

图 2.9 判断视频序列包含“宽高比变化”的过程

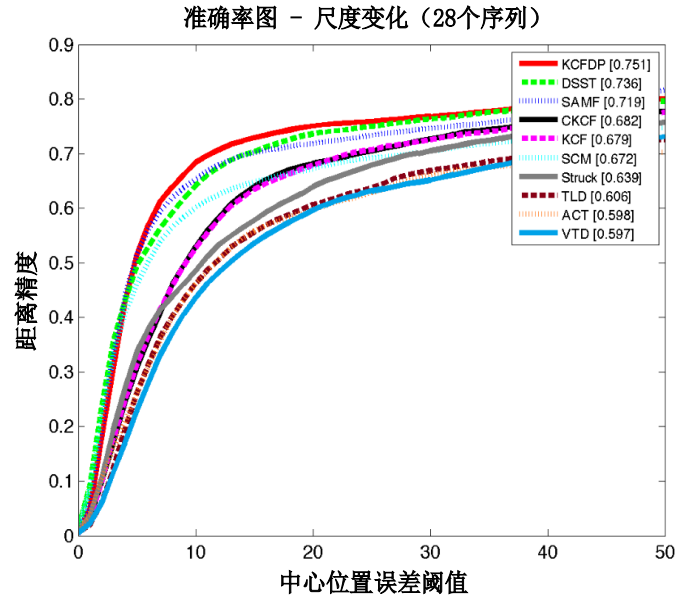
OTB 测试集除了提供了 50 个视频序列和正确的跟踪结果，还额外提供了 29 个跟踪器在它之上的运行结果。这 29 个跟踪器中不乏 SCM^[13]、Struck^[20]、TLD^[23]、VTD^[10] 和 ALSA^[11] 等当前最为先进的跟踪器。本节将直接利用这些跟踪器的运行结果，与 KCFDP 的结果进行对比评测。此外，还有 5 个基于相关滤波的跟踪器被包含在对照组中，即 KCF^[9]、DSST^[15]、ACT^[8]、SAMF^[16] 和 CKCF。其中 CKCF 就是去除了目标候选生成器的 KCFDP，也即是加入了图像特征整合和鲁棒更新的 KCF。综上，对照组中将共有 34 个跟踪器。

2.7.1.2 评价标准

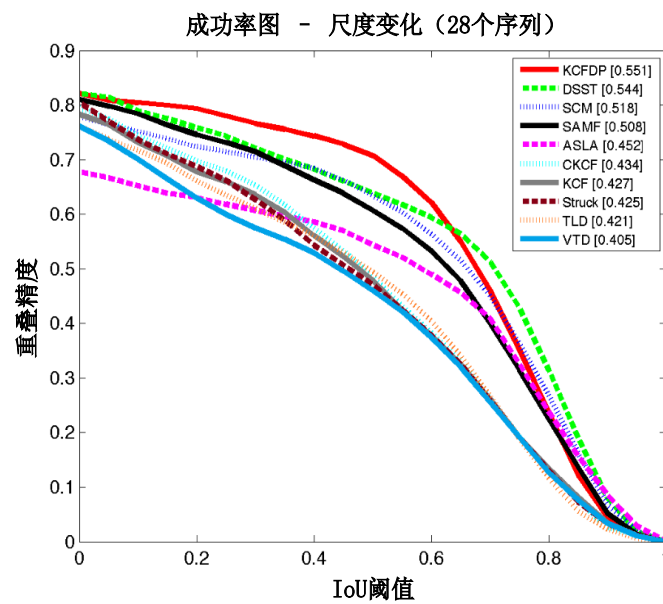
在一些早期的工作中，通常使用平均中心位置误差 (Averaged CLE) 或者平均重叠率 (Averaged IoU) 来评价跟踪器的准确性。平均中心位置误差即是在视频序列的所有帧中，计算跟踪器结果和正确跟踪结果中心间的距离，并进行平均。显然，中心位置误差无法反映出尺度和宽高比是否准确。平均重叠率进了一步，它用跟踪器结果和正确跟踪结果的 IoU 来替代中心位置误差。但是，这两者最大的问题在于，当跟踪器丢失目标时，会对跟踪结果进行随机的惩罚，从而都无法准确反映跟踪器的准确性。

因此，本节将沿用 [3] 中提出的准确性评价标准，即使用准确率图 (Precision Plot) 和成功率图 (Success Plot)。要得到准确率图，需要先计算距离精度 (Distance Precision)。距离精度的定义为：视频序列中中心位置误差小于某个阈值的帧所占的比例。而成功率图则依赖于重叠精度 (Overlap Precision)，其定义为：视频序列中 IoU 大于某个阈值的帧所占的比例。准确率图展示的是，在不同的中心位置误差阈值下 (沿横坐标)，跟踪器在所有序列上的平均距离精度 (沿纵坐标)。类似的，成功率图展示的是在不同的 IoU 阈值下 (沿横坐标)，跟踪器在所有序列上的

平均重叠精度（沿纵坐标）。对于准确率图，跟踪器将按照以 20 像素为阈值时的距离精度进行排名。而对于成功率图，跟踪器将按照“曲线下方面积”（Area Under Curve, AUC）进行排名。

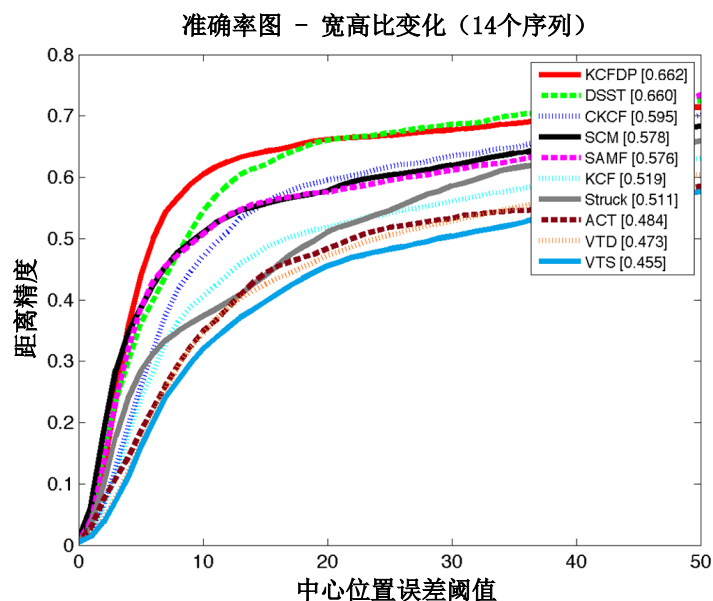


(a) 针对尺度适应力的准确率图

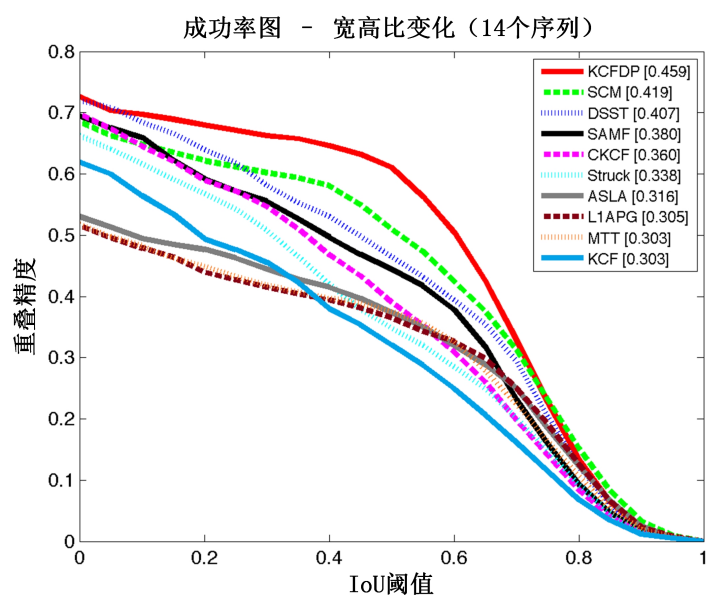


(b) 针对尺度适应力的成功率图

图 2.10 针对尺度适应力的实验结果



(a) 针对宽高比适应力的准确率图



(b) 针对宽高比适应力的成功率图

图 2.11 针对宽高比适应力的实验结果

2.7.2 尺度和宽高比适应力评测

为了评测本章方法的尺度适应力, 本节将所有的 34 个跟踪器, 加上 KCFDP, 应用于包含“尺度变化”这一跟踪障碍的 28 个视频序列上进行跟踪。得到的准确率图和成功率图见图2.10。同理, 为了评测 KCFDP 的宽高比适应力, 本节将所有跟踪器应用于上一节提取的, 包含“宽高比变化”障碍的 14 个视频序列上, 得到

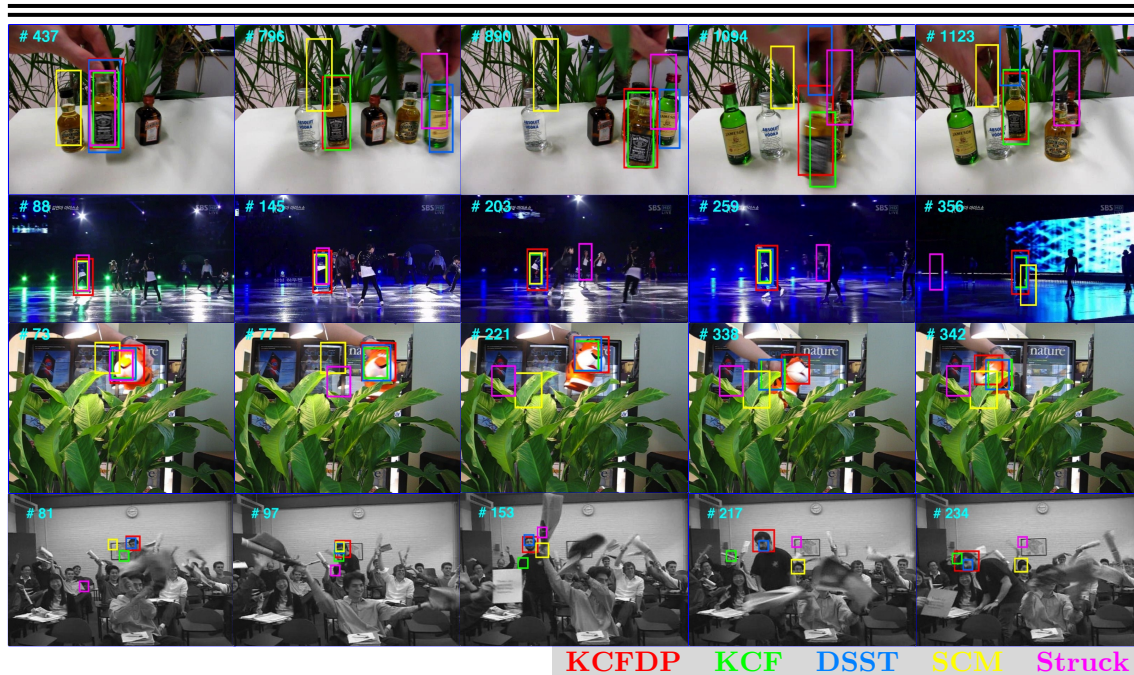


图 2.12 KCFDP 与 4 个跟踪器的可视化对比

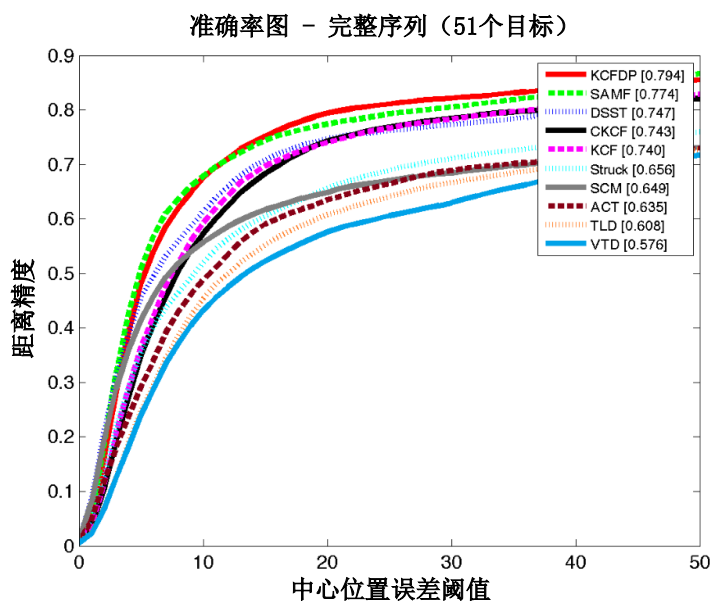
了如图2.11所示的准确率图和成功率图。在图2.10和图2.11中，为了显示清晰，只给出了排名前 10 的跟踪器的对应曲线。各图中，右上角的图例中的排名即是跟踪器的准确性排名，方括号内的数值是排名的依据，即以 20 像素为阈值时的距离精度（准确率图中），或者曲线下方的归一化面积（成功率图中）。

从图2.10和图2.11中可以清晰地看出，无论对于尺度变化还是宽高比变化，本章提出的跟踪方法均具有最佳的适应力。此外，KCFDP 在成功率图中的优势更为明显。这证明了目标候选方法主要提高了跟踪结果和正确结果间的重叠率，即是说，跟踪结果的尺度和宽高比更加准确了。

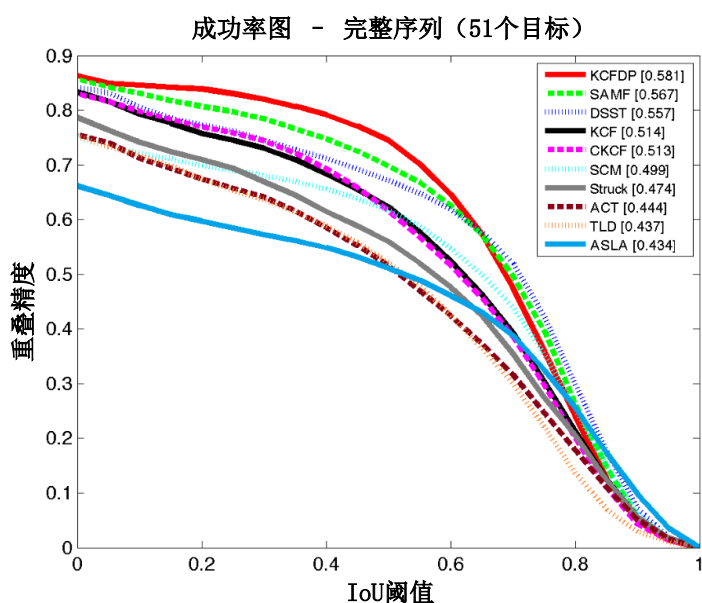
2.7.3 整体性能评测

为了评测 KCFDP 对于各种跟踪障碍的适应力和鲁棒性，以及得出其跟踪效率，本节将在完整的 50 个视频序列（包含 51 个跟踪目标）下对比所有的跟踪器。图2.12展示了在 4 个较为困难的视频序列（*Liquor*, *Skating1*, *Tiger1* 和 *Freeman4*）上，排名前 5 的跟踪器的直观可视化对比。可以清楚地看出，KCFDP 在准确地适应尺度和宽高比变化的同时，还能保持很高的跟踪精度。

将所有跟踪器运行于完整的 50 个视频序列上所得的准确率图和成功率图见图2.13。显然，在各种跟踪障碍下，KCFDP 仍然超过了所有其它跟踪器和相关滤波跟踪器变型。更进一步地分析这两节的实验结果，可以看出在图2.10、2.11、2.13中，KCFDP 与 CKCF 之间都有着明显的性能差距。这充分显示了在跟踪过程中加入目标候选所带来的准确性提升。尽管 CKCF 和 KCF 在“尺度变化”障碍下



(a) 完整的 50 个视频序列上的准确率图



(b) 完整的 50 个视频序列上的成功率图

图 2.13 完整的 50 个视频序列上的实验结果

和整个测试集下的性能区别很小，但是对于“宽高比变化”，CKCF 具有明显的性能优势。因此可以推断，整合后的混合特征仅在候选图像块出现密集形状变化时才具有较大优势。这也是本章在 KCFDP 中使用混合特征特原因，即用于辨别尺度和宽高比不断变化的目标候选。此外，相比在整个测试集和“尺度变化”子集下，KCFDP 在“宽高比变化”子集下的优势明显最大。原因在于，每个序列都包含着多个跟踪障碍，而一些障碍会导致目标候选生成器失效或者出错，从而部分

抵消掉它带来的性能提升。例如“运动模糊”会导致目标物体的边界变得模糊，严重影响依赖物体边界来生成目标候选的 **EdgeBoxes**。并且在“尺度变化”子集下，**DSST** 和 **SAMF** 等跟踪器也具有特定的尺度适应力优化，故 **KCFDP** 的相对优势略有降低。

除了很高的跟踪精度以外，**KCFDP** 的跟踪效率也令人满意，在整个测试集上达到了平均 20.8 FPS 的跟踪速度。考虑到 **KCFDP** 主要使用较低效的 **Matlab** 实现，且运行于低功耗处理器上，本节认为对实现稍作优化后，就完全可以达到实时跟踪的需求。作为对比，在完全相同的实验环境下，**DSST** 和 **SAMF** 这两个基于相关滤波的跟踪器的跟踪速度分别为 28.9 FPS 和 12.0 FPS，但是他们均不具有宽高比适应力。**KCFDP** 跟踪效率较高的原因在于，相关滤波器实际上仅需要额外辨别很少量的目标候选，且经过调优后的目标候选生成过程也十分高效。

2.8 小结

本章提出了一个在视觉物体跟踪中提高尺度和宽高比适应力的全新方法。通过对一个“物体类别无关”的目标候选生成器进行调优，并配合目标候选过滤过程，本章的方法能够为相关滤波器提供不同尺度和宽高比的候选输入图像块。为了能够准确辨别这些目标候选，本章还利用图像特征整合和更鲁棒的模型更新策略来加强了相关滤波器。得益于目标候选的高度灵活性和相关滤波器的准确辨别力，本章得到的跟踪器在一个大型公开测试集上展现出了很强的鲁棒性和适应力，同时还达到了令人满意的跟踪速度。本章的方法是一个通用方法，在下一章中，该方法将用于把更多的目标候选生成器和跟踪器相结合，以深入研究目标候选在视觉物体跟踪中的作用和潜力。此外，本章并未对目标候选生成算法进行优化，下一章将针对跟踪任务的需求，对目标候选生成器进行深入改进。