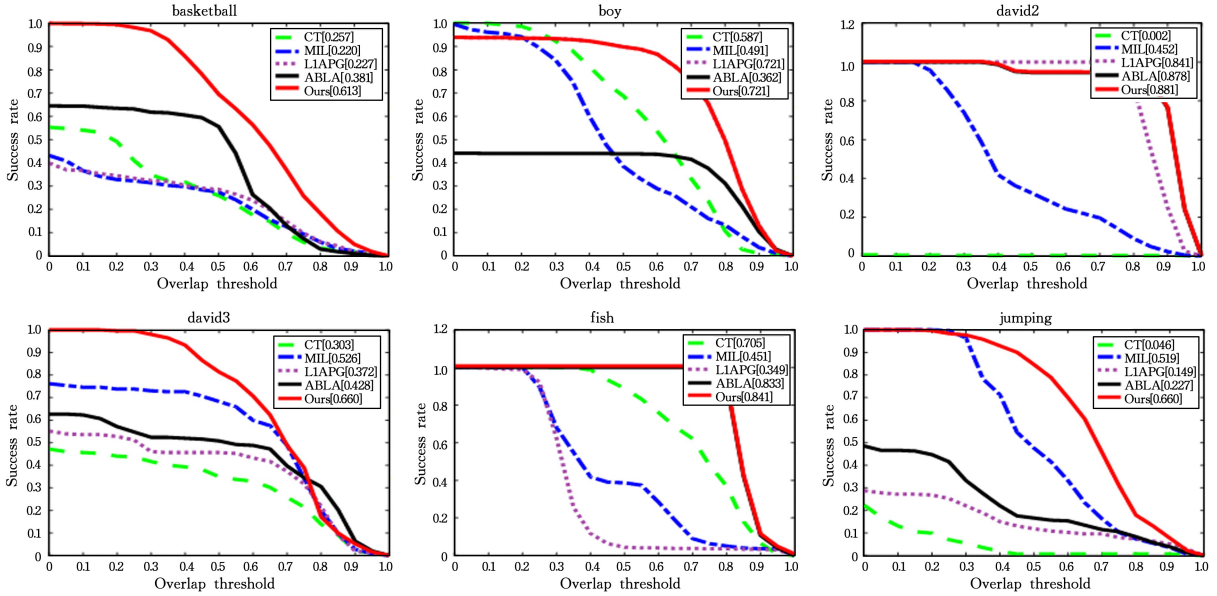


图 3 重合度成功率

Fig. 3 Success rate of coincidence degree

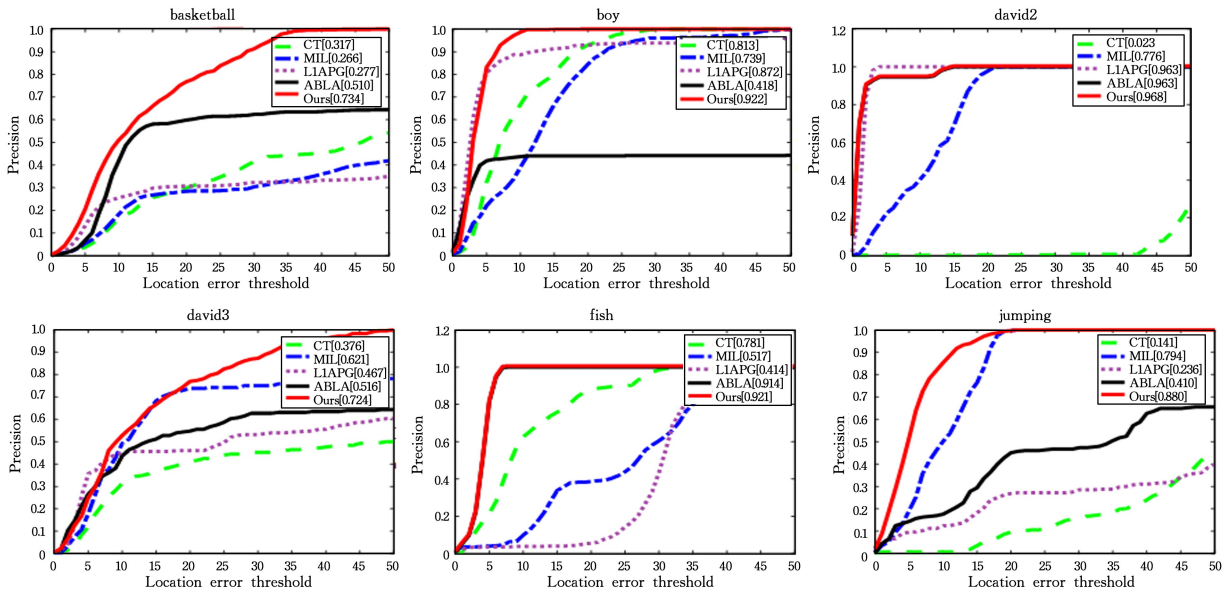


图 4 中心距离误差成功率

Fig. 4 Success rate of center distance error

Basketball: 该测试视频序列中, 目标区域中包含复杂背景区域, 为非凸的, 如目标运动时腿间区域以及头部两侧区域。非目标区域在目标跑动时随着背景改变呈现不规则变动, 造成跟踪困难。L1APG 算法由于采用整体目标作为模板, 当第 105 帧时, L1APG 算法发生了目标偏移; ASLA 算法由于采用固定尺度获取目标的局部信息, 在 199 帧时发生漂

移; 在 236 帧时, 算法 ASLA 和 MIL 丢失目标; 在 293 帧时, CT 和 L1APG 均漂移到非目标队员身上, 这主要是 CT 中贝叶斯分类器无法提取到有效的判别性特征, 而 L1APG 采用整体目标稀疏表示时, 其目标模板与背景队员颜色接近, 产生选择干扰。从测试结果可知, 本文算法能够在存在背景干扰的环境下保证跟踪精度, 表现出了算法的鲁棒性。

Boy:该测试序列中存在目标快速移动形成的运动模糊、相机抖动形成的图像整体模糊,目标外观形变造成了跟踪困难。从图5可以看出,第131帧由于目标外观变化,CT和MIL无法提取到目标信息而发生漂移;第266帧时,ASLA和L1APG模型丢失目标,这主要是因为目标快速移动产生的运动模糊使得稀疏表示得到的琐碎系数不具备稀疏性,影响算法的跟踪效果,而本文算法能够较为准确地跟踪到目标,表明提出的算法能够有效处理运动模糊和形变等跟踪难题。

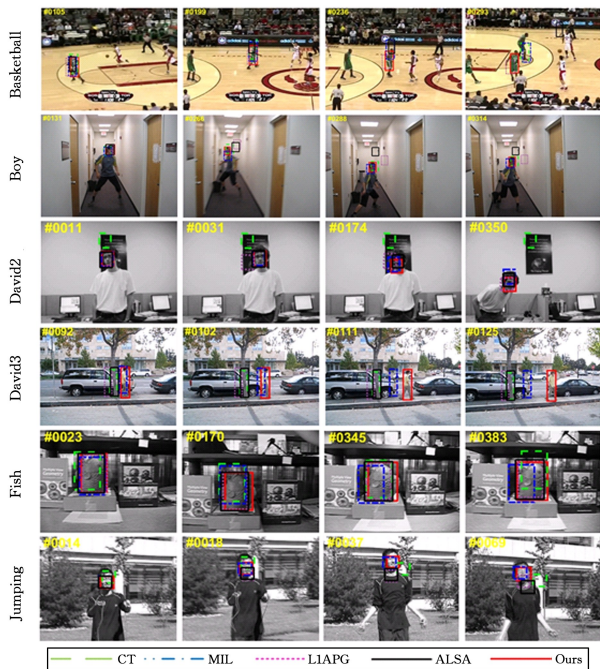


图5 跟踪结果图

Fig.5 Tracking result graph

David2:该测试视频是在室内环境下采集的。视频特点为:跟踪区域较小,难以提取判别性的特征;目标背景与目标区域相似。从图3可以看出,本研究能够较为准确地跟踪到目标。由图5可知,在11帧时,CT由于无法投影到有效的哈尔特征作为分类器输入,从而丢失目标;在31帧时,L1APG丢失目标,这主要是因为目标区域与背景区域接近,L1APG无法由稀疏表示选出有效的候选样本;在170~400帧的过程中,受到目标区域发生旋转等形变的影响,MIL和ASLA发生了严重偏移,这主要是因为ASLA算法在局部稀疏表示过程中选取固定尺度的分块,无法获取到不同尺度下的目标局部信息,发生了偏移。

David3:该测试序列的主要特点在于:目标在经过树木时,目标区域在短时间内存在较大面积遮挡。由图5可知,在第92帧时,CT,L1APG,ASLA算法均漂移到树木上,丢失了目标,这主要是因为CT,L1APG算法在目标被遮挡时,无法提取到有效哈尔特征并且整体稀疏表示不能获取到有效目标区域信息,而ASLA虽然采用了分块稀疏表示的方法,来获取目标局部信息,但遮挡区域较大,使得无法获取到充足的局部信息,丢失目标,在111帧时,MIL算法丢失目标。相比ASLA算法,本文算法采用一种多尺度的目标区域分块方式,使得不同尺度下获取的目标局部信息互补,在跟踪目标区域大面积遮挡时仍然能够获取有效的局部信息,对不同分块稀疏系数和不同尺度下的池化向量赋予自适应权重,避免了跟踪遮挡物体过程中发生的遮挡问题。整体性能上可看出,本

文算法能够有效跟踪存在遮挡的目标。

Fish:该测试序列特点:视频拍摄相机在跟踪过程中一直晃动,且目标区域发生了两次剧烈光照变化,从图3可以看出L1APG,ASLA和本文算法能够比较准确地跟踪到目标,且本文算法AUC值达到了0.85,高于L1APG和ASLA算法。而CT和MIL算法在第170,345,383帧发生了严重偏移。由实验数据可知,本文算法在处理视频中的相机抖动与运动模糊时具有鲁棒性。

Jumping:本测试序列在室外条件下采集,视频中存在大比例运动模糊帧。从图5可看出,CT算法在第14帧发生偏移,并在37帧时完全丢失目标,而L1APG和ASLA算法均能在第18帧发生严重偏移,且在37帧时丢失目标,从图3可知,MIL和本文算法能够准确跟踪到目标。

结束语 针对ASLA算法在目标跟踪过程中存在的不足,本文提出了一种基于多尺度自适应权重的稀疏表示目标跟踪算法。采用多尺度分块思想并结合自适应权重,解决原算法采用单一固定权重方法导致的目标局部互补信息缺失的问题,提高复杂跟踪环境中的跟踪准确率。在模板更新阶段,提高最新结果模板的权重,使算法能够尽快适应最新结果,提高算法的鲁棒性。实验结果表明,本文算法在处理复杂环境时较ASLA等其他算法有更高的准确率,能够完成实时目标跟踪。在后续研究中,将会在现有基础上结合深度学习特征提高跟踪效率。

参考文献

- [1] ZHANG M, WANG J. An Interactive Likelihood Target Tracking Algorithm Based on Deep Learning[J]. Computer Science, 2019, 46(2): 279-285.
- [2] DANELLJAN M, SHAHBAZ KHAN F, FELSBERG M, et al. Adaptive color attributes for real-time visual tracking[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1090-1097.
- [3] ZHANG T, JIA K, XU C, et al. Partial occlusion handling for visual tracking via robust part matching[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1258-1265.
- [4] WANG Y X, ZHAO Q J, CAI Y M, et al. Target tracking based on self-reconstructed particle filter algorithm[J]. Chinese Journal of Computers, 2016, 39(7): 1294-1306.
- [5] JIANG W T, LIU W J, YUAN W, et al. Visual Quantum Target Tracking Method[J]. Journal of Software, 2016, 27(11): 2961-2984.
- [6] GRABNER H, BISCHOF H. On-line boosting and vision[C]// 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). IEEE, 2006: 260-267.
- [7] KALAL Z, MATAS J, MIKOLAJCZYK K. Pn learning: Bootstrapping binary classifiers by structural constraints[C]// 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010: 49-56.
- [8] JIANG W T, LIU W J, YUAN W. Target tracking based on soft feature theory[J]. Chinese Journal of Computers, 2016, 39(7): 1334-1355.
- [9] XUE M G, ZHU H, YUAN G L. Online Robust Discriminant Dictionary Learning Visual Tracking[J]. Chinese Journal of Electronics, 2015, 44(4): 838-845.
- [10] GAO S, CHIA L T, TSANG I W H, et al. Concurrent single-la-

- bel image classification and annotation via efficient multi-layer group sparse coding [J]. *IEEE Transactions on Multimedia*, 2014, 16(3):762-771.
- [11] HU W, LI W, ZHANG X, et al. Single and multiple object tracking using a multi-feature joint sparse representation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(4):816-833.
- [12] JEPSON A D, FLEET D J, EL-MARAGHI T F. Robust online appearance models for visual tracking [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(10):1296-1311.
- [13] MEI X, LING H, WU Y, et al. Minimum error bounded efficient ℓ_1 tracker with occlusion detection [C] // *CVPR 2011*. IEEE, 2011:1257-1264.
- [14] REN X, RAMANAN D. Histograms of sparse codes for object detection [C] // *2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2013:3246-3253.
- [15] TONG Y, FEI S Z, SHEN J. Fast Target Tracking Method Based on TLD Framework [J]. *Application Research of Computers*, 2018, 35(1):317-320.
- [16] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(7):1409-1422.
- [17] JIA X, LU H, YANG M H. Visual tracking via adaptive structural local sparse appearance model [C] // *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012:1822-1829.
- [18] MEI X, LING H. Robust visual tracking using ℓ_1 minimization [C] // *2009 IEEE 12th International Conference on Computer Vision*. IEEE, 2009:1436-1443.
- [19] ZHANG K, ZHANG L, YANG M H. Real-time compressive tracking [C] // *European Conference on Computer Vision*. Springer, Berlin, Heidelberg, 2012:864-877.
- [20] BABENKO B, YANG M H, BELONGIE S. Visual tracking with online multiple instance learning [C] // *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2009:983-990.
- [21] BAO C, WU Y, LING H, et al. Real time robust ℓ_1 tracker using accelerated proximal gradient approach [C] // *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012:1830-1837.
- [22] WU Y, LIM J, YANG M H. Online object tracking: A benchmark [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2013:2411-2418.
- (上接第 171 页)
- [9] WANG P, CHEN P, YUAN Y, et al. Understanding convolution for semantic segmentation [C] // *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2018:1451-1460.
- [10] ZAGORUYKO S, KOMODAKIS N. Wide residual networks [J]. *arXiv*:1605.07146, 2016.
- [11] KRIZHEVSKY A, HINTON G. Learning multiple layers of features from tiny images [R]. University of Toronto, 2009.
- [12] COATES A, NG A, LEE H. An analysis of single-layer networks in unsupervised feature learning [C] // *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. 2011:215-223.
- [13] SWERSKY K, SNOEK J, ADAMS R P. Multi-task bayesian optimization [C] // *Advances in Neural Information Processing Systems*. 2013:2004-2012.
- [14] DUNDAR A, JIN J, CULURCIELLO E. Convolutional Clustering for Unsupervised Learning [J]. *Computer Science*, 2015:1143-1151.
- [15] DOSOVITSKIYA, FISCHER P, SPRINGENBERG J T, et al. Discriminative Unsupervised Feature Learning with Exemplar Convolutional Neural Networks [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2014, 38(9):1734-1747.
- [16] ZHAO J, MATHIEU M, GOROSHIN R, et al. Stacked what-where auto-encoders [J]. *arXiv*:1506.02351, 2015.
- [17] DEVRIES T, TAYLOR G W. Improved regularization of convolutional neural networks with cutout [J]. *arXiv*:1708.04552, 2017.



CHENG Zhong-jian. His main research interests include data mining and so on.



ZHOU Shuang-e, professor, is a senior member of China Computer Federation. Her main research interests include big data analysis technology, computer application technology, information security.



WU Hao-hao, born in 1995, master. His main research interests include deep learning and so on.



WANG Fang-shi, born in 1969, Ph.D., professor, is a member of China Computer Federation. Her main research interests include visual information processing and pattern recognition.