**基于无人驾驶场景的流数据**

**三维物体检测框架**

无人驾驶是近二十多年来最有望取得突破性进展的技术。目前，无论是学术界还是工业界，对无人驾驶技术的研究都如火如荼。无人驾驶应用十分广泛，无论是民用还是军工领域。 机器取代人类驾驶车辆，一方面能让人们出行效率更高， 另一方面，机器之间的通讯比人类效率更高，这使得车辆能够更好的规划自己的行驶路径，降低交通事故发生概率。作为智慧城市的不可缺少的组成部分， 无人驾驶也能在非常时期发挥重要作用。譬如在面对重大疫情时，人员之间要尽量减少接触，此时无人驾驶能在保证安全的情况下，极大方便人们的出行。因此，无人驾驶作为可预见的未来技术，一直是各国在高科技领域技术的“必争之地”。

无人驾驶主要包括感知、决策和规划三个重要模块，其中感知模块是无人驾驶技术的基础，也是难点。感知模块的目标是赋予机器以人的环境感知能力，使机器能够准确捕获复杂道路场景的有用信息。环境感知的一大核心任务是三维物体检测，即确定环境中障碍物的位姿，包括三维坐标、形状以及朝向。近几年来越来越多研究者将物体检测从二维扩展到三维，受三维数据的稀疏性以及计算量的限制，目前的三维物体检测算法在精度和速度上都离二维物体检测有很大差距。另外，目前绝大多数三维物体检测方法都是基于单帧数据的，即每次只能检测一帧数据。由于这些方法没有利用帧与帧之间的时序信息，因此无法高效的应用到真实驾驶环境中的连续三维数据流，也就无法直接应用到无人驾驶的感知模块上。

本项目计划提出一套基于无人驾驶场景的流数据三维物体检测框架。该框架为现阶段主流的三维物体检测方法提供了统一的实现方式，极大的方便了不同三维检测算法的对比研究以及新算法的开发。另外，本框架能够直接作用与三维流数据，这极大地方便了算法的落地与调试。该框架类似于Google 公司的自动驾驶平台 Waymo 以及百度公司的阿波罗（Apollo）自动驾驶框架中的环境感知模块，与前两者不同的是，该框架能够高效整合基于深度学习技术的三维物体检测前沿技术，并将这些基于单帧的技术推广到流数据场景，能够极大提高其运行效率，扫除了算法落地的障碍。据调研，目前国内外没有类似的工作。

我们目前已经完成项目的大部分准备工作。双路物体检测与跟踪框架（DODT）是我们提出的原始版本，该框架结构图如图1所示。DODT是基于关键帧进行流数据物体检测的，有上下两个相同的物体检测模块负责检测两关键帧的物体。对流数据进行采样检测一方面能够很大程度上减少算法在整个流数据上的耗时，另一方面也能够很好的捕获帧间的时序信息。该时序信息可以进一步提高物体检测的精度，同时也可以将关键帧的检测结果传播到其他非关键帧。时序信息的捕获由图中的时序模块（Temporal Module）进行，该模块的输出是上下两关键帧中物体的对应关系以及相对偏移。由于时序模块能够得到帧间物体的对应关系，因此得到物体检测结果后，追踪结果也能够很方便得到。该成果目前已被ICRA2020录取。

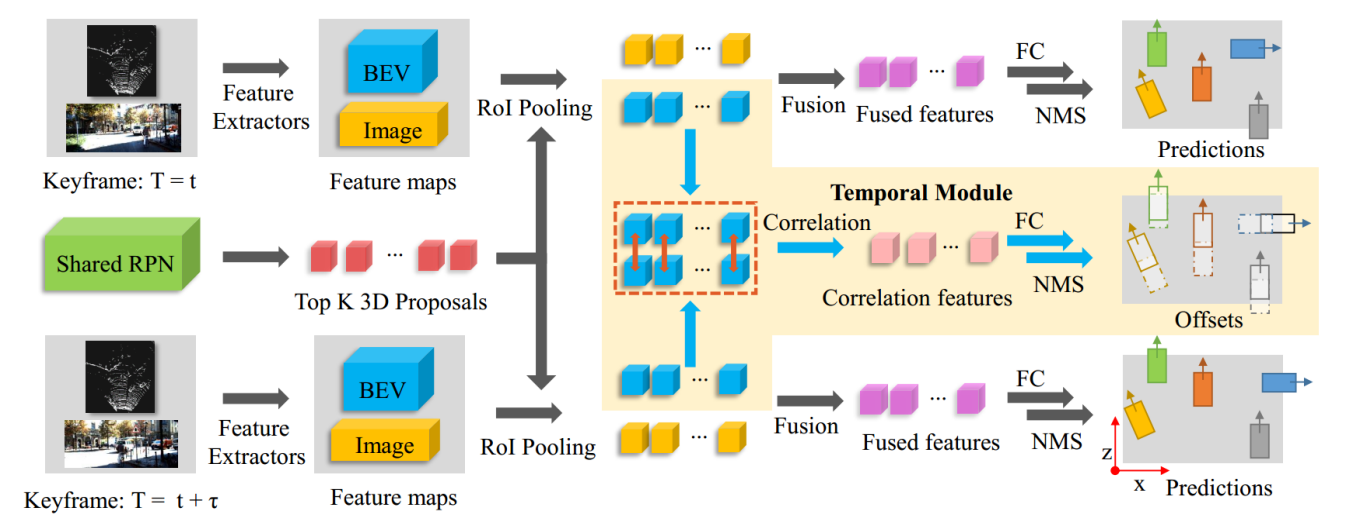


图1 双路物体检测与跟踪框架结构图

项目后续工作计划在该框架中整合目前主流的三维物体检测算法，并探索不同的时序信息获取方法，使得该框架能够服务于学术界以及工业界，为无人驾驶技术中的环境感知模块提供统一的技术平台。