上海大学无人艇工程研究院

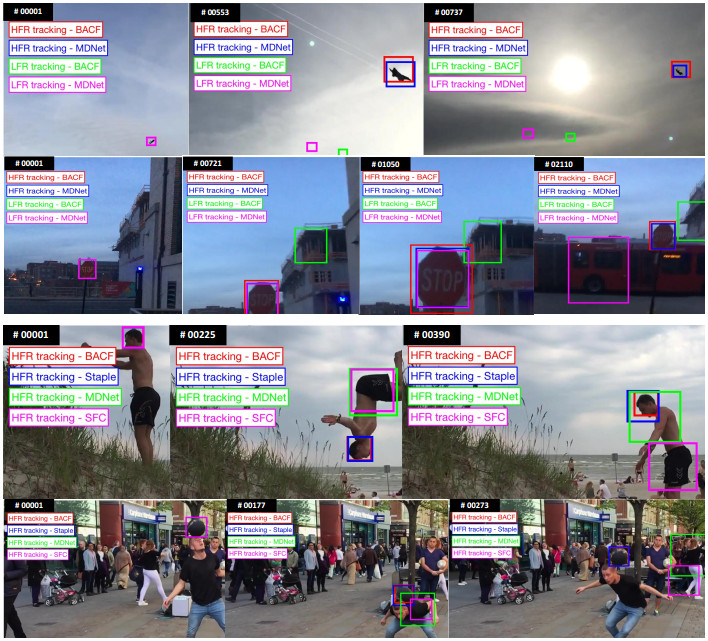
——环境感知组

# Need for Speed: A Benchmark for Higher Frame Rate Object Tracking—NfS

作者：**Hamed Kiani Galoogahi** Ashton Fagg Chen Huang Deva Ramanan

主页：<http://ci2cv.net/nfs/index.html>

出处：2017年ICCV——数据开源



注：**加粗**的作者为重点关注研究者

图注：本算法的核心示意图

Date：2017.10.28

## 版本更新记录

|  |  |  |  |
| --- | --- | --- | --- |
| **日 期** | **更新人** | **主要更新内容描述** | **版本号** |
| 2017年03月28日 | 陈加宏 | 完成大致框架搭建 | V1.0.0 |
| 2017年10月28日 | 陈加宏 | 完成算法细节的总结 | V1.0.1 |

目 录

[Need for Speed: A Benchmark for Higher Frame Rate Object Tracking—NfS 1](#_Toc496954359)

[版本更新记录 2](#_Toc496954360)

[1、概述 3](#_Toc496954361)

[1.1 前言 3](#_Toc496954362)

[1.2 创新点 3](#_Toc496954363)

[2、相关研究 3](#_Toc496954364)

[2.1 跟踪数据集 3](#_Toc496954365)

[2.2 跟踪算法 4](#_Toc496954366)

[3、实验 5](#_Toc496954367)

[3.1 数据集 5](#_Toc496954368)

[3.2 评估方法 5](#_Toc496954369)

[3.3 更新/学习系数 6](#_Toc496954370)

[3.4 跟踪效果 6](#_Toc496954371)

[4、总结 6](#_Toc496954372)

## 1、概述

在传统跟踪中定义的跟踪实时性是在25FPS，目前提出的大量数据集以及跟踪算法都是按照这个跟踪实时性标准研究的，但是目前硬件采集设备的图像帧率已经大幅度提升，所以高帧率数据集对于今后的跟踪算法的研究具有显著的意义。

### 1.1 前言

截止2017年，视觉目标跟踪在相关滤波和深度学习技术快速发展的大背景下跟踪精度和速度都有了明显的进步，很多算法可以在保证精度的前提下做到实时跟踪，但是这里的实时指的的传统的25到30FPS，然而目前的大多数图像采集设备，包括智能相机、平板电脑相机、无人机、机器人，其采集的视频帧率达到了240FPS，在这样的硬件设备快速发展的背景下，重新定义视觉目标跟踪实时性显得很有必要。

在相关滤波大幅度的提升跟踪速度以及深度学习大幅度的提升跟踪精度的前提下，同时保持跟踪速度和精度是理想跟踪算法的基本需求，深度学习之所以可以得到很好的跟踪精度得益于其强大的特征表示能力以及判别能力。视觉跟踪任务的本质是跟踪器可以很好的应对目标在视频各帧中形态的变化，在视频帧率提高之后，意味着视频图像的前后帧目标的变化更小，这样的话，不使用深度学习这样需要大量数据以及需要巨大计算量的方法就可以得到很好的跟踪效果。

结合上述两个技术背景，在一些硬件处理能力有限的应用场合可以在高帧率相机的支持下使用简单的跟踪算法就可以同时保持跟踪的精度和实时性。这就是本文的出发点。本文提出了240FPS的视频数据集以及跟踪评估平台，用于研究高帧率图像序列的跟踪问题。

### 1.2 创新点

本文的创新点

1、收集了240FPS高帧率图像数据集，并且设计了一个对应的跟踪评估平台用于评估各种跟踪算法的跟踪性能；

2、在本文的跟踪评估平台上测试目前优秀的CF类和DL类跟踪算法；

3、实验结果表明高帧率视频下相对简单有效的跟踪算法，比如CF类算法可以表现的更好，这样就可以获得同时满足高速和高精度的要求；

## 2、相关研究

该部分主要讲述的是本文的研究基础，包括对跟踪数据集以及优秀跟踪算法的研究，其中的跟踪算法包括了目前最流行最实用的CF类和DL类算法。

### 2.1 跟踪数据集

OTB50和OTB100：这两个数据集是跟踪领域最经典的数据集，也是最早提出的数据集之一，分别包括50和100个跟踪序列，序列中每帧的目标都被bounding boxes标注了出来，同时将所有跟踪序列分成了11个属性，用于评估不同场景下的跟踪效果。目前我们实验室已经完全掌握了这个数据集的使用。

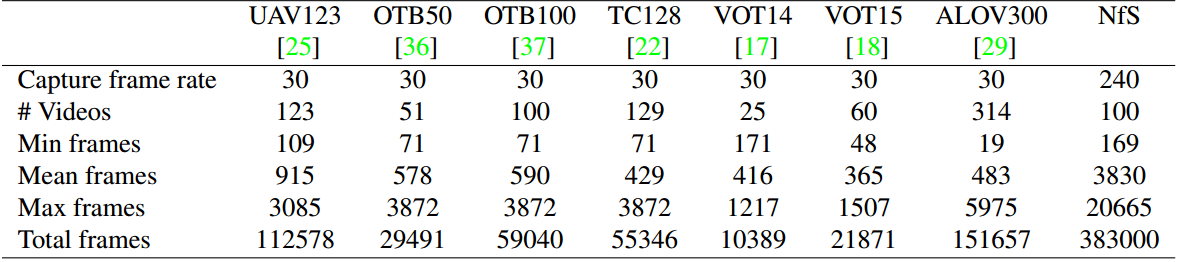
TC128：包括了128个跟踪数据序列，标注方法和序列划分和OTB基本相同，但是主要针对颜色特征的研究；

VOT14和VOT15：分别包含了25和30个跟踪序列，基本都是从OTB中选取的，但是分辨率更高，同时标注了可以旋转的标注框；

ALOV300：包含了314个跟踪序列，都是从OTB、VOT和TC128中选取出来的，所有序列被分成14个跟踪属性；

UAV123：包含了128个跟踪序列，115个是从无人机视野下采集的，剩余的8个序列是由无人机模拟器生成的，所有目标都被标注出来了，同时被分成了12个跟踪属性；

NfS：本文提出的数据集，包括集包含100个视频图像序列，总帧数超过380K帧，数据量是目前最大数据集ALOV300的两倍以上。这些视频图像都是使用帧率为240FPS的高速摄像机在现实世界中拍摄得到的，同时作者还对每帧图像进行了精确的目标标注，给出目标的bounding boxes框的坐标信息，并且还将各个视频数据根据拍摄场景的不同分成了9个属性，它们分别是：遮挡、快速运动、背景杂波等。本文的一大贡献就是在高帧率视频图像下评估各种跟踪算法的性能，评估是按照跟踪的精度和跟踪速度两个方面评估的。文中通过实验得出的结论是相对简单的跟踪器，如相关滤波跟踪算法在高速目标的跟踪效果是好于基于深度方法的跟踪算法的。下表显示了上述跟踪数据集的详细数据信息：



### 2.2 跟踪算法

目前最主流的跟踪算法包含两类，相关滤波跟踪算法和深度学习跟踪算法。

CF类算法：相关滤波跟踪算法在2010年MOSSE提出之后，如春笋般涌现，为了增强跟踪的鲁棒性，使用了更好的特征表示，包括使用了HOG的KCF，结合HOG和颜色直方图的Staple，也有为了增强跟踪尺度自适应的LCT和SAMF算法，为了克服边界效应给跟踪带来的影响，提出剪裁样本的CFLB及其增强的BACF算法，和空间约束的SRDCF。在深度学习席卷视觉各个领域的前提下，基于深度特征的CF类算法也被提出，HCF和HDT直接使用卷积神经网络的卷积特征，但是这样的跟踪算法因为计算复杂度高，跟踪速度较慢。

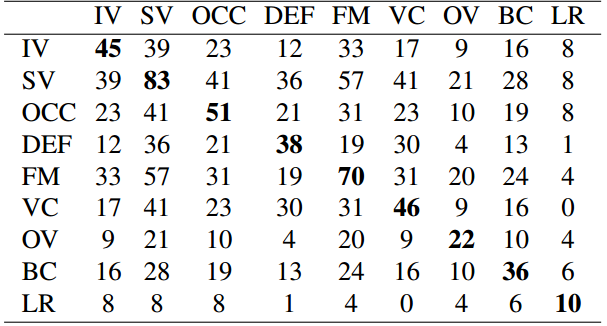
DL类算法：深度学习跟踪算法的实现思路与CF类不同，其核心思路在于如何利用大规模数据集进行适合跟踪领域的训练以及检测，MDNet对于训练的处理将传统的目标检测识别数据集可以预训练网络，在具体跟踪时微调网络以达到很好的跟踪性能，其获得了VOT15的跟踪冠军。之后的FCNT利用全卷积网络完成跟踪的训练检测，但是上述两个方法速度都有很慢，对于硬件计算性能要求较高。目前速度快的DL跟踪算法有GOTURN和SFT，都是基于孪生网络Siamese的，只进行离线训练，没有在线微调的过程，过意跟踪速度极快，可以得到100PFS和75FPS，但是在跟踪精度上有限，在跟踪速度上CPU表现明显不如GPU。

## 3、实验

该部分主要讲述算法实现代码的主要流程、实验环境及效果分析、算法优缺点的总结，最后提出后续可改进的方面。实验是检验真理的唯一标准，那么对实验结果详细的分析以及结合算法的原理对算法本质上的一些思考有利于之后研究工作的开展，也是今后工作的一个研究突破点。

### 3.1 数据集

NfS数据集包括的100个序列中，有75个是使用iphone 6和ipad pro采集的，同时保留了数据集采集过程中的惯导和陀螺仪的传感器信息。剩余的25个数据集是从youtube上采集的，这100个序列都是240FPS的高帧率视频数据。所有序列被分成9个属性。详细信息如下表所示：

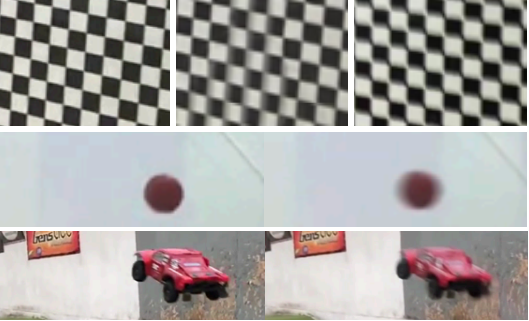


### 3.2 评估方法

本文使用的用于评估的跟踪算法有15个，按使用的特征分成三类：传统特征的CF类BACF/SRDCF/Staple/DSST/KCF/LCT/SAMF/CFLB、深度特征的CF类HCF/HDT以及DL类算法MDNet/SiameseFC/FCNT/GOTURN。还有一个基于SVM的使用传统特征的MEEM。

评估方法是基于重叠区域IOU的，基于IOU对跟踪性能进行排序。同时还提出了一种相对准确度度量方式，是用提升的精度与低帧率精度做比值：

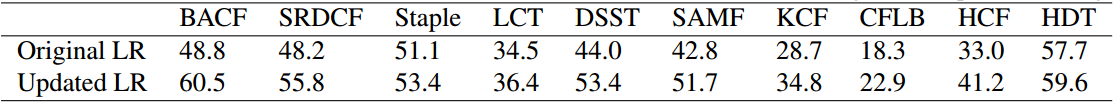
为了对比240FPS高帧率和30FPS低帧率的跟踪性能，需要从240FPS的序列中每8帧采样得到30FPS的视频数据，但是高帧率的视频数据的运动模糊现象很少，所以需要在低帧率的场景下模拟出运动模糊的场景，作者通过视频处理工具AE模拟出运动模糊，具体效果如下图所示



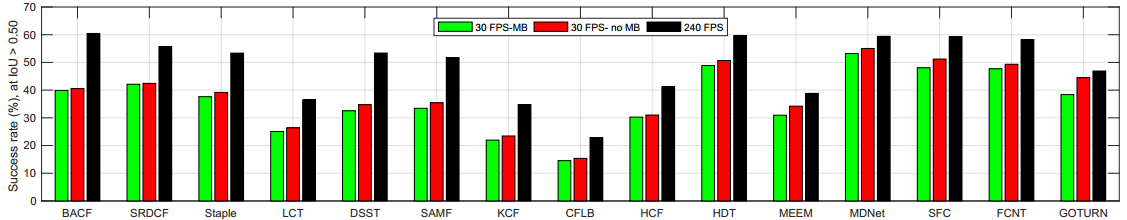
### 3.3 更新/学习系数

为了更好的是原来的跟踪算法的更新频率在高帧率的场景下表现的更好，较小了各个跟踪算法的学习系数，即：

完全可以证明修改后的学习速率可以提高跟踪算法在高帧率下的跟踪效果，从实验结果来看，确实提升明显：

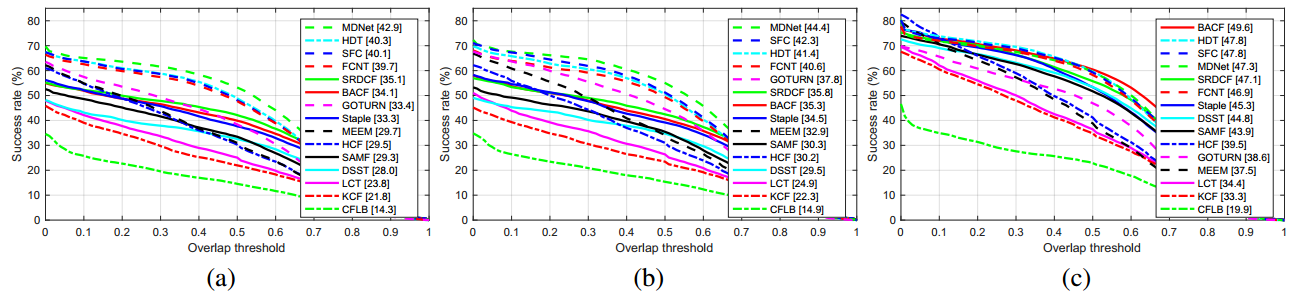


### 3.4 跟踪效果

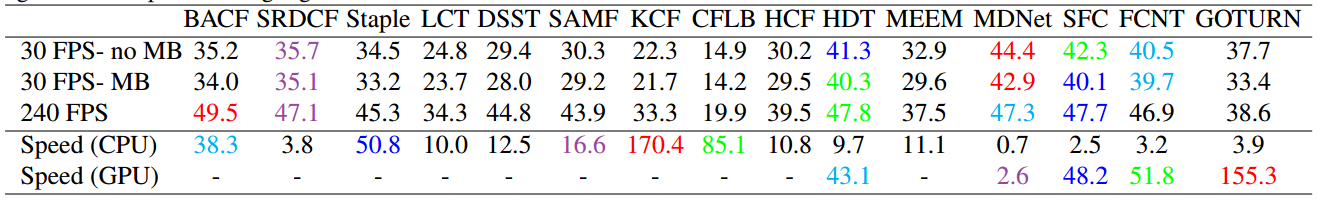


可以从上图中明显的看到更新后的跟踪效果的提升，以及在高帧率下的跟踪效果也提升明显。

下图是整体的跟踪性能对比图：



成功率对比以及跟踪速度的对比图如下：



## 4、总结

以下是这篇论文可以为我所用的几点总结：

1、100个序列的数据集；

2、评估平台；

3、标注工具the VATIC toolbox