

文章编号 1004-924X(2015)04-1161-10

增量深度学习目标跟踪

程 帅¹, 孙俊喜^{2*}, 曹永刚^{1,3}, 赵立荣^{1,3}

- (1. 长春理工大学 电子信息工程学院, 吉林 长春 130022;
2. 东北师范大学 计算机科学与信息技术学院, 吉林 长春 130117;
3. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033)

摘要: 由于现有目标跟踪算法在复杂环境下易发生目标漂移甚至跟踪丢失, 故本文提出了以双重采样粒子滤波为框架, 基于增量深度学习的目标跟踪算法。该算法在粒子滤波中引入粒子集规模自适应调整的双重采样来解决粒子衰减及贫化问题, 并利用无监督特征学习预训练深度去噪自编码器以克服跟踪中训练样本的不足。将深度去噪自编码器应用到在线跟踪中, 使提取的特征集能够有效表达粒子图像区域。在深度去噪自编码器中添加了增量特征学习方法, 得到了更有效的特征集以适应跟踪过程中目标外观变化。该方法还用线性支持向量机对特征集进行分类, 提高对粒子集合的分类精度, 以得到更精确的目标位置。在复杂环境下对不同图片序列进行的实验表明: 该算法的跟踪综合评价指标为 94%、重叠率为 74%, 平均帧率为 13 frame/s。与现有的跟踪算法相比, 本算法有效地解决目标漂移甚至跟踪丢失问题, 并且对遮挡、相似背景、光照变化、外观变化具有更好的鲁棒性及精确度。

关键词: 目标跟踪; 粒子滤波; 深度去噪自编码器; 支持向量机; 增量特征; 深度学习;

中图分类号: TP391.4 **文献标识码:** A **doi:** 10.3788/OPE.20152304.1161

Target tracking based on incremental deep learning

CHENG Shuai¹, SUN Jun-xi^{2*}, CAO Yong-gang^{1,3}, ZHAO Li-rong^{1,3}

- (1. School of Electronic Information Engineering, Changchun University of Science and Technology, Changchun 130022, China;
 2. School of Computer Science and Information Technology, Northeast Normal University, Changchun 130117, China;
 3. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China)
- * Corresponding author, E-mail: junxi_sun@126.com

Abstract: As current tracking algorithms lead to target drift or target loss in the complex environment, a tracking algorithm based on the incremental deep learning was proposed under a double-resampling particle filter framework. To solve the problem of particle degradation and depletion, the double-resampling method was introduced to adapt to the particle size in particle filtering and a Stacked Denoising Autoencoder (SDAE) was pre-trained by the unsupervised feature

收稿日期: 2014-10-09; 修订日期: 2014-12-03.

基金项目: 吉林省科技厅资助项目 (No. 20090512; 20100312)

learning to alleviate the lack of training samples in visual tracking. Then, the SDAE was applied to online tracking, so that the extracted feature sets could express the region image representations of the particles effectively. The incremental feature learning was introduced to the encoder of SDAE, the feature sets were optimized by adding new features and merging the similar features to adapt to appearance changes of the moving object. Moreover, a support vector machine was used to classify the features then to improve the classification accuracy of the particles and to obtain a higher tracking precision. According to the results of experiments on variant challenging image sequences in the complex environment, the F -measure and the overlapping ratio of the presented algorithm are 94%, 74%, respectively and the average frame rate is 13 frame/s. Compared with the state-of-the-art tracking algorithms, the proposed method solves the problems of target drift and target loss efficiently and has better robust and higher accuracy, especially for the target in the occlusions, background clutter, illumination changes and appearance changes.

Key words: target tracking; partical filter; stacked denoising autoencoder; support vector machine; incremental feature; deep learning

1 引言

目标跟踪是计算机视觉研究领域的重要组成部分,广泛应用在视频监控、行为识别及国防安全等领域^[1]。目标跟踪也面临一系列挑战,如遮挡、相似背景、光照变化、外观变化等^[2]因素常导致目标漂移甚至跟踪失败。

作为一种典型的跟踪算法,基于外观模型的跟踪算法^[3-5]使用的特征模型不能反映目标本质信息,也不是目标的有效表达,在跟踪过程中常出现误匹配问题,而且复杂的外观模型计算成本高。针对这个问题,有学者在外观模型中引入了机器学习算法^[6-10],将跟踪视为二值分类问题。分类模型可以有效地利用目标和背景信息,将跟踪目标从背景中分离出来。此类算法虽改进了跟踪效果,但也存在一些问题,即如果没有足够的训练样本学习分类模型,得到的分类模型不精确。另外,需标记大量训练样本,成本高。

跟踪过程中目标信息随时间不断发生变化,目前需要解决的问题是基于深度学习的跟踪算法^[11]如何得到最优特征集。因为特征数太多,会导致目标模型过拟合或收敛速度过慢;而特征数太少,又会因缺少相关联特征导致模型欠拟合。该算法仅使用 sigmoid 函数作为分类模型,在相似背景条件下常出现错误分类的情况。

为解决上述问题,本文提出以双重采样粒子滤波为框架,基于增量深度学习的目标跟踪算法。该

算法的双重采样过程可以自适应调整粒子集的规模,解决了粒子衰减和贫化问题。其通过无监督特征学习得到的深度去噪自编码器(Stacked Denoising Autoencoder, SDAE)^[12]提取粒子区域特征,以表达图像的本质信息;选用增量特征学习优化粒子区域特征集来适应目标在跟踪中外观变化,以便更有效地表达粒子区域;利用线性支持向量机(Support Vector Machine, SVM)^[13]对优化后的特征集进行分类,得到粒子置信度;同时微调深度网络,将粒子置信度最高的特征作为跟踪结果。

2 双重采样粒子滤波

粒子滤波是基于蒙特卡洛和递推贝叶斯估计的非线性滤波方法^[14],被广泛应用在目标跟踪中,其用于目标跟踪时存在粒子衰减和贫化问题。为解决上述问题,对粒子集进行双重采样。第一次采样为在顺序重要性采样过程中引入重采样算法,用于抑制或者剔除小权重粒子,保留大权重粒子,根据粒子权重重新分配计算资源,得到等权重粒子集。第二次采样为粒子稀疏化聚合采样^[15],基于粒子空间尺度网格划分对粒子稀疏化加权聚合,以缩减粒子集规模,实现粒子集规模与粒子空间规模的“自适应”一致。

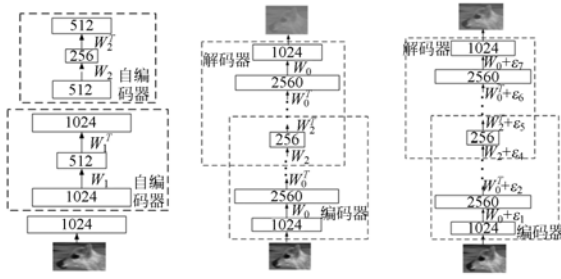
3 增量深度分类神经网络

增量深度分类神经网络由三部分构成:(1)

SDAE 编码器用于提取粒子集合特征;(2)增量特征学习,用于自适应调整、优化特征集数量;(3)线性 SVM 分类器,用于得到粒子置信度,将置信度最高的粒子作为跟踪结果,同时进一步优化网络参数。

3.1 深度去噪自编码器

SDAE 从施加噪声的训练集中学习、重构原始数据,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^[16]预训练多个自编码器,预训练过程如图 1(a)所示。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 的网络结构如图 1(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的误差,对整个网络进行微调,微调后的网络结构如图 1(c)所示。由图 1(b)、1(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。



(a) SDAE 预训练 (b) 网络结构 (c) 微调
(a) Pretraining (b) Network architecture (c) Fine-tuning
图 1 SDAE 构建图

Fig. 1 Establishment diagram of SDAE

3.2 线性 SVM 分类器

利用线性 SVM 分类器对特征集进行二值分类,同时利用反向传播算法进一步微调整个网络,以提高分类精度,解决相似目标条件下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 $(x_n, y_n), n=1, \dots, N, x_n \in \mathcal{R}^D, y_n \in \{-1, +1\}$ 优化无约束条件的 SVM 问题如下:

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

其中: C 为规则参数,式(1)为 hinge loss 函数二阶范数的优化问题(L2-SVM)。

预测样本 x 的类标记公式如下:

$$\operatorname{argmax}_y (w^T x) y. \quad (2)$$

利用反向传播算法微调整个网络,将等式(1)作为目标函数 $l(w)$,倒数第二层的激活值作为输入,优化公式如下:

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

3.3 增量特征学习

为自适应地调整、优化 SDAE 编码器提取到特征的数量,在编码器和线性 SVM 之间添加了增量特征学习层,优化特征集以获得紧凑的特征表示,适应目标在跟踪过程中外观变化。

增量特征学习^[17]由添加新特征和整合相似特征两个组件构成,图 2 为增量特征学习框图。利用混合模型的目标函数优化添加和整合特征过程中的网络参数。混合模型目标函数由生成模型目标函数和判别模型目标函数组成:

$$L_{\text{hybrid}}(x, y) = L_{\text{disc}}(x, y) + \eta L_{\text{gen}}(x), \quad (4)$$

其中: $L_{\text{gen}}(x)$ 为输入数据和重构数据之间的交叉熵函数; $L_{\text{disc}}(x)$ 为 L2-SVM 函数。

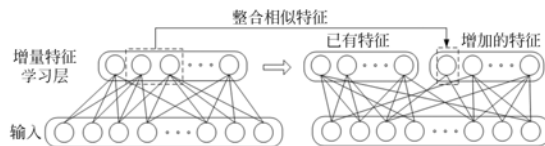


图 2 增量特征学习框图

Fig. 2 Diagram of incremental feature learning

如图 2 所示,添加新特征过程中,保持原有特征集的参数不变,仅训练新添加的特征及其参数,最小化生成模型和判别模型的目标函数来优化网络参数。整合相似特征过程中,在已有的特征集中选择最为相似两个特征作为整合的候选特征集,同时从特征集中删除这两个相似特征。利用混合模型目标函数优化网络参数,在参数集合中,添加新整合特征的参数,同时删除两个相似特征的参数。表 1 为增量特征学习执行过程。

表 1 增量特征学习的执行过程

Tab. 1 Processing of incremental feature learning

算法 1: 增量特征学习

1. 计算输入数据 x 的目标函数 $L(x)$ 。
2. 为集合 B 收集难例样本, 如果 $L(x) > \mu$, $B \leftarrow B \cup \{x\}$ 。
3. 如果 $|B| > \tau$
 - a. 选择 $2\Delta M$ 个候选特征, 整合为 ΔM 个特征。
 - b. 通过贪婪优化算法添加 ΔN 个新特征。
 - c. 令 $B = \phi$, 更新 $\Delta N, \Delta M$ 。
4. 在目前数据集中微调所有特征
5. 重复以上步骤直到收敛

4 算法实现

本文算法流程如图 3 所示。

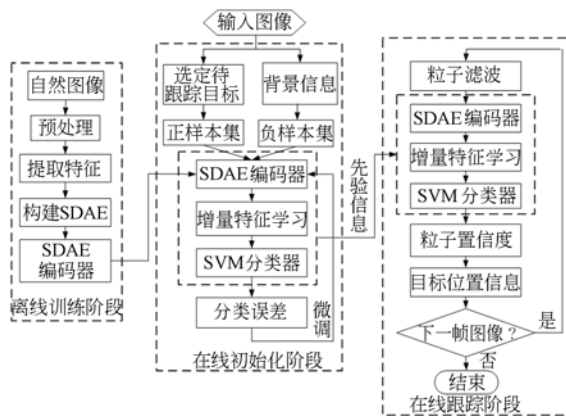


图 3 算法流程图

Fig. 3 Flow diagram of proposed algorithm

离线训练阶段, 将 32×32 自然图像集^[18]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。由于跟踪过程中可用训练样本数量很少, 利用少量的样本训练 SDAE, 会导致网络模型欠拟合。自然图像集可实现 SDAE 网络参数的优化, 解决训练样本不足导致的欠拟合问题。将预训练得到的 SDAE 编码器应用到在线跟踪过程中, 使提取的特征能够更有效地表达图像数据。为更好地提取图像结构信息, 在网络结构中添加超完备滤波层(超完备层), 超完备滤波器如图 4 所示。滤波器大部分为局部边缘检测器。



图 4 超完备滤波器

Fig. 4 Overcomplete filters

为加快预训练过程, 将 32×32 自然图像分为 5 个大小为 16×16 的图像块, 分别在图像左上角、右上角、左下角、右下角及中心位置, 然后训练这 5 个有 512 个隐藏单元的自编码器, 再根据编码器的权重值初始化第一层的自编码器。

在线初始化阶段, 首先在第一帧图像中手动选取待跟踪目标, 并通过对目标图像进行扭曲、旋转、缩放建立完整的目标全视角角数据库, 以此作为正样本, 这对目标旋转、缩放、外观局部变化等条件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增量深度分类神经网络。正负样本集通过 SDAE 编码器提取样本特征, 经增量特征学习层优化特征数量, 利用提取特征集和类标记训练线性 SVM 分类器, 利用分类误差微调网络参数。

在线跟踪阶段, 当新一帧图像到来时, 根据粒子滤波算法分布粒子集, 通过前向传播经增量深度分类神经网络得到每个粒子的置信度。如果所有粒子的最大置信度均大于设定的阈值 τ , 将置信度最大的粒子作为目标的位置信息。反之, 所有粒子的最大置信度小于预先定义的阈值 τ 时, 表明跟踪过程中外界环境导致目标的外观发生改变, 此时依然将置信度最大的粒子作为跟踪结果, 将当前跟踪结果及周围的背景信息进行旋转、缩放, 来更新正负样本集, 用得到的正负样本集微调整量深度分类神经网络, 以更新网络参数。跟踪过程中如果阈值设定太小, 跟踪算法不能很好地适应外观变化; 如果设定太大, 算法可能将背景错误地认为是目标, 从而导致目标漂移问题。

算法实现细节如下: 离线训练 SDAE 阶段, 在自然图像中添加方差为 0.000 4 的高斯噪声。粒子滤波中设定粒子数为 1 500, 置信度阈值 $\tau = 0.85$; 增量特征学习层节点单元数初值为 256, 采用 cosinDistance 函数计算两特征间的相似性, $\Delta N, \Delta M$ 及更新规则的选择参考文献[17]; 混合模型目标函数中 $\eta = 0.3$ 时, 算法可得到最好结果。SDAE 为非线性多层网络结构, 由于每层都包含大量的网络节点及参数, 使得在特征提取、优化特征集合、微调网络参数时比较耗时。为提高

处理速度,离线训练阶段和在线跟踪阶段均采用并行计算和图形处理单元加速技术。

5 实验结果与分析

对本文算法与 10 个主流跟踪算法在不同图片序列上进行测试。对比跟踪算法包括:IVT^[3], MIL^[5], OAB^[7], SBT^[8], TLD^[10], DLT^[11], CT^[19], CXT^[20], Frag^[21], KMS^[22], 这些算法参数均使用默认值,具体参考相应文献。选用 Animal 等 8 组图片序列^[2]进行测试,8 组图片序列特点如表 2 所示。

表 2 图片序列特点

Tab. 2 Features of image sequences

图片名称	特点
Animal	剧烈运动、相似背景
Car11	光照、相似背景
Football	相似背景、剧烈运动、旋转
Girl	遮挡、旋转
MountainBike	剧烈运动、姿态变化
Shaking	剧烈运动、光照、相似背景
Subway	遮挡、相似背景
Woman	遮挡、尺度、视角变化

使用重叠率 overlap、综合评价指标 F 衡量算

法优越性,计算公式如下:

$$overlap = \frac{area(BB_T \cap BB_G)}{area(BB_T \cup BB_G)}, \quad (5)$$

其中: BB_T 为跟踪结果, BB_G 为目标真实位置, $area$ 为区域面积,重叠率大于 0.5 认为跟踪成功。

$$F = \frac{2PR}{P+R}, \quad (6)$$

其中: P 为准确率、 R 为召回率。

试验中每个算法在每组图片上均进行 20 次测试,取 20 次结果的平均值作为算法的最终测试结果,跟踪结果 F 、 $overlap$ 如表 3 所示。

由表 3 可知,每组图片序列有两行测试结果: F (第一行)和 $overlap$ (第二行)。在 8 组序列中 IDL 算法有 7 组都为最优,其余一组为次优/次优。mean 行为每种算法 8 组序列 F 和 $overlap$ 的均值,IDL 算法的 F 和 $overlap$ 分别为 0.94、0.74,明显高于其他算法。由此可见在遮挡、光照变化、尺度变化、相似背景等复杂条件下,IDL 具有更高的准确度和鲁棒性。

中心误差表示跟踪结果与真实目标中心位置的欧氏距离,距离越小表示跟踪结果越有效、越精确。图 5 为 11 种跟踪算法每组图片序列的平均中心误差柱状图。从图 5 可以看出,IDL 中心误差总体上小于其他算法,这表明其跟踪结果具有较高的精确度,可以有效解决目标漂移问题。

表 3 不同图片序列下本文算法和 10 种主流跟踪方法的平均 F 、 $overlap$ 值

Tab. 3 Average F and overlaps of proposed method and 10 main stream methods in differen image sequences

序列	IDL	CT	CXT	DLT	Frag	IVT	KMS	TLD	MIL	OAB	SBT
Animal	<u>1.00</u>	0.03	0.20	0.15	0.15	0.03	0.51	0.73	0.13	0.96	0.80
	<u>0.75</u>	0.02	0.35	0.20	0.12	0.03	0.41	0.60	0.13	0.72	0.67
Car11	<u>1.00</u>	0.00	0.60	0.80	0.28	0.69	0.39	0.54	0.18	0.95	0.93
	0.88	0.03	0.56	0.69	0.32	0.62	0.36	0.45	0.20	0.79	0.84
FootBall	0.84	0.19	0.66	0.34	<u>0.92</u>	0.75	0.02	0.43	0.74	0.36	0.23
	<u>0.65</u>	0.34	0.55	0.31	<u>0.70</u>	0.57	0.07	0.50	0.59	0.34	0.22
Girl	<u>0.96</u>	0.06	0.69	0.69	0.58	0.21	0.40	0.78	0.30	0.95	0.56
	<u>0.73</u>	0.31	0.59	0.54	0.47	0.17	0.34	0.58	0.41	<u>0.73</u>	0.49
MountainBike	<u>0.93</u>	0.53	0.28	0.31	0.14	0.91	0.57	0.26	0.58	0.92	0.59
	<u>0.74</u>	0.41	0.23	0.49	0.13	0.73	0.47	0.20	0.46	0.64	0.45
Shaking	<u>0.88</u>	0.03	0.12	0.02	0.09	0.01	0.15	0.40	0.23	0.01	0.08
	<u>0.72</u>	0.14	0.11	0.04	0.09	0.03	0.21	0.39	0.43	0.01	0.26
Subway	<u>0.99</u>	0.63	0.23	0.32	0.51	0.06	0.25	0.23	0.81	0.22	0.38
	<u>0.74</u>	0.52	0.18	0.42	0.45	0.07	0.19	0.19	0.66	0.17	0.29
Woman	<u>0.93</u>	0.16	0.26	0.09	0.18	0.19	0.12	0.17	0.19	0.62	0.37
	<u>0.75</u>	0.13	0.24	0.25	0.14	0.15	0.10	0.13	0.16	0.49	0.35
mean	<u>0.94</u>	0.20	0.38	0.34	0.36	0.37	0.30	0.44	0.40	0.62	0.49
	<u>0.74</u>	0.24	0.35	0.36	0.30	0.30	0.27	0.38	0.38	0.49	0.45

注:划线为最好结果,斜体加粗为次好结果。

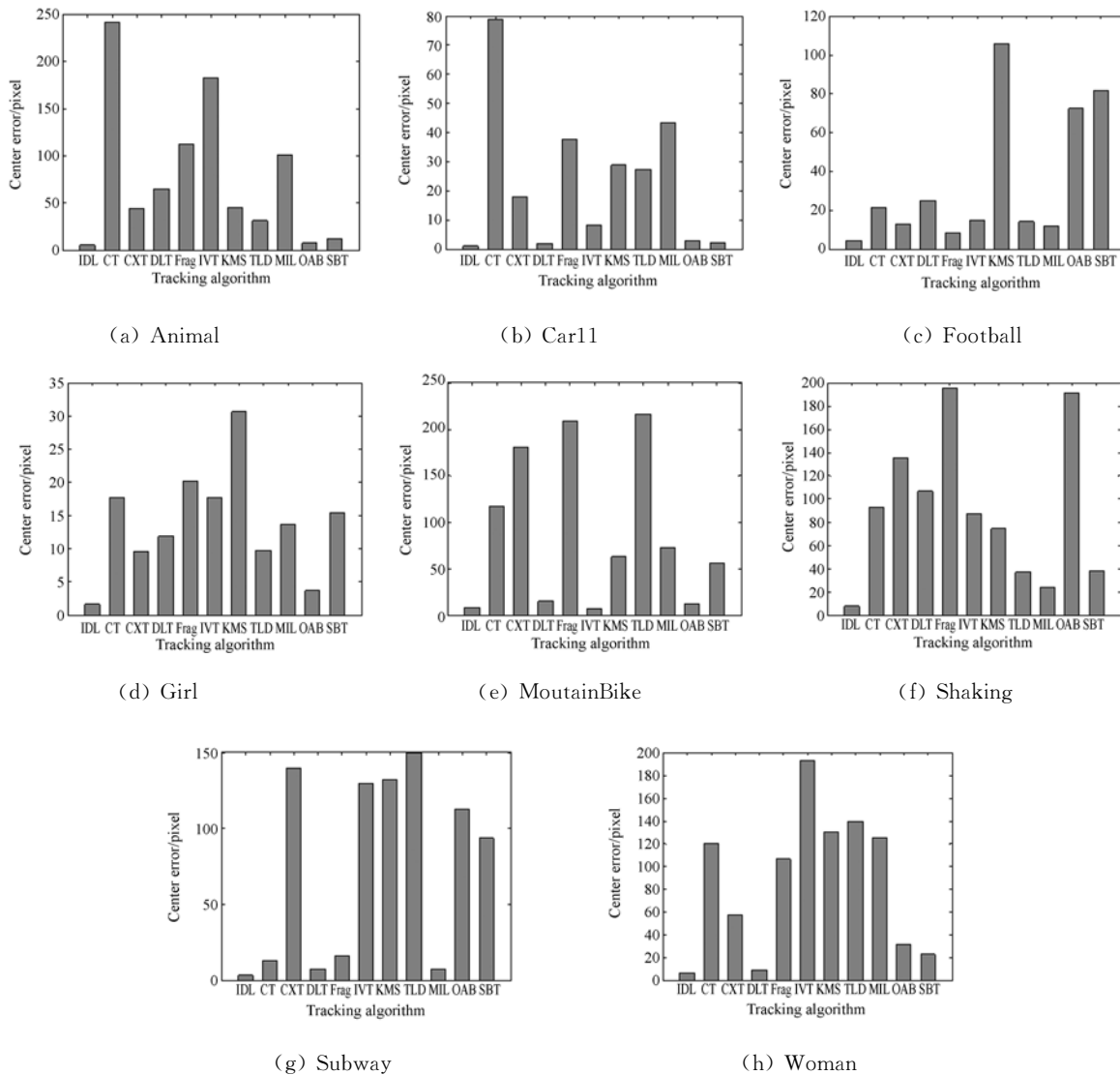


图 5 中心误差柱状图

Fig. 5 Histograms of central errors

表 2 给出了目标跟踪过程中不同因素的影响,目标所处的环境不同,会导致跟踪算法在不同序列的处理速度上有所差异,表 4 为跟踪算法对 8 组图片序列的处理速度。

由表 4 可得,IDL 算法的最快处理速度、最慢处理速度、平均处理速度依次为 20.9、8.6、13.2 frame/s。由于光照、遮挡等因素导致图片序列中目标外观频繁改变,IDL 算法需要不断调整 SDAE 网络参数、添加新特征和整合相似特征,故大大增加了计算成本。反之,在目标外观变化平缓的图片序列中,减少了参数调整过程,IDL 算法计算成本有所增加,再通过采用并行计算与图像处理单元加速技术提高处理速度。由表 4 可知,

IDL 算法的处理速度仅次于 CT 和 DLT 算法。IDL 算法与 DLT 算法均采用并行计算和图形处理单元加速技术,相比与 DLT 算法,IDL 算法中预训练 SDAE、构建正负样本数据库、训练线性 SVM 分类器及微调 SDAE 网络参数都是在离线和初始化阶段完成时,这些环节都不影响在线跟踪过程中算法的处理速度,多出的计算时间大多用在优化特征集合,并且线性 SVM 分类器的分类与微调过程的时间复杂度也高于 sigmoid 函数。CT 算法的 3 种处理速度明显高于其他算法,这是因为 CT 算法是基于压缩感知理论,利用符合压缩感知条件的随机感知矩阵对多尺度高维图像特征进行降维,降维后的特征依然可以高概

率地重构出高维特征,很大程度上降低了特征提取的时间复杂度,而且分类过程仅采用简单的朴素贝叶斯分类器,从而提高了算法的运算速度。

素贝叶斯分类器,从而提高了算法的运算速度。

表4 算法处理速度对比

Tab.4 Comparison of processing speeds for several algorithms

速度/ (frame · s ⁻¹)	IDL	CT	CXT	DLT	Frag	IVT	KMS	TLD	MIL	OAB	SBT
最快	20.9	47.6	16.2	25.5	7.8	8.3	11.3	18.9	15.7	15.3	10.4
最慢	8.6	21.1	2.2	10.8	1.5	6.1	1.6	2.7	7.2	2.3	1.9
平均	13.2	30.2	8.8	15.8	3.8	7.5	3.1	9.9	10.7	8.9	7.1

图6(彩图见期刊电子版)为11种跟踪算法在8组图片序列的跟踪结果。

在Animal序列中跟踪目标是快速运动的小鹿,在跟踪过程中伴有运动模糊和相似背景的干扰。在第11、14 frame,除IDL算法外,DLT、CXT算法均出现目标漂移现象,而其他算法都已跟踪失败。在第25 frame中出现了由快速运动引起的模糊拖影等现象,只有IDL能够正确跟踪。在相似背景的干扰下,如第30、50 frame,只有IDL和OAB算法能正确跟踪目标,可见,在快速运动、模糊拖影及相似背景环境下,IDL算法能够实现稳定的目标跟踪。

在Car11序列中小车在行驶过程中光照不断变化,如第169 frame,导致CT、MIL、Frag算法出现目标漂移,而且在跟踪后期无法正确跟踪。第369 frame,由于光线不足,只有IDL、DLT算法能够稳定跟踪,其他算法均出现不同程度的漂移甚至跟踪失败。由此可知IDL算法可有效处理光照变化,实现稳定的跟踪。

在Football序列中运动员之间常发生碰撞和遮挡后分离等现象。OAB、KMS算法在跟踪前期就已跟踪失败,其他算法在运动员之间发生多次碰撞和遮挡后依然可以完成目标跟踪,但会发生不同程度的漂移,IDL算法能够正确跟踪。可见,IDL算法在碰撞和遮挡条件下可以实现鲁棒的跟踪。

在Girl序列中女孩头部不断出现旋转、姿态变化、部分遮挡和尺度变化。由于目标旋转、倾斜和摇摆导致大部分算法出现目标偏移问题,如第109、237、247、329、332 frame。在第434 frame,出现人脸的部分遮挡时,DLT算法目标偏移严

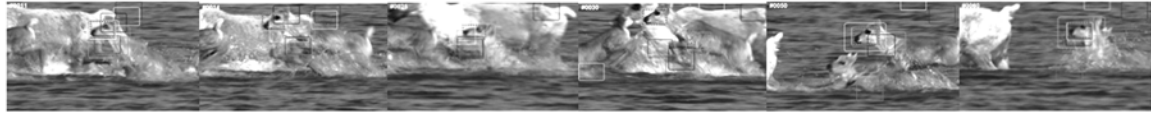
重,Frag、MIL和CT算法甚至跟踪失败,只有IDL算法能够正确的跟踪。因此IDL算法对目标旋转、部分遮挡和尺度变化具有很好的鲁棒性。

在MountainBike序列中车手在空中运动时背景、姿态和尺度都发生变化。CT、CXT、Frag、SBT算法在跟踪前期就已经失败,IDL算法跟踪效果最好,其他算法虽然都能完成整个过程的跟踪,但均有不同程度的漂移。结果表明IDL算法在动态背景等条件下的跟踪效果最好。

在Shaking序列中跟踪目标频繁晃动而且所处环境灯光急剧变化。在第35、58、60、61 frame光照突然变化时大部分算法都已跟踪失败,如OAB、TLD、CXT算法。这是因为光照突然变化会导致学习样本不能及时更新,缺乏先验知识。头部频繁摇晃,如第132、194 frame,只有IDL算法能进行稳定跟踪,其他算法都已跟踪失败。这表明IDL算法在光照突变和频繁晃动的情况下具有较强的鲁棒性。

Subway序列目标在行走过程中不断被其他行人遮挡,导致部分算法出现错误跟踪,如第33、49、75、99 frame中的IVT、OAB、CXT算法。在后续帧中只有IDL、CT、DLT算法能完成整个序列的跟踪,但DLT出现严重的漂移问题。

在Woman序列中行走的目标总是被车辆部分遮挡,且尺度和观测视角不断变化。目标被遮挡后再次出现在视场中,如第298 frame,只有IDL、DLT算法能跟正确跟踪,第560 frame观测角度和尺度都发生了变化,CXT、DLT、OAB算法虽完成了跟踪但没有适应尺度变化,IDL算法在部分遮挡、尺度和观测角度发生变化时,具有较好的跟踪稳定性。



(a) Animal



(b) Car11



(c) Football



(d) Girl



(e) MountainBike



(f) Shaking



(g) Subway



(h) Woman

■ ■ ■ IDL — CT — CXT — DLT — Frag — IVT — KMS — TLD — MIL — OAB — SBT

图 6 不同算法的跟踪结果

Fig. 6 Comparison of different tracking algorithms

6 结 论

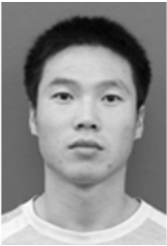
本文提出一种以双重采样粒子滤波为框架的增量深度学习跟踪算法。该方法利用深度学习实现图像的有效表达,并将离线训练得到的信息应用到在线跟踪过程。利用增量特征学习和 SVM 分类器的微调过程适应跟踪过程中目标的外观变化。将深度学习、增量学习、SVM 相结合,并应用到目标跟踪中。

参考文献:

- [1] 郭敬明,何昕,魏仲慧. 基于在线支持向量机的 Mean Shift 彩色图像跟踪[J]. 液晶与显示,2014,29(1): 120-128.
GUO J M, HE X, WEI ZH H. New mean shift tracking for color image based on online support vector machine [J]. *Chinese Journal of Liquid Crystals and Displays*, 2014,29(1): 120-128. (in Chinese)
- [2] WU Y, LIM J, YANG M H. Online object tracking: A benchmark [C]. *2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Portland, 2013: 2411-2418.
- [3] ROSS D A, LIM J, LIN R S, et al. Incremental learning for robust visual tracking [J]. *International Journal of Computer Vision*, 2008, 77(1-3): 125-141.
- [4] 李静宇,王延杰. 基于子空间的目标跟踪算法研究[J]. 液晶与显示,2014,29(4): 617-622.
LI J Y, WANG Y J. Subspace based target tracking algorithm [J]. *Chinese Journal of Liquid Crystals and Displays*, 2014, 29(4): 617-622. (in Chinese)
- [5] BABENKO B, YANG M H, BELONGIE S. Robust object tracking with online multiple instance learning [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(8): 1619-1632.
- [6] 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 D CH, ZHU M, GAO W, et al.. Real-time object tracking via online weighted multiple instance learning [J]. *Opt. Precision Eng.*, 2014, 22(6): 1661-1667. (in Chinese)
- [7] GRABNER H, GRABNER M, BISCHOF H. Real-time tracking via on-line boosting [C]. *Proceedings of British Machine Vision Conference, Edinburgh, Scotland: BMVA*, 2006: 47-56.
- [8] GRABNER H, LEISTNER C, BISCHOF H. Semi-supervised on-line boosting for robust tracking[C]. *Proceedings of European Conference on Computer Vision, Berlin, Germany: Springer*, 2008: 234-247.
- [9] 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 J, WU M Y. On-line boosting based target tracking under occlusion[J]. *Opt. Precision Eng.*, 2012, 20(2): 439-446. (in Chinese)
- [10] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(7): 1409-1422.
- [11] WANG N Y, YEUNG D Y. Learning a deep compact image representation for visual tracking [C]. *Proceedings of Twenty Seventh Annual Conference on Neural Information Processing Systems, Lake Tahoe, USA: Nevada*, 2013: 5-10.
- [12] 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: 3371-3408.
- [13] BURGESS C J C. A tutorial on support vector machines for pattern recognition[J]. *Data Mining and Knowledge Discovery*, 1998, 2(2): 121-167.
- [14] ROB H ALAN F. Discriminatively trained particle filters for complex multi-object tracking [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL: IEEE*, 2009: 240-247.

- [15] 李天成, 孙树栋. 采用双重采样的移动机器人 Monte Carlo 定位方法[J]. 自动化学报, 2010, 36(9): 1279-1286.
LI T CH, SUN SH D. Double-resampling based Monte Carlo localization for mobile robot[J]. *Acta Automatica Sinica*, 2010, 36(9): 1279-1286. (in Chinese)
- [16] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. *Science*, 2006, 313(5786): 504-507.
- [17] ZHOU G Y, SOHN K, LEE H. Online incremental feature learning with denoising autoencoders [J]. *Journal of Machine Learning Research-Proceedings Track*, 2012: 1453-1461.
- [18] TORRALBA A, FERGUS R, 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.
- [19] ZHANG K, ZHANG L, YANG M H. Real-time compressive tracking [C]. *Proceedings of European Conference on Computer Vision, Florence, Italy, 2012*, 3:864-877.
- [20] DINH T B, VO N, MEDION G. Context tracker: Exploring supporters and distracters in unconstrained environments [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI; IEEE, 2011*: 1177-1184.
- [21] ADAM A, RIVLIN E, SHIMSHONI I. Robust fragments-based tracking using the integral histogram[C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, New York, USA; IEEE, 2006*: 798-805.
- [22] COMANICIU D, RAMESH V, MEER P. Kernel based object tracking[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(5): 564-577.

作者简介:



程 帅(1987—),男,内蒙古呼伦贝尔市人,博士研究生,2009年于长春工业大学获得学士学位,2012年于长春理工大学获得硕士学位,主要从事图像处理、目标跟踪、深度学习等方面的研究。
E-mail: chengshuai_pd@126.com

导师简介:



孙俊喜(1971—),男,河北唐山人,博士,教授,2004年于上海交通大学生物医学工程专业获得工学博士学位,2007年于中国科学院长春光学精密机械与物理研究所光学工程专业完成博士后研究,主要从事快速运动目标的检测与跟踪、嵌入式车牌识别系统、医学图像处理与分析。E-mail:junxi_sun@126.com

(版权所有 未经许可 不得转载)