

【Paper Reading】 RoadMap

RoadMap: A Light-Weight Semantic Map for Visual Localization towards Autonomous Driving

论文阅读个人总结

1. 现实原因

HD-Map和激光雷达 成本高

激光点云地图消耗了大量内存

HD-Map制作需要大量人力,难以及时更新

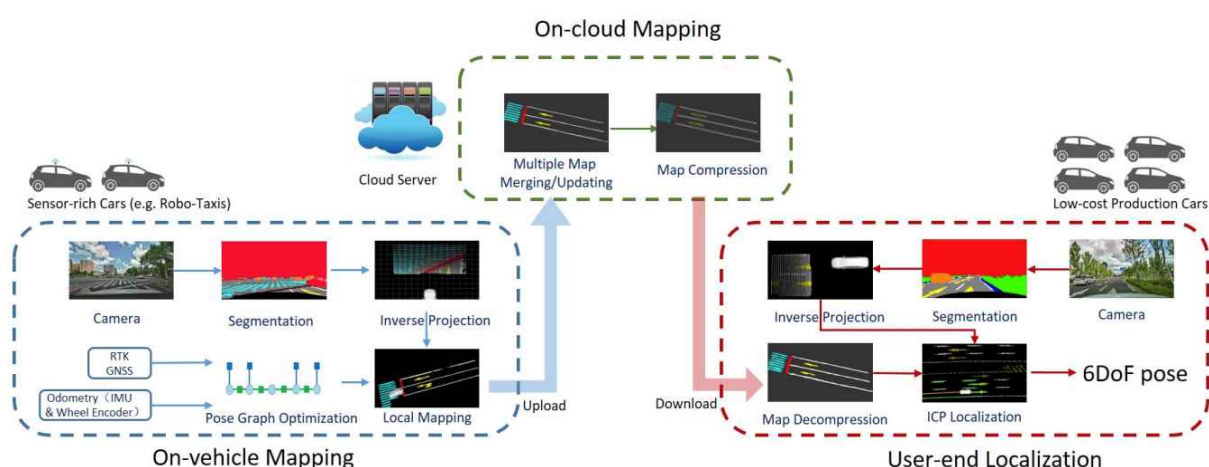
解决思路 ==> 只依赖低成本传感器和紧凑型地图的思路

2. 整体方案

2.1 传感器

丰富传感器的采集车(e.g. robot-taxi): 前视相机 + RTK-GPS + IMU + 轮速计

低成本的量产车: 多相机 + 低精度 GPS + IMU + 轮速计



2.2 PartA 车端建图

1. CNN语义分割出语义元素ground、lane line、stop line、road marker

2. 选相机前12mx8m的ROI逆透视变换并转换到车体坐标系下

针孔相机模型:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

$$Z \begin{bmatrix} u \\ z \\ 1 \end{bmatrix} = K P_{camera}$$

$$Z \begin{bmatrix} u \\ z \\ 1 \end{bmatrix} = T K P_{world}$$

可推出

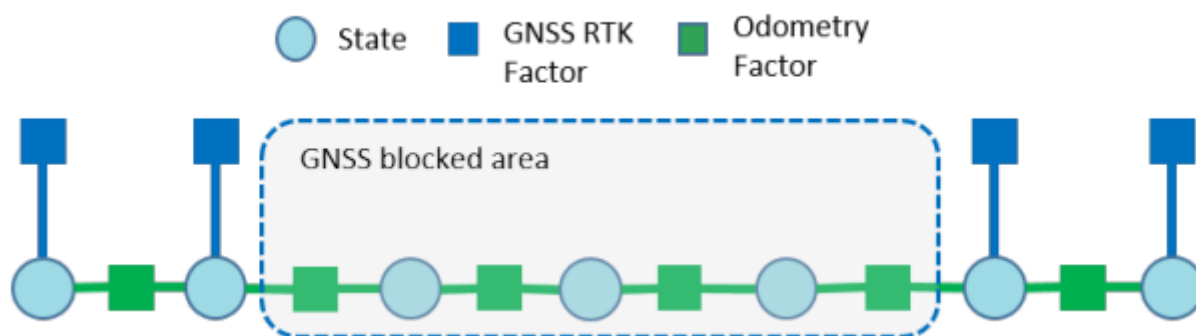
$$\frac{1}{\lambda} P_{vehicle} = T^{-1} \pi^{-1} P_{camera}$$

3. 图优化

RTK-GNSS开阔地可提供厘米级定位

imu+wheel encoder=doometry 有飘移

两者进行图优化



设计残差:

$$\min_{s_0 \dots s_n} \left\{ \sum_{i \in [1, n]} \|r_o(s_{i-1}, s_i, \hat{m}_{i-1, i}^o)\|^2 + \sum_{gnss\ useful} \|r_g(s_i, \hat{m}_i^g)\|^2 \right\}$$

具体是:

$$r_o(s_{i-1}, s_i, \hat{m}_{i-1, i}^o) = \begin{bmatrix} R(q_{i-1})^{-1} (p_i - p_j - \delta \hat{p}_{i-1, i}^o) \\ (q_{i-1}^{-1} q_i)^{-1} \delta \hat{q}_{i-1, i}^o \end{bmatrix}$$

$$r_g(s_i, \hat{m}_i^g) = p_i - \hat{m}_i^g$$

odom的位置旋转残差、旋转残差 + gnss的位置残差

4. 局部建图

$$\begin{bmatrix} x^w \\ y^w \\ z^w \end{bmatrix} = R \begin{bmatrix} x^v \\ y^v \\ 0 \end{bmatrix} + p_i$$

车体坐标系的单帧语义元素坐标转换到世界坐标系

struct Grid

```
{
    position;
    semantic labels;
    counts of each labels;
}
```

划分3D网格三轴方向10cm分辨率，在不断插入语义点至网格时，通过数据统计(多帧投票之类的方法)确定，选取得分最高的语义标签作为该网格的标签

2.3 PartB 云端建图

1. 地图合并与更新

为节省宽带只上传占据网格

云端同样3D网格10cm分辨率

占据标签得分也会被加入全局地图

2. 地图压缩

只向用户分发元素组的轮廓

2.4 PartC 用户端定位

1. 地图解压

从平面图像恢复至世界坐标

填充轮廓为实心

2. ICP定位

车体坐标系下的观测-->配准-->世界坐标系下的局部语义地图

$$q^*, p^* = \arg_{q,p} \min \sum_{k \in S} ||T_{wv} P_v - P_w||^2$$

得到的T即车体相对于世界坐标系的位置，即车体定位

最后imu+wheel encoder融出来的odomEKF再融合

3. 实验结论

3.1 地图生产

22km 路测 原始16.7MB 压缩后0.786MB 也就是36kb/km
和谷歌地图比较



3.2 定位精度

Method	x error [m]		y error [m]		yaw error [degree]	
	average	90%	average	90%	average	90%
vision	0.043	0.104	0.040	0.092	0.124	0.240
Lidar	0.121	0.256	0.091	0.184	0.197	0.336

4. 文本展望

拓展至3D semantic map