

指导教师： 杨涛

提交时间： 20160320

CVPR2015 Paper Translation

No : 01

姓名 : 梁琛

学号 : 2013302600

班号 : 10011306



基于颜色的自由模板的识别追踪

Thomas Mauthner* Horst Bischof Horst Possegger*
格拉茨技术大学 计算机图像视觉学会
 {possegger,mauthner,bischof}@icg.tugraz.at

摘要

在本文中，我们解决的问题是基于颜色特征的对在线自由模型的识别跟踪。根据最近的基准评估的结果，以往的追踪系统往往只追踪到和目标具有相似的外观的区域。为了克服这个限制，我们提出了一个有效的判别对象模型，它可以事先确定潜在的目标区域。此外，我们利用相关知识来事先适应对象，这样可以使干扰抑制和漂流的风险显著降低。我们在跟踪基准数据库中评估我们的方法，评估展示是国家的最先进的结果。特别是，我们的方法和最近的别的跟踪算法相比有更好的精度和鲁棒性。此外，我们所提出的方法可以进行在线的实时目标跟踪。

介绍

视觉目标跟踪是计算机视觉应用中的一项基础性工作。如视觉监视，机器人，人机交互，增强现实等都需要可靠的估计目标在整个图像序列的位置。尽管近几年来取得了重大进展，的创建一个通用的对象跟踪器仍然是相当具有挑战性的，因为现实世界的现象，如光照变化，背景杂波和物体的快速运动都会对追踪产生影响。

虽然在一些应用领域，我们能够假设强大的易于识别的目标（例如行人跟踪[11]、[36]、[37]、[40]），但通常需要建立一个通用的跟踪器可以用于任意的对象追踪。这种自由模板既适用于学习对象模型并且开发特

定类的先验知识。因此，一个具有代表性的对象模型必须用一个给定的一个输入框（可能是嘈杂的）标记初始对象，如一个轴对齐包围框。

在早期的跟踪方法中，颜色直方图法（例如[12]、[34]、[35]）是外观追踪的常用方法。然而，在过去的十年里，研究的重点已经转移到基于特定功能的追踪如 HOG（方向梯度直方图）（例如[13]、[23]），相关滤波器（例如[10]、[22]），和更复杂的颜色特征追踪，如颜色属性[14]，这种跟踪器有极好由基准评价的追踪性能（例如，[27]），而与更复杂的颜色特征相比，基于标准颜色模型的追踪性能较差。

与上一段所说的相反，我们认为，基于标准的颜色的追踪系统仍能达到国家的最先进的性能。我们利用基于颜色的追踪可以观察到相似外表附近地区。再使用自适应对象模型，这样可以有效的干扰抑制，并可以显著地减少漂流的问题，产生强大可靠的追踪结果。由于我们方法具有良好的代表性和简单性，它非常适合应用于关键的时间，如监视和机器人。

我们的贡献如下。我们提出了一个有判别力的对象模型，它能够从背景中区分出目标对象。虽然它依赖于标准的颜色直方图，但在各种具有挑战性的序列的识别中已经达到国家的最先进的性能标准。我们扩展了这个方法，使其能识别和抑制不准确的区域，显著提高了追踪的鲁棒性。此外，我们

提出了一个有效的规模估算方案，使我们能够获得准确的追踪结果，如图 1 所示。最后，我们在最近的基准数据

集中广泛评估我们的方法，结果和各种先进的跟踪器相比，我们的都展示了更良好的性能。

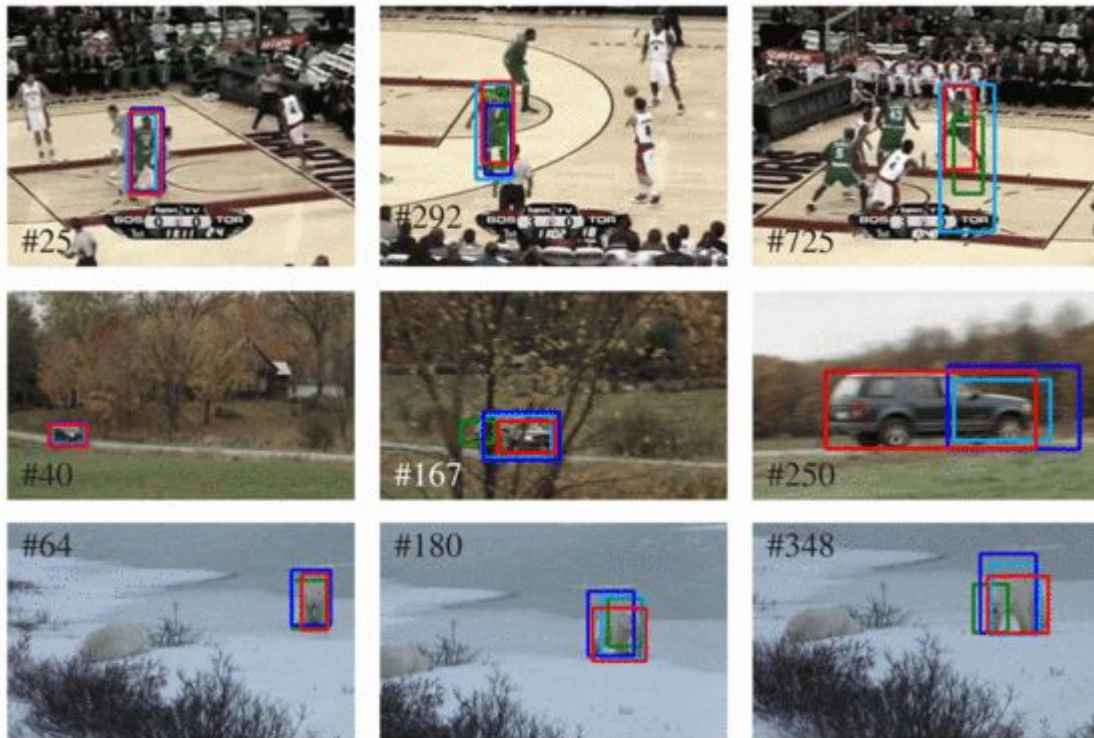


图 1: 比较该方法 ACT,与国家的最先进的跟踪方法, DSST 和 KCF。例如 vot14 下打篮球, 汽车, 和北极熊的序列框。

相关工作

为了模拟对象的外观，追踪方法可以由具有判别力的陈述而定。生成的方法可以产生一个和目标最相似的参考模型。跟踪器通常是基于模板（例如[2]、[20]、[31]），基于颜色表示（例如[12]、[34]、[35]），基于子空间模型（例如[9]、[38]），和基于稀疏表示（例如[6]、[24]、[30]）等。然而，典型的由模型捕获的矩形包围框总是也包含目标的初始化背景信息。虽然有几种利用分割的方法（例如[7]、[17]）提高生成模型的精确度，但这些方法依然不能很好的将对象从周围的背景中分离出来。

最近的基准评估（例如[27]、[44]）表明，利用我们的判别方法生成的模板往往更加优秀，其中二进制分类器能够区分目标和背景。跟踪器利用模板（例如[22]），颜色线索（例如，[8]、[14]、[32]），Haar-like 特征（例如，[4]、[21]、[47]），HOG 特征（如[13]、[23]），和二进制模式（例如[15]、[25]）来识别对象模型的外观。这种模型可以在一个整体的方式表示（例如，[22]、[47]）或部分（例如[42]、[16]）和片面的方式表示（例如[29]、[30]、[32]）已被证明这种方法有利于追踪高度非刚性物体或考虑部分被遮挡的物体。由于判别模型的成功，各种各样的合适的与视觉跟踪相关的研究方法已经被提

出, 如结构化输出向量机算法[21], 排序支持向量机算法[5],boosting 算法, [3], [18], 支持向量回归算法 [23], [22], 和多实例学习算法[4]。

为了进一步提高性能, 有些追踪器可以将上下文的信息整合 (例如, [15]、[19]、[45]、[48])。这种方法可以区分的背景中集中和分散的地区, 如[19], [15]虽然目标有不同的外观, 但这种方法可以找出共同发生的, 有价值的线索, 这样可以克服遮挡问题。另一方面, 有的物体具有和目标相似的外观, 因此可能与目标混淆。通常情况下, 整合上下文信息的追踪器如[45], [48]认为干扰和目标是同一对象类 (例如行人), 则会一直跟踪的最先的目标, 这样可以防止干扰漂移。在这些方法中, 我们没有在对象类上假设干扰。此外, 我们的方法能提前识别潜在的干扰区域从而抑制, 没有明确的跟踪和干扰是必需的。

我们关于自由模板追踪的方法对以两个基本要求为基础: 第一, 要考虑到后续帧, 使用的模型必须能够将对象从当前的背景区分出来。其次, 要减少漂移的风险, 具体表现在能区分类似的外观的目标, 这样的区域必须事先确定, 并且抑制, 以确保强大的追踪性能。因此, 我们提出了判别对象模型, 解决了这些关键的要求。见 3.1。基于这种方法, 3.2 展示示对象可以是整个视频序列。此外, 我们的模型可以高效且大规模的追踪, 将在 3.3 中具体讨论。

3.1. 干扰感知的对象模型

为了从周围的背景像素中区分出对象的像素 $x \in O$, 我们采用了基于颜色直方图的贝叶斯分类器对输入图像进行处理。让 $H_{\Omega}^I(b)$ 表示非归一化直方图,

让 H 表示在区域 b 仓计算的数值 $\Omega \in I$ 。此外, 让 b_x 表示本分配到 b 区域的颜色分量 $I(x)$ 。给定一个矩形区域 O (即假设标注或电流跟踪初始边界框) 及其周边地区 (见图 2a), 我们应用贝叶斯法则得到目标位置可能在:

$$P(x \in O | O, S, b_x) \approx \frac{P(b_x | x \in O) P(x \in O)}{\sum_{\Omega \in \{O, S\}} P(b_x | x \in \Omega) P(x \in \Omega)}. \quad (1)$$

特别是, 我们可能直接从颜色直方图估计, 即在 $|\cdot|$ 为基数的情况下, 利用 $P(b_x | x \in O) \approx H_O^I(b_x) / |O|$ 和 $P(b_x | x \in S) \approx H_S^I(b_x) / |S|$ 得到。此外, 先验概率可以近似为 $P(x \in O) \approx |O| / (|O| + |S|)$ 。这时, 方程 (1) 就可以简化为:

$$P(x \in O | O, S, b_x) = \begin{cases} \frac{H_O^I(b_x)}{H_O^I(b_x) + H_S^I(b_x)} & \text{if } I(x) \in I(O \cup S) \\ 0.5 & \text{otherwise,} \end{cases} \quad (2)$$

看不见的像素值被分配为最大熵先验值 0.5。此时的判别模型已经使我们能够区分目标和背景像素, 见图 2a。

然而, 依然还有一个基于颜色的在线跟踪器最常见的问题存在。即这样的算法可能会产生漂移现象, 即追踪到附近区域内与目标对象表现出类似外观的对象, 为了克服这个问题, 我们扩展了对象模型, 以抑制这种分散的区域。由于计算对象似然值的公式 Eq (2) 可以通过一个有效的查找表来实现, 用非常低的计算成本计算一个大的搜索区域的似然值。将在 3.2 节详细讨论, 这样, 我们能够提前识

别潜在的分散地区并对他们进行处理。

现在，让我们假设我们已给出当前对象的假设 O 和一组潜在分散的区域 D ，如图 2b 所示。我们可以利用这个信息来建立一个模型能够区分对象和干扰像素。因此，类似于 Eq (2)。我们定义物体干扰模型为：

$$P(x \in O|O, D, b_x) = \begin{cases} \frac{H_O^I(b_x)}{H_O^I(b_x) + H_D^I(b_x)} & \text{if } I(x) \in I(O \cup D) \\ 0.5 & \text{otherwise.} \end{cases} \quad (3)$$

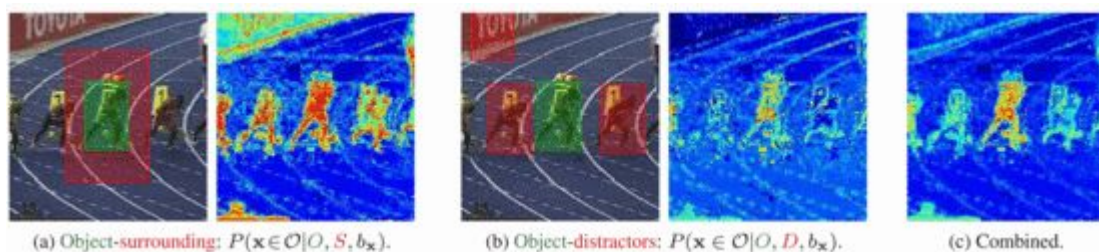


图 2

3.2. 定位

我们采用了广泛使用的跟踪检测原理来定位目标对象在一个新的时间帧 t 内的位置。特别是，我们提取矩形搜索区域上先前物体的位置 O_{t-1} 和获得新目标位置 O_t^* ：

$$O_t^* = \arg \max_{O_{t,i}} (s_v(O_{t,i}) s_d(O_{t,i})),$$

$$s_v(O_{t,i}) = \sum_{x \in O_{t,i}} P_{1:t-1}(x \in O|b_x),$$

$$s_d(O_{t,i}) = \sum_{x \in O_{t,i}} \exp\left(-\frac{\|x - c_{t-1}\|^2}{2\sigma^2}\right),$$

其中 $s_v(\cdot)$ 表示基于组合对象模型的投票评分， $s_d(\cdot)$ 表示到目标中心 c_{t-1} 欧氏

将对象的背景模型与上述干扰感知的模型相结合，得到最终的对象模型为：

$$P(x \in O|b_x) = \lambda_p P(x \in O|O, D, b_x) + (1 - \lambda_p) P(x \in O|O, S, b_x)$$

λ_p 是一个预定义的加权参数。组合对象模型（见图 2C）中错误的对象像素会产生高似然值同时降低分散地区的影响。为了适应对象的外观和光照条件的改变，我们的目标模型能够更新定期，并且可利用学习率 η 进行线性插值：

$$P_{1:t}(x \in O|b_x) = \eta P(x \in O|b_x) + (1 - \eta) P_{1:t-1}(x \in O|b_x)$$

距离。这个计算相邻时间帧运动距离的方法，类似以高斯余弦窗的相关跟踪等方法 [10] [13]，[23]。

我们假设对密集区域内重叠部分进行搜索并计算每个选取点 $O_{t,i}$ 的加权和距离值。这使我们能够有效地获得新的目标定位 O_t^* ，也能让我们区分视觉上相似的干扰物，如部分加权值很高的点。如果一个点 $O_{t,i}$ 满足 $s_v(O_{t,i}) \geq \lambda_v s_v(O_t^*)$ 并且 $\lambda_v \in [0, 1]$,

我们考虑这个点是分心点。为了防止选择模糊干扰（例如由于对象本身增加规模引起）我们采用迭代非极大值抑制策略，即在选点（无论是 O_t^* 或分心点）时，重叠的假设都会被丢弃。获得新的目标定位和干扰项的参数后，对象模型可以根据 Eqs (2) 和 (3) 是通过抑制背景干扰区域从

而在稍后的时间内减少漂流的风险。

3.3. 尺度估计

类似于最近先进的尺度自适应追踪器 [13], 我们首先把目标对象定位在一个新的框架内, 随后进行规模估算。我们利用我们的对象模型对目标对象通过阈值进行自适应尺度分割。然而, 选择一个预先定义的阈值可能会阻碍由于背景杂波或快速光照变化的规模适应。因此, 我们提出了一个自适应阈值如下。

让 L 表示在搜索区域中通过组合对象模型评价得到的每一个位置的似然图, 如图 3a 所示。然后, 我们通

$$c_O^L(b) = \sum_{i=1}^b H_O^L(i) / |O|$$

$$\text{和 } c_S^L(b) = \sum_{i=1}^b H_S^L(i) / |S|$$

计算在目标区域 O 和周边地区 S 累积直方图 (分别在图 3b 所示)。我们可以利用这些累积直方图计算自适应阈值 (以 τ^* 分割) :

$$\tau^* = \arg \min_{\tau} (2c_O^L(b_{\tau}) - c_O^L(b_{\tau} + 1) + c_S^L(b_{\tau})), \quad (\text{s.t. } c_O^L(b_{\tau}) + c_S^L(b_{\tau}) \geq 1.$$

这一方法适用于通过累计直方图计算平坦的区域内目标对象 c_O^L 阈值。例如在在图 3D 中, 阈值范围内 $[0.2, 0.3]$ 。得到的阈值可以显著降低背景噪声的同时产生一个足够大的目标对象的像素数。

为适应当前对象的假设规模 O_t^* , 我们定义了一个安全的前景区域 (即 O_t^* 内 80% 区域), 并且基于分割结果对自适应阈值后进行连通域分析 (见图 3c)。连接的组件会产生一个高于平均对象的可能性值和相交的安全前景区域被

标记为对象区域。在当前帧 O_t^s 内, 对这些地区进行封闭框规模估算计算。如果当前和以前的帧之间的估计规模是一个合理的比例变化, 我们就放弃了不可靠的分割尺寸。否则, 我们用 $O_t = \lambda_s O_t^s + (1 - \lambda_s) O_t^*$ 来更新对象

的假设尺寸。最近的尺度自适应方法如 [13], [23], 我们的规模估计方法不局限于一个固定的长宽比, 如图 1 所示。

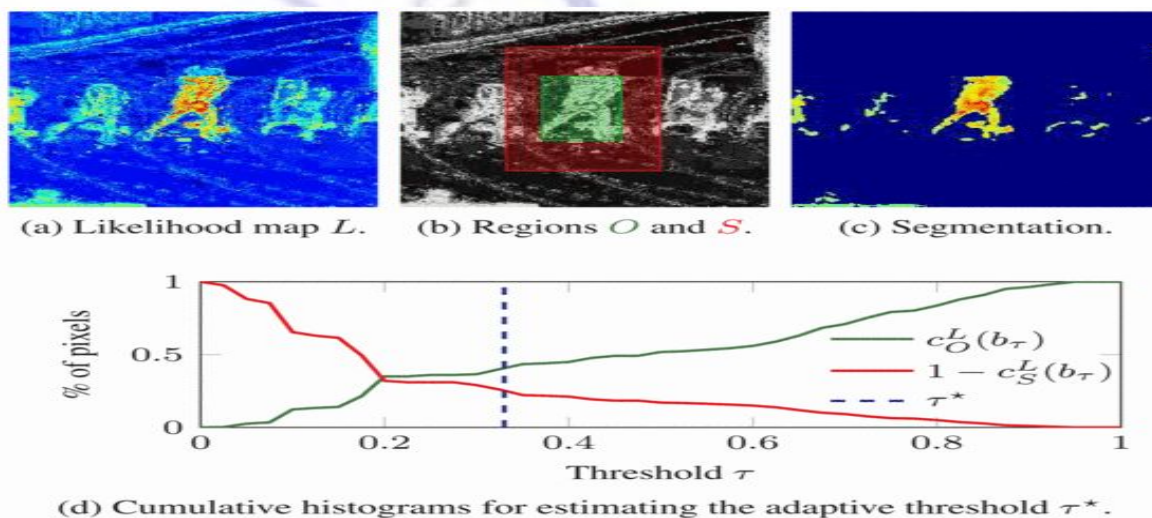


图 3

评估

我们在两个公开的基准数据集评估我们的分心-注意跟踪方法，即视觉目标跟踪(VOT)挑战数据集 vot14 [27] 和 vot13 [26]。考虑到提交的类似跟踪方法的数量很多，这些挑战最多的是自由跟踪模型的基准日期。在下面，我们专注于将我们的方法与国家的最先进的跟踪算法一个详细的比较。附加的视觉结果在补充材料中展示。

数据集特征

序列中包含的数据集已经对知名的 VOT 进行收集跟踪评估，如基于在线跟踪 (OTB) [44]，对阿姆斯特丹图书馆普通视频调查实验 (ALOV) [41]，以及最近由不同作者出版的的视频序列 (包括 [1], [17]、[28]、

[29]、[38])。特别是，VOT 委员会提出了一序列的选择方法来编制涵盖各种现实生活中的视觉现象的同时保持合理的低数序列的数据集。总的来说，该数据集包括 16 (VOT 13) 和 25 (vot14) 序列捕获严重的光照变化，物体的变形和外观的变化，突然的运动变化，显著的尺度变化、摄像机运动、和闭塞等。

评估协议

我们遵循基于时间的挑战协议，即追踪器在第一帧的视频使用的地面真理注释作为初始化并且在偏离目标时重新初始化一次。VOT 框架用两个弱相关的性能指标提供一个标准化的分析，即精度 (包围框平均重叠值) 和 稳定性 (重新初始化数) 分析。

Tracker	Accuracy		Robustness		Combined
	Score [↑]	Rank [↓]	Score [↓]	Rank [↓]	Rank [↓]
ACT [14]	0.53	7.66	1.48	7.38	7.52
CMT [33]	0.48	8.89	2.64	9.14	9.02
DSST [13]	0.62	4.78	1.16	6.44	5.61
FoT [43]	0.51	8.37	2.28	9.54	8.95
IIVT [46]	0.47	9.94	3.19	9.66	9.80
KCF [23]	0.62	4.48	1.32	6.76	5.62
LGT [42]	0.46	9.33	0.66	6.20	7.77
MIL [4]	0.39	11.69	2.27	8.96	10.32
OGT [32]	0.54	7.23	3.34	9.86	8.55
PT [16]	0.44	11.02	1.40	7.20	9.11
SPOT [48]	0.48	9.92	2.16	9.36	9.64
Struck [21]	0.51	8.31	2.16	8.84	8.57
noDAT	0.55	6.43	3.68	9.72	8.08
DAT	0.56	6.80	1.08	5.72	6.26
DATs	0.61	5.05	0.84	5.14	5.09

(a) Results VOT14.

Tracker	Accuracy		Robustness		Combined
	Score [↑]	Rank [↓]	Score [↓]	Rank [↓]	Rank [↓]
ACT [14]	0.60	6.43	0.94	8.12	7.28
CT [47]	0.47	12.52	1.77	9.59	11.06
DFT [39]	0.60	6.56	1.31	9.16	7.86
FoT [43]	0.63	6.05	1.38	8.38	7.21
HT [17]	0.48	11.37	3.64	8.97	10.17
IVT [38]	0.61	6.57	1.81	9.09	7.83
KCF [23]	0.61	6.10	0.88	7.62	6.86
LGT [42]	0.54	8.09	0.28	5.97	7.03
MIL [4]	0.52	10.30	1.48	8.59	9.45
PLT [26]	0.58	7.30	0.00	4.66	5.98
SPOT [48]	0.56	9.40	1.46	8.44	8.92
Struck [21]	0.51	8.49	3.94	7.97	8.23
TLD [25]	0.59	8.27	6.69	12.00	10.13
DAT	0.60	7.04	0.38	6.19	6.61
DATs	0.63	5.64	0.12	5.25	5.45

(b) Results VOT13.

表 1

此外，VOT 框架提供了一个基于这些指标排名分析。这一排名考虑的性能的差异统计显着性，以确保一个公平的比较。如果只有从实际的角度来看存在微小差别追踪器也具有同样出的

排名。对详细的描述和排序方法的评价，有兴趣的读者可以参考 [26]，[27]。

根据的 VOT 评估协议，我们把整个实验参数固定。我们的颜色在 RGB 颜色

立方体的联合分布并且使用每信道 10 像素的颜色直方图建模。此外，我们使用模型的加权因子 $\lambda_p = 0.5$ 和更新速度比 $\eta = 0.1$ 。搜索区域 O_{t-1} 为一个假设对象的三倍尺寸和周边区域 O_t 两倍尺寸的区域，为了确定分散的地区，我们采用投票的因子 $\lambda_v = 0.5$ 。对规模进行更新使用 $\lambda_s = 0.2$ 。

vot14 结果基准

为了确保公平和公正的比较，我们使用由相应的作者或研究委员会提交给 vot14 原有的结果进行挑战（基于相应的公开可用的实现）。我们将我们的方法和国家最新最先进的算法比较，包括对 vot14 挑战获胜方法 DSST [13]，和两个在线跟踪基准 [44] 顶尖的跟踪器方法，即 Struck [21] 和 CSK [22]。我们使用其最近的扩展 KCF [23]（即尺度自适应的结果提交给 vot14）。此外，我们比较的方法包括基于颜色属性 ACT [14]，基于关键点的 CMT [33] 和 IIVT [46]，基于部分 LGT [42]，OGT [32]，和 PT [16]，基于判别 MIL [4] 和 FOT [43]。此外，我们提供的结果为上下文感知的点 [48] 使用他们的在线可用的实现。

总体结果

从表 1A 和图 4a 的排名结果看，我们的分心注意跟踪器（DAT）及其尺度自适应版本（DATS）的相对精度和鲁棒性的跟踪级别最高。在综合排名中，由于其良好的鲁棒性，我们的方法优于所有的竞争对手。我们得到的分数有成为国家的最先进的尺度自适应跟踪方法的竞争力（即相比于 DSST [13] 和 [23]），同时显著降低了漂移的问题，可以从表 3 详细的鲁棒性评分看出。更关键的是，我们的方法基于

颜色的跟踪时，判别对象的表现也显著优于其他方法，如 ACT [14] 和 OGT [32]，以及基于图像梯度和颜色信息的追踪、如 PT [16]。此外，注意前 3 名的追踪方法（DAT，DSST，和 KCF）使用规模运算，而没有使用规模运算的 DAT 排在第 4 位，战胜了其余的竞争对手。

分心-意识的优点

为了证明表征分心感知对象的重要性，我们仅仅通过运用差异性的背景模板基线跟踪（即式（2））来比较我们的（DAT）法和 NODAT 法。总的来说，由于区别模型导致的竞争准确性差异，noDAT 实现了其中端性能

（见 NODAT 达到表 1A 和图 4a）。然而，没能彻底挖掘有关潜在干扰物相关方面的缘故，NODAT 有漂移现象，详细可见鲁棒性评分表 3。包括分心感知的表示（DAT）可显著降低这个限制序列类似的视觉区域，如篮球、螺栓、鱼类、慢跑、滑冰。NODAT 平均达到了 3.68 分的鲁棒性，并且包括分心感知性提高至 1.08 而产生模糊精度性。因此，分心感知的对象模型是建立一个强大的在线追寻系统的重要线索。

喧闹的初始化及鲁棒性

VOT 框架提供了附加随即变动初始边框的实验装置。根据 VOT 使用方法，我们进行 15 次的噪音初始化并报告了平均结果。为了表征在本次噪音实验的前五者，我们将表 2 与我们的方法进行了比较。尽管初始化并不完全可信，DAT 和 DATS 都超越了以下跟踪装置：DSST [13] [23]，KCF，LGT [42]，ACT [14]，和 Struck [21]。以上的表征表明我们可以恢复这些初始化的错误，并在精确度和鲁棒性方面进行有力的改良。

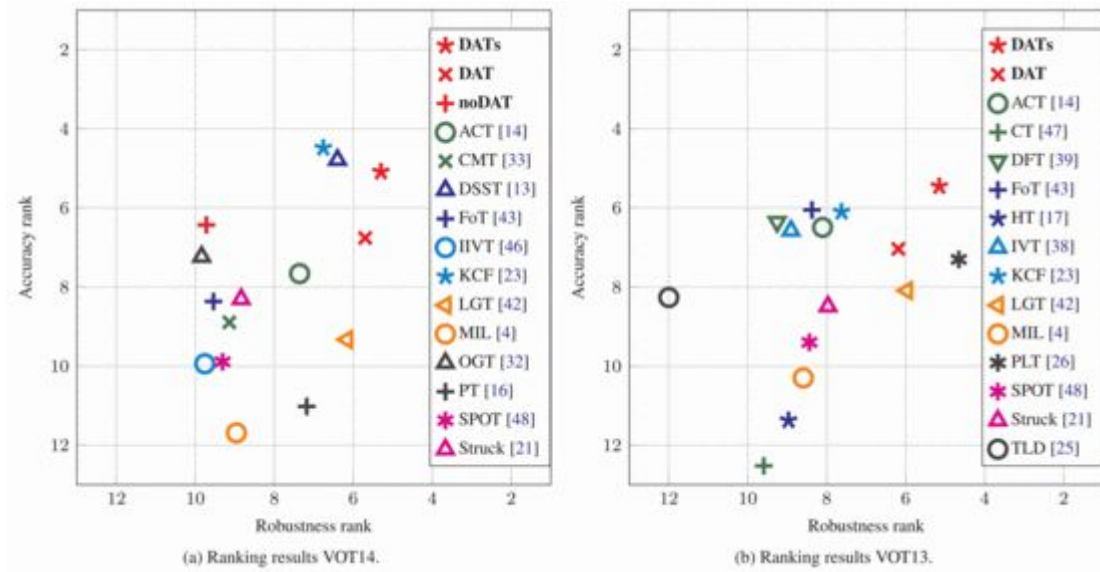


图 4

运行时性能

包括分心感知的对象在内只消耗相当低的计算成本。PC 上的一个 3.4 GHz 的英特尔 CPU 纯 MATLAB 原型数据在 17 帧运行，而不干扰信息跟踪

(NODAT) 平均能达到 18 帧。规模估计步骤是非常有效的，作为完整的跟踪方法 (DATs) 整个过程平均为 15 帧。因此，该数据跟踪已经可以用于时间的应用领域，如视频监控、机器人。

Tracker	Accuracy		Robustness		Combined Rank \downarrow
	Score \uparrow	Rank \downarrow	Score \downarrow	Rank \downarrow	
ACT [14]	0.49	5.02	1.77	4.56	4.79
DSST [13]	0.57	3.10	1.28	3.98	3.54
KCF [23]	0.57	3.44	1.51	4.28	3.86
LGT [42]	0.46	5.12	0.64	3.54	4.33
Struck [21]	0.48	5.42	2.22	5.00	5.21
DAT	0.55	3.20	1.06	3.38	3.29
DATs	0.58	2.70	1.03	3.26	2.98

表 2

4.2. vot13 结果基准

此外，我们对于我们的方法在 vot13 基准数据集的应用进行了评估。类似以前的评价，我们又由 VOT 委员会验证原始的 vot13 挑战结果。我们将我们的方法来和 vot13 挑战冠军 PLT 的延伸 [26] 比较。此外，我们还和 CT [47], DFT [39], FoT [43], HT [17], IVT [38], LGT [42], MIL [4], TLD [25]。我们还报告了最近的 ACT [14], KCF [23], SPOT [48] 方法，这些方法都可以实现公开可用的追踪。

平均性能指标和排名的结果如表格 1B 图 4B 所示。再一次，我们的方法在排名当中，占据着顶尖的位置。特别是，我们的尺度自适应数据优于 vot13 挑战冠军 PLT，而单表数据排名第三。这说明我们的方法有效的感知相当复杂的色彩表现模型，（例如

ACT [14]）而且性能可与先进的基于 HOG 的追踪器相媲美。

总结

我们提出了一个基于有效判别色彩模型的通用对象跟踪方法。为了克服国家最先进方法所遇到的基于颜色的追踪漂移的问题，我们提前适应对象，确定了分散的地区以抑制这些地区。我们详细评估了最近的基准数据集发现，基于颜色的跟踪可以实现现实世界中具有挑战性的序列，而且准确性高。此外，使用该判别器感知对象模型显著提高了追踪的鲁棒性，即使是可用的只有初始化的嘈杂。总体而言，我们的基于颜色的方法表现出产生良好的性能，最近的国家的最先进的基于更复杂的功能跟踪器，实现了高帧速率下，适用于时间的关键型应用程序。

Sequence	DATs		DAT		noDAT		ACT [14]		DSST [13]		KCF [23]		LGT [42]		OGT [32]	
	Acc [†]	Rob [‡]	Acc [†]	Rob [‡]	Acc [†]	Rob [‡]	Acc [†]	Rob [‡]	Acc [†]	Rob [‡]	Acc [†]	Rob [‡]	Acc [†]	Rob [‡]	Acc [†]	Rob [‡]
ball	0.87	0.00	0.67	0.00	0.67	0.00	0.41	0.00	0.56	1.00	0.75	1.00	0.31	1.13	0.72	0.00
basketball	0.61	1.00	0.73	1.00	0.68	26.00	0.66	0.00	0.64	1.00	0.64	0.00	0.50	0.80	0.55	8.93
bicycle	0.54	0.00	0.48	1.00	0.48	1.00	0.46	1.00	0.58	0.00	0.62	0.00	0.53	0.93	0.65	0.20
bolt	0.52	0.00	0.49	1.00	0.47	2.00	0.54	1.00	0.56	1.00	0.49	3.00	0.38	0.67	0.67	23.20
car	0.80	0.00	0.41	0.00	0.41	0.00	0.53	1.00	0.73	0.00	0.70	0.00	0.51	0.80	0.49	0.00
david	0.64	0.00	0.64	0.00	0.61	1.00	0.72	0.00	0.80	0.00	0.82	0.00	0.56	0.00	0.50	0.13
diving	0.41	1.00	0.35	1.00	0.44	5.00	0.20	4.00	0.44	1.00	0.25	4.00	0.33	1.27	0.23	4.40
drunk	0.53	0.00	0.48	0.00	0.44	4.00	0.46	0.00	0.55	0.00	0.53	0.00	0.52	0.00	0.55	1.00
fernando	0.45	0.00	0.44	0.00	0.42	0.00	0.43	1.00	0.34	1.00	0.41	1.00	0.47	0.47	0.35	1.67
fish1	0.72	0.00	0.58	2.00	0.41	6.00	0.44	0.00	0.32	1.00	0.42	3.00	0.36	0.93	0.51	2.13
fish2	0.41	2.00	0.36	3.00	0.43	4.00	0.31	5.00	0.35	4.00	0.26	6.00	0.28	1.80	0.23	5.73
gymnastics	0.58	0.00	0.57	0.00	0.57	0.00	0.51	2.00	0.63	5.00	0.53	1.00	0.48	1.00	0.56	2.73
hand1	0.59	1.00	0.61	0.00	0.63	0.00	0.41	5.00	0.21	2.00	0.56	3.00	0.55	0.00	0.57	1.33
hand2	0.59	0.00	0.56	3.00	0.56	3.00	0.39	8.00	0.52	6.00	0.49	6.00	0.49	1.20	0.49	8.87
jogging	0.72	0.00	0.75	0.00	0.80	6.00	0.70	1.00	0.79	1.00	0.79	1.00	0.35	1.00	0.61	1.73
motocross	0.35	4.00	0.34	4.00	0.34	8.00	0.47	3.00	0.42	4.00	0.36	2.00	0.41	1.00	0.20	5.40
polarbear	0.82	0.00	0.55	0.00	0.57	0.00	0.51	0.00	0.63	0.00	0.78	0.00	0.65	0.00	0.65	0.00
skating	0.51	5.00	0.50	9.00	0.52	14.00	0.50	0.00	0.59	0.00	0.68	1.00	0.32	0.40	0.56	6.93
sphere	0.78	0.00	0.71	0.00	0.71	0.00	0.72	0.00	0.92	0.00	0.90	0.00	0.64	0.00	0.50	0.00
sunshade	0.57	0.00	0.58	0.00	0.54	1.00	0.79	0.00	0.78	0.00	0.76	0.00	0.55	0.40	0.74	0.00
surfing	0.76	0.00	0.85	0.00	0.87	3.00	0.83	0.00	0.90	0.00	0.79	0.00	0.57	0.00	0.70	0.00
torus	0.82	0.00	0.75	0.00	0.65	1.00	0.79	0.00	0.81	0.00	0.85	0.00	0.63	0.00	0.74	0.53
trellis	0.50	0.00	0.51	0.00	0.51	1.00	0.58	2.00	0.80	0.00	0.79	0.00	0.48	0.00	0.68	1.40
tunnel	0.47	7.00	0.34	2.00	0.43	4.00	0.31	0.00	0.80	0.00	0.68	0.00	0.36	1.47	0.50	3.53
woman	0.65	0.00	0.63	0.00	0.68	2.00	0.66	3.00	0.79	1.00	0.74	1.00	0.36	1.13	0.63	3.67
Mean	0.61	0.84	0.56	1.08	0.55	3.68	0.53	1.48	0.62	1.16	0.62	1.32	0.46	0.66	0.54	3.34

Table 3: Detailed results for the VOT14 benchmark dataset comparing our approach to the top 5 competitors. For each sequence, we report the average accuracy (Acc) and robustness (Rob) scores. Best, second best, and third best results have been highlighted. For the non-deterministic trackers LGT and OGT, we report the results averaged over 15 runs.

表 3

致谢

这项工作是由奥地利科学基金资助（FWF）下的学习先进跟踪和检测技术的项目。

(I535-N23)

参考书籍

1. A. Adam, E. Rivlin and I. Shimshoni

"Robust Fragments-based Tracking using the Integral Histogram"

Proc. CVPR,

2. N. Alt, S. Hinterstoisser and N. Navab

"Rapid Selection of Reliable Templates for Visual Tracking"

Proc. CVPR,

3. S. Avidan

"Ensemble Tracking"

PAMI, vol. 29, no. 2, pp. 261-271, 2007

4. B. Babenko, M.-H. Yang and S. Belongie

"Robust Object Tracking with Online Multiple Instance Learning"

PAMI, vol. 33, no. 7, pp. 1324-1338, 2011

5. Y. Bai and M. Tang

"Robust Tracking via Weakly Supervised Ranking SVM"

Proc. CVPR,

6. C. Bao, Y. Wu, H. Ling and H. Ji

"Real Time Robust L1 Tracker Using Accelerated Proximal Gradient Approach"

Proc. CVPR,

7. V. Belagiannis, F. Schubert, N. Navab and S. Ilic

"Segmentation Based Particle Filtering for Real-Time 2D Object Tracking"

Proc. ECCV,

8. C. Bibby and I. Reid

"Robust Real-Time Visual Tracking using Pixel-Wise Posteriors"

Proc. ECCV,

9. M. J. Black and A. D. Jepson

"EigenTracking: Robust Matching and Tracking of Articulated Objects Using a View-Based Representation"

IJCV, vol. 26, no. 1, pp. 63-84, 1998

10. D. S. Bolme, J. R. Beveridge, B. A. Draper and Y. M. Lui

"Visual Object Tracking using Adaptive Correlation Filters"

Proc. CVPR,

11. M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier and L. Van Gool

"Online Multi-Person Tracking-by-Detection from a Single, Uncalibrated Camera"

PAMI, vol. 33, no. 9, pp. 1820-1833, 2011

12. D. Comaniciu, V. Ramesh and P. Meer

"Kernel-Based Object Tracking"

PAMI, vol. 25, no. 5, pp. 564-577, 2003

13. M. Danelljan, G. Häger, F. S. Khan and M. Felsberg

"Accurate Scale Estimation for Robust Visual Tracking"

Proc. BMVC,

14. M. Danelljan, F. S. Khan, M. Felsberg and J. van de Weijer

"Adaptive Color Attributes for Real-Time Visual Tracking"

Proc. CVPR,

15. T. B. Dinh, N. Vo and G. Medioni

"Context Tracker: Exploring Supporters and Distracters in Unconstrained Environments"

Proc. CVPR,

16. S. Duffner and C. Garcia

"PixelTrack: a fast adaptive algorithm for tracking non-rigid objects"

Proc. ICCV,

17. M. Godec, P. M. Roth and H. Bischof

"Hough-based Tracking of Non-Rigid Objects"

Proc. ICCV,

18. H. Grabner, C. Leistner and H. Bischof

"Semi-supervised On-line Boosting for Robust Tracking"

Proc. ECCV,

19. H. Grabner, J. Matas, L. Van Gool and P. Cattin

"Tracking the Invisible: Learning Where the Object Might be"

Proc. CVPR,

20. G. D. Hager and P. N. Belhumeur

"Efficient Region Tracking With Parametric Models of Geometry and Illumination"

PAMI, vol. 20, no. 10, pp. 1025-1039, 1998

21. S. Hare, A. Saffari and P. H. S. Torr

"Struck: Structured Output Tracking with Kernels"

Proc. ICCV,

22. J. F. Henriques, R. Caseiro, P. Martins and J. Batista

"Exploiting the Circulant Structure of Tracking-by-detection with Kernels"

- Proc. ECCV, "Visual Tracking Decomposition"
23. J. F. Henriques, R. Caseiro, P. Martins and J. Batista
Proc. CVPR,
29. J. Kwon and K. M. Lee
"High-Speed Tracking with Kernelized Correlation Filters"
PAMI, 2015
"Highly Non-Rigid Object Tracking via Patch-based Dynamic Appearance Modeling"
24. X. Jia, H. Lu and M.-H. Yang
PAMI, vol. 35, no. 10, pp. 2427-2441, 2013
"Visual Tracking via Adaptive Structural Local Sparse Appearance Model"
Proc. CVPR,
30. B. Liu, J. Huang, L. Yang and C. Kulikowsk
"Robust Tracking Using Local Sparse Appearance Model and K-Selection"
25. Z. Kalal, K. Mikolajczyk and J. Matas
Proc. CVPR,
31. I. Matthews, T. Ishikawa and S. Baker
"Tracking-Learning-Detection"
PAMI, vol. 34, no. 7, pp. 1409-1422, 2012
"The Template Update Problem"
26. M. Kristan, R. Pflugfelder, A. Leonardis, J. Matas, F. Porikli, L. Cehovin, G. Nebehay, G. Fernandez and T. Vojír
PAMI, vol. 26, no. 6, pp. 810-815, 2004
32. H. Nam, S. Hong and B. Han
"The Visual Object Tracking VOT2013 challenge results"
Proc. VOT (ICCV Workshop),
"Online Graph-Based Tracking"
27. M. Kristan, R. Pflugfelder, A. Leonardis, J. Matas, L. Cehovin, G. Nebehay, T. Vojír and G. Fernandez
Proc. ECCV,
Show Context
33. G. Nebehay and R. Pflugfelder
"The Visual Object Tracking VOT2014 challenge results"
Proc. VOT (ECCV Workshop),
"Consensus-based Matching and Tracking of Keypoints for Object Tracking"
28. J. Kwon and K. M. Lee
Proc. WACV,
34. K. Nummiaro, E. Koller-Meier and L. Van Gool

"An Adaptive Color-Based Particle Filter"

IVC, vol. 21, no. 1, pp. 99-110, 2002

35. P. Pérez, C. Hue, J. Vermaak and M. Gangnet

"Color-Based Probabilistic Tracking"

Proc. ECCV,

36. H. Possegger, T. Mauthner, P. M. Roth and H. Bischof

"Occlusion Geodesics for Online Multi-Object Tracking"

Proc. CVPR,

37. H. Possegger, S. Sternig, T. Mauthner, P. M. Roth and H. Bischof

"Robust Real-Time Tracking of Multiple Objects by Volumetric Mass Densities"

Proc. CVPR,

38. D. A. Ross, J. Lim, R.-S. Lin and M.-H. Yang

"Incremental Learning for Robust Visual Tracking"

IJCV, vol. 77, no. 1, pp. 125-141, 2008

39. L. Sevilla-Lara and E. Learned-Miller

"Distribution Fields for Tracking"

Proc. CVPR,

40. G. Shu, A. Dehghan, O. Oreifej, E. Hand and M. Shah

"Part-based Multiple-Person Tracking with Partial Occlusion Handling"

Proc. CVPR,

41. A. W. M. Smeulders, D. M. Chu, R. Cucchiara, S. Calderara, A. Dehghan and M. Shah

"Visual Tracking: an Experimental Survey"

PAMI, vol. 36, no. 7, pp. 1442-1468, 2014

42. L. Cehovin, M. Kristan and A. Leonardis

"Robust Visual Tracking using an Adaptive Coupled-layer Visual Model"

PAMI, vol. 35, no. 4, pp. 941-953, 2013

43. T. Vojir and J. Matas

"Robustifying the Flock of Trackers"

Proc. CVWW,

44. Y. Wu and M.-H. Yang

"Online Object Tracking: A Benchmark"

Proc. CVPR,

45. M. Yang, Y. Wu and G. Hua

"Context-Aware Visual Tracking"

PAMI, vol. 31, no. 7, pp. 1195-1209, 2009

46. K. M. Yi, H. Jeong, B. Heo, H. J. Chang and J. Y. Choi

"Initialization-Insensitive Visual
Tracking Through Voting with Salient
Local Features"

Proc. ECCV,

48. L. Zhang and L. van der Maaten

Proc. ICCV,

"Preserving Structure in Model-Free
Tracking"

47. K. Zhang, L. Zhang and M.-H. Yang

PAMI, vol. 36, no. 4, pp. 756-769, 2014

"Real-time Compressive Tracking"

