



DOI: 10.12086/oe.2018.180030

基于动态捕获区域的 DC-TLD 目标跟踪算法

何俊衡, 刘 曙, 狄红卫*

暨南大学理工学院光电工程系, 广东 广州 510632



摘要: 为提升 TLD 目标跟踪算法的处理速度, 以达到在更高分辨率视频中跟踪目标的实时性要求, 在 TLD 算法框架的基础上, 提出了一种基于动态捕获区域的 TLD 目标跟踪算法(DC-TLD)。算法采用前一帧目标位置作为当前帧目标位置的预测值, 减小了目标位置的预测误差。研究了检测区域负样本出现需满足的条件, 分析了检测区域大小对算法鲁棒性的影响。针对样本的访问方式, 提出基于索引的访问方法, 极大地减少了访问时间。实验结果表明, 该方法不仅有效降低了 TLD 算法的样本检测时间, 而且提高了算法的鲁棒性。

关键词: 目标跟踪; TLD; 动态捕获区域; 负样本; 索引访问

中图分类号: TP394.1; TH691.9

文献标志码: A

引用格式: 何俊衡, 刘曙, 狄红卫. 基于动态捕获区域的 DC-TLD 目标跟踪算法[J]. 光电工程, 2018, 45(8): 180030

TLD target tracking algorithm based on dynamic capture

He Junheng, Liu Shu, Di Hongwei*

Department of Optoelectronic Engineering, College of Science & Engineering, Jinan University, Guangzhou, Guangdong 510632, China

Abstract: In order to improve the ability of processing video image with high frame rate and meet the real-time requirement in higher resolution video tracking goal, DC-TLD target tracking algorithm, which is based on TLD algorithm and the theory of capturing sample dynamically, is proposed. Firstly, DC-TLD takes the target location of the previous frame as the predicted value of the current frame. Secondly, it calculates the threshold r of negative samples for ensuring that the number of negative samples is sufficient. Thirdly, it accesses the samples by index. The results of experiments show that DC-TLD is more robust and more efficient.

Keywords: target tracking; tracking-learning-detection; dynamic capture; negative sample; index access

Citation: He J H, Liu S, Di H W. TLD target tracking algorithm based on dynamic capture[J]. *Opto-Electronic Engineering*, 2018, 45(8): 180030

1 引言

基于计算机视觉的目标跟踪技术被广泛应用于民

用领域和军事领域^[1], 如交通监控、安防监控、无人机跟踪和人机交互等。然而, 目标跟踪在实际的应用

收稿日期: 2018-01-18; 收到修改稿日期: 2018-04-29

基金项目: 广东省科技计划项目(2013B010204055)

作者简介: 何俊衡(1994-), 男, 硕士研究生, 主要从事目标跟踪和图像处理的研究。E-mail: hjh381@163.com

通信作者: 狄红卫(1969-), 男, 博士, 教授, 主要从事图像与视频处理、目标检测和跟踪的研究。E-mail: tdhw@jnu.edu.cn

中依然面临诸多的挑战,如目标外观变化、高速移动、遮挡消失和算法实时性等,使得大多数目标跟踪算法不能够有效地应用到实际场景中。

近年来目标跟踪算法发展迅速。目前主流的目标跟踪方法可分为三类。第一类是基于相关滤波的目标跟踪算法,如 DCF(tracking with discriminative correlation filter)、KCF(tracking with kernelized correlation filters)^[2]等,这类算法实时性强,但在目标被遮挡、目标形变等复杂情况下,容易丢失跟踪目标。第二类是基于深度学习的目标跟踪方法,如 MDNet(multi-domain convolutional neural networks)^[3]、TCNN(CNNs in a tree structure)^[4]等,这类算法鲁棒性好,跟踪精度高,但由于算法和网络架构通常比较复杂,导致算法速度慢,实时性较差,且现阶段难以利用基于深度学习的目标跟踪算法进行长时间的目标跟踪。第三类是基于生成方法或判别方法的目标跟踪算法,如 TLD(tracking learning detection)^[5]、Struck(structured output tracking with kernels)^[6]等,这类算法综合性能较强,但算法复杂度普遍较高,难以保证跟踪的实时性。考虑到 TLD 算法具有较好的鲁棒性,适合应用到实际跟踪过程中,本文对 TLD 算法进行研究并提出改进和优化。

TLD 目标跟踪算法^[5]是英国萨里大学 Kalal 提出的一种单目标长时间跟踪算法。该算法与传统跟踪算法的显著区别在于把跟踪算法和检测算法相结合来解决被跟踪目标在被跟踪过程中发生的形变、部分遮挡和快速移动等问题。同时,通过在线学习模块不断更新跟踪模块和检测模块的目标模型及相关参数,从而使得跟踪效果更加稳定、鲁棒和可靠。但 TLD 目标跟踪算法复杂度高,难以达到目标跟踪的实时效果。本文通过分析 TLD 算法特点,提出基于动态捕获区域的 DC-TLD(TLD target tracking algorithm based on dy-

namic capture)目标跟踪算法。首先通过改变目标位置预测方法,利用前一帧的位置预测目标在当前帧的位置,减少目标样本的检测时间;然后改进样本选取方法,在尽量小的样本选取范围内获取足够的正负样本;最后改善样本访问方法,对样本进行编号,采取索引访问的方式,提高样本的访问速度。上述措施降低了 TLD 目标跟踪算法的复杂度,提高了算法的实时性和鲁棒性。

2 TLD 目标跟踪算法原理

TLD 目标跟踪算法由跟踪器、检测器、学习模型和整合器四个模块组成。算法框架如图 1 所示。在算法初始化时,把视频帧数据流输入机器学习模块中,由该模块构造目标模型作为训练样本,并以此模型更新跟踪器和检测器,同时构造正负样本来更新训练集。在整个视频图像序列中,目标跟踪器和目标检测器是同时进行工作的,并把各自的结果和通过整合器返回的结果输入到机器学习模块中,学习模块通过以上数据更新目标模型,并依据目标模型更新跟踪器和检测器。最后结果通过整合器输出。

2.1 跟踪器

TLD 的跟踪器由光流法改进而来的中值光流法^[7-8]进行目标跟踪。在目标跟踪过程中,目标图像框内均匀撒点,并通过中值过滤筛选出可靠的目标像素点,然后采用前向后向误差法,即通过两次光流法有效地得出成功跟踪的像素点,从而生成成功跟踪目标的矩形框。

2.2 检测器

检测器基于滑动窗口^[9]进行检测。模块采用级联分类器架构。级联分类器由方差分类器、集成分类器和最近邻分类器组成。当通过最近邻分类器的样本和

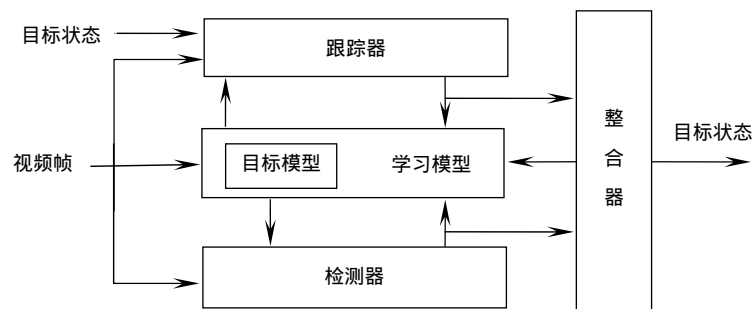


图 1 TLD 算法框架^[5]

Fig. 1 TLD algorithm framework^[5]

目标模型的相对相似度大于一定值时,则判断样本为成功检测目标。成功检测目标将和成功跟踪目标在整合器中进行比较整合,确定当前帧的最优输出结果。

2.3 学习模型

学习模块的核心思想为 P-N(positive-negative)学习^[10]。此模块通过生成正负样本来约束目标检测模块和目标跟踪模块。正样本是指包含跟踪目标的前景样本,负样本是指不包含跟踪目标的背景样本。P 约束寻找视频序列中时域上的结构性特征,并假设目标的运动是按照一定的轨迹进行的。N 约束寻找视频序列中空域的结构特征,并假设目标在一个视频中只能出现在一个位置上。两种约束皆会生成正负样本,学习模块利用新生成的正负样本更新目标模型。

2.4 整合器

整合器接收目标跟踪模块和目标检测模块的结果,通过比较和目标模型的相似度,得出哪个模块的结果更可靠,并以该模块的结果为最终输出结果,同时更新目标模型。

3 基于动态捕获区域的目标跟踪算法(DC-TLD)

在逐帧视频中,跟踪目标在相邻两帧的位移一般比较小。因此,本文提出的 DC-TLD 目标跟踪算法主要针对检测器中跟踪目标在小位移下的情况。DC-TLD 目标跟踪算法先通过利用上一帧成功跟踪目标的位置作为当前帧的目标预测位置中心,并通过目标尺寸计算获取搜索样本区间的边长,结合所得的中心与边长构造正方形搜索区间来获取正负样本,且在获取样本的过程中对样本编号,最后选择符合条件编号的样本输入检测器,算法流程如图 2 所示。当跟踪目标位移较大导致目标丢失时,DC-TLD 目标跟踪算法沿用 TLD 全屏检测的方法重新获取目标位置。

3.1 目标运动状态预测

在 TLD 目标跟踪算法里,每帧的检测样本通过滑动窗口扫描整帧图像得到。这种方式会生成大量多余的样本,使得算法效率偏低。为了提高算法运行速率,一些研究者采用 Kalman 滤波的方法来预测目标在下一帧的位置^[11-12]。本文提出直接利用上一帧成功跟踪到的目标位置作为当前帧的目标预测位置,在当前帧以该目标位置为中心,将一定范围内的图片样本作为当前帧需要检测的样本。

为了比较 Kalman 滤波方法和利用前一帧目标预测当前帧目标位置的效果,本文利用 Visual Tracker Benchmark^[13]中的视频集进行测试。Visual Tracker Benchmark 是 OTB2013 的扩展,也被称为 OTB2015,发表于 2015TPAMI,包含 100 个视频序列,都经过人工标注,包含十一个跟踪难点:光照变化、尺度变化、遮挡、形变、运动模糊、快速运动、平面内旋转、平面外旋转、完全消失、背景杂波(背景相似目标干扰)和低分辨率,是目前使用最广泛最权威的数据库。下文所有测试序列均从 OTB2015 中获取。

图 3 是 BlurFace 测试视频被跟踪目标位置的预测分析。其中纵坐标 $pixels_x$ 和 $pixels_y$ 分别表示跟踪目标在视频图像中 x 轴方向和 y 轴方向上的坐标,曲线 $true_t$ 和 $true_{t-1}$ 分别表示目标在当前帧与前一帧的真实位置,曲线 Kalman 表示用 Kalman 滤波方法预测得到的在当前帧的目标位置,曲线 $error_Kal$ 表示 Kalman 预测的当前帧目标位置与当前帧目标真实位置的误差绝对值, $error_true$ 表示当前帧相对前一帧目标真实位置的偏移绝对值。可以看出,Kalman 算法预判的位置和当前帧目标物体的真实位置有较大偏差。当直接使用前一帧成功跟踪目标所在的位置作为目标当前帧位置的预测值时,其误差(图中曲线 $error_true$ 所示)远小于使用 Kalman 算法预判位置和真实位置的偏差(图中曲线 $error_Kal$ 所示)。在图 4 中的 Jumping

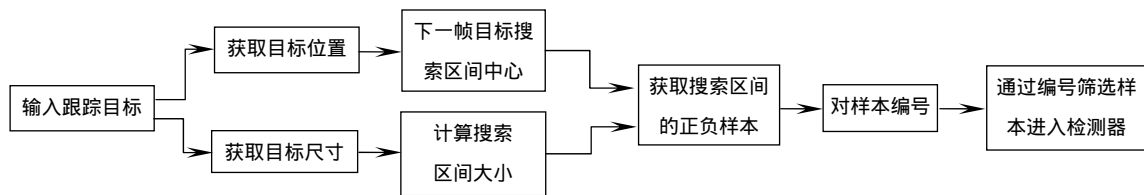


图 2 DC-TLD 算法流程

Fig. 2 DC-TLD algorithm framework

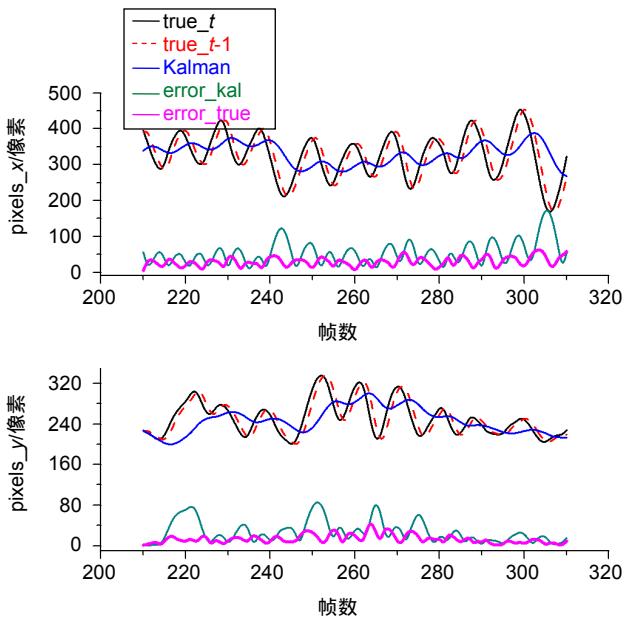


图3 BlurFace 目标位置预测与误差
Fig. 3 BlurFace target location prediction and error

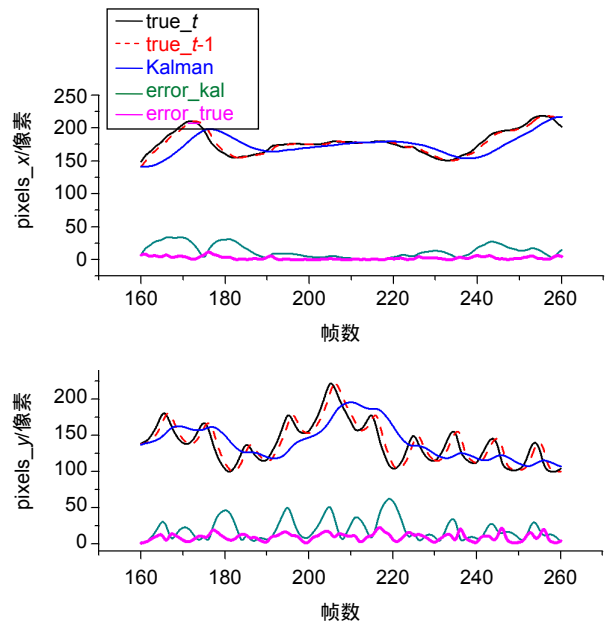


图4 Jumping 目标位置预测与误差
Fig. 4 Jumping target location prediction and error

测试视频中，被跟踪目标主要在 y 轴上往返跳跃，所以在 x 轴上的位移变化相对平缓。综上所述，以前一帧成功跟踪目标所在位置为中心，在当前帧一定范围内获取检测图像样本的方法是可行且有效的。

3.2 基于临界参考点的范围选取方法

确定了下一帧目标检测中心后，再在目标检测中心附近的一定范围内获取样本。当跟踪目标在视频中的相对位置变化比较平稳时，获取样本的范围不是越小越好，而是存在一较佳临界值。当缩小获取样本范围时，算法运行速度会明显提升，但会导致鲁棒性变差。原因在于 TLD 目标跟踪算法是在线学习的更新算法，目标检测模块中的分类器需要正负样本来进行训练，才可以使得检测模块适应跟踪环境的变化。当目标范围选取过小时，即使仍可保证有足够多的样本，但 TLD 目标跟踪算法只会把和目标模型重叠率小于 0.2 的样本标为负样本，来进行相应的学习和更新。因此在较小范围内获取样本，会使得负样本数量不足，使得学习效果变差。下面讨论如何确定样本范围。

设两个面积均为 s 的样本 e 和 e' ，重叠部分面积为 s_0 ，当重叠率为 0.2 时，有：

$$\frac{s_0}{2s - s_0} = 0.2 \quad (1)$$

解得 $s_0 = \frac{1}{3}s$ 。

设样本 e 的边长分别为 a 和 b ，如图 5 实线矩形所示。现在考虑样本 e' 在其左下方的重叠情况，并将样本 e 左下方的顶点作为坐标轴原点建立坐标系。当样本 e' 和样本 e 的重叠率刚好为 0.2 时，样本 e' 的右上角坐标点必然满足方程：

$$xy = \frac{1}{3}ab \quad (2)$$

以样本 e 的右上角顶点 p 作为观察点，来获取以样本 e 为中心的样本。考察选取样本从左下方与样本 e 重叠的情况，令这些样本的右上顶点落在以 p 为中

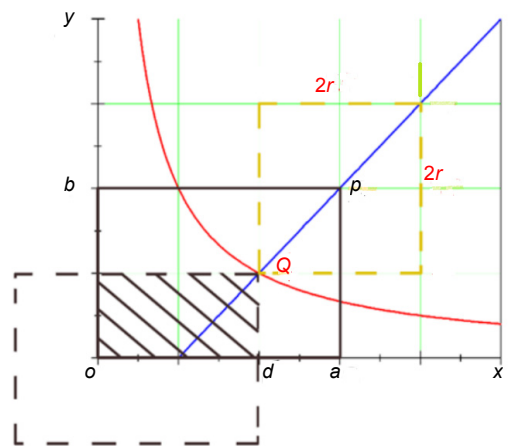


图5 临界点分析图
Fig. 5 Critical point analysis diagram

心点的正方形区域, 边长设为 $2r$, 如图 5 中 $2r \times 2r$ 区域。当正方形区域中有部分像素点在双曲线 $xy = ab/3$ 的左下方时, 则判断为重叠率小于 0.2 的负样本。当正方形区域和双曲线只有一个交点时, 交点必然在过 p 点且斜率为 1 的直线上, 即:

$$y = x - a + b \quad (3)$$

设交点 Q 的横坐标是 d , 则有:

$$r = a - d \quad (4)$$

将 $x = d = a - r$ 代入式(2)、式(3), 可得:

$$r = \frac{a+b}{2} - \frac{1}{2} \sqrt{a^2 + b^2 - \frac{2}{3}ab} \quad (5)$$

因此负样本出现的阈值 r 可根据每一帧中跟踪到的目标区域的长宽(即 a 和 b) 计算得出。

以 r 值为基准, 改变样本检测区域的范围, 设改变后的正方形边长为 $2R$ 。为了便于观察 R 值的大小对算法鲁棒性的影响, 在 Visual Tracker Benchmark 中选取视频 Fish、BlurFace、Jumping 进行测试, 结果如图 6 所示。其中, bad_box 是负样本的数量, success 为算法成功跟踪目标的帧数占测试视频总帧数的百分比, error 是算法跟踪目标失败的帧数占测试视频总帧数的百分比。

由图 6 可知, 从 $R=0.8r$ 开始, 负样本数开始缓慢

增加。当 $R > 0.8r$ 时, 随着 R/r 的增大, 负样本数快速增加, 成功跟踪目标的帧数和跟踪目标失败的帧数呈现出上下波动的状态。当 $R/r=2.9$ 时, 三个视频测试集同时出现了较高的成功率。

定义评估算法性能的品质因数 Q :

$$Q = \frac{E_s}{(E_f+1)(\sqrt{D_s}+1)(\sqrt{D_f}+1)} \quad (6)$$

式中: E_s 和 $\sqrt{D_s}$ 是跟踪目标成功样本的期望和标准差, E_f 和 $\sqrt{D_f}$ 是跟踪目标失败样本的期望和标准差, 为了避免出现分母等于零的情况, 把分母各项因式加 1, 把三个视频测试集结果代入式(6)中, 计算出品质因数 Q 和 R/r 的关系, 如图 7 所示。

由图 7 可得, R/r 数值的大小和品质因数 Q 的关系曲线不是近似单调递增曲线, 当 $R/r=2.9$ 时, 品质因数 Q 出现最大值。

3.3 基于索引访问的样本访问方法

选定检测范围后, 可在该范围内访问和获取检测样本。TLD 算法每帧涉及到数万个样本, 对每一个待检测样本都需要通过计算重叠率来判断样本是否在设定的样本选取范围内, 这无疑大大增加了算法的运行时间。本文根据索引访问分析的方法来改进算法,

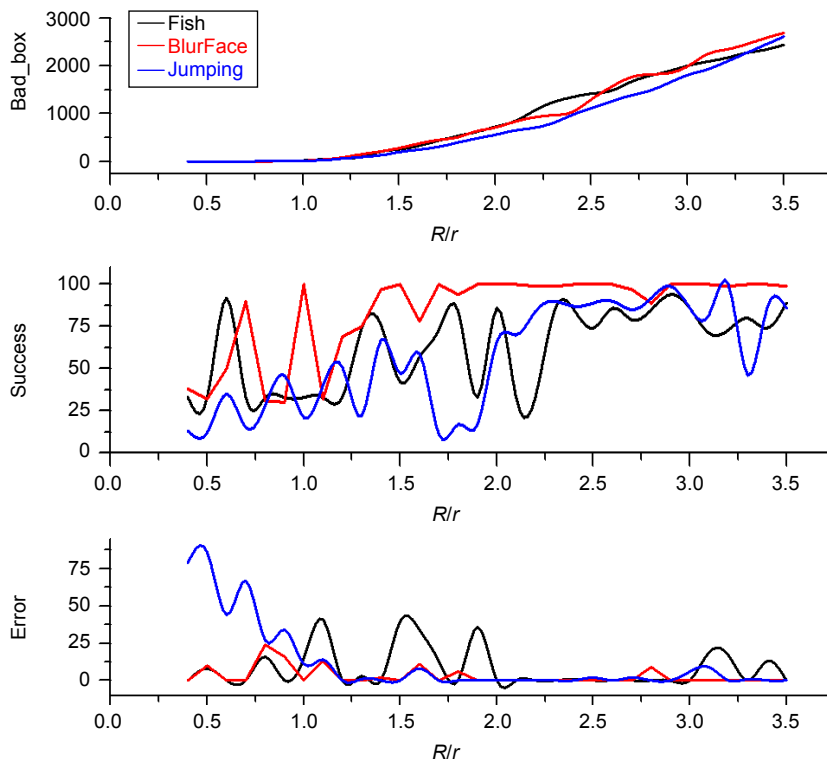


图 6 R/r 和跟踪成功率、失败率及负样本数的关系

Fig. 6 R/r and tracking success rate, failure rate, negative sample size relationship

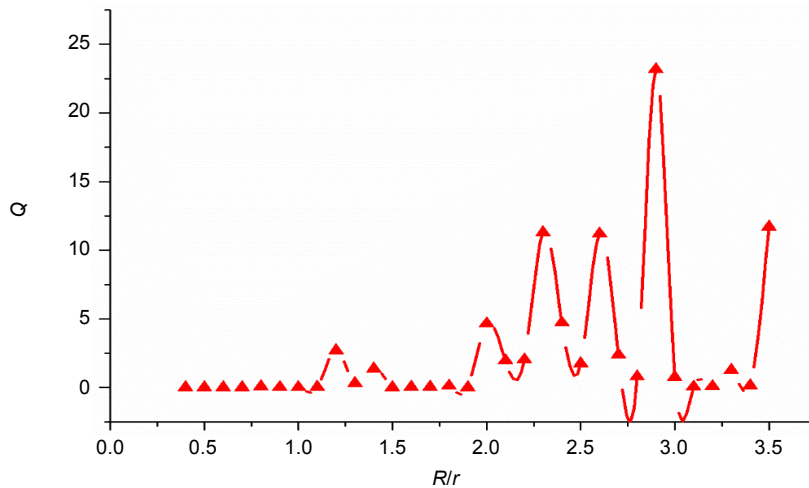


图 7 R/r 和品质因数 Q 的关系
Fig. 7 R/r and Q relations

计算效率得到较大的提高。

在 TLD 算法初始化阶段,会创建样本集合。本文利用算法的该特点,若每个样本都在集合中,可通过对样本进行编号,利用样本编号作为下标进行访问。样本选取时,滑动窗口每次移动的步长是样本大小的 0.1 倍。为了对样本进行编号,本文在算法中添加一个结构体,用以记录样本信息,如样本每一尺度的长宽、每一尺度的起始编号、滑动步长像素大小等。然后把结构体内的信息结合前一帧目标中心点的位置信息,通过计算可以直接得出相应的样本编号。DC-TLD 算法则通过编号作为下标去访问集合中的样本。

为了避免每个样本都要利用计算重叠率的方式来判断样本是否为在设定的选取范围内,且考虑到当待检测样本和当前帧目标中心样本相距甚远时,它们的重叠率也会很低。因此,在算法的循环中,结合样本编号,只需要选取和当前帧目标中心样本相近的几个尺度的待检测样本即可。

运用基于索引访问的方法,选取连续的 7 个尺度作为检测尺度。利用 Crowds、Football、Human6、

Walking 视频测试集对两种方法进行测试,表 1 是测试结果。

其中加速比 γ 是重叠率方法和索引访问方法选取单位样本的时间对比,即:

$$\gamma = \frac{N_i / T_i}{N_o / T_o} \quad (7)$$

式中: N_i 是索引访问选取样本数, T_i 是索引访问耗时, N_o 是重叠率选取样本数, T_o 是重叠率耗时。

由表 1 可得,基于索引访问方法可以有效提高算法效率。

4 实验结果和数据分析

本文根据以上分析提出基于动态捕获检测区域的 TLD 目标跟踪算法(DC-TLD)。将前一帧跟踪到的目标位置作为当前帧的检测范围的中心位置,以 $R=2.9r$ 作为获取样本范围,使用基于索引访问的样本访问方法进行目标跟踪。

仿真实验的软件环境: VisualStudio 2012, OpenCV 2.4.9, Windows 10 系统; 硬件平台: Inter(R) Core(TM)

表 1 重叠率方法和索引访问方法运行时间对比

Table 1 Overlap ratio method and index access method running time comparison

测试集	重叠率选取样本数 N_o	索引访问选取样本数 N_i	重叠率耗时 T_o/ms	索引访问耗时 T_i/ms	加速比 γ
Crowds	2523	5278	2.79	0.0282	207
Football	3069	4054	2.87	0.0221	172
Human6	2861	7608	3.15	0.0597	140
Walking	3849	14000	3.55	0.0741	174

i5-3470 CPU @ 3.20 GHz, 内存 : 8.00 GB。实验共用了 6 个测试集对 TLD 及 DC-TLD 算法进行测试。

4.1 跟踪时间

以 Crowds 和 BlurFace 为测试视频, 把 TLD 和 DC-TLD 进行运行对比。得到数据后, 为便于观察采用了 Savitzky-Golay 滤波的方式对数据图像进行处理。最终结果如图 8、图 9 所示。

观察图 8 和图 9, 由于 Crowds 视频测试集中第 50 帧前后帧和 BlurFace 视频测试集中第 151 帧前后帧中, 目标特征发生了较大变化, 所以两种算法曲线该处都出现了一个较大波动。但波动前后, 两种算法在每帧的运行速度稳定在一个范围内, 所以两种算法跟踪目标较为平稳。

分别在 Crowds、Football、Jumping、Walking、Fish、BlurFace 视频测试集运行 TLD 算法和 DC-TLD 算法,

测试跟踪目标所需时间, 表 2 是测试结果。

由表 2 可得, 改进后的 DC-TLD 目标跟踪算法起到明显的提速效果。相对于其他测试序列, Crowds 中的跟踪目标比其他测试序列的跟踪目标都要小一到两倍, 算法对 Crowds 的提速效果最明显。因此对于跟踪较小目标而言, DC-TLD 算法提速效果显著。

4.2 跟踪精确度

图 10 为 Football、Crowds、Walking、Jumping、Fish、BlurFace 测试序列的部分截图, 图 10 中黄色框为本文改进算法的跟踪目标框, 蓝色框为原始 TLD 算法跟踪到的目标框。改进后的 DC-TLD 算法在 TLD 算法跟踪失败的帧里没有出现跟踪失败的情况。

表 3 是测试视频精确度的结果, 以成功跟踪目标帧数为算法精确度的度量。当把 R 值设定为 $2.9r$ 时, DC-TLD 成功跟踪目标的帧数比 TLD 有所提升。

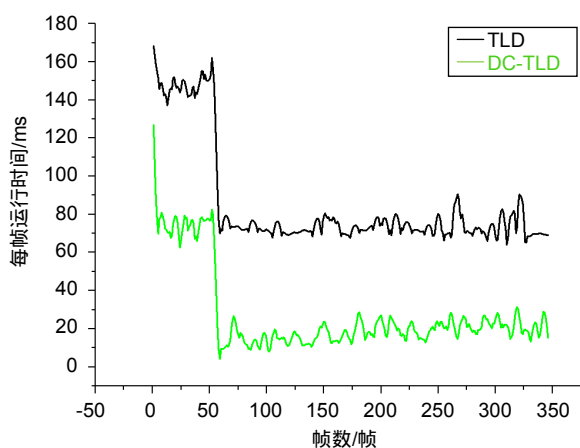


图 8 Crowds 测试集算法每帧运行时间

Fig. 8 Crowds test set algorithm running time per frame

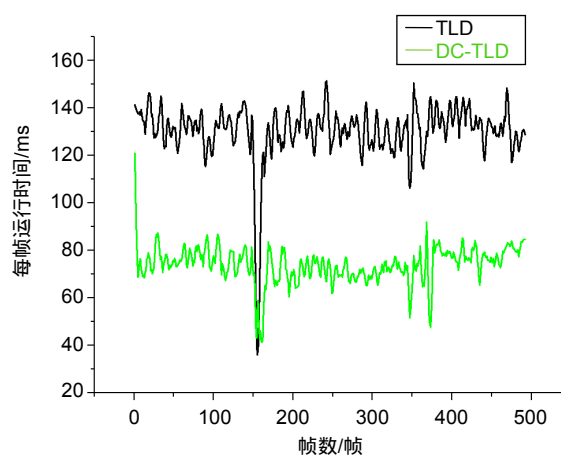


图 9 BlurFace 测试集算法每帧运行时间

Fig. 9 BlurFace test set algorithm running time per frame

表 2 视频测试实验跟踪时间结果

Table 2 Results of video test tracking time

测试集	视频帧数	TLD/ms	DC-TLD/ms
Crowds	347	85.42	27.47
Football	312	104.37	55.28
Jumping	313	45.16	20.34
Walking	412	149.63	76.47
Fish	475	71.22	34.43
BlurFace	493	130.30	73.85

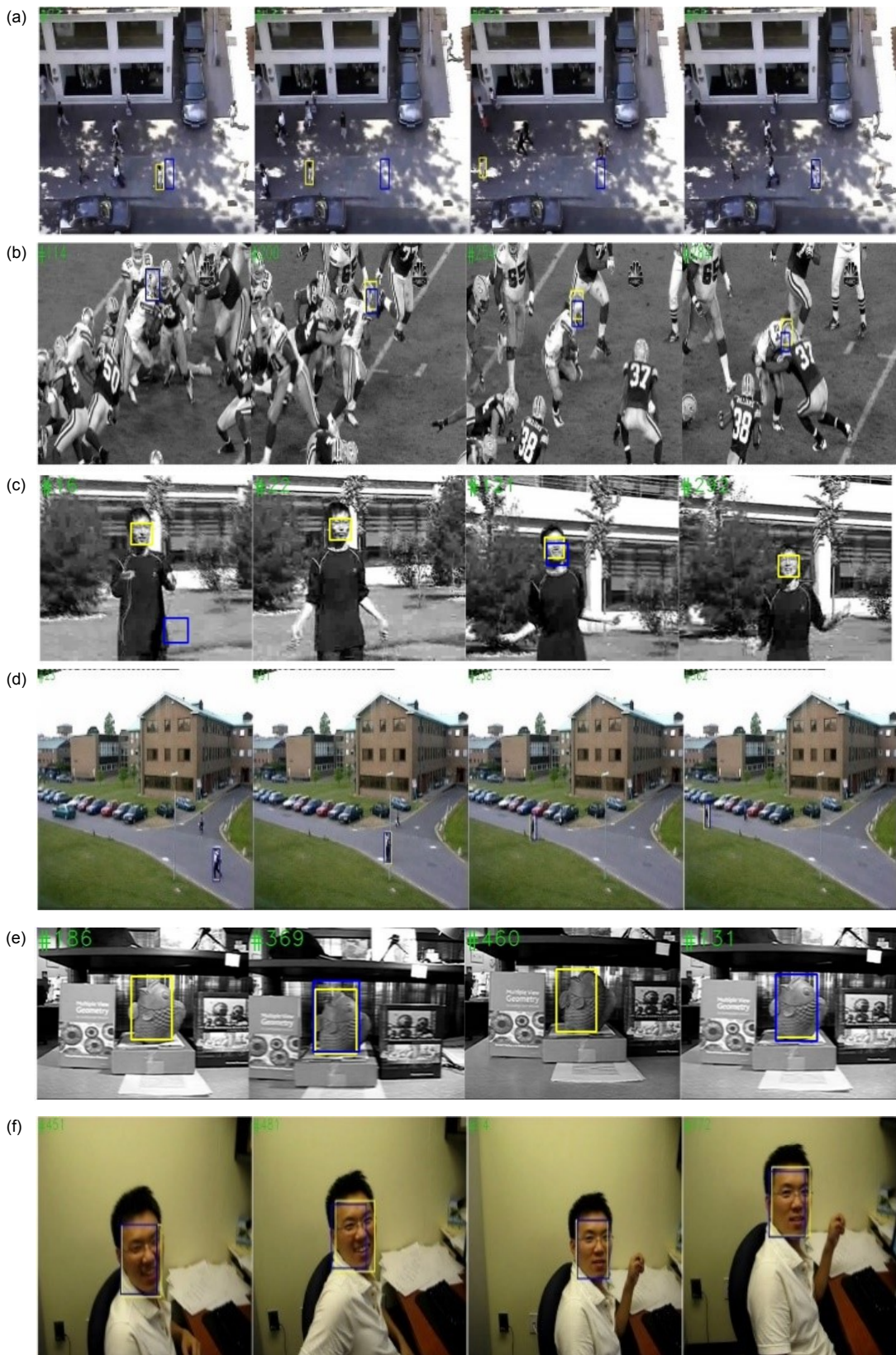


图 10 视频序列测试图。(a) Football; (b) Crowds; (c) Walking; (d) Jumping; (e) Fish; (f) BlurFace
Fig. 10 Video sequence test diagram. (a) Football ; (b) Crowds ; (c) Walking ; (d) Jumping ; (e) Fish ; (f) BlurFace

表 3 测试视频精确度
Table 3 Test video accuracy

测试集	视频帧数	TLD 成功跟踪帧数	DC-TLD 成功跟踪帧数
Crowds	347	76	232
Football	312	286	288
Jumping	313	217	309
Walking	412	412	412
Fish	475	399	448
BlurFace	493	493	493

5 结 论

本文对 TLD 目标跟踪算法中的目标检测区域选取方法进行了研究，提出了基于动态捕获区域的 DC-TLD 目标跟踪算法。首先采用前一帧成功跟踪的目标位置作为目标当前帧位置的预测值。然后通过分析得到负样本出现的阈值 r ，以 r 为基准，改变样本检测区域范围，得出 $R=2.9r$ 时，DC-TLD 算法的鲁棒性最好。最后针对检测样本的访问方式，提出了基于索引的访问方法。实验结果表明，提出的 DC-TLD 算法减少了时间复杂度，提高了跟踪精度。

DC-TLD 算法仍存在不足之处，当跟踪目标较大时，DC-TLD 算法和 TLD 算法的跟踪效果相近，且暂时无法对多目标进行跟踪。

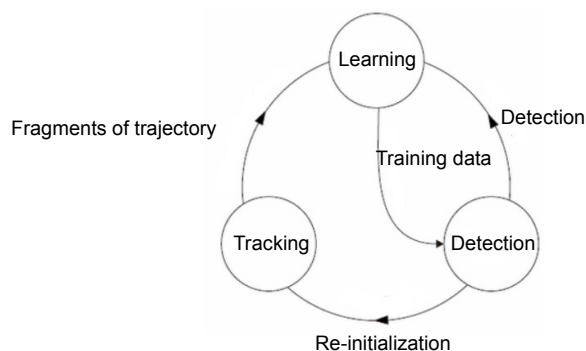
参考文献

- [1] Cheng W L, Wang X J, Wan Z J, et al. Research and implementation of target tracking algorithm in compression domain on miniaturized DSP platform[J]. *Opto-Electronic Engineering*, 2017, **44**(10): 972–982.
程卫亮, 王向军, 万子敬, 等. 压缩域目标跟踪算法在小型化 DSP 平台上的研究与实现[J]. *光电工程*, 2017, **44**(10): 972–982.
- [2] Henriques J F, Rui C, Martins P, et al. High-speed tracking with kernelized correlation filters[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(3): 583–596.
- [3] Nam H, Han B. Learning multi-domain convolutional neural networks for visual tracking[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 4293–4302.
- [4] Nam H, Baek M, Han B. Modeling and propagating CNNs in a tree structure for visual tracking[EB/OL]. <https://arxiv.org/abs/1608.07242v1>, 2016.
- [5] Kalal Z, Mikolajczyk K, Matas J. Tracking-learning-detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(7): 1409–1422.
- [6] Hare S, Saffari A, Torr P H S. Struck: structured output tracking with kernels[C]//*Proceedings of 2011 IEEE International Conference on Computer Vision*. Barcelona, Spain: IEEE, 2012: 263–270.
- [7] Brox T, Bruhn A, Papenberg N, et al. High accuracy optical flow estimation based on a theory for warping[C]//*Proceedings of 2004 Computer Vision-ECCV*. Berlin, Heidelberg: Springer, 2004: 25–36.
- [8] Kalal Z, Mikolajczyk K, Matas J. Forward-backward error: automatic detection of tracking failures[C]//*Proceedings of 2010 20th International Conference on Pattern Recognition*. Istanbul, Turkey: IEEE, 2010: 2756–2759.
- [9] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]//*Proceedings of 2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Kauai: IEEE, 2003.
- [10] Kalal Z, Matas J, Mikolajczyk K. P-N learning: bootstrapping binary classifiers by structural constraints[C]//*Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco: IEEE, 2014: 49–56.
- [11] Jiang B. Research on TLD algorithm based on Kalman filter[D]. Xi'an: Xi'an University of Science and Technology, 2013.
江博. 基于 Kalman 的 TLD 目标跟踪算法研究[D]. 西安: 西安科技大学, 2013.
- [12] Wang M J, Wu X J. New IMM smoothing algorithm suitable for nonlinear maneuvering target tracking[J]. *Opto-Electronic Engineering*, 2016, **43**(10): 18–24.
王美健, 吴小俊. 适于非线性机动目标跟踪的新 IMM 平滑算法[J]. *光电工程*, 2016, **43**(10): 18–24.
- [13] Wu Y, Lim J, Yang M H. Object tracking benchmark[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(9): 1834–1848.

TLD target tracking algorithm based on dynamic capture

He Junheng, Liu Shu, Di Hongwei*

Department of Optoelectronic Engineering, College of Science & Engineering,
Jinan University, Guangzhou, Guangdong 510632, China



The TLD algorithm framework

Overview: The target tracking technology based on computer vision is widely used in civil and military fields, such as traffic monitoring, security monitoring, uav tracking and human-computer interaction. However, target tracking is hardly applied to the actual scene. The main reasons are as followed. To begin with, target may suffer from deformation, illumination variation and background clutter during the tracking, which will make the tracking system lose the target. What's more, target may be vague due to the fast motion. Last but not least, target may be blocked by something. In addition, many target tracking algorithms are too complicated to complete real-time tracking.

The tracking-learning-detection algorithm is a new single-target long time tracking algorithm proposed by Zdenek Kalal. The tracking-learning-detection (TLD) algorithm is different with the algorithm which is based on the conventional tracking algorithm. The TLD algorithm combines the tracking algorithm and the detection algorithm to solve the problem about the shape change, fast moving and partial shade of the target during the tracking. At the same time, the target model and related parameters of tracking module and detection module are constantly updated through online learning module, which makes the tracking effect more stable, robust and reliable. However, TLD target tracking algorithm has high complexity and it is difficult to achieve the real-time effect of target tracking. Because the TLD algorithm has good robustness and is suitable for application to actual tracking process, this paper studies how to improve and optimize the TLD algorithm.

By analyzing the characteristics of the TLD algorithm, this paper proposes DC-TLD target tracking algorithm based on dynamic capture region. Firstly, by changing the target location prediction method, the position of the previous frame is used to predict the position of the target in the current frame, which will reduce the detection time of the target sample. Secondly, the sample selection method is improved to obtain enough positive and negative samples in the small sample selection. Thirdly, the sample access method is improved by numbering the sample, and the access speed of the sample is improved by index access. The above measures successfully reduce the complexity of the TLD target tracking algorithm and improve the robustness of the algorithm, which make the algorithm closer to real-time requirements. However, when the tracking target is large, the tracking effect is similar as DC-TLD algorithm and TLD algorithm. And it is hard for DC-TLD to track multiple targets.

Citation: He J H, Liu S, Di H W. TLD target tracking algorithm based on dynamic capture[J]. *Opto-Electronic Engineering*, 2018, 45(8): 180030

Supported by Guangdong Science and Technology Project (2013B010204055)

*E-mail: tdhw@jnu.edu.cn