

技术洞察之一—— 基于人形机器人的SLAM

Guan Weipeng

Email: wpguan@connect.hku.hk



The University of Hong Kong

Department of Mechanical Engineering

Adaptive Robotic Controls Lab (ArcLab)



香港大學
THE UNIVERSITY OF HONG KONG

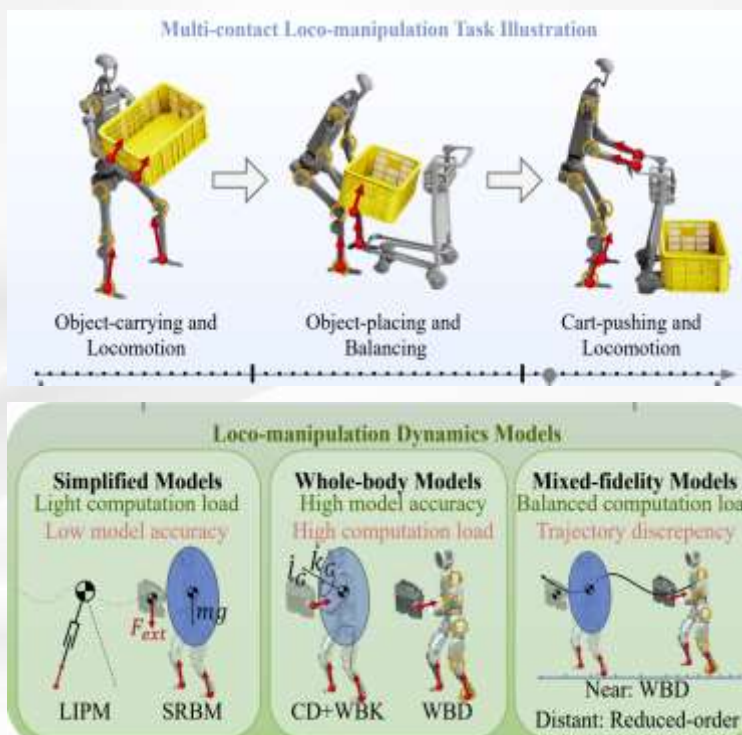
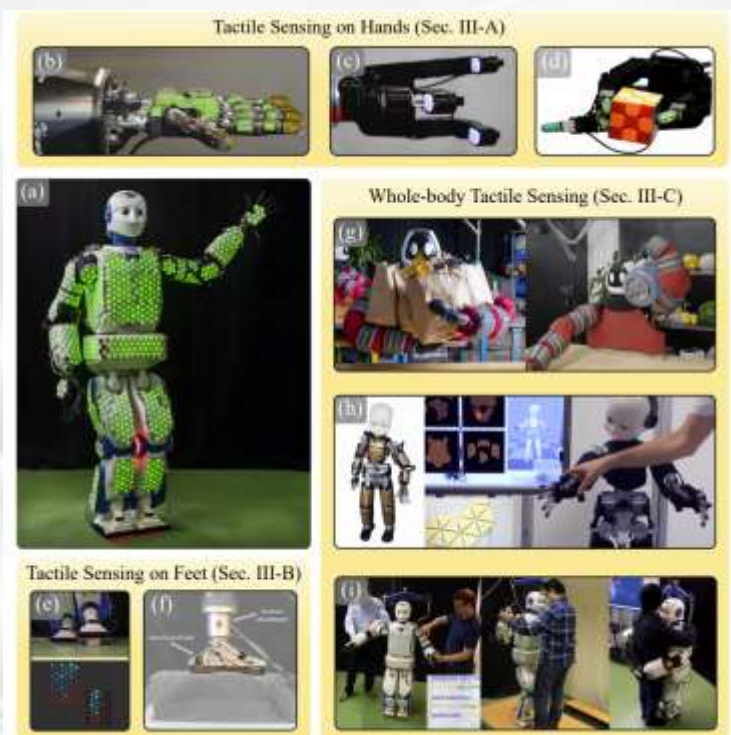
目 录

CONTENTS

- ◆ 人形机器人
- ◆ 3D LiDAR/Image SLAM
- ◆ 基于足式机器人的SLAM

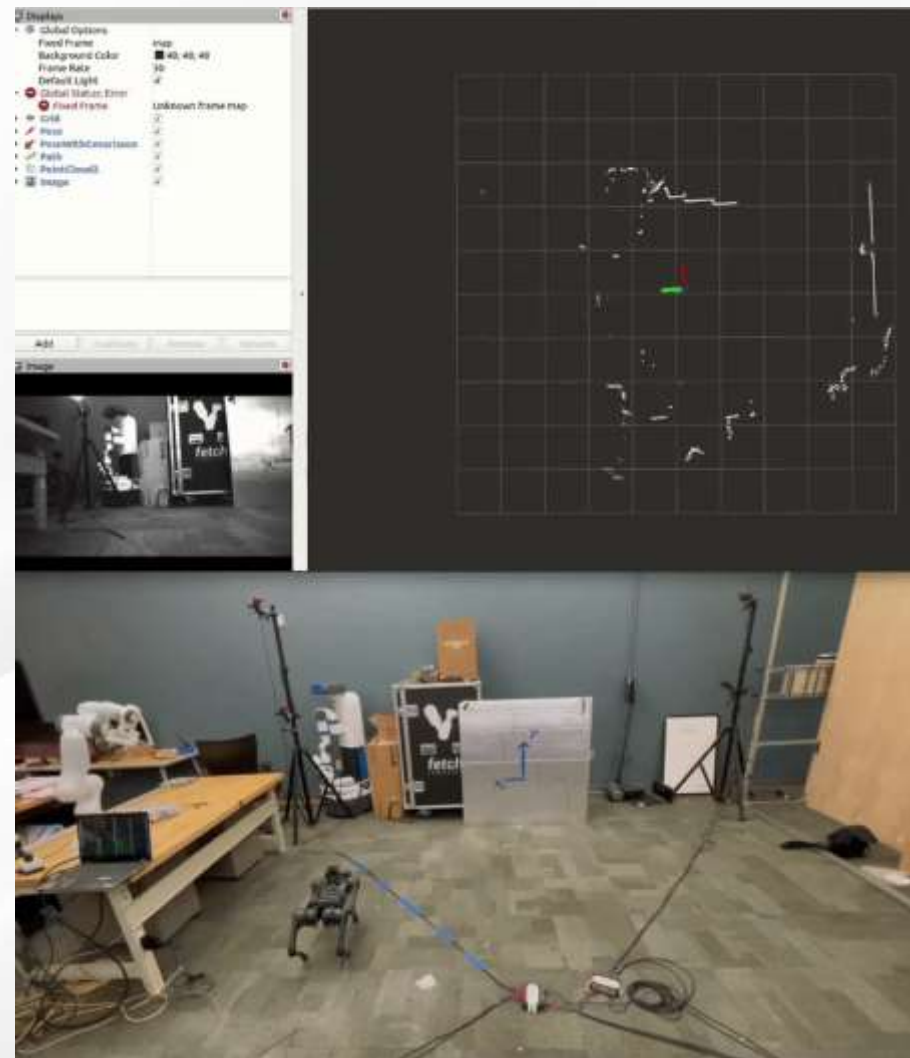
Embodied Intelligence

- 人形机器人主要的研究热点有：仿人机器人的运动控制与规划、感知与交互、双足行走与导航、全身操作与触觉传感、多接触规划、模型预测控制、全身控制、技能学习；
- 人形机器人涉及了大量与环境的交互任务，因此，对于自身的姿态估计以及周围环境物体的感知是其应用的最基本要素；
- 解决基于人形机器人的SLAM问题是人形机器人系列研究热点的基础性问题，更是人形机器人从实验室走到实际应用的最关键因素；



Motivation

- Simultaneous localization and mapping(SLAM), 包括6自由度的姿态估计及三维重建, 以实现对自身的姿态估计及周围场景感知的相关任务;
- 3D SLAM主要以手持、无人机为研究载体, 可迁移至足式机器人(机械狗、人形等);
- 针对足式机器人的SLAM算法的设计需要能够应对腿式机器人的挑战性条件, 如高振动、远程激光雷达退化和z轴漂移;
- 此外, 复杂的地形, 如斜坡滑动、可变形的地形(草地、泥土)也会给足式机器人的SLAM问题带来巨大的挑战;
- 四足与人形的运动模态非常相似, 只有落地足可以产生观测量, 因此, 虽然基于人形的SLAM研究较少, 但是基于四足机器人的SLAM算法基本可以借鉴;
- 本技术洞察从3DSLAM到基于足式机器人的SLAM为主, 其中, 3D-SLAM系列技术大部分有实测经历(以博客或B站视频的形式给出);
- 本报告附带的, 基于人形/四足机器人的SLAM相关论文调研请见: [链接](#);

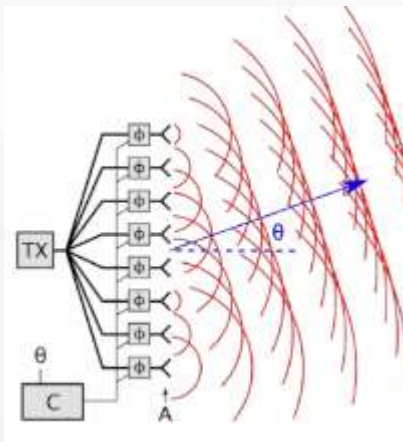
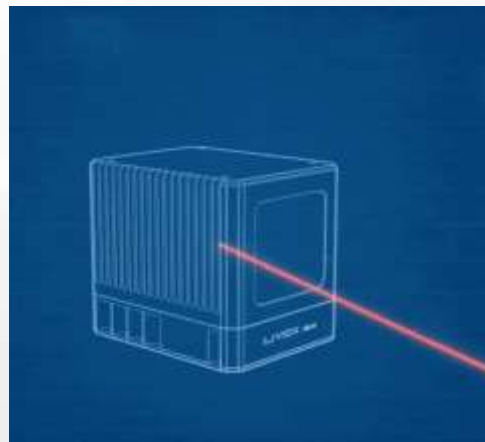
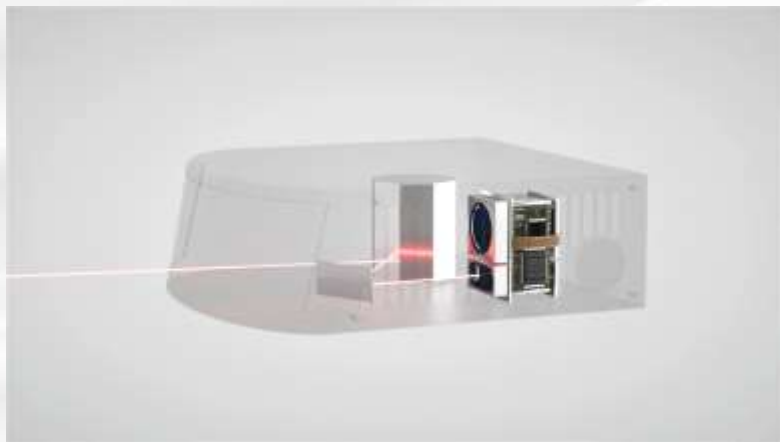


From: https://github.com/jerredchen/A1_SLAM

3D LiDAR-based SLAM

激光雷达

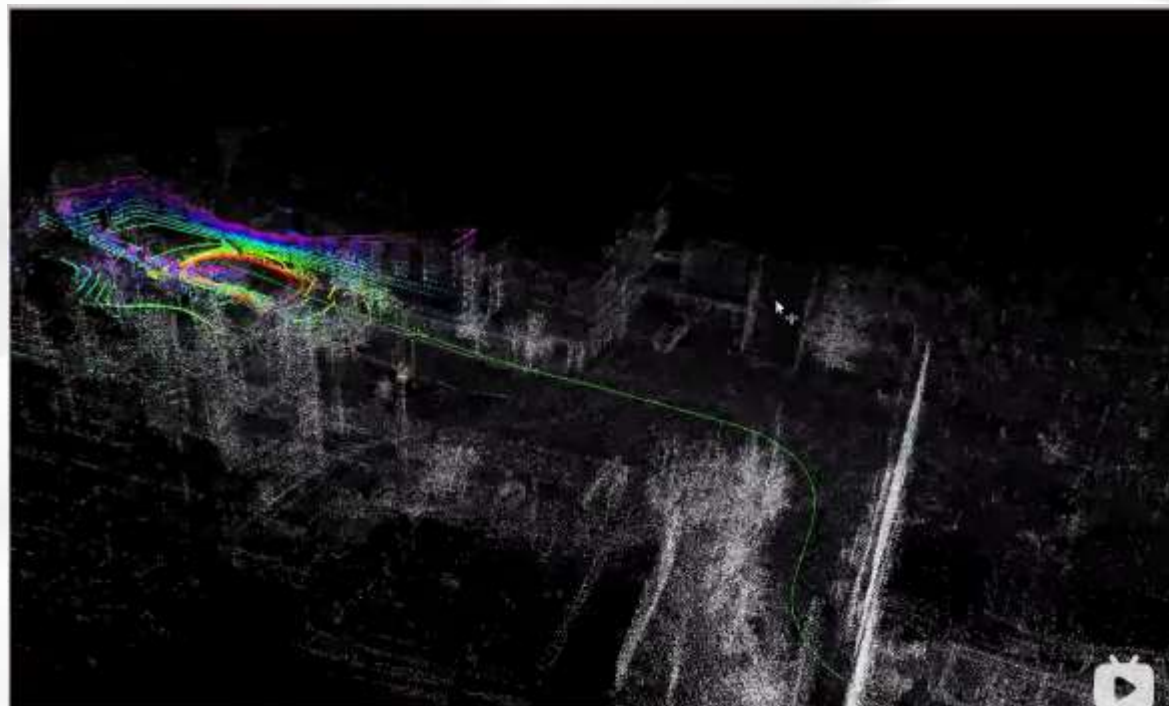
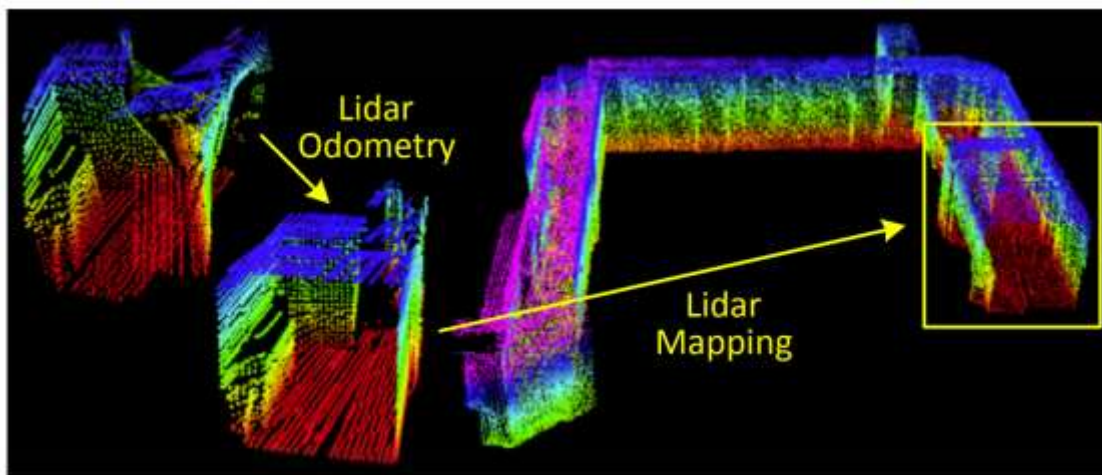
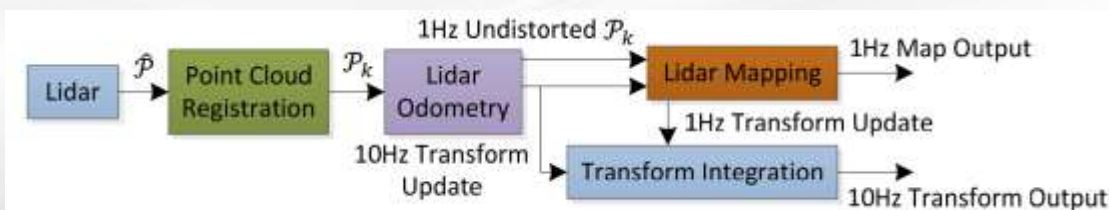
- 激光雷达 (LiDAR, Light Detection and Ranging) 其主要的工作原理是红外光束从光源发射并被场景中的物体反射回探测器, 通过测量光束飞行时间 (Time of Flight, 简称ToF), 可以推算出场景内物体的距离, 并生成距离地图。
- 机械激光雷达: 发射和接收模块被电机电动进行360度旋转。在竖直方向上排布多组激光线束, 发射模块以一定频率发射激光线, 通过不断旋转发射头实现动态扫描;
- 半固态—MEMS式激光雷达: 将原本激光雷达的机械结构通过微电子技术集成到硅基芯片上;
- 半固体—棱镜式激光雷达 (Livox): 激光在通过第一个楔形棱镜后发生一次偏转, 通过第二个楔形棱镜后再一次发生偏转。控制两面棱镜的相对转速便可以控制激光束的扫描形态。
- 固态—OPA激光雷达: 通过控制相控阵雷达平面阵列各个阵元的电流相位, 利用相位差可以让不同的位置的波源会产生干涉;



LOAM: Lidar odometry and mapping in real-time

- Lidar 3D SLAM 的经典之作，一直以来都霸占着 KITTI 的前列
- 主要的四大模块：
 - ❑ Point Cloud Registration: 点云不是同一时刻获取的，每一个帧点云，其中的每一个点，都是不同时刻获取的，因此把它进行运动补偿；获取每个点的时间戳，位姿插值，把点云先投影到同一时刻；提取特征点。
 - ❑ Lidar Odometry: 估计两帧点云之间的位姿变换，获得两个时刻之间的相对位姿，频率较高 10Hz Lidar
 - ❑ Mapping: 建图模块，把连续10帧的点云数据和整个地图匹配，获得世界坐标系下的位姿，频率较低 1Hz。
 - ❑ Transform Intergration: 实时利用世界坐标系下的位姿和两个时刻之间的相对位姿，更新各个时刻世界坐标系下的位姿。

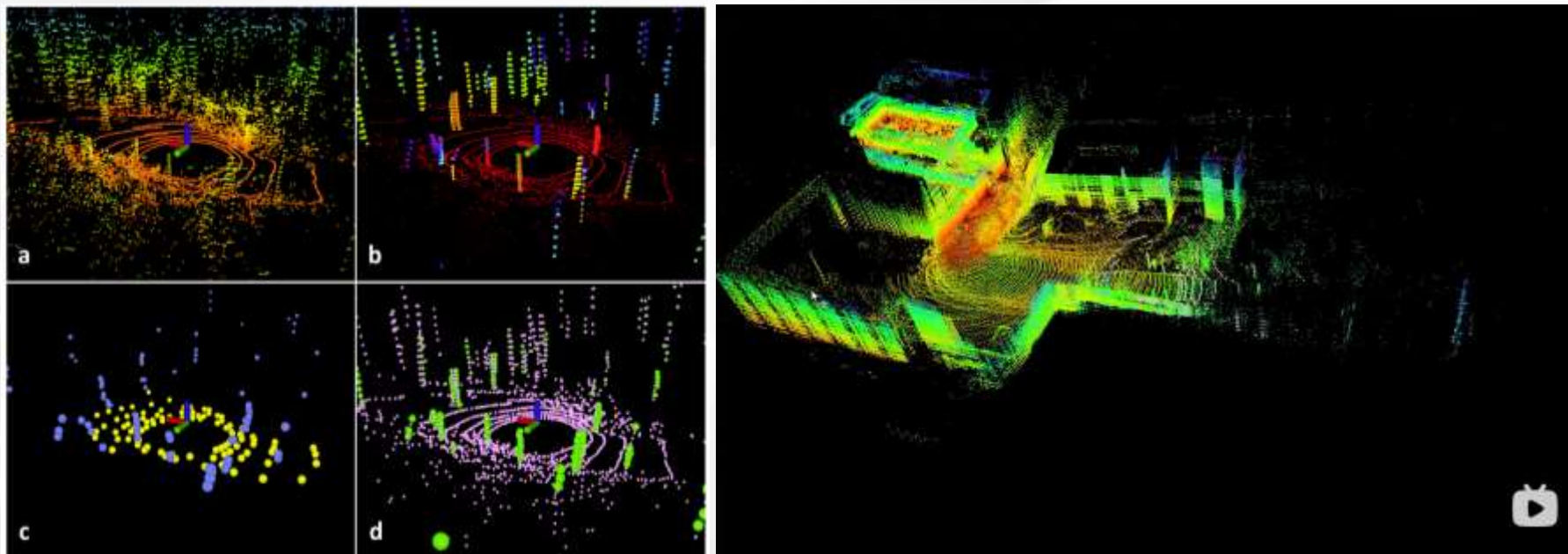
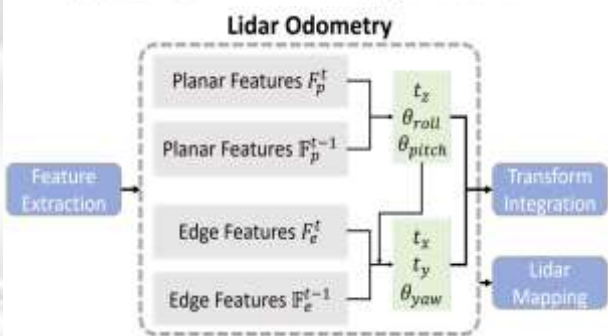
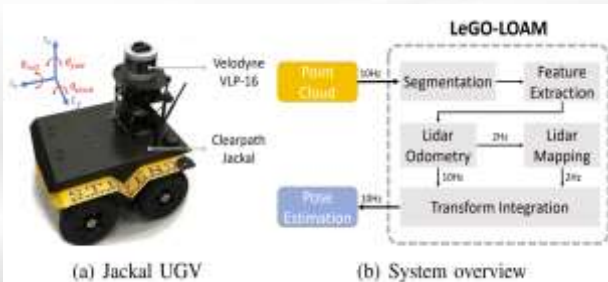
点击图片观看B站视频展示



Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain

- 在原有LOAM基础上，进行轻量级和地面优化两个层面的改进
- 主要框架：
 - ❑ Segmentation: 对点云进行分割，分割为地面和非地面区域；
 - ❑ Feature Extraction: 基于分割后的点，和LOAM类似的算法提取出边缘点和平面点；
 - ❑ Lidar Odometry: 基于提取的特征点，scan-to-scan推算两帧激光之间的相对位姿变换（使用两次LM优化），频率较高（10Hz）；
 - ❑ Lidar Mapping: scan-to-map，构建全局地图，获得世界坐标系下的位姿，频率较低（2Hz）；
 - ❑ Transform Integration: 与LOAM相同，实时利用世界坐标系下的位姿和两个时刻之间的相对位姿，更新各个时刻世界坐标系下的位姿

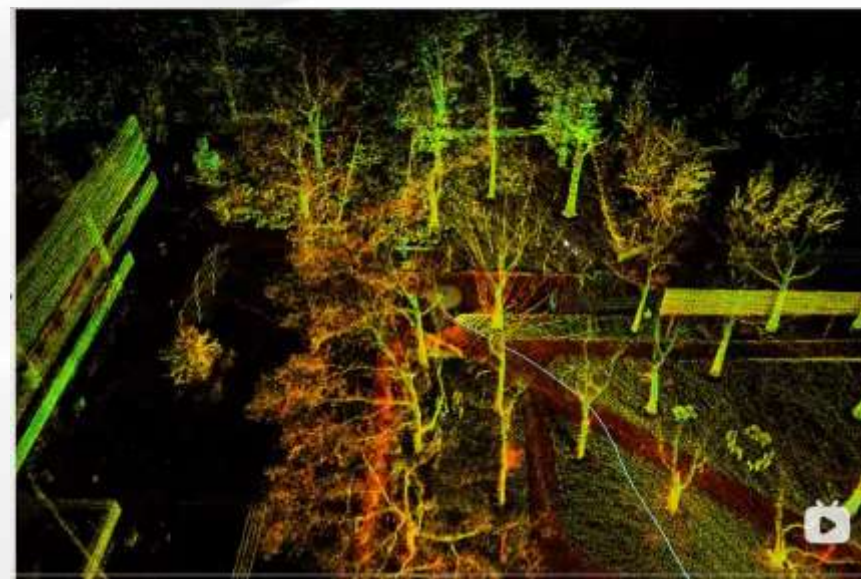
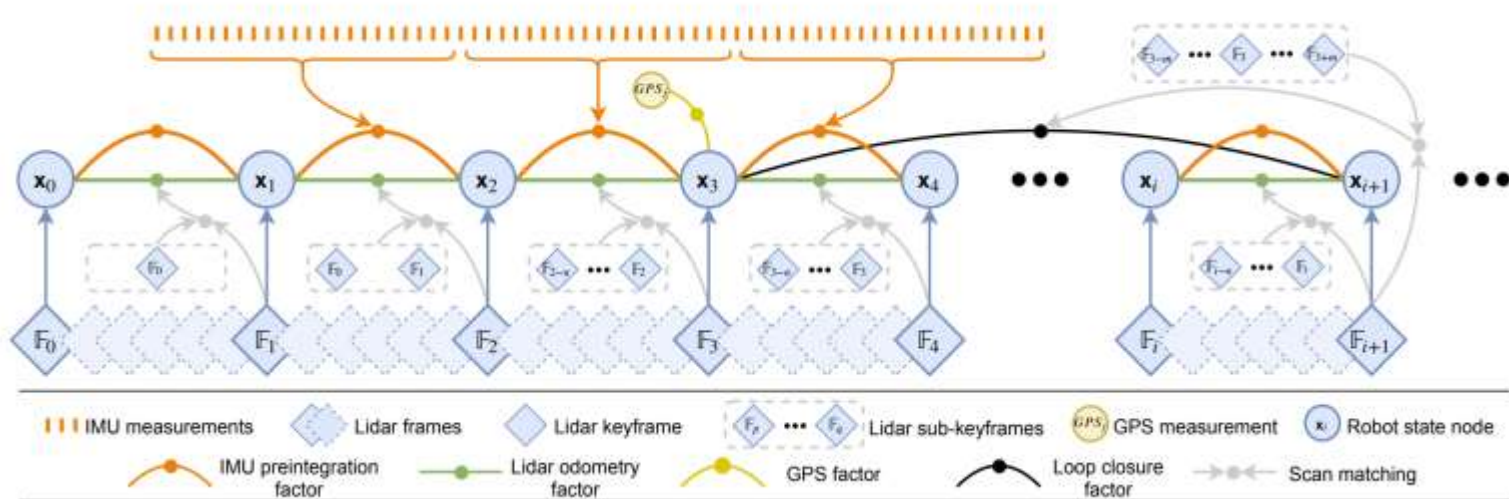
点击图片观看B站视频展示



Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping

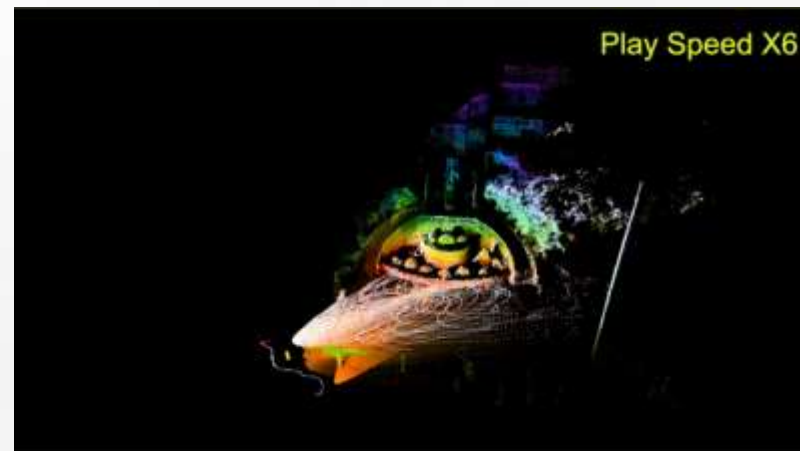
- LIO-SAM是在LeGO-LOAM的扩展。添加了IMU预积分因子和GPS因子：前端使用紧耦合的IMU融合方式，替代原有的帧间里程计，使得前端更轻量；后端沿用LeGO-LOAM，在此基础上融入了GPS观测。同时前端后端相互耦合，提高系统精度；
- LIO-SAM代码简明易读，在Fast-LIO2出来前是LIO相关工作的主要baseline；
- 同期的工作还有LIO-mapping与SC-LIO-SAM，均为基于LiDAR特征点法的图优化框架LIO系统；

点击图片观看B站视频展示

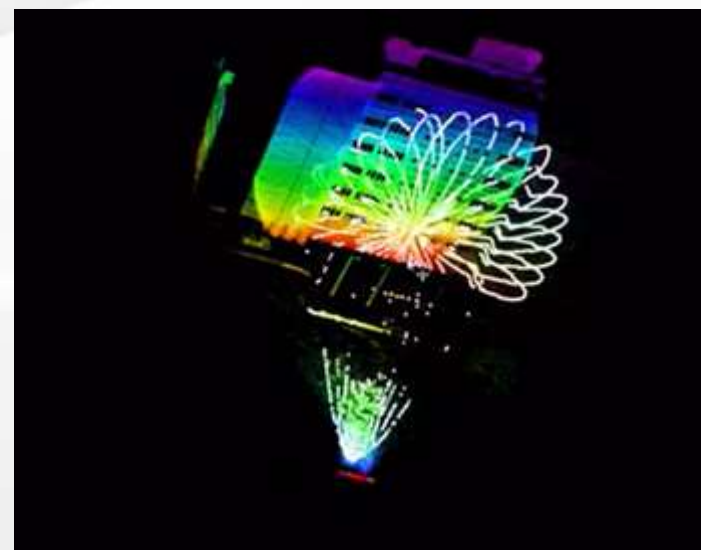
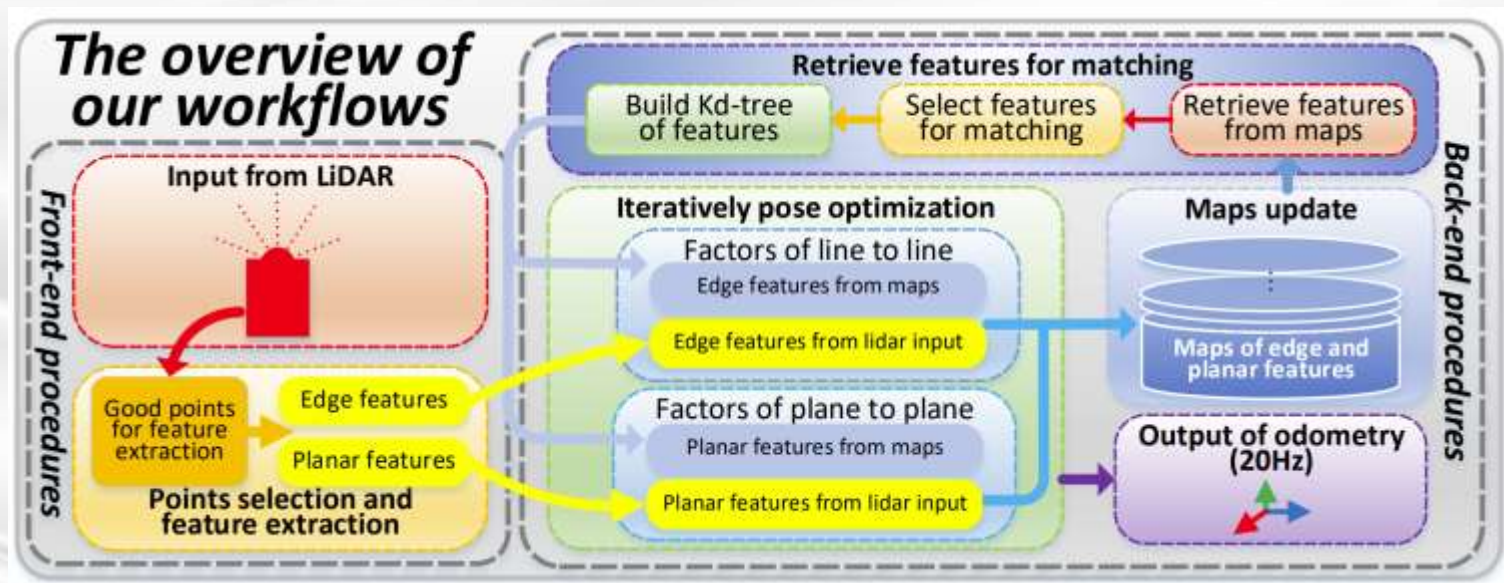


Loam livox

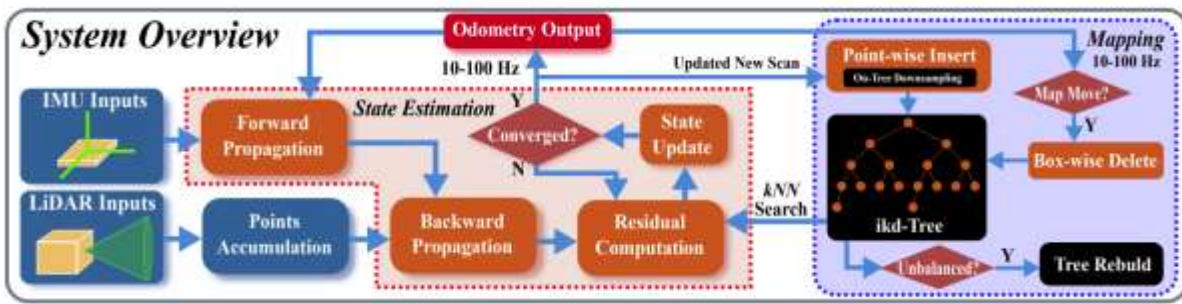
- Livox激光雷达在SLAM上应用的开创性工作；
- 原版loam是基于机械雷达的，当使用固态雷达的时候，loam则不兼容。
- 固态雷达的主要问题是视角比较小，原版loam提取特征则不容易得到高精度里程计，因此主要做了以下改进：
 - 1.基于固态雷达的特征提取和里程解算流程
 - 2.去畸变、特征点的筛选
- 当时livox还是窄角度的（只有不到40度），目前的livox已经有360度的版本了；



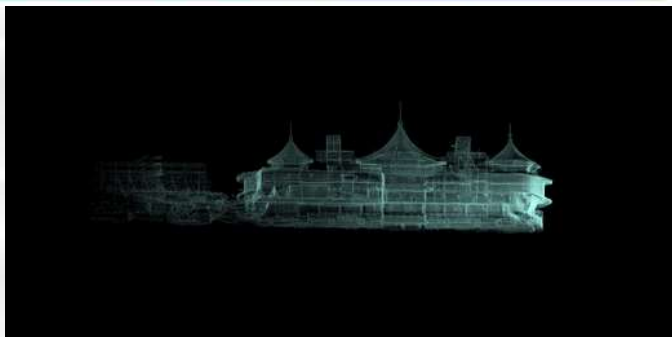
点击图片观看B站视频展示



- 基于高效的紧密耦合迭代卡尔曼滤波器的lidar惯性里程计；
- FAST-LIO2在FAST-LIO的基础上添加了ikd-tree以及去除了特征匹配，采用直接匹配的方式构建残差；
- 直接将原始点注册到地图而不提取特征点。这可以全部利用环境中的所有特征，因此提高了准确性；此外，取消手工设计的特征提取模块也使该框架可以自然适应不同的LiDAR上；
- MARS-LAB团队的开创性工作，后续大量工作都是基于这个架构开发的；
- 在自动驾驶及手持场景下的测试请见([链接](#))

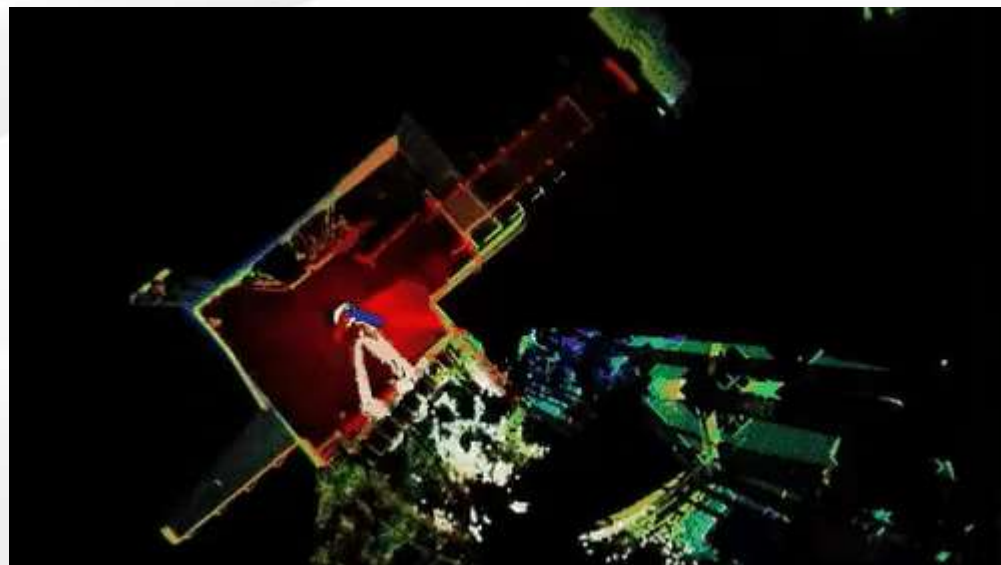
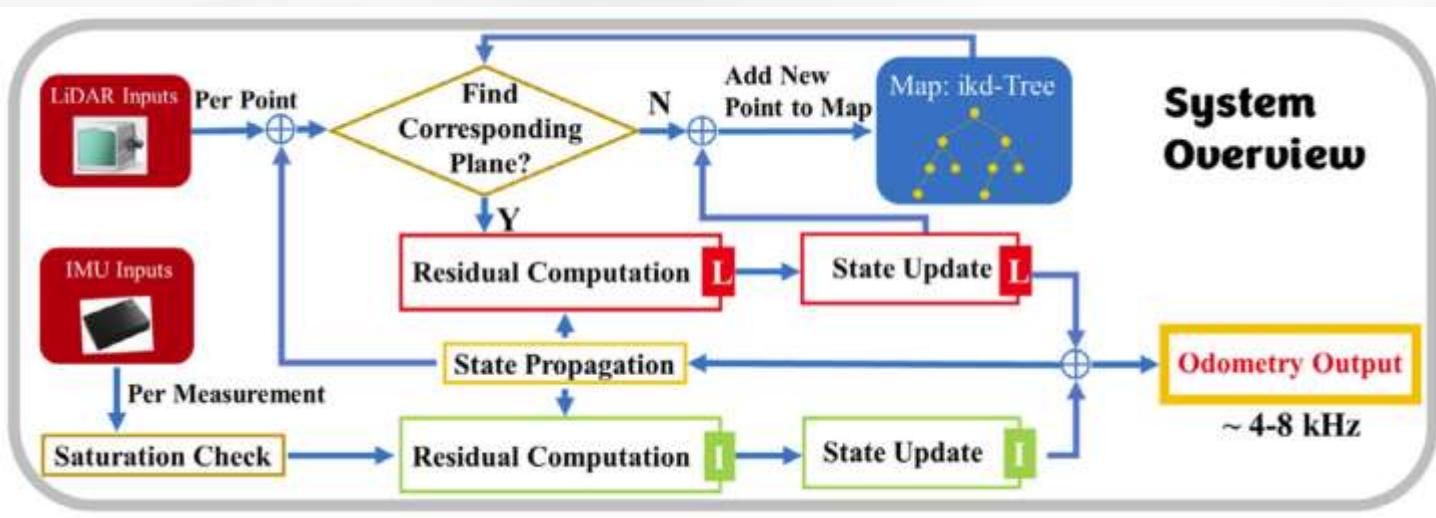


流形科技文
物保护案例



● Point-LIO

- 能在非常剧烈运行下，仍然可以提供稳定姿态估计的LiDAR惯性里程计，提供精确的高频里程计（4-8kHz）；
- Point-LIO有两个关键：
 1. 逐点LIO框架，更新与每个LiDAR点测量对应的状态，该框架允许极高频率的里程表输出，显著增加了里程表带宽，并从根本上消除了人为的帧内运动畸变；
 2. 随机过程增强的运动学模型，将IMU测量作为输出建模。这种新的建模方法即使在IMU测量在运动中间饱和的情况下，也能对猛烈运动进行准确的定位和可靠的建图。
- Point-LIO在自动驾驶场景下的测试([链接](#))；



高速旋转下仍然保存稳定的姿态估计

Immesh: An immediate lidar localization and meshing framework

- 一种用于实时同步定位和网格划分的新型LiDAR-IMU SLAM。
- ImMesh主要由四个模块组成：接收器、定位、网格和广播器。
 - 定位模块处理传感器数据，估计姿态并动态更新地图。
 - 网格划分模块使用高效的体素结构(Voxel map)，从LiDAR扫描中增量重建三角形网格。
 - 体素网格化操作可以提高效率，将3D点投影到2D平面并逐步重建三角网格。
 - 通过广播发布实时里程计、地图和网格。仅使用普通CPU即可实现大规模实时网格重建。
- 超大场景下测试ImMesh ([链接](#))

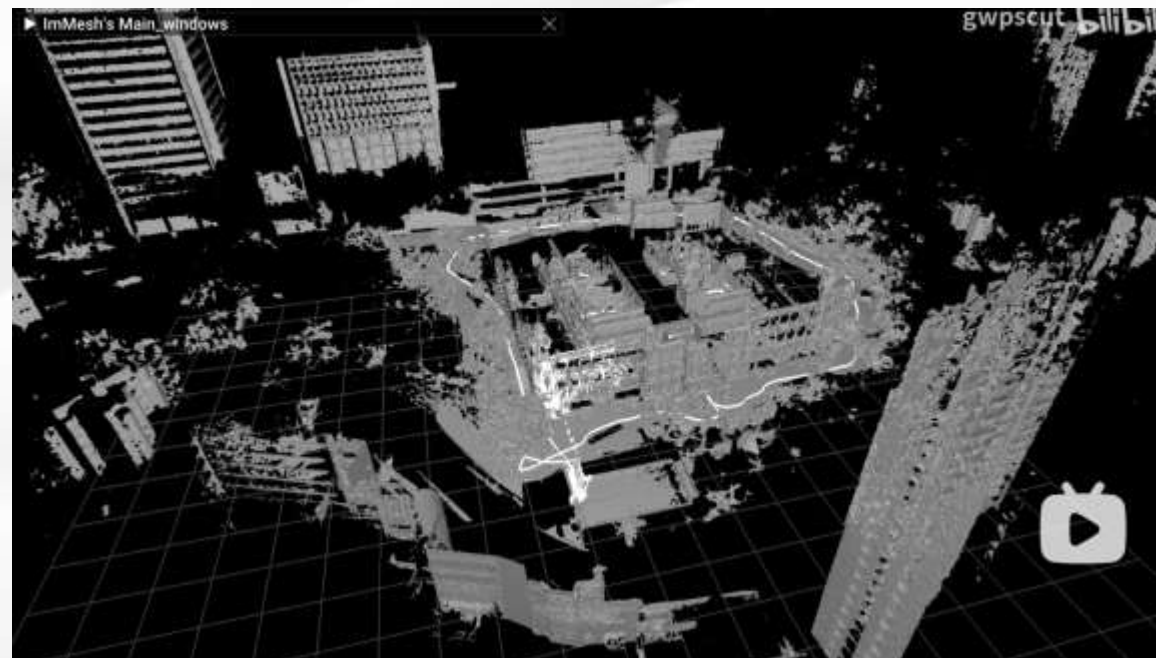
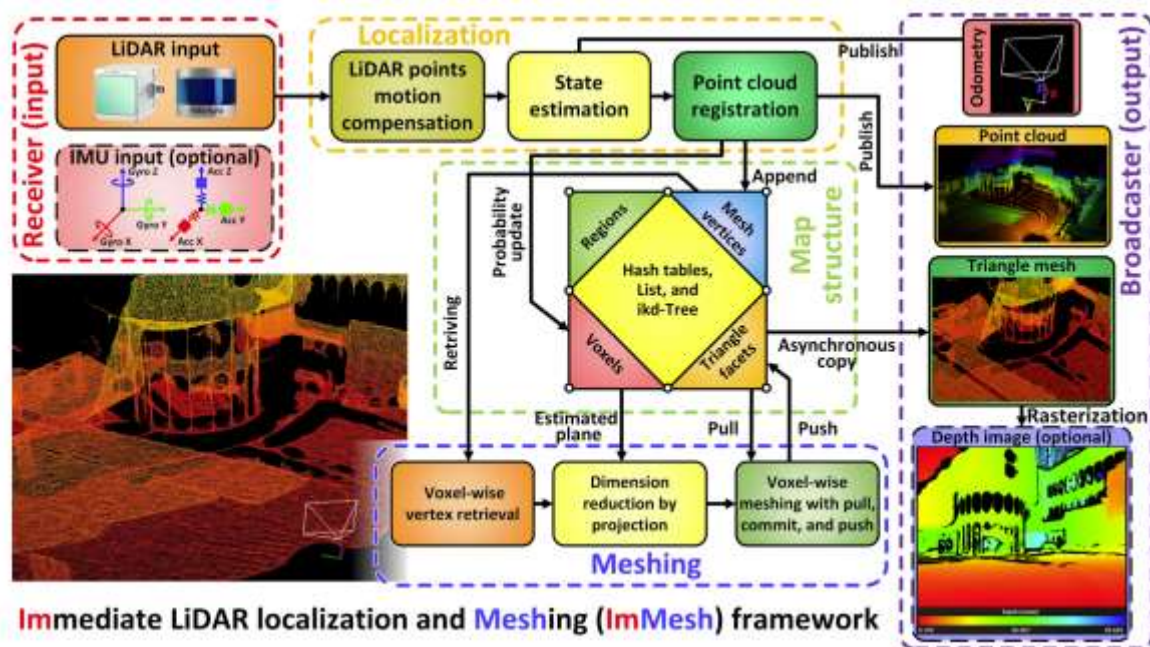


Image-based 3D SLAM

图像传感器

- 图像传感器：单目、双目、单目结构光、双目结构光、ToF、鱼眼相机等；
- 其他视觉传感器：事件相机、红外相机



单目



双目



深度相机



Military ▲



Security ▲



Outdoors ▲



Rescue ▲



Hunting ▲

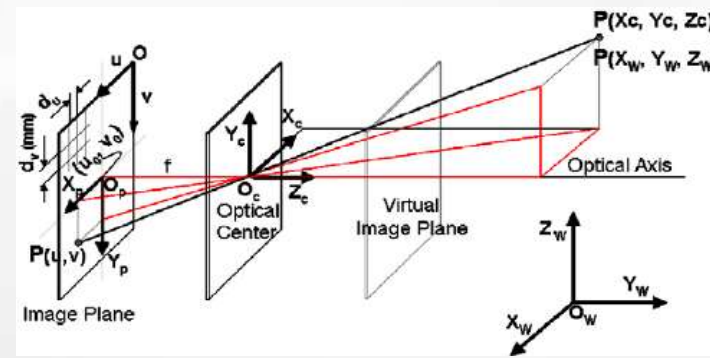


Guard ▲

红外相机

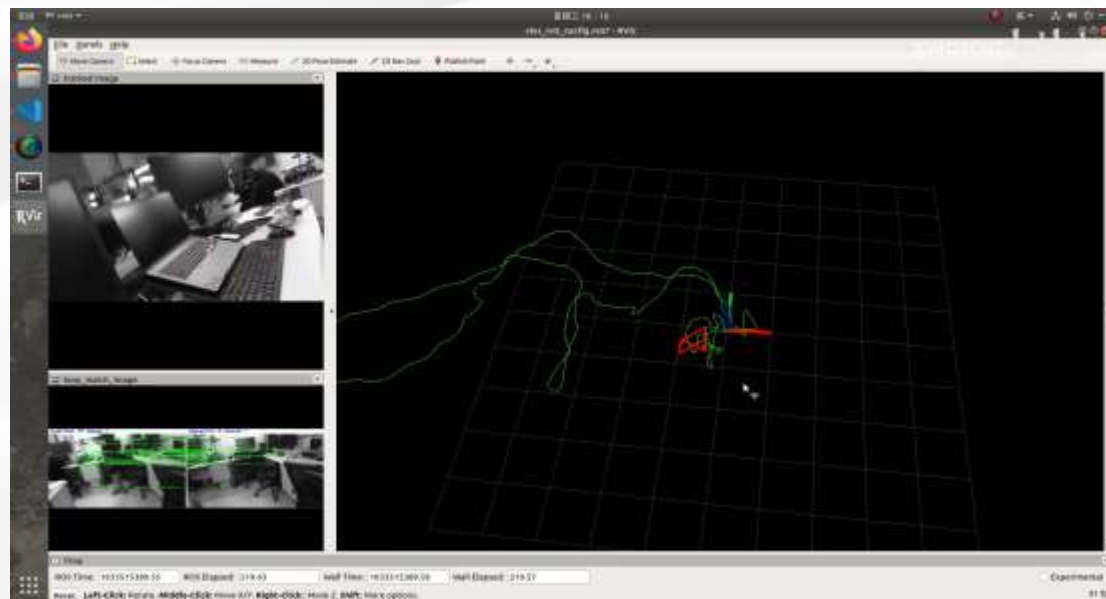
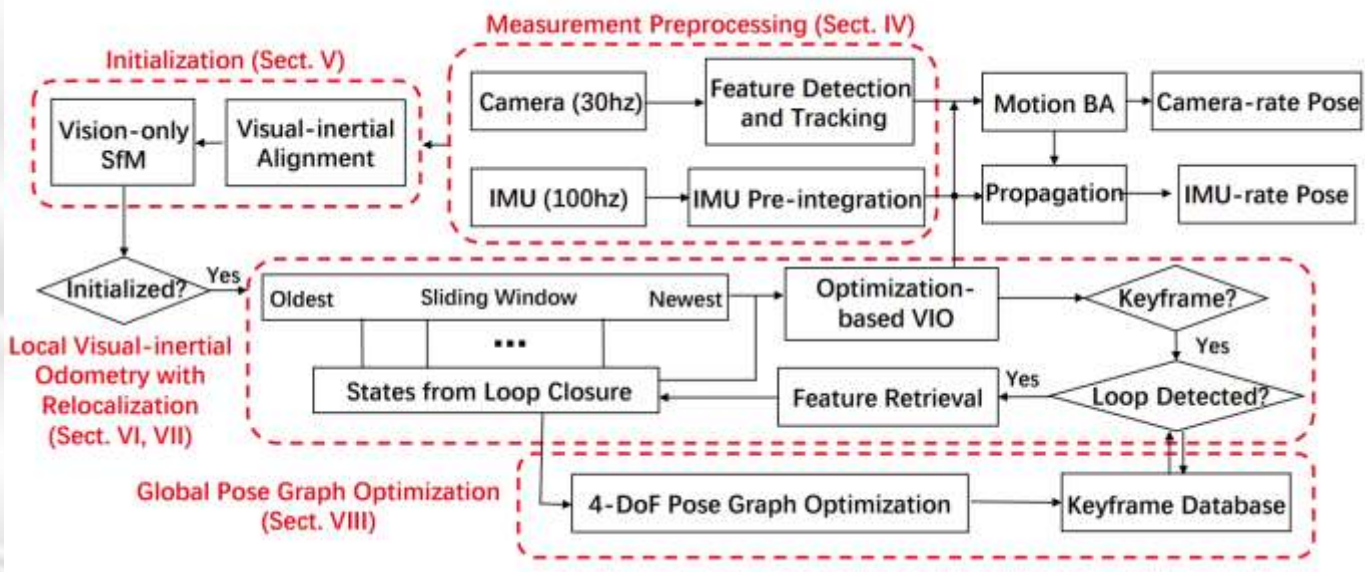


事件相机



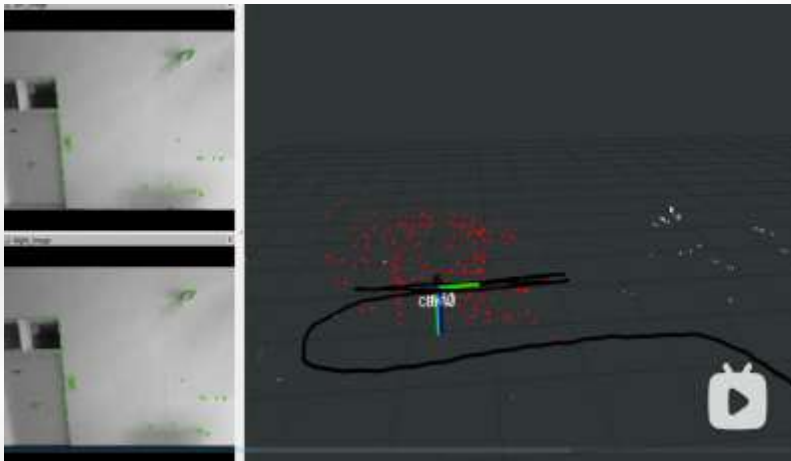
投射投影模型与相机标定

- 最基本也是最经典的VIO框架，由港科大学开源的一个VIO算法，是一种具有鲁棒性和通用性的单目视觉惯性状态估计器；通过紧耦合方法实现的，通过单目+IMU恢复出尺度，效果非常好；网上有大量的解读，如[博客](#)；
- 双目版本：VINS-Fusion；GPS+单目版本：GVINS；
- VINS-Mono [测试](#)；VINS-Fusion [测试](#)；GVINS [测试](#)；基于VINS-MONO开发的单目稠密SLAM系统（[代码链接](#)）；

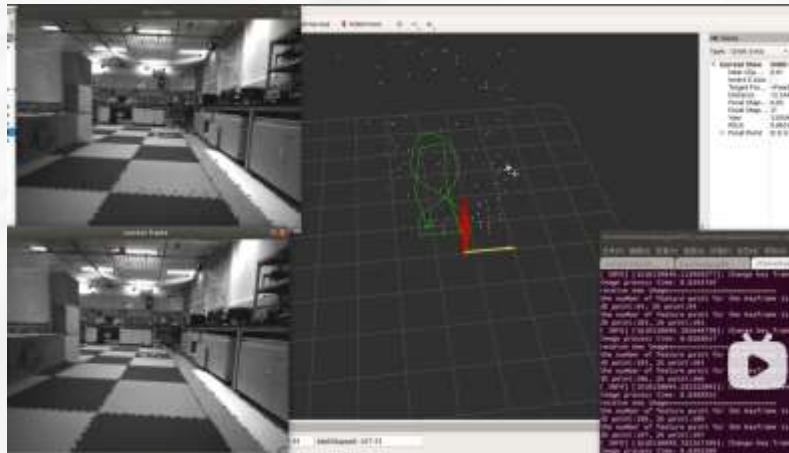


● 其他视觉SLAM方法

- 基于特征点法的视觉SLAM，如，ORB-SLAM1~3, OPEN-VINS等；
- 基于直接法的视觉SLAM，如SVO, DSO, DM-VIO, EN-VIO等；



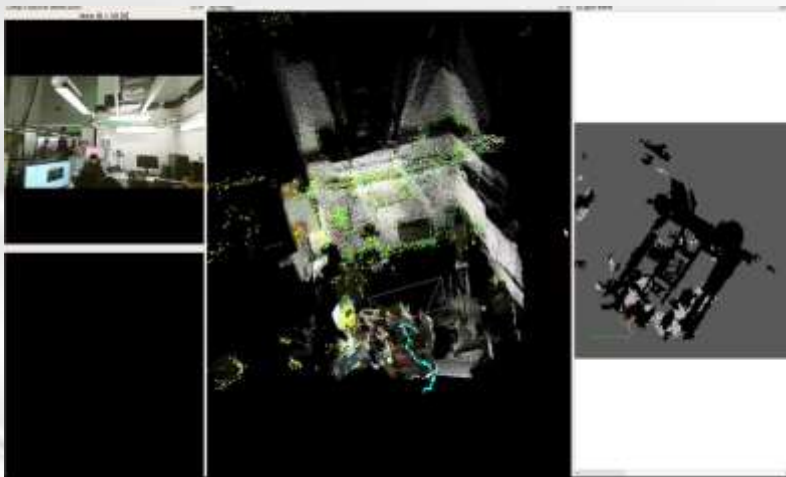
ENVI0, 白墙下稳定定位, 点击观看B站视频



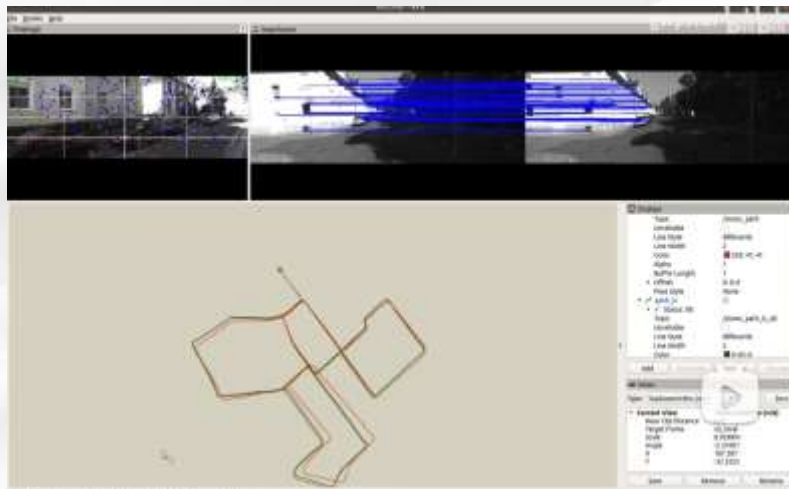
手写的双目视觉惯性里程计, 点击观看B站视频



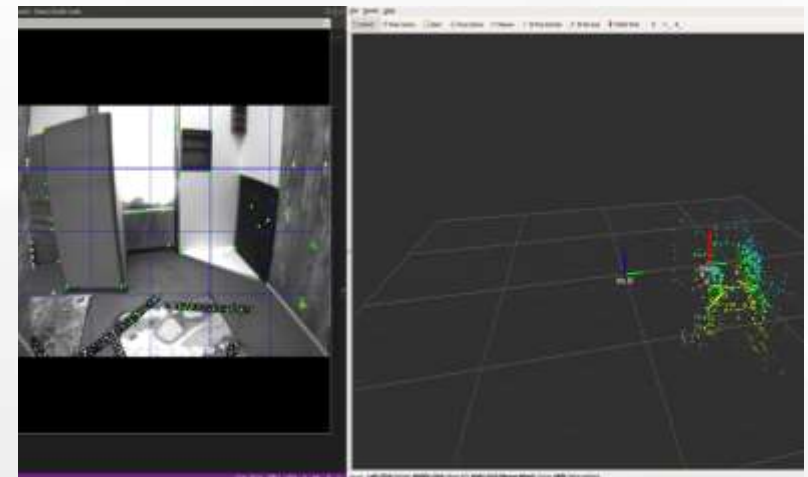
SVO, 点击观看B站视频



基于Kinect的RTAB-Map, 点击观看B站视频



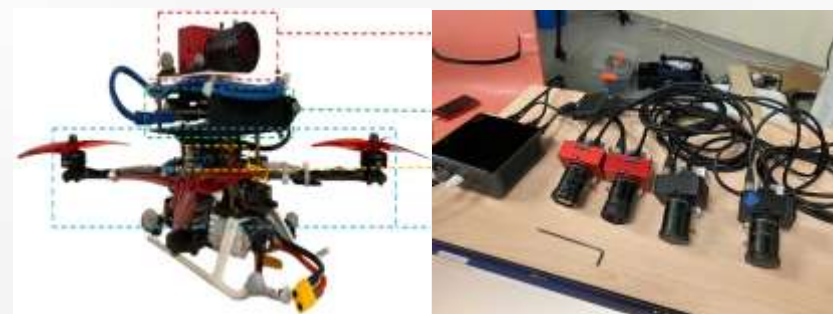
FLVIS, 点击观看B站视频



MSCKF-VIO, 点击观看B站视频

Event Camera

- 传统相机在运动模糊以及强光或者弱光等挑战性的场景下，难以捕捉有效的信息；
- 事件相机只要有光强变化即可异步触发感知信息，具有响应速度快，高动态范围等特点优势，适用于快速运动、光照不好等情况下的SLAM问题；
- 而事件相机的motion-active的特性也是非常适合用于足式机器人高振动、跳跃等剧烈运动的场景的；
- 针对机械狗、无人机等场景已有大量的使用事件相机来做感知的工作；
- 更多事件相机相关的介绍请见[博客](#)；



Latency & Motion blur

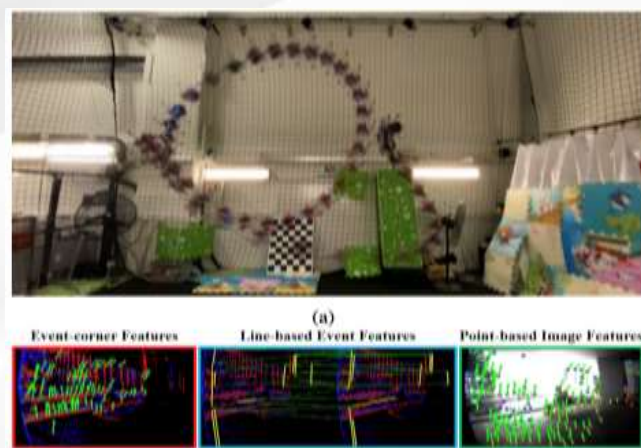
Dynamic Range



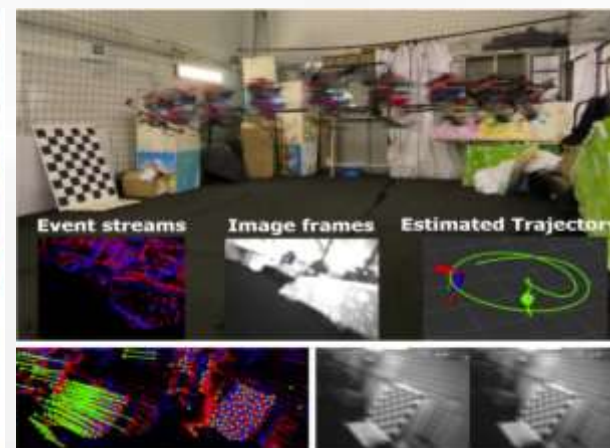
H. Rebecq [BMVC17]



Ultimate-SLAM [RAL 18]



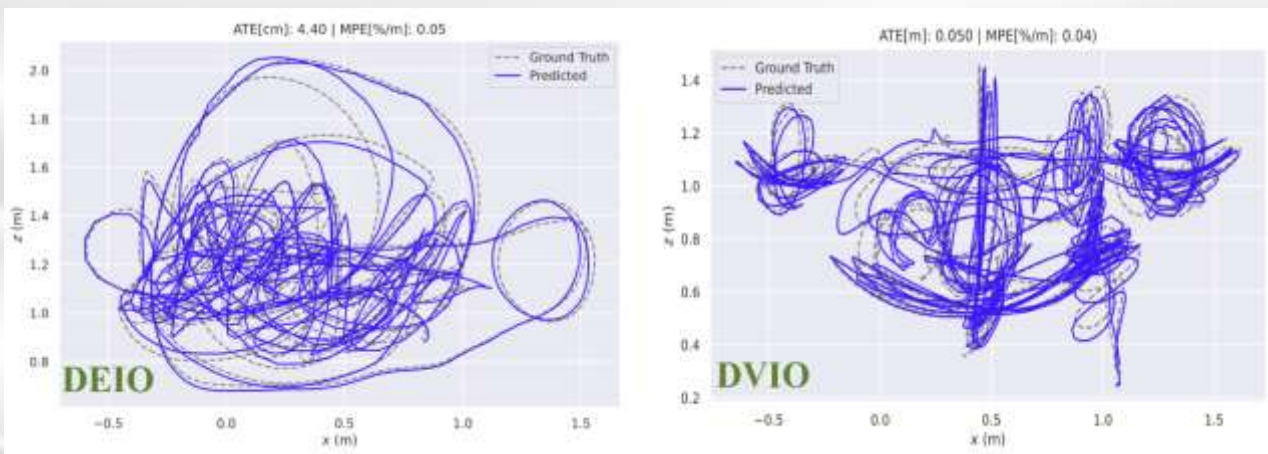
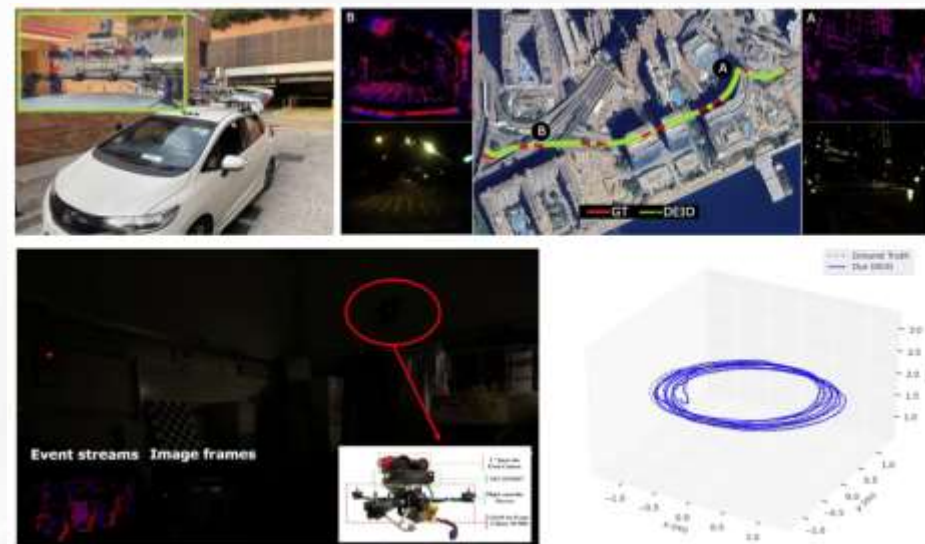
PL-EVIO [TASE 23]



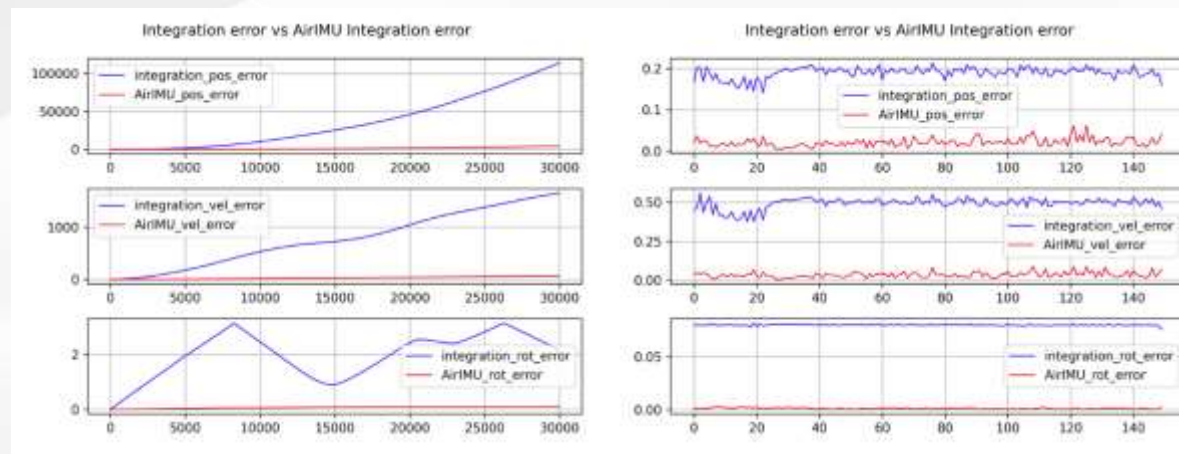
ESVIO [RAL 23]

● Learning-based VO, VIO, IO

- 利用深度学习来对视觉数据进行数据关联，提升VIO的鲁棒性及精度；此外也可以对IMU的噪声进行补偿；
- DEIO世界首个基于深度学习的事件惯性里程计，将深度学习与传统非线性图优化相结合，在**10**个挑战性数据集下超越**20**多种方法方法；
- 该深度学习与传统非线性图优化相结合的框架也适配基于图像的视觉惯性里程计；
- 仅在仿真的数据集训练，迁移到真实场景，具有强大的泛化能力；
- 更多关于基于深度学习的VO、VIO、IO的介绍请见[博客](#)；



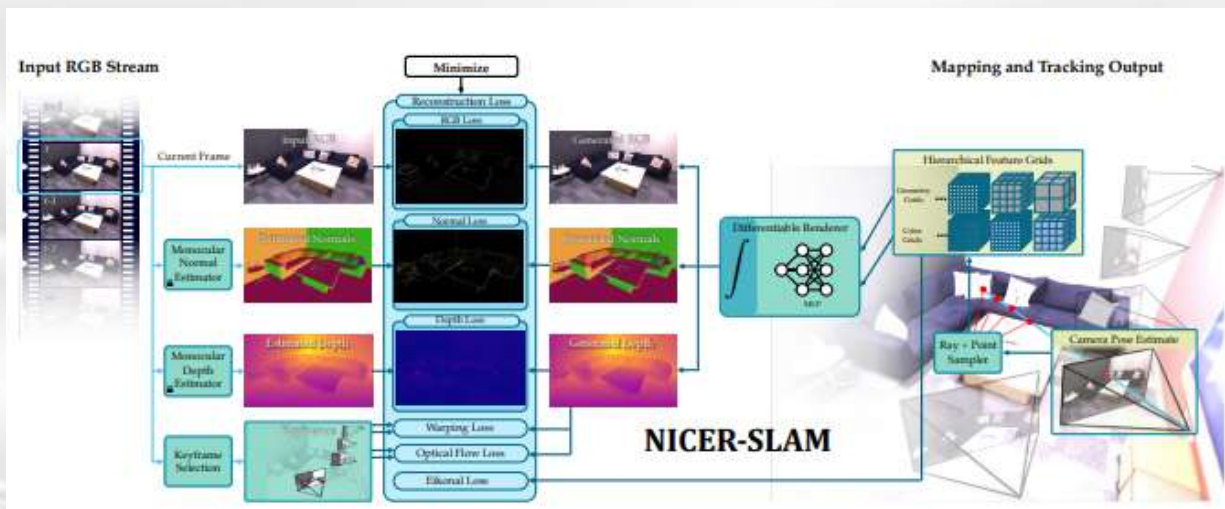
HKU ArcLab开发的基于图像/事件相机的惯性里程计



基于深度学习对IMU漂移进行补偿

● NeRF-based SLAM

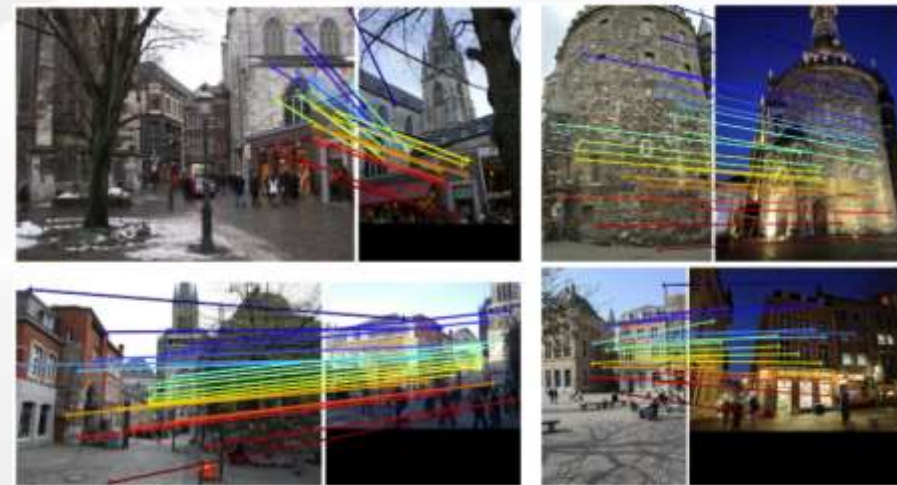
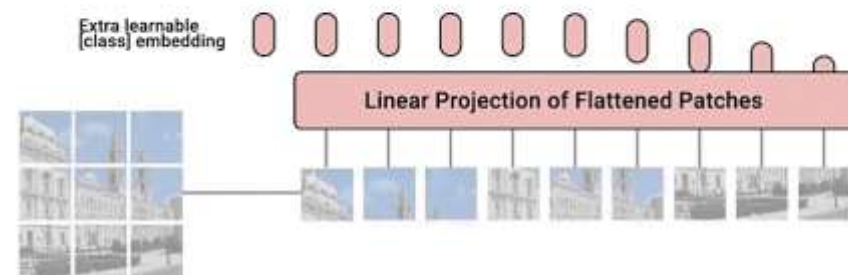
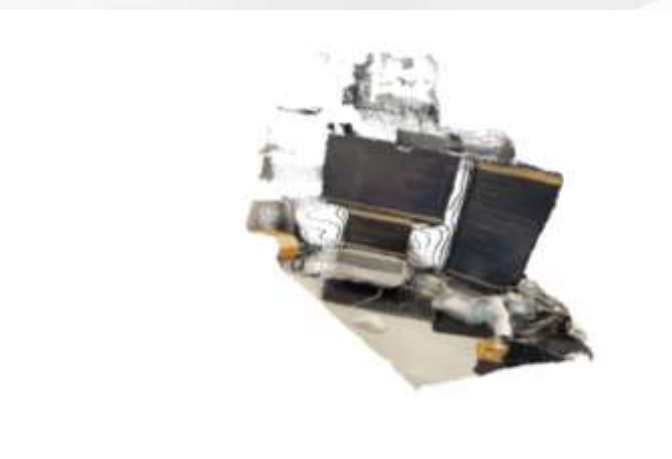
- NeRF 所做的任务（新视角合成），即在若干已知视角下对场景进行一系列的观测（相机内外参、图像、姿态等），合成任意新视角下的图像。
- NeRF 的思想:通过输入视角的图像每个像素的射线对于密度（不透明度）积分进行体素渲染，然后通过该像素渲染的 RGB 值与真值进行对比作为 Loss。由于设计的体素渲染是完全可微的，因此该网络可学习；
- 基于神经辐射场的SLAM是一种结合神经辐射场（NeRF）和SLAM的先进技术，用于实时地构建三维环境地图并同时估计相机的姿态。目前NeRF-SLAM主要有以下两个方向：
 - SLAM为NeRF训练提供位姿，然后建立稠密细腻的三维场景。简而言之就是NeRF只做mapping；
 - 在NeRF里建立各种损失函数反过来优化pose和depth。简而言之就是full slam；
- 更多关于基于神经辐射场的SLAM的工作介绍请见[博客](#)；



Result from NICE SLAM

Transformer-based SLAM

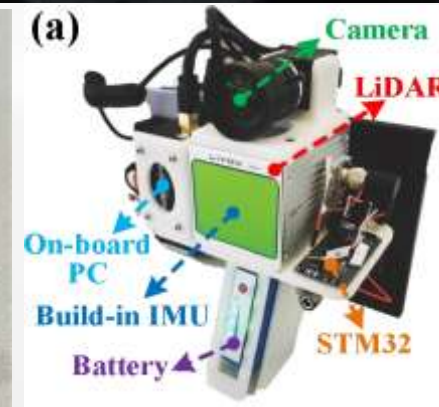
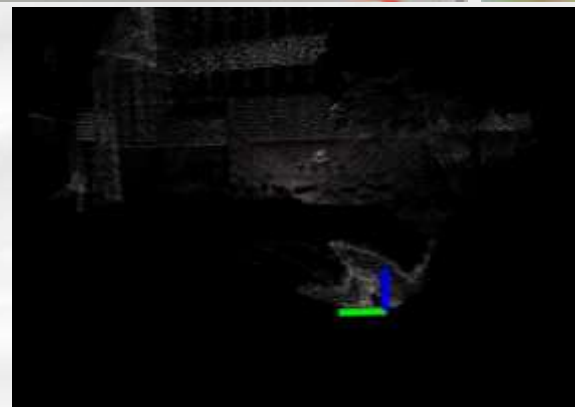
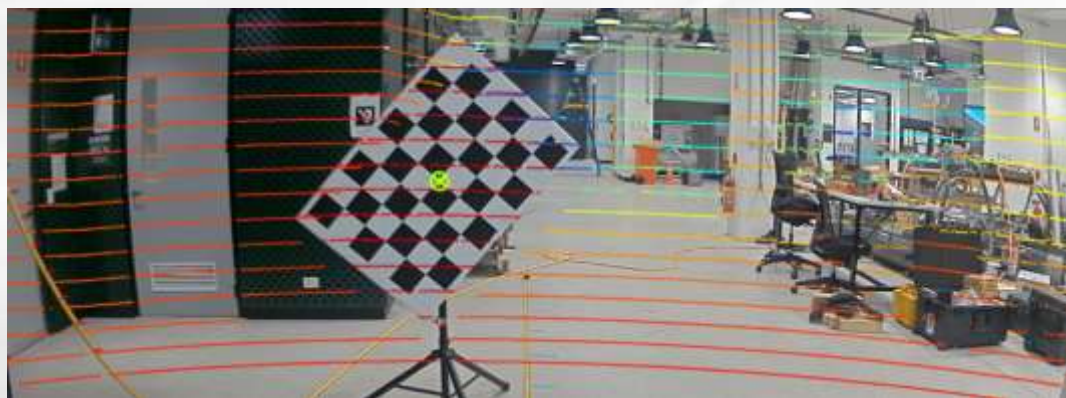
- 从NLP到CV具有大量研究的Transformer结构，利用其时间与空间维度的关联能力，在三维重建、位姿估计、特征匹配等领域具有巨大的潜力；
- 在极端视点变换的情况下，仍然可以实现良好的匹配效果；
- 2张图片即可实现三维重建；
- 仅图片/视频输入即可估算full SLAM（无需内参）；
- 从三维空间上实现点的匹配，构建SLAM系统的新范式；
- Transformer是目前视觉SLAM种的佼佼者，更多介绍请见[博客](#)；



LiDAR + Image 3D SLAM

Motivation

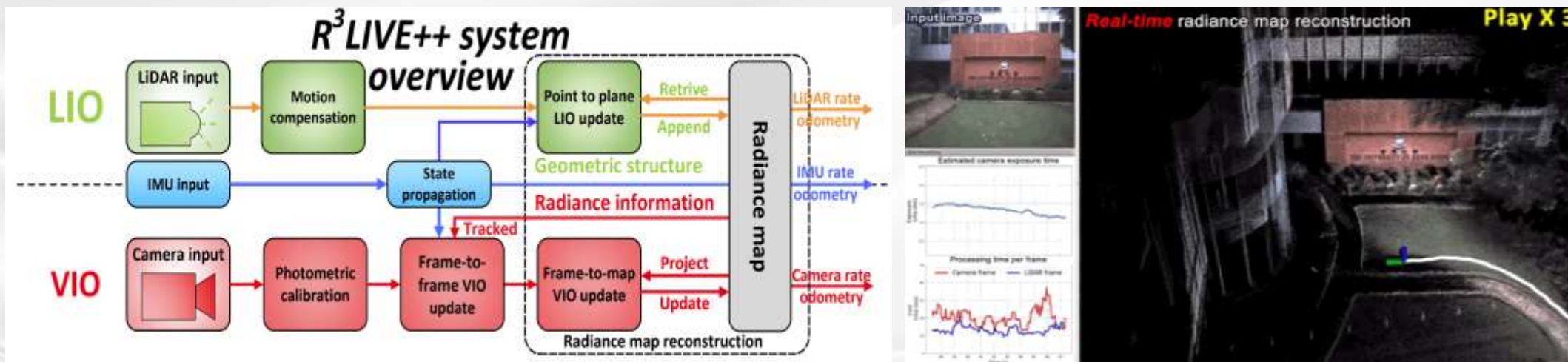
- LiDAR-图像融合SLAM是结合了LiDAR和视觉传感器的优势，提供高度准确和强大的定位和建图功能。
- LiDAR-图像融合利用了激光雷达的精确距离测量和摄像头捕获的丰富环境细节，从而在各种具有挑战性的环境中提高了性能；
- 传感器的标定是关键；



Result from SR LIVO

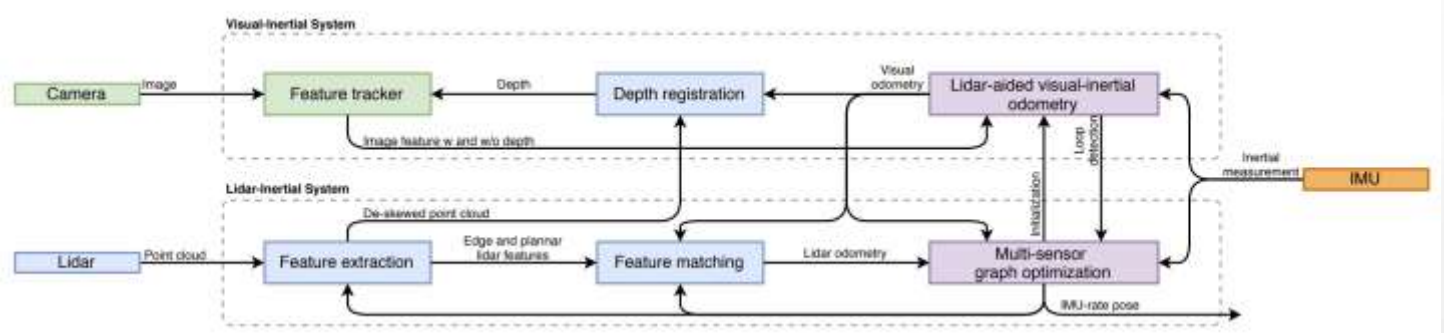
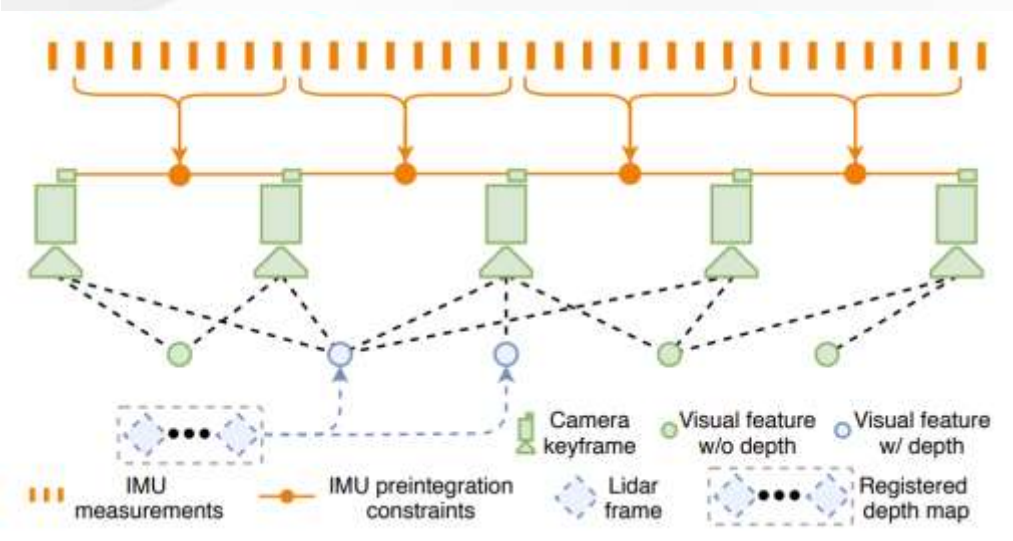
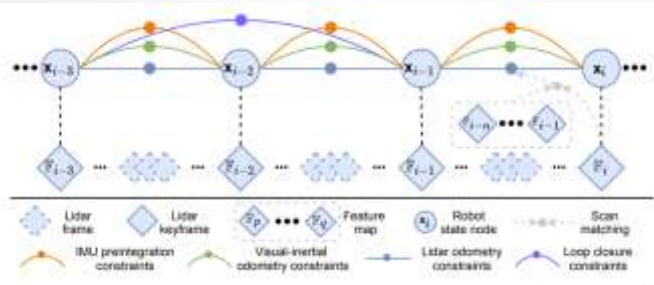
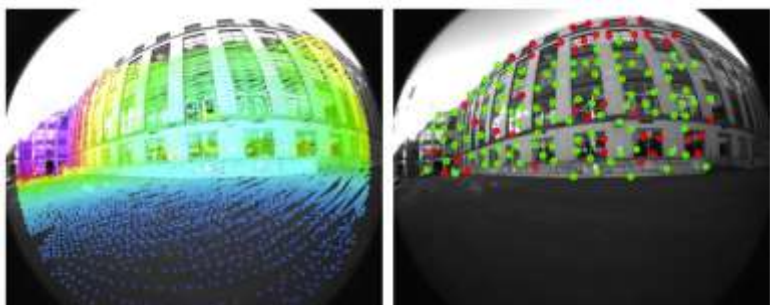
R2LIVE, R3LIVE, R3LIVE++

- 经典的雷达-视觉-惯性SLAM框架，用于鲁棒和精确的状态估计和实时辐射图重建。集成了用于几何结构的实时LiDAR惯性测距（LIO）和用于辐射信息的视觉惯性测距（VIO）；
- R2LIVE系列是首个提出雷达-视觉-惯性SLAM框架，并且在LVI建图上具有开创性意义；
- R2LIVE测试([链接](#))； R3LIVE测试([链接](#))； R3LIVE++测试([链接](#))；



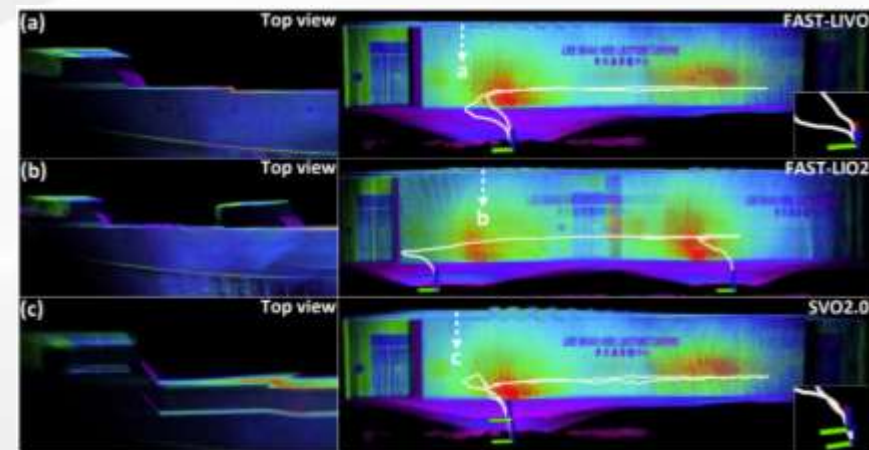
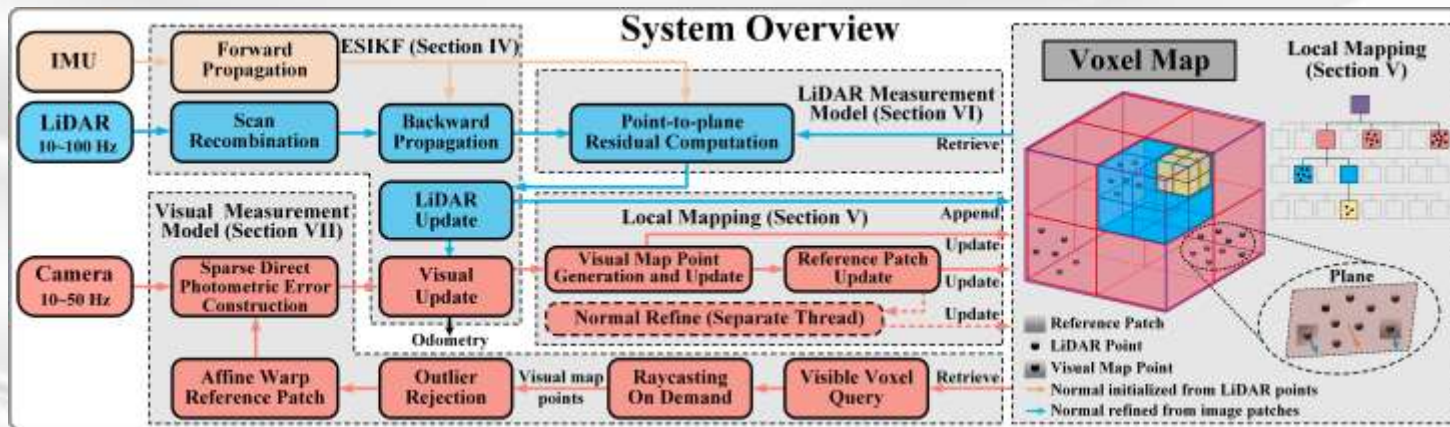
Lvi-sam: Tightly-coupled lidar-visual-inertial odometry via smoothing and mapping

- 雷达-视觉-惯性里程计：LIO系统为LIO-SAM，VIO系统为VINS-mono；
- 实际代码实现种，两者个子系统是松融合的；视觉系统会利用激光雷达提供的深度信息，而LIO系统利用VIO提供的信息来进行畸变去除及回环等；
- LVI-SAM测试：作者提供的数据集([链接](#))；自动驾驶场景([链接](#))；



FAST-LIVO and FAST-LIVO2

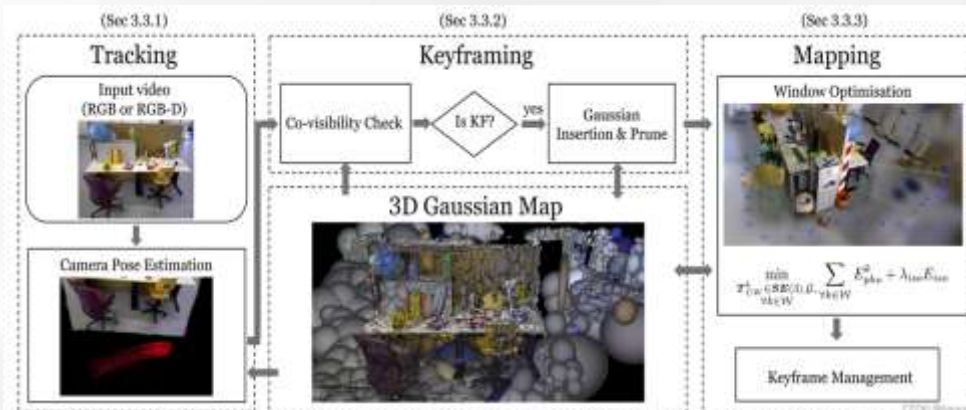
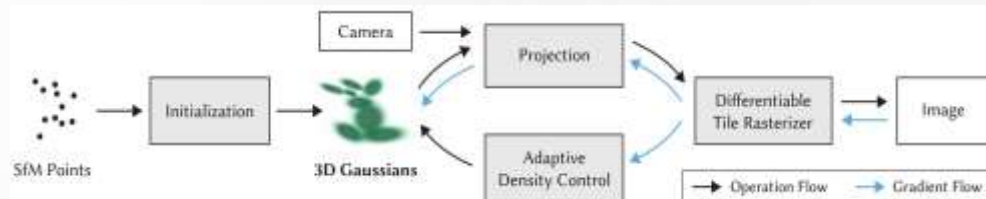
- 基于直接法的雷达视觉惯性里程计；
- 通过利用LiDAR与image的互补性，很好的处理纯Lidar或者存image下的退化场景；
- 基于误差状态迭代卡尔曼滤波（ESIKF）和顺序更新策略融合IMU、LiDAR和图像观测信息。图像与Lidar都是采用直接法（避免额外提取特征的开销同时也可以保证在少纹理或结构的场景提供更多的信息）；此外，引入Voxelmap（体素图）进行几何结构和图像对齐。利用Lidar平面先验和动态参考补丁更新来提高图像对齐精度。
- FAST-LIVO测试([博客](#))； FAST-LIVO2测试([博客](#))；



退化场景测试

3DGS-based SLAM

- 3D Gaussian Splatting (3DGS)可以理解作为一种新的地图表征形式，它的特点在于重建质量高的情况下还能接入传统光栅化进行实时渲染，且优化速度快。
- 其拓展有2DGS, 4DGS, CVPR2025上最新的GS相关工作已经宣称得到1000+ FPS的动态场景的渲染；
- 3DGS在三维重建，基于LiDAR、图像、深度相机等SLAM领域均有大量的研究。详细请见博客：
 1. [3D Gaussian Splatting及其在SLAM与自动驾驶上的应用调研](#)
 2. [基于图像、LiDAR、事件相机相关的3DGS-SLAM](#)
 3. HKU ArcLab研发的基于3D高斯的LiDAR-视觉惯性SLAM的演示请见[网页](#)；
- 3DGS的应用瓶颈主要有：内存消耗、训练视觉与新视觉的差异、高斯球的各项异性、硬件需求等；



基于足式机器人的SLAM



Comparative Evaluation of RGB-D SLAM Methods for Humanoid Robot Localization and Mapping

- 通过对比评估了三种RGB-D SLAM算法(ORB-SLAM3, RTAB-MAP, OpenVSLAM)在SURENA-V人形机器人的定位和地图构建任务中的性能;
- 在定位精度方面, ORB-SLAM3表现最佳, 其ATE为0.1073, 次之为RTAB-Map (0.1641) 和OpenVSLAM (0.1847);

TABLE I: Comparison of representative visual slam methods

	RTAB-Map	ORB-SLAM3	OpenVSLAM
ROS Compatible	✓	✓	✓
Feature Matching	GFTT, BRIEF	FAST, BRIEF	FAST, BRIEF
Map Output	Dense, OctoMap, Occupancy Grid	Sparse	Sparse
Loop Closure	✓	✓	✓
Camera Tracking	Indirect	Indirect	Indirect
Implementation	C++	C++	C++
Publication Year	2019	2021	2019
Sensor Type	Mono, Stereo, RGB-D, Fisheye, LiDAR, External Odometry	Mono, Stereo, RGB-D, Fisheye, IMU	Mono, Stereo, RGB-D, Fisheye, Equirectangular



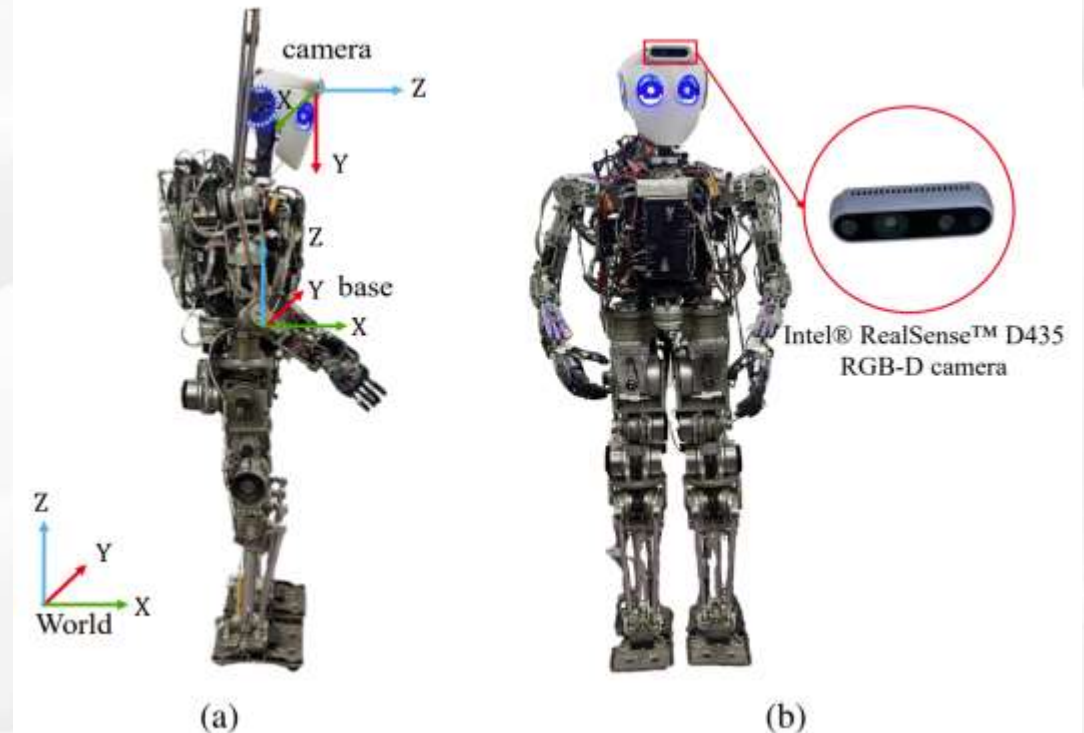
(a)



(b)



(c)

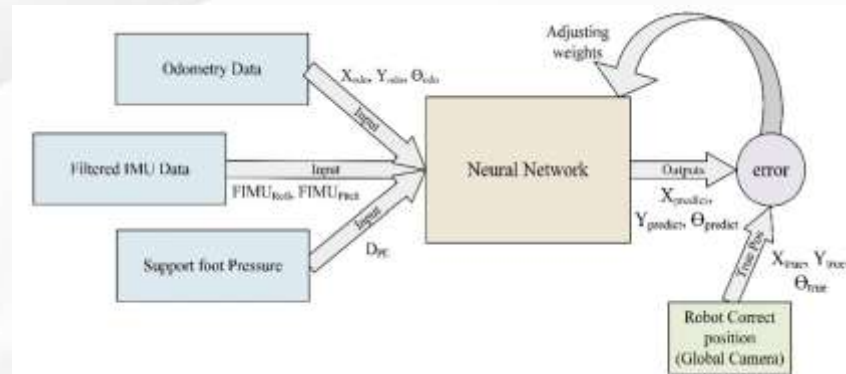
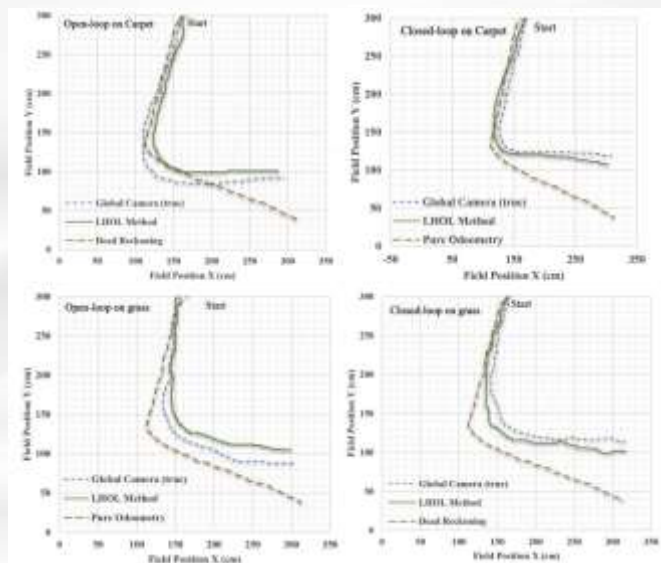
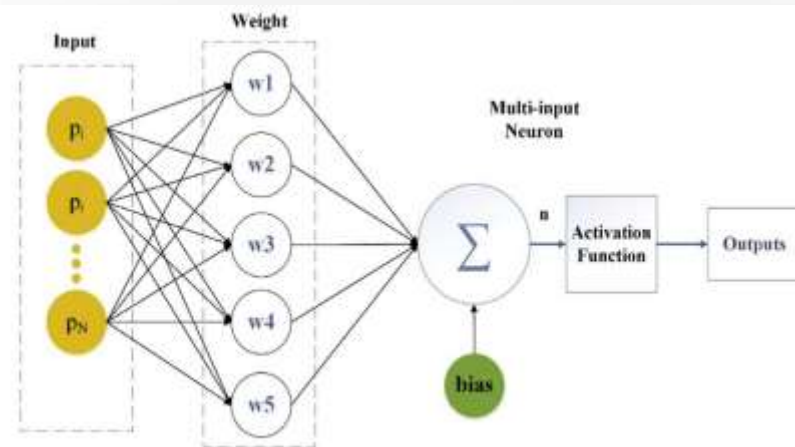


(a)

(b)

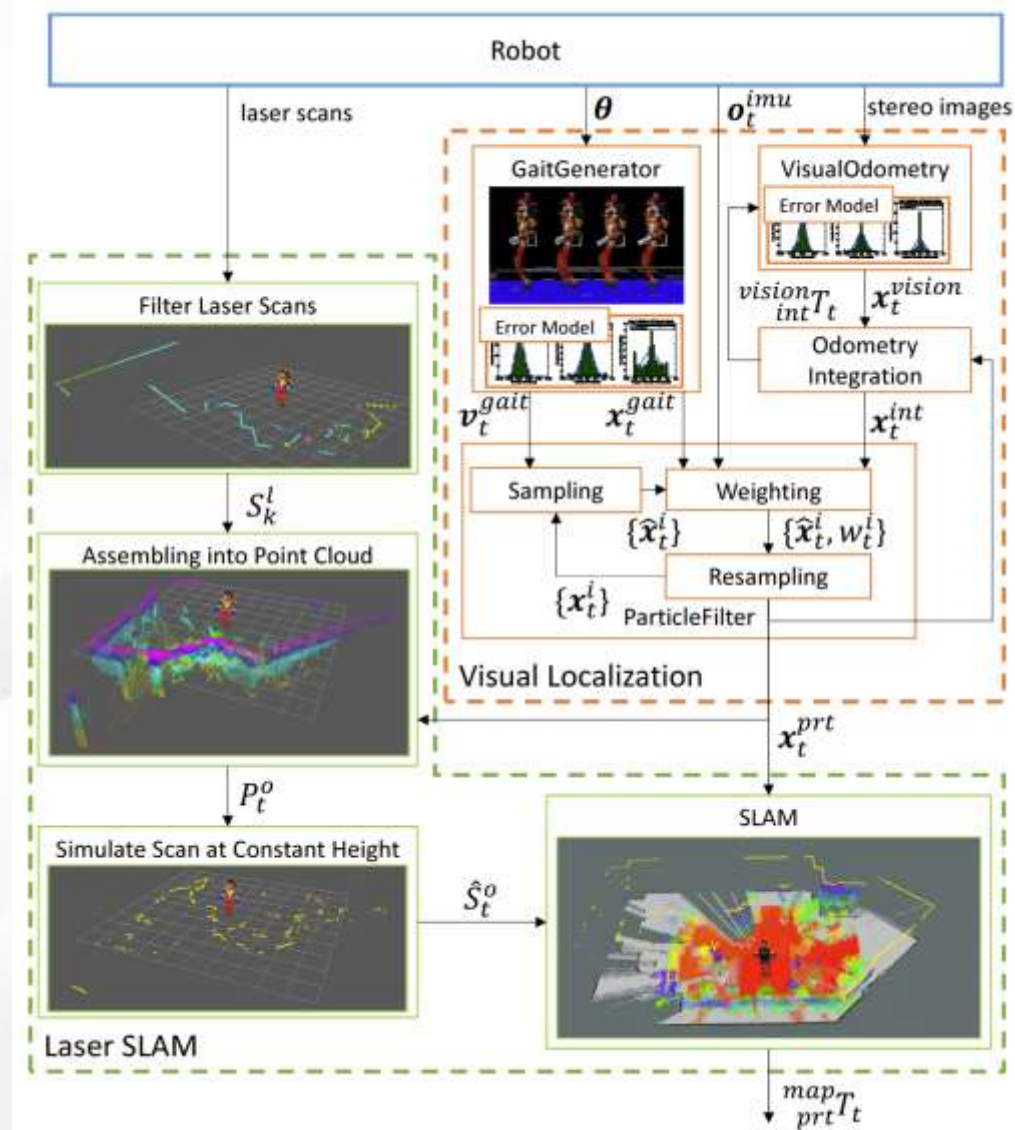
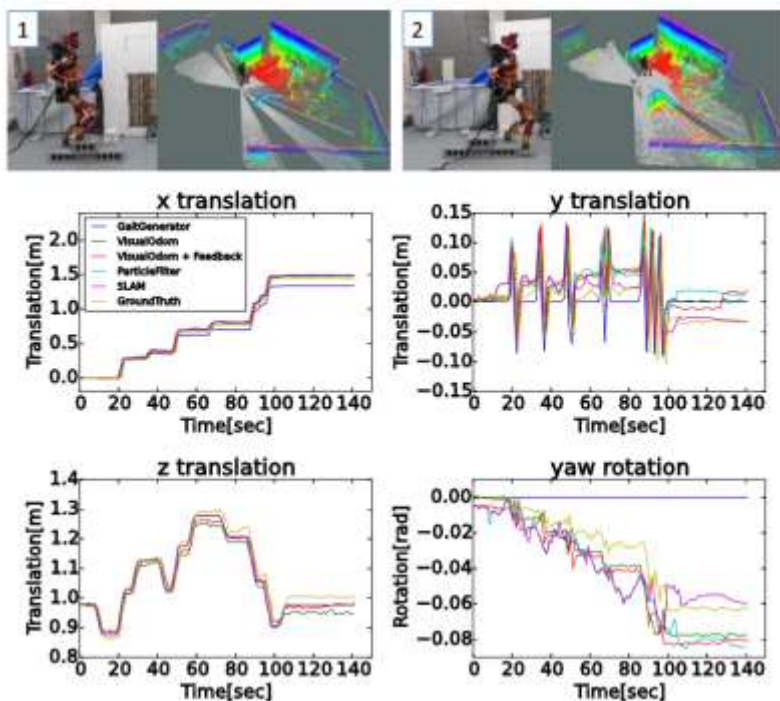
Novel lightweight odometric learning method for humanoid robot localization

- 本文提出的就是基于人形机器人的惯性里程计 (inertial odometry), 不依赖于其他外部的感知, 仅仅用IMU, 并且采用ANN来进行运动学计算。
- 采用的网络是最基本的MLP, 输入为来自人形机器人上的IMU、里程计、腿式压力等数据, 输出直接为全局坐标系下的位置信息。
- 基于深度学习的惯性里程计可以实现对机器人运动模态的数据关联, 进而实现端到端的感知, 更多关于IO的调研请见[博客](#);



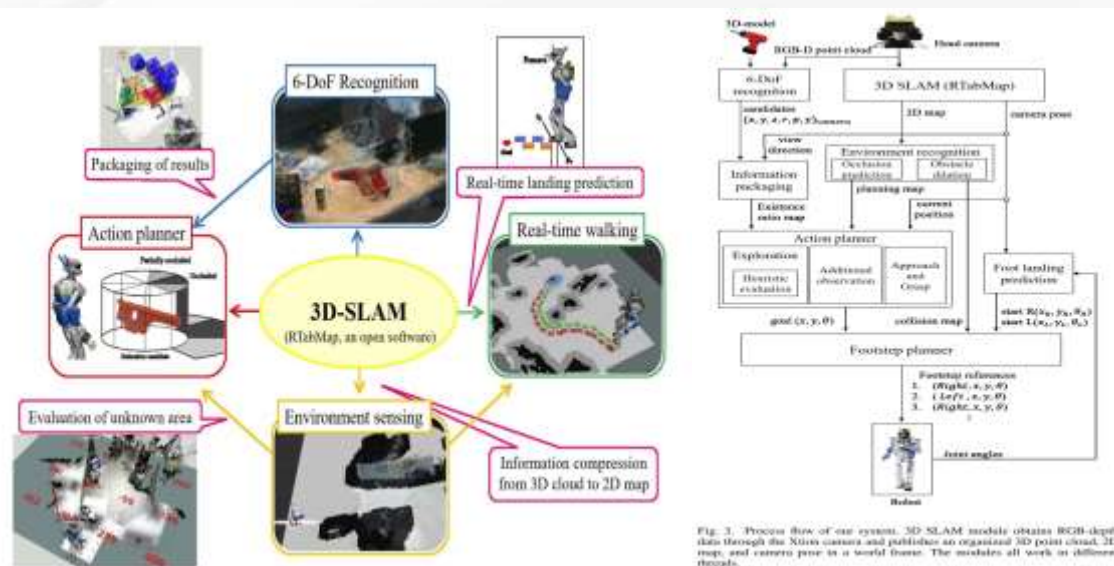
Achievement of Localization System for Humanoid Robots with Virtual Horizontal Scan Relative to Improved Odometry Fusing Internal Sensors and Visual Information

- 将视觉里程计、步态发生器的前馈命令以及惯性传感器的方位信息来提升里程计的性能。
- 进一步地，将该里程计用于从连续激光扫描的累积中生成3D点云，然后对所获得的3D点云进行适当的切片，以创建高度固定的水平虚拟激光扫描。
- 虚拟的激光扫描进一步放到Gmapping中来实现更高精度的SLAM；



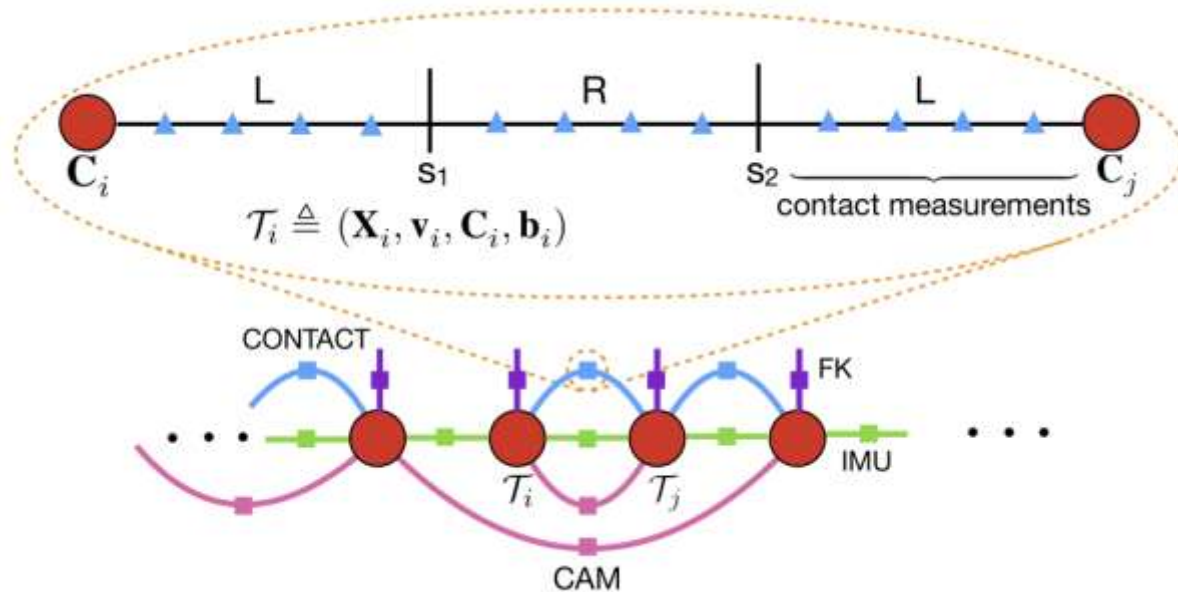
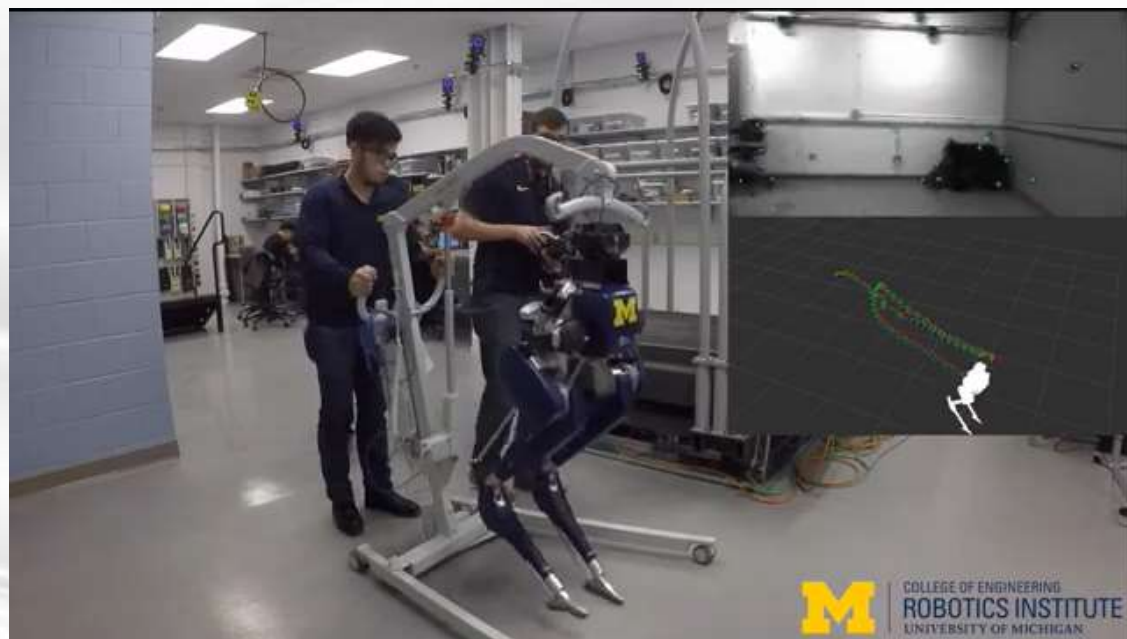
Online Object Searching by a Humanoid Robot in an Unknown Environment

- 传统的人形机器人寻找目标物体的方法都是需要一个静态的地图，同时需要有关于对象位置、区域大小限制或每次观察的离线视点规划时间等提示信息。而本文则是单纯基于目标物体的3D模型，在未知环境中通过3D SLAM寻找这个物体；
- 要完成一项任务需要四个主要部分：目标识别，环境感知，运动规划以及双足行走；
- SLAM用的是RTAB-Map(但是生成的是2D map)，同时结合目标物体的3D点云识别(ICP)，路径规划(A*)、避障以及自主探索以及足迹规划等。
- 但论文只是把上述各个模块融入到RTAB-Map中，使其可以完成完整的任务，并没有对SLAM算法做改进；



Hybrid contact preintegration for visual-inertial-contact state estimation using factor graphs

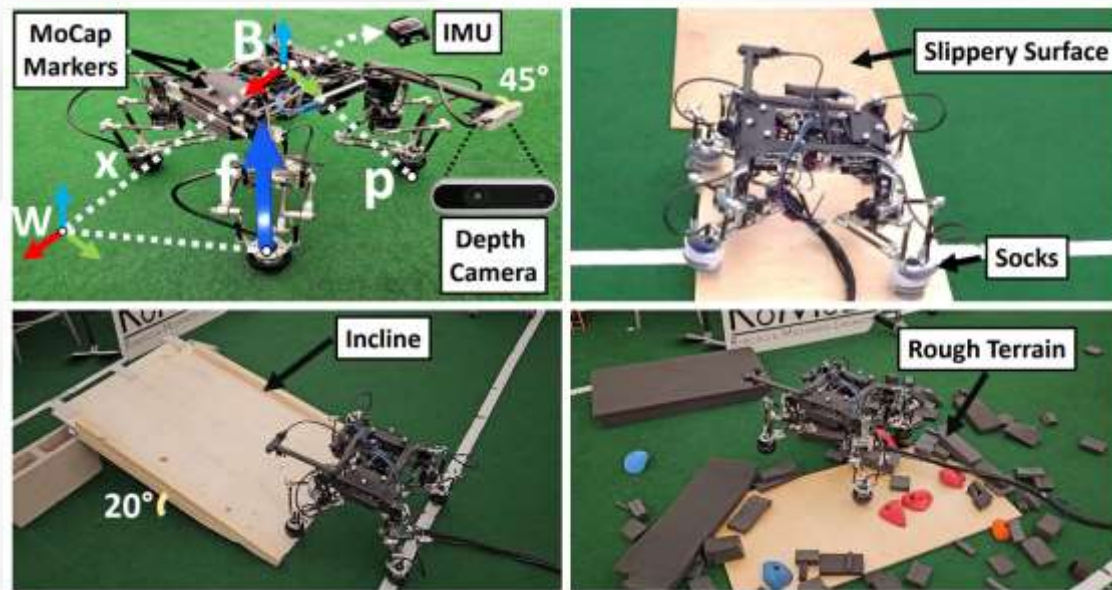
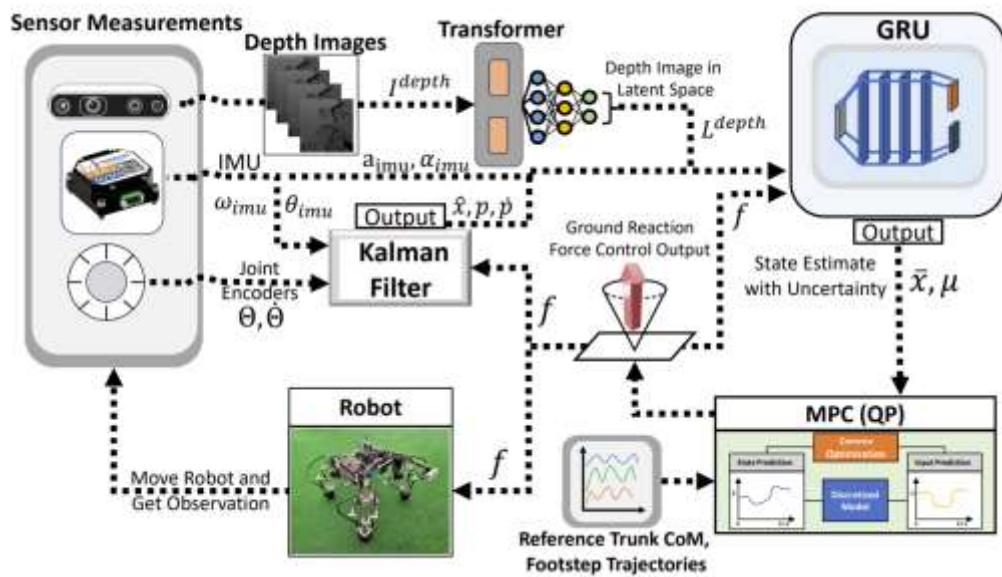
- 该工作提出了基于图优化框架的视觉、imu、关键编码器
和腿式接触传感器融合定位系统。
- IMU部分采用预计分构建的因子，视觉里程计采用的是
SVO2；
- 至于腿式接触传感器则是开发了一种通过任意数量的接触
来实现预积分接触信息的方法。这样以预积分的形式加入
因子图中可以减少优化的变量；



OptiState: State Estimation of Legged Robots using Gated Networks

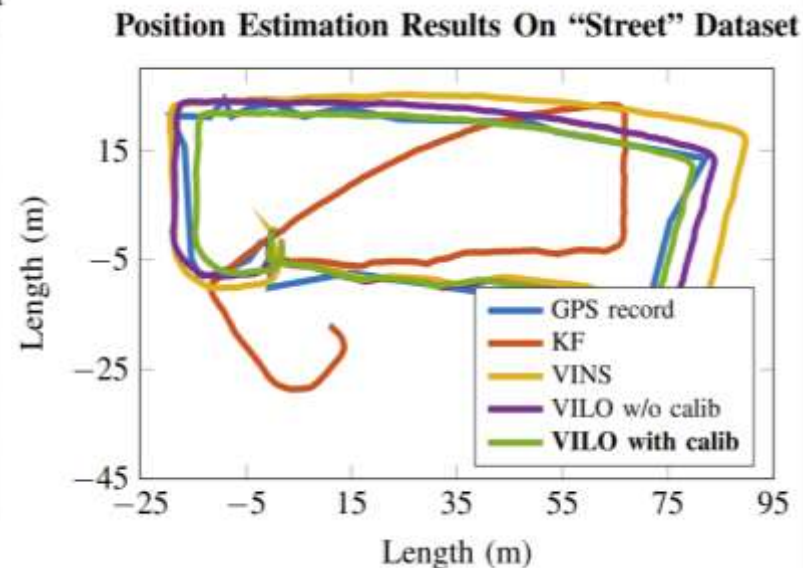
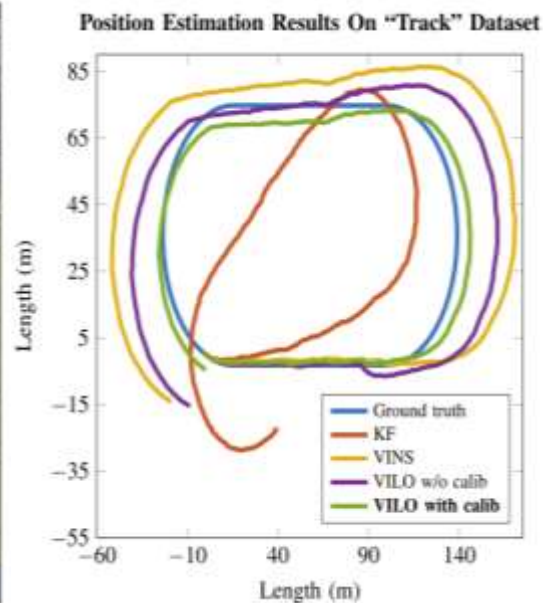
with Transformer-based Vision and Kalman Filtering

- 对于足式机器人而言，踏步容易引起相机的运动模糊，使得视觉数据不可靠。而由于机械腿与地面的不断接触和断开（还存在滑动），基于IMU等运动学信息也容易出错，特别是在有大干扰（比如障碍物）或柔性表面的情况。
- 该工作通过融合卡尔曼滤波、优化以及深度学习方法，提出了一个融合内部与外部传感器信息的机器人躯干状态估计算法（最终通过实验验证精度比t265输出的要好3倍以上）。
 - 1.首先采用卡尔曼滤波来对关节编码器以及IMU测量量进行状态预测，通过复用来自凸模型预测优化方法输出的地面反作用力来进行状态的更新。而卡尔曼滤波的输出会输入GRU网络。
 - 2.此外，GRU网络还会有来自Vision Transformer处理的深度图的隐特征，而GRU输出的姿态信息将会与动捕提供的真值数据求loss来实现网络的训练。



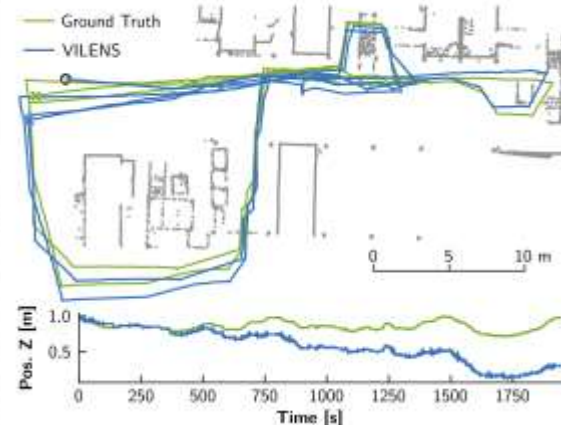
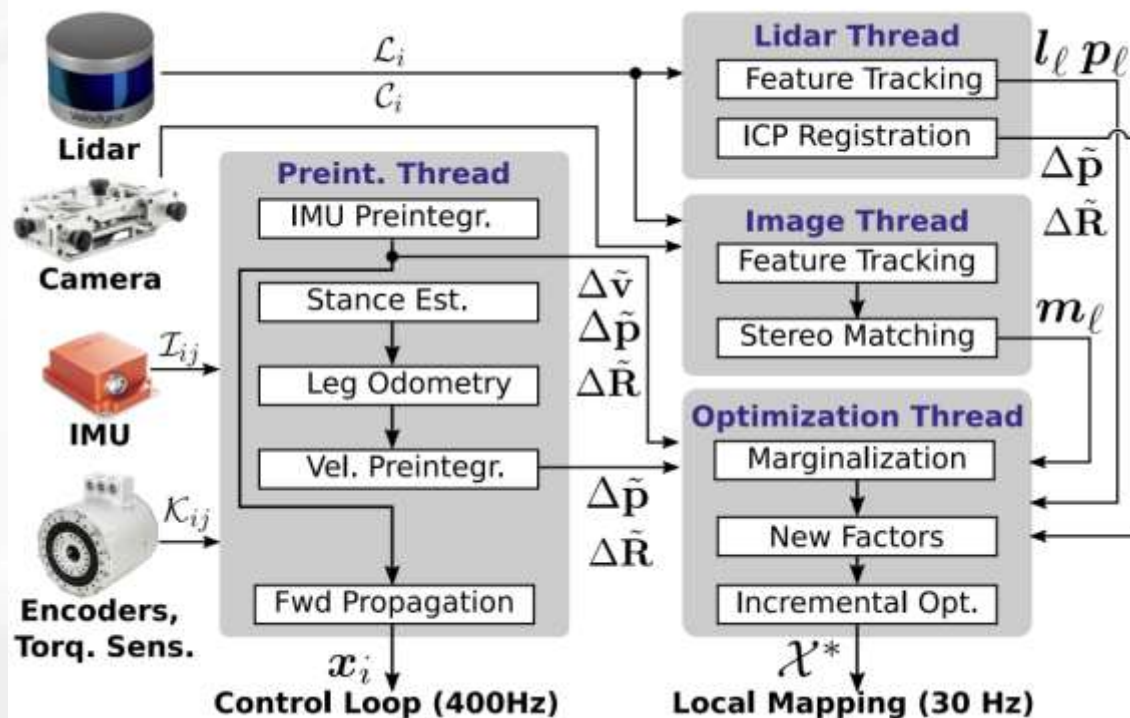
Cerberus: Low-Drift Visual-Inertial-Leg Odometry For Agile Locomotion

- 该工作提出了一个视觉-惯性-腿式里程计（包含的传感器有：双目图像、IMU、关节编码器、腿部接触传感器）；
- 实现了在线的运动学参数的校正以及接触传感器的异常值剔除，进而减少了定位的漂移。通过室内与室外的实验（450米远、平均运动速度0.5m/s室内，1m/s室外）也证明了所采用的运动学参数估算可以将累积误差减少的1%以内；
- 此工作有开源的数据集和代码，后续可以测试看看实际效果如何；



VILENS: Visual, Inertial, Lidar, and Leg Odometry for All-Terrain Legged Robots

- 基于因子图的紧耦合的视觉-IMU-lidar-腿式里程计，并通过对线速度的bias进行在线估计来减少腿式里程计的漂移；
- Bias的在线估计是通过因子图融合四种传感器的信息获得的；
- 一共2小时，1.8km的实验来验证系统的性能，并且在实验的场景中还包含了松散岩石、斜坡、泥土、滑动、地形形变等挑战；
- 主要包括IMU及运动学的预积分、camera特征跟踪（用的时FAST corner以及KLT），lidar的特征跟踪及ICP匹配，最后就是融合前面因子的图优化框架；



Leg-KILO: Robust Kinematic-Inertial-Lidar Odometry for Dynamic Legged Robots

- 足式机器人上高频的运动(如小跑步态)会引入频繁的脚步撞击, 进而导致IMU的退化以及lidar的运动失真。而直接用IMU观测的数据会容易导致明显的漂移(特别是在Z轴上);
- 那么针对这些局限性, 论文提出了基于图优化的腿式里程计、雷达里程计以及回环检测三者紧耦合的系统, 进一步地提出机器人位中心的增量式建图方式来提高地图的维护效率;
 - leg odometry: 提出了基于on-manifold error-state Kalman filter的运动惯性里程计, 通过结合接触高度检测的约束来进一步减少高度方向的波动;
 - lidar odometry: 设计了一种自适应激光束切片和拼接的方法, 以减轻高动态运动的影响;
- 在下右图的实验中也发现, 在平地运动时, baseline方法均有较大的高度变化, 而Leg-KILO则是在10cm以内。

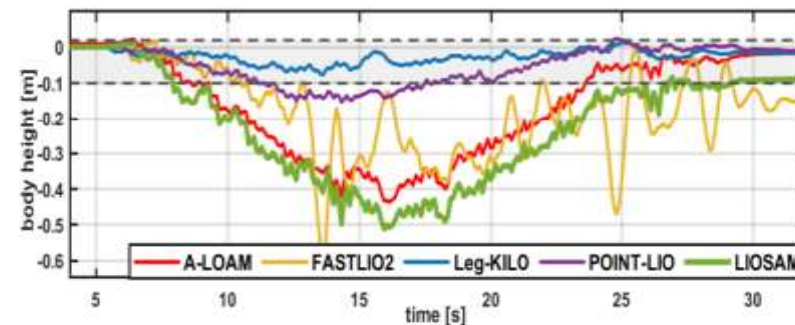
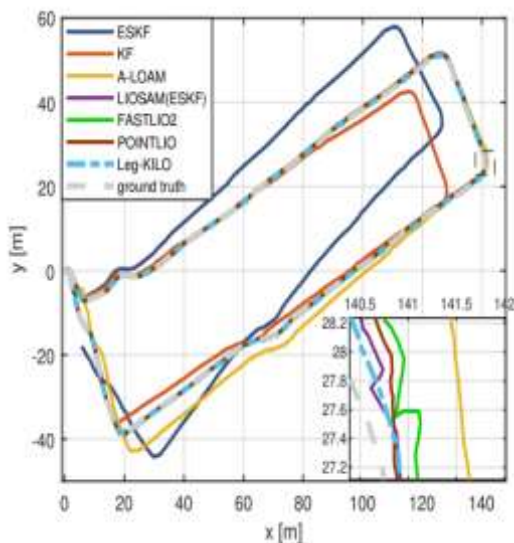
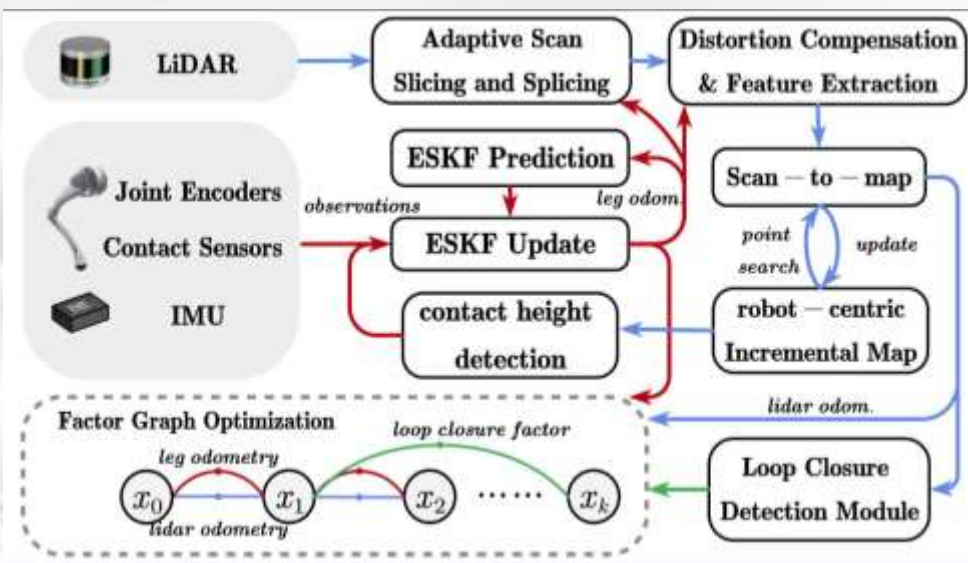
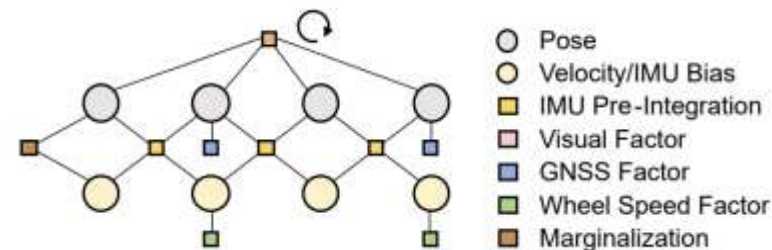
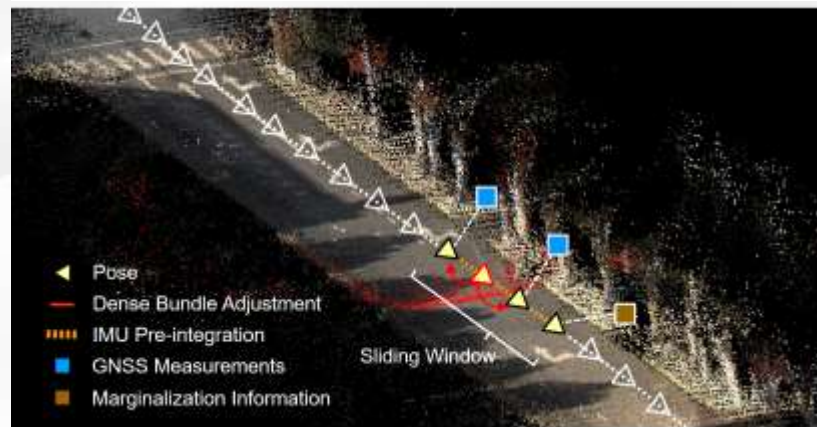
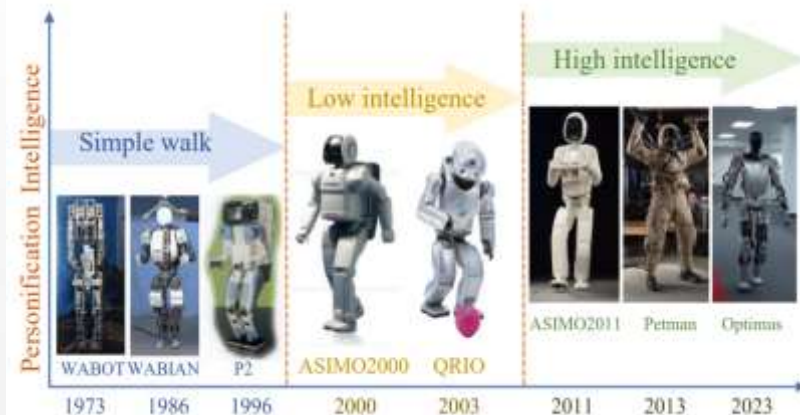


Fig. 8. Comparison of body height's variation for various methods while running on flat ground. The body height will drop slightly when running, and will not exceed the gray area in the figure.

结束语

- 目前人形机器人3D SLAM大多数局限于仅仅移植已有3D SLAM（如RTAB-Map）算法结合运动规划来完成自主作业的任务，缺乏像四足机器人一样，针对机器人的运动特性来提升SLAM算法；
- 多传感器融合是唯一解决方案，激光雷达、视觉为主，惯导、里程计、腿式压力计等其他传感器为辅；
- 数据集是关键；目前四足机器人SLAM数据集较多，但是基于人形机器人的还较少；
- 3D SLAM主流解决方案以无人机或手持为载体，迁移到足式机器人上需根据机器人的实际运动模式进行调整；
- SLAM中位姿估计是基础性技术，而三维重建应结合实际应用场景需求，如：重定位、语义信息、目标跟踪与分割、避障规划、纹理细节、mesh结构等；



- Awesome-LiDAR-Visual-SLAM: <https://github.com/sjtuyinjie/awesome-LiDAR-Visual-SLAM>;
- Awesome-LiDAR-Camera-Calibration: <https://github.com/Deephome/Awesome-LiDAR-Camera-Calibration>;
- MARS-LAB: <https://mars.hku.hk/>;
- 一文看懂激光雷达LIDAR基本工作原理: <https://zhuanlan.zhihu.com/p/602055107>;
- 人形机器人综述论文《Humanoid locomotion and manipulation: Current progress and challenges in control, planning, and learning》;
- Quadruped SLAM using the A1's onboard sensors: https://github.com/jerredchen/A1_SLAM;
- Multi-Sensor State Estimator for Legged Robots: <https://github.com/ori-drs/pronto>;

Thank you