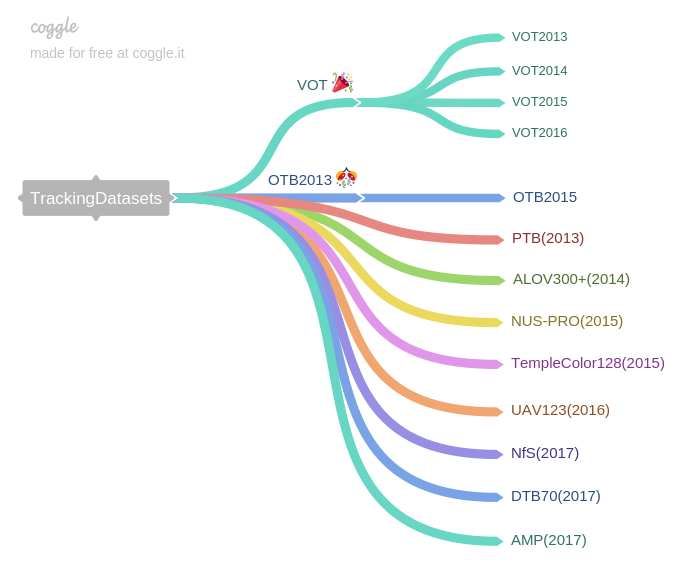
上海大学无人艇工程研究院

——环境感知组

# 视觉跟踪算法之数据测试平台



Date：2017.10.28

注：以下**加粗**的作者为重点关注学者

## 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 | 框架搭建并填充大纲 | V1.0.0 |
| 2017年03月29日 | 陈加宏 | 完成OTB部分的总结 | V1.0.1 |
| 2017年10月28日 | 陈加宏 | 完成NfS部分的总结 | V1.0.2 |
| 2017年10月28日 | 陈加宏 | 完成TC128部分的总结 | V1.0.3 |
| 2017年10月29日 | 陈加宏 | 完成UAV123部分的总结 | V1.0.4 |
| 2017年10月29日 | 陈加宏 | 完成DTB70部分的总结 | V1.0.5 |

目 录

[视觉跟踪算法之数据测试平台 1](#_Toc497038718)

[版本更新记录 2](#_Toc497038719)

[1、OTB 3](#_Toc497038720)

[2、VOT 7](#_Toc497038721)

[3、PTB 8](#_Toc497038722)

[4、ALOV300+ 9](#_Toc497038723)

[5、NUS-PRO 10](#_Toc497038724)

[6、TempleColor128 11](#_Toc497038725)

[7、UAV123 12](#_Toc497038726)

[8、NfS 13](#_Toc497038727)

[9、DTB70 14](#_Toc497038728)

[10、AMP 15](#_Toc497038729)

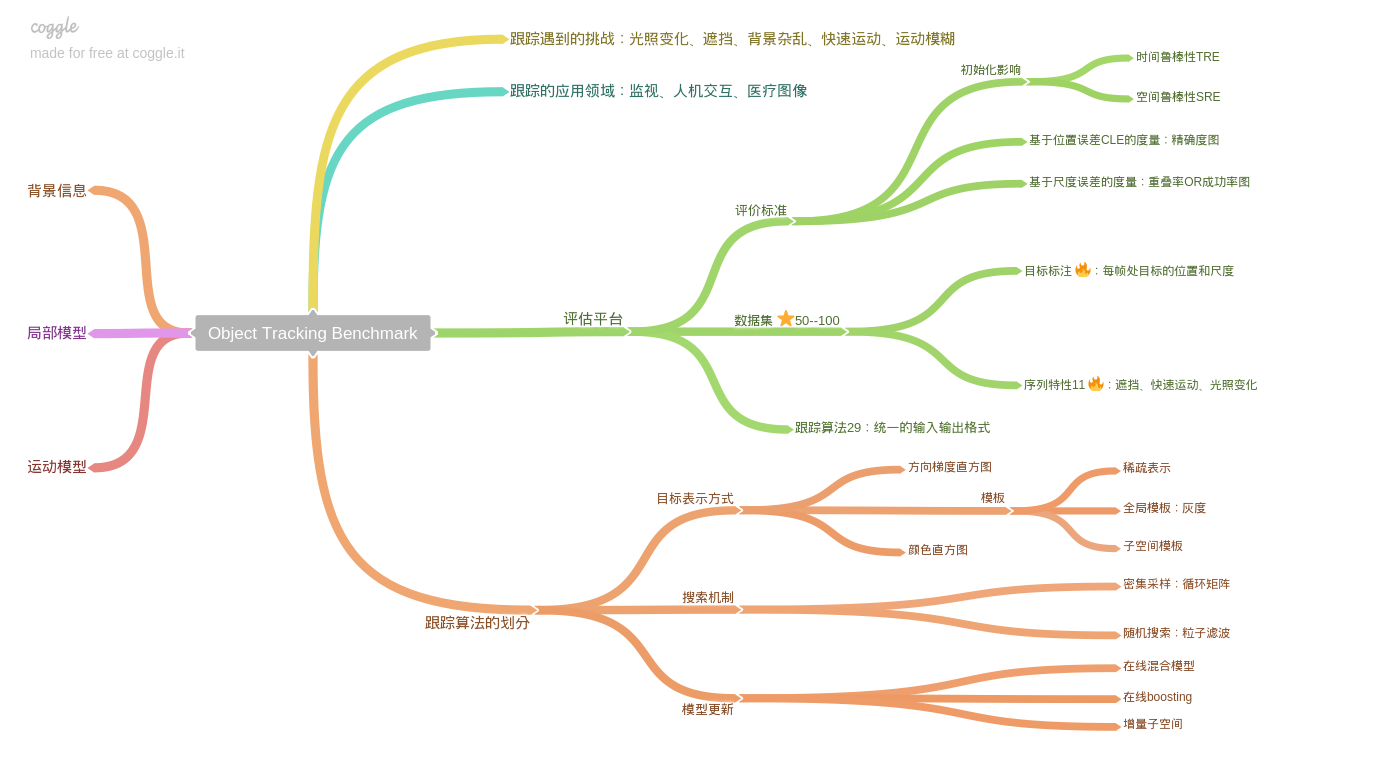
## 1、OTB

论文：Online Object Tracking: A Benchmark

作者：Yi Wu Jongwoo Lim Ming-Hsuan Yang

主页：<http://cvlab.hanyang.ac.kr/tracker_benchmark/index.html>

出处：2013CVPR——数据开源



概述：Object Tracking Benchmark是一个比较各种优秀跟踪器的评价平台。首先需要明确的是：什么是视觉目标跟踪。视觉目标跟踪是在给定图像序列首帧中目标的位置和尺度，能够实现在接下来各帧中确定目标位置和尺度的算法。视觉目标跟踪技术应用广泛，主要包括：人机交互、自主监视、医疗图像等。那么视觉跟踪算法遇到的主要难点在于，目标在视频序列中会遭遇光照变化、遮挡、背景杂乱等问题。

如上所述，为了更好的研究跟踪问题，需要一套库和标准来评估跟踪器的性能，从而能够展示跟踪算法的优点和缺点，可以明确未来的研究方向。这样的评估平台由三个部分组成：数据集、跟踪算法和评估方法。

对于综合性能评估，收集具有代表性的数据集是极其重要的。对收集好的原始数据资料 进行目标的标注，同时需要对视频序列进行特性标注，这样可以更好更有针对性的分析各跟踪算法的优缺点。

原作者公开的跟踪算法在数据输入和输出格式上并不相同，这样，对于算法的评估造成一定的困难，所以一个好的评估平台需要明确和调整跟踪算法的数据输入、输出格式相同。

评价标准从跟踪算法对目标初始化的敏感性出发，分别从空间和时间两个方面验证算法对初始化的鲁棒性。同时从位置误差和尺度重合率两个方面对跟踪算法进行整体性能的评估。

视觉跟踪的主要模块：1、表示方式：全局模板（原始灰度值）、子空间、稀疏表示、颜色直方图、方向梯度直方图、协方差区域描述算子、类哈尔特征、判别模型（机器学习中的分类算法）、分块（解决姿态变化和部分遮挡问题）；2、搜索机制：梯度下降法、密集抽样方法、随机性搜索（如粒子滤波，对局部最小值不敏感）3、模型更新：在线混合模型、在线boosting、增量子空间、上下文信息融合。

下面就平台内所有的**跟踪算法**进行一个统计，主要区分各算法的表示方式、搜索策略、模型更新、实现代码和跟踪速率。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Method** | **Representation** | **Search** | **MU** | **Code** | **FPS** |
| CPF [44] | L, IH | PF | N | C | 109 |
| LOT [43] | L, color | PF | Y | M | 0.70 |
| IVT [47] | H, PCA, GM | PF | Y | MC | 33.4 |
| ASLA [30] | L, SR, GM | PF | Y | MC | 8.5 |
| SCM [65] | L, SR, GM+DM | PF | Y | MC | 0.51 |
| L1APG [10] | H, SR, GM | PF | Y | MC | 2.0 |
| MTT [64] | H, SR, GM | PF | Y | M | 1.0 |
| VTD [33] | H, SPCA, GM | MCMC | Y | MC-E | 5.7 |
| VTS [34] | L, SPCA, GM | MCMC | Y | MC-E | 5.7 |
| LSK [36] | L, SR, GM | LOS | Y | M-E | 5.5 |
| ORIA [58] | H, T, GM | LOS | Y | M | 9.0 |
| DFT [49] | L, T | LOS | Y | M | 13.2 |
| KMS [16] | H, IH | LOS | N | C | 3,159 |
| SMS [13] | H, IH | LOS | N | C | 19.2 |
| VR-V [15] | H, color | LOS | Y | MC | 109 |
| Frag [1] | L, IH | DS | N | C | 6.3 |
| OAB [22] | H, Haar, DM | DS | Y | C | 22.4 |
| SemiT [23] | H, Haar, DM | DS | Y | C | 11.2 |
| BSBT [50] | H, Haar, DM | DS | Y | C | 7.0 |
| MIL [5] | H, Haar, DM | DS | Y | C | 38.1 |
| CT [63] | H, Haar, DM | DS | Y | MC | 64.4 |
| TLD [31] | L, BP, DM | DS | Y | MC | 28.1 |
| Struck [26] | H, Haar, DM | DS | Y | C | 20.2 |
| CSK [27] | H, T, DM | DS | Y | M | 362 |
| CXT [18] | H, BP, DM | DS | Y | C | 15.3 |

表1.1 被评估的跟踪算法

**目标表示**：L: local（局部），H: holistic（整体），T: template（模板），IH: intensity histogram（灰度直方图），BP: binary pattern（二值模式），PCA: principal component analysis（主成分分析），SPCA: sparse PCA（稀疏PCA），SR: sparse representation（稀疏表示），DM: discriminative model（判别模型），GM: generative model（生成模型）；

**搜索机制**：PF: particle filter（粒子滤波），MCMC: Markov Chain Monte Carlo（马尔科夫链蒙特卡洛方法），LOS: local optimum search（局部最优搜索），DS: dense sampling search（密集采样搜索）。

**模型更新**：N: No，Y: Yes。

**实现代码**：M: Matlab，C:C/C++，MC: Matlab和C/C++混合编程， suffix E: executable binary code（可执行二进制代码），MU: model update, FPS: frames per second。

详细的说明一下对视频序列进行特性标注的情况。该平台的视频序列被标注成11个特性，从11个方面来分别对跟踪算法的有效性进行验证。

|  |  |
| --- | --- |
| **属性** | **描述** |
| IV | 光照变化 - 目标区域内的光照剧烈变化 |
| SV | 尺度变化 - 第一帧中和当前帧中的边界框尺寸之比的范围超过[1/ts, ts]， ts >1 (ts=2) |
| OCC | 遮挡 - 目标被部分或全部遮挡 |
| DEF | 形变 - 非刚体目标发生形变 |
| MB | 运动模糊 - 目标或摄像机的运动导致的目标区域变模糊 |
| FM | 快速运动 - ground truth的运动大于tm个像素点(tm=20) |
| IPR | 平面内旋转 - 目标在图像平面内发生旋转 |
| OPR | 平面外旋转 - 目标在图像平面外发生旋转 |
| OV | 超出视野 - 目标的一部分离开视野 |
| BC | 背景杂斑 - 目标附近的背景具有和目标类似的颜色或纹理 |
| LR | 低分辨率 - ground-truth边界框内的像素点个数少于tr(tr =400) |

表2. 测试序列所标注的各种属性的清单。表中也列出了测试中使用的阈值。

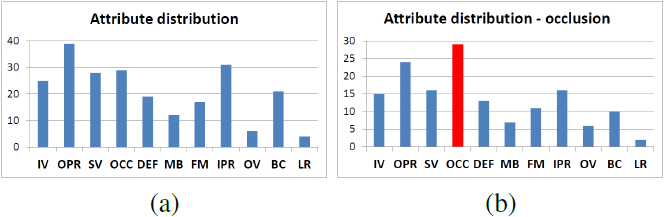


图1.1 (a)为整个测试数据集的属性分布 (b)为遮挡（OCC）特性在序列中的分布。

该平台使用精度和成功率做定量的评估跟踪算法的性能。可视化是通过精确度图和成功率图来实现的。跟踪器的性能评估标准：平均中心位置误差（CLE）、平均成功率（SR）、平均距离精度（DP）。下面详细的叙述：

**精确度图**：**精确度图能够显示出评估的位置在给定的准确值的阈值距离之内的帧数占总帧数的百分比。**在跟踪精度评估中，一个被广泛使用的标准是**中心位置误差**CLE，其被定义为跟踪目标的中心位置和手工标定的准确位置之间的平均欧氏距离。一个序列中所有帧的平均中心位置误差被用于概括跟踪算法对该序列的总体性能。但是，当跟踪器丢失目标时，输出的跟踪位置是随机的，此时的平均误差值可能无法正确估量跟踪的性能。近年来，精确度图已经被用于测量跟踪的整体性能。对于每个跟踪器具有代表性的精度评分，我们使用的分数阈值等于**20个像素点**。

**成功率图**：另一个评估标准是边界框的**重叠率**。为了估量算法在一系列帧中的性能，我们计算重叠率大于给定的阈值的成功帧的数量。成功率图给出了此阈值从0到1变化时成功帧所占的比例SR。使用某一特定阈值（比如0.5）下的一个成功率来评估跟踪器可能并不公平或具有代表性。我们使用每一个**成功率图的曲线下面积（AUC）**作为替代，用于给跟踪算法进行排序。

**鲁棒性评估：**评估跟踪器的传统方式是，根据第一帧中的准确位置进行初始化，然后在一个测试序列中运行算法，最后得出平均精确度或成功率的结果报告。我们把这种方法称为**一次通过的评估（OPE）**。然而跟踪器可能对初始化非常敏感，并且在不同的初试帧给予不同的初始化会使其性能变得更差或更好。因此，我们提出两种方式来评估跟踪器对初始化的鲁棒性，即在时间上（即在不同帧开始跟踪）和空间上（即以不同的边界框开始跟踪）扰乱初始化。这两个测试分别称为**时间鲁棒性评估（TRE）**和**空间鲁棒性评估（SRE）**。所提出的测试场景大部分就存在于现实世界中实际应用的场合，而跟踪器通常通过目标检测器来初始化，检测器在位置或尺寸方面可能会给跟踪器引入初始化误差。另外，在不同时刻的实例中，检测器可能被用于重新初始化跟踪器。通过研究跟踪器在不同鲁棒性评估中的特点，我们可以对跟踪算法进行更为深入的理解和分析。

**时间鲁棒性评估：**给定一个标记了目标准确边界框的初试帧，跟踪器被初始化并运行直到序列结束，即整个序列的一部分。跟踪器会在每一个序列的片段上进行评估且整体的统计数据也会被记录下来。

**空间鲁棒性评估：**我们在第一帧中通过移动或缩放准确的ground truth来抽取初始化的边界框。在这里，我们使用8种空间位置上的偏移，包括4种中心偏移和4种角偏移，以及4种尺度变化（见补充材料）。偏移量为目标尺寸的10%，尺度比例变化可取准确值的0.8、0.9、1.1和1.2。因此，针对SRE我们对每个跟踪器评估12次。

**小结：**本文通过大规模的实验对近年来在线跟踪算法的性能进行了评估。基于本文的评估结果和观察，强调了对提高跟踪性能必需的一些模块。首先，**背景信息**对有效的跟踪是至关重要的，我们可以通过在判别模型中隐含地使用先进的学习技术对背景信息进行编码（如Struck），或显式地用作跟踪上下文信息（如CXT）来利用背景信息。其次，**局部模型对跟踪是非常重要的**，正如局部稀疏表示（如ASLA和SCM）比全局稀疏表示（如MTT和L1APG）更能提升性能。它在目标表观发生部分变化时尤其有用，比如部分遮挡或形变。再者，**运动模型或者说动态模型在目标跟踪中也是至关重要的**，尤其是当目标的运动幅度较大或突然运动时。然而，本文所评估的大部分跟踪器并未注意到这个部分。基于动态模型的位置预测能够减少搜索范围，因此可以提高跟踪的有效性和鲁棒性。改进这些部分将能进一步提高在线目标跟踪算法现有的技术。评估结果展示了近十年来目标跟踪领域所取得的巨大进步。我们提出并演示了从多个方面深入分析跟踪算法的评估标准。这些大量的性能评估有助于更好地理解当前先进的在线目标跟踪方法，也给新算法的测评提供了一个平台。

## 2、VOT

论文：

作者：

主页：

出处：

源代码：

概述：

## 3、PTB

论文：

作者：

主页：

出处：

源代码：

概述：

## 4、ALOV300+

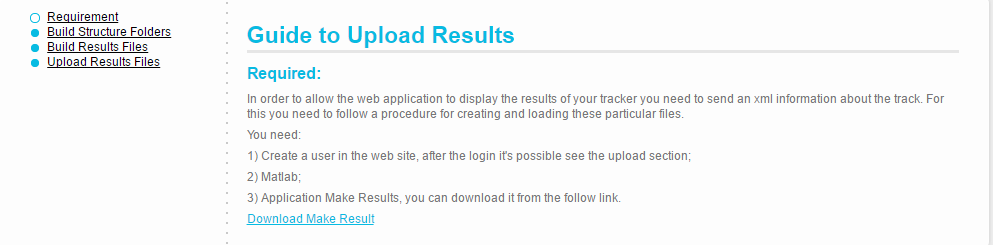
论文：Visual Tracking: An Experimental Survey

作者：Arnold W. M. Smeulders Dung M. Chu Rita Cucchiara

主页：<http://tracking.cs.princeton.edu/>

出处：2014年PAMI

源代码：



概述：

ALOV300旨在涵盖尽可能多的环境：照明，透明度，镜面反应，与类似物体的混淆，杂乱，遮挡，变焦，严重的形状变化，运动模式，低对比度等。在组合ALOV300数据集时，优先考虑许多较短的短视频。在这些方面的每一个方面，我们收集视频序列，从轻松到困难，重视视频难度。ALOV300就在光，反照率，透明度，运动平滑度，混乱等方面向上兼容其他跟踪基准，遮挡和摇动相机方面，还包含了最近的跟踪标准中经常使用的标准跟踪视频序列。数据集由315个视频序列组成。数据的主要来源是来自YouTube的现实视频，具有64种不同类型的目标，从人脸，人物，球，章鱼，微观细胞，塑料袋到罐子。该集合被分为十个方面的困难与许多难以非常硬的视频，如舞者，音乐会的摇滚歌手，完整的透明玻璃，章鱼，群鸟，伪装的士兵，完全遮挡物体和极端缩放的视频引进目标的突然动作。

为了最大化多样性，大多数序列很短。这些方面的平均长度为9.2秒，最长为35秒。一个额外的类别包含十长的视频，持续时间在一到两分钟之间。他们包含来自\ cite {Kalal2010}的三个视频，沙漠中有一个快速移动的摩托车，高速公路上的低对比度记录汽车和追逐车; 来自3DPeS数据集的三个视频\ cite {Baltieri2011}具有不同的照明条件和复杂运动对象; 来自\ cite {Benfold2011}的数据集中的一个视频，监视多个人; 和YouTube的三个复杂视频。

ALOV300中的帧总数为89364. ALOV300中的数据每五帧由沿着主轴方向的灵活尺寸的矩形边界框注释。在极少数情况下，当运动很快时，注释更加频繁。已经通过线性插值获得了中间帧的标准边界框。第一帧中的标准边界框被指定给跟踪器。跟踪器可用的目标特定信息的唯一来源。

（Mirror Link:ALOV300++ Dataset

<http://crcv.ucf.edu/people/phd_students/afshin/ALOV300/Frames.zip>

Mirror Link:ALOV300++ Groundtruth

<http://crcv.ucf.edu/people/phd_students/afshin/ALOV300/GT.zip> ）

## 5、NUS-PRO

论文：NUS-PRO: A New Visual Tracking Challenge

作者：Annan Li Min Lin **Yi Wu Ming-Hsuan Yang** Shuicheng Yan

主页：<https://sites.google.com/site/li00annan/nus-pro>

出处：2015年PAMI——数据开源

<https://drive.google.com/file/d/0B6eYf2Rj8c79bUlLM0pDWURubTQ/view?usp=sharing>

概述：在过去很多目标跟踪的算法已经被推出了，但是大多数的跟踪算法都是在有一定限制的视频序列以及标注上进行评估的。为了全面对跟踪算法进行评估，我们提出了大数据集，这个数据集包含了356个行人以及刚性物体的视频序列。这个数据集包含了12种目标种类，并且大部分的视频序列都是来自于**移动相机采集**的。每一个序列为了评估都标注了目标位置以及遮挡程度。一个对于现如今最先进的20种跟踪算法全方面的评估都通过不同的网格展现出。

（数据集下载：

第一部分：

<https://drive.google.com/folderview?id=0B6eYf2Rj8c79Smk4Q1BxU1ROS28&usp=sharing>

第二部分：

<https://drive.google.com/folderview?id=0BwFzRq8t3gu5VWFRNGp5dlBkSU0&usp=sharing> ）

## 6、TempleColor128

论文：Encoding Color Information for Visual Tracking: Algorithms and Benchmark

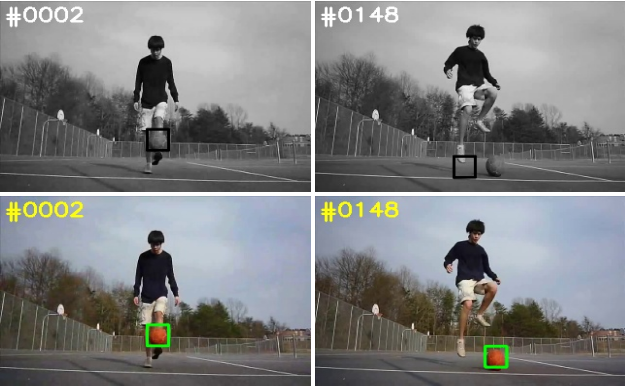
作者：Pengpeng Liang Erik Blasch Haibin Ling

主页：<http://www.dabi.temple.edu/~hbling/data/TColor-128/TColor-128.html>

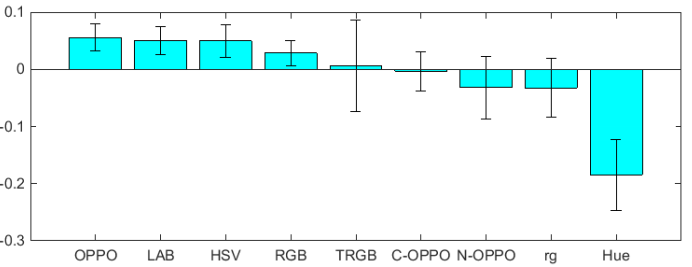
出处：2015年TIP——数据开源

概述：Temple Color 128 - Color Tracking Benchmark

尽管众所周知颜色信息能为视觉判断提供丰富的可识别的线索，但是现如今大多数的视觉跟踪器却只限制在灰度空间。尽管最近在跟踪算法上已经有对合成颜色空间的研究，但是却有一系列对于颜色信息所能起到的作用并不能充分理解，我们针对这个问题，在算法以及评价标准方面，进行了一些列的科学研究。在算法方面，我们精心选择现如今最先进的视觉跟踪器，而后将10种颜色模型拓展到16种跟踪器中。在评价基准方面，我们构建了一个128个彩色序列，这128个彩色序列都标注好了标准边界框以及其所具有的影响因素（例如：遮挡）。**将所有基于色彩信息进行编码的跟踪器与最近提出的两种色彩跟踪器进行全面的评估。**基于RGBD跟踪标准，将会进行这个更深度的评估。这个结果清楚地显示了利用色彩信息进行跟踪的优点。并且我们也在一些特殊的方法上进行细节分析，例如将不同颜色模型以及不同的跟踪器的结合，每一个序列最极端的跟踪情况下，以及对于不同影响因素对于跟踪影响程度。



结论：使用这些颜色特征表示会使跟踪性能有提升：Opponent、HSV、 LAB、RGB。



## 7、UAV123

论文：A Benchmark and Simulator for UAV Tracking

作者：Matthias Mueller Neil Smith Bernard Ghanem

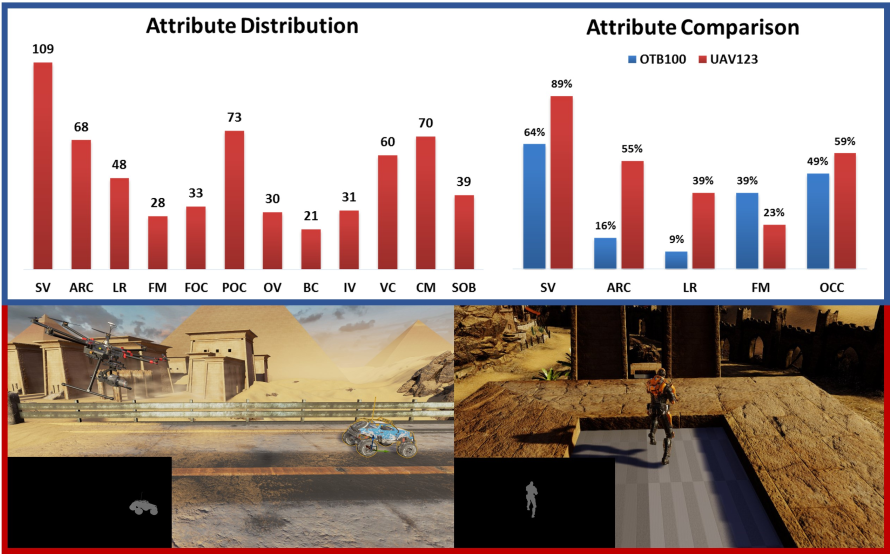
主页：<https://ivul.kaust.edu.sa/Pages/pub-benchmark-simulator-uav.aspx>

出处：ECCV 2016——数据开源

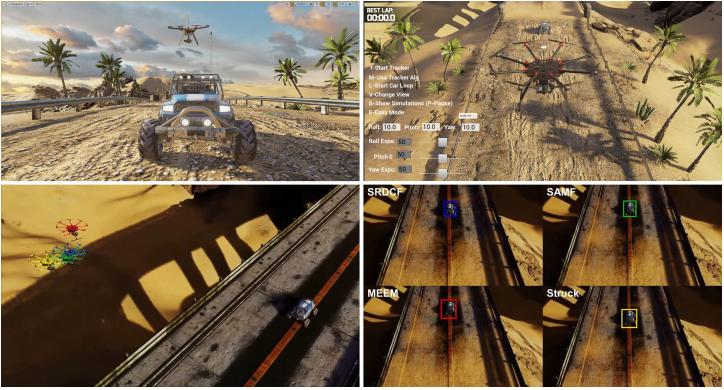
代码开源地址：<https://github.com/IVUL-KAUST/UAVSimUE4>

数据下载地址：<https://ivul.kaust.edu.sa/Pages/Dataset-UAV123.aspx>

概述：本文提出了对于低空无人机目标跟踪一个新的空中视频数据集和基准，以及一个可以处理目标跟踪方法的逼真的UAV模拟器。我们的基准测试将从低空空中视角捕获的123个新的并完全注释的高清视频序列，提供现如今许多最先进和流行的跟踪器，从而进行首次评估。在比较跟踪器中，我们就每个跟踪器追踪精度和运行时间从而确定哪些是最适合无人机跟踪。该模拟器可用于在将其部署在现场的无人机上前在实际情况下评估跟踪算法，以及生成合成的但逼真的带有标准边界框的跟踪数据集，以轻松扩展现有的真实的数据集。



本文最终的贡献在于作者提出了一个模拟视觉跟踪的模拟器，用于更好更直观的研究视觉目标跟踪，这也是我们需要努力的方向——构建一个无人艇的视觉模拟器。



## 8、NfS

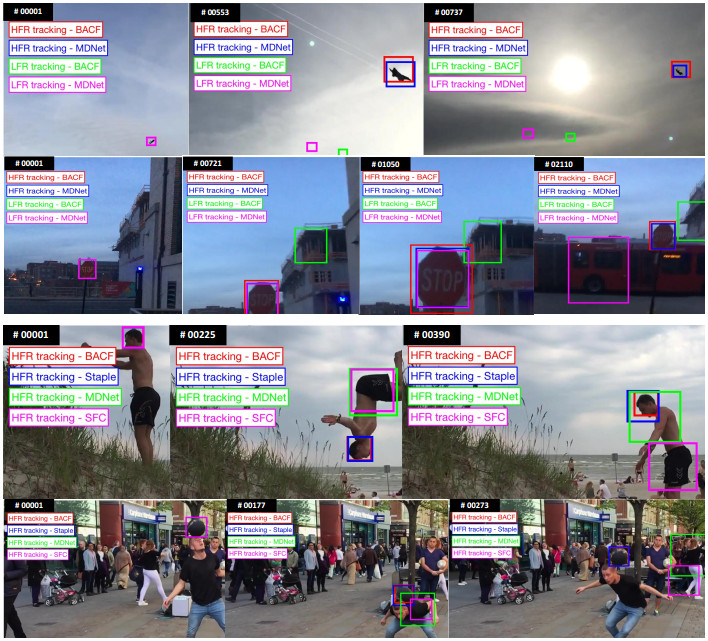
论文：Need for Speed: A Benchmark for Higher Frame Rate Object Tracking

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

概述：本文作者提出了一个帧率更高的跟踪数据集和评价平台。其中的数据集包含100个视频图像序列，总帧数超过380K帧，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。



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

## 9、DTB70

论文：Visual Object Tracking for Unmanned Aerial Vehicles: A Benchmark and New Motion Models.

作者：Siyi Li Dit-Yan Yeung

主页：<https://github.com/flyers/drone-tracking>

出处：2017年AAAI——数据开源

概述：这也是一个关于无人机视野下的视觉目标跟踪数据集，但是与UAV123不同的是这篇论文的重点在于研究跟踪算法的运动模型。作者研究的是在无人机上装有摄像头，这样的应用场景会面临特点的挑战：相机自身会发生运动。作者为了证明基础算法在无人机视野在的跟踪效果，也是为了之后更好的研究无人机上的视觉目标跟踪，提出了一个包含70个视频序列的数据集，同时作者研究了在跟踪算法中加入了运动模型。



全文的创新点有三个：1、构建一个统一的跟踪数据集用于详细的分析跟踪算法的性能；2、设计一个基础算法，将相机的运动模型加入到各种跟踪系统中；3、通过实验证明相机运动模型对于跟踪效果的影响，同时为运动模型之后的研究奠定一个坚实的基础。

## 10、AMP

论文：Beyond Standard Benchmarks: Parameterizing Performance Evaluation in Visual Object Tracking

作者：Luka Čehovin Zajc; Alan Lukežič; Aleš Leonardis; Matej Kristan

主页：

出处：2017年ICCV——暂无开源

源代码：

概述：