首发干

语义定位与建图

ICRA2021华为奥迪阿里自动驾驶道路语义建图与定位文章总结



wanghy

自动驾驶定位SLAM

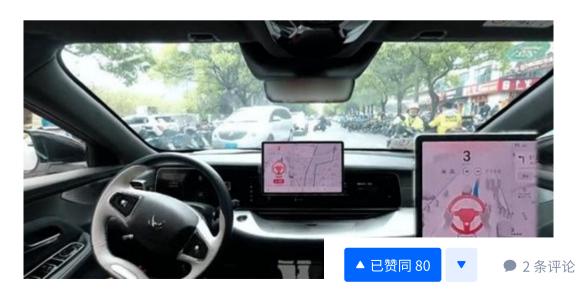
已关注

80 人赞同了该文章





由于传统的视觉特征定位方法会受到季节、天气、光照、视角、遮挡影响,基于视觉特征的长时 间、长航程精确鲁棒定位是很难实现的。为了解决传统视觉特征的问题,近些年,基于道路语义特 征的定位方法受到了学术界和工业的广泛关注。道路语义特征主要包括路面的车道线、停止线、路 面标志(箭头等)、杆状物、红绿灯和标牌等。相比干传统的视觉特征,这些道路语义特征广泛且 长期稳定存在于城市道路场景,对季节、天气、光照、视角等变化比较鲁棒。随着深度学习的广泛 应用,道路语义特征易于检测和提取,且具有紧凑的表达形式。高精度地图采用矢量表达道路语义 特征,只需要占用极小的存储资源。这可以解决无人驾驶车辆在有限存储条件下的高精度定位。因 此特对ICRA2021自动驾驶道路语义建图与定位的文章进行总结。



知平 首发干 语义定位与建图

作者: Huayou Wang, Changliang Xue, Yanxing Zhou, Feng Wen and Hongbo **Zhang**

单位: 华为诺亚方舟实验室自动驾驶项目组

摘要:





高精度和鲁棒定位能力对于自动驾驶车辆而言是及其重要的。基于传统视觉的方法在光照、天气、 视角和外观变化场景容易跟丢。因此,本文提出了一种**创新的视觉语义定位算法,该方法基于 HDmap和具有紧凑表达的语义特征**。语义特征广泛存在于城市道路场景,且对光照、天气、视角 和外观变化比较鲁棒。语义特征的重复结构、视觉感知的误检测、漏检测、环境发生变化地图更新 不及时,且需要同时对多个感知语义特征与地图与语义元素进行关联,这使得数据关联具有高度奇 异性,并成为**语义定位的最大挑战**。为了解决该问题,**文本提出了一个综合考虑局部结构一致性、 全局模式一致性和时序一致性的数据关联方法**。然后,本文提出了一个融合关联结果和里程计观测 的滑动窗口因子图优化框架,且该方法不需要HDmap landmark具有及其精确的高度信息。

我们同时在仿真和真实城市道路场景数据上对本文提出的定位算法进行了评估。实验结果表明,本 文提出的方法能够实现精确的定位,具有**0.43m的平均纵向误差、0.12m的平均横向误差和0.11°** 的平均航向角误差。

主要贡献:

1. 一种基于视觉语义特征和轻量级高精度地图的精确、 landmark具有及其精确的高度信息











- 3. 一个紧耦合数据关联结果和里程计观测的**因子图优化框架**;
- 4. 一系列的仿直和实际道路场景实验来验证数据关联方法和有效性和定位结果的准确性。

系统框架:

全局定位问题可以被定义为:给定一系列传感器观测 $\mathcal{Z} = \{\mathbf{z}_k\}_{k=1}^K$ 和HD map landmark $\mathcal{L} = \{\mathbf{l}_m\}_{m=1}^M$ 估计自车位姿序列 $\mathcal{X} = \{\mathbf{x}_t\}_{t=1}^T$ 。该定位问题能够被表征为如下最大后验推理问 题:



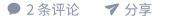
$$\hat{\mathcal{X}} = rg \max_{\mathcal{X}} p(\mathcal{X}|\mathcal{Z},\mathcal{L})$$



该最大后验推理问题可以被分为两步:数据关联过程和基于数据关联结果的位姿估计。上式隐含了 数据关联过程,基于先验位姿 χ^0 ,构建观测和地图路标的关联 $\mathcal{D} = \{\mathbf{d}_t\}_{t=1}^T$ 是必要的。因此, MAP推理问题可以被重新定义为:

$$\begin{split} \hat{\mathcal{X}}, \hat{\mathcal{D}} &= \operatorname*{arg\,max}_{\mathcal{X}, \mathcal{D}} p(\mathcal{X}, \mathcal{D}|\mathcal{Z}, \mathcal{L}) \\ &= \operatorname*{arg\,max}_{\mathcal{X}, \mathcal{D}} p(\mathcal{X}|\mathcal{Z}, \mathcal{L}, \mathcal{D}) p(\mathcal{D}|\mathcal{X}^0, \mathcal{Z}, \mathcal{L}) \end{split}$$

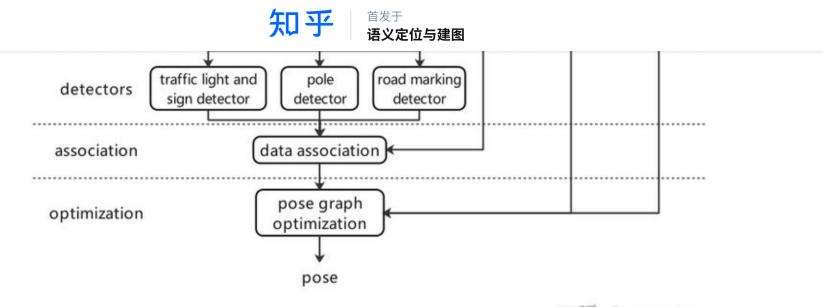












已赞同 80



Fig. 2. Overview of the localization framework.

整个定位框架主要分为四个部分: 传感器与地图、检测层、关联层和优化层。

传感器与地图:

一个相机、一个IMU、两个论速记、GNSS接收器。相机用来提取语义特征、IMU和论速记构成里程计、GNSS提供初始位姿估计;

检测层: 提取道路环境中语义元素

本文采用流行的卷积神经网络方法来检测特征。一个检测的标牌 $s_t = (s_t^l, s_t^c, s_t^b)$ 由检测类别 s_t^l 、检测置信度 s_t^c 和bounding box s_t^b 表征。标牌在HD map中存储四个轮廓点,并且每个点的高度都是相对于当前位置地面的高度。一个检测的杆 $s_t = ($

▲ 已赞同 80

■ 2条评论

▼ 分享

★ 收藏

知乎 首发干 语义定位与建图

关联层:建立感知语义特征与地图语义特征的匹配关系

由干语义特征的重复结构、视觉感知的误检测、漏检测、环境发生变化地图更新不及时,且需要同 **时对多个感知语义特征与地图与语义元素进行关联**,这使得**数据关联**具有高度奇异性,并成为**语义 定位的最大挑战**。为了解决该问题,**文本提出了一个综合考虑局部结构一致性、全局模式一致性和** 时序一致性的数据关联方法,且该方法能够给出确定性匹配和不确定的概率匹配关系。该算法流程 如下:





- 1. 基于里程计递推先验位姿,将局部地图语义特征重投影到图像平面;
- 2. 单帧数据关联:综合考虑匹配数量、语义特征匹配相似性和语义特征局部结构相似性、将单帧数 据关联问题构建为一个**多级图匹配问题**来实现最优的全局一致匹配:

其中, N 和 M 分别表示感知和重投影语义特征数量。 N_{--} 表示匹配数量。 N_{--} 表示两个语义特 征构成边的数量。 ω_1 、 ω_2 和 ω_3 是超参数。 $x_{ii'}$ 表













知乎 | 首发于 语义定位与建图

该表达式前两部分可以基于感知结果进行离线学习,第三部分表示结构相似性。对于牌,该项主要包含位置 x_p 和大小 x_s 相似性(表示为指数函数,将任何非负输入转化到[0,1]区间):





对于杆,结构相似性表示为位置和倾斜角度的相似性。

 $s_{ij,i'j'}$ 表示边 e_{ij} 和边 $e_{i'j'}$ 的相似性,

其中, 边 e_{ij} 和边 $e_{i'j'}$ 表示分别如下图3所示:

知乎 | 首发于 语义定位与建图





该多级图匹配优化问题可以通过通用的随机re-weighted游走算法求解。

3. 单帧匹配置信度:

单帧匹配置信度通过度量匹配元素之间的特征相似度和局部结构相似度来进行评估:

▲ 已赞同 80











知乎 语义定位与建图

连续帧之间发生漏检时的误匹配问题。

5. **时序平滑**: 该步骤构建连续帧感知语义特征和地图语义特征之间的最优一致匹配。时序平滑步骤 通过对滑动窗口中每一帧的匹配 $D_{1:T}$ 和匹配置信度 $c_{i,i}$ 进行加权来得到地图元素 r^{i} 对应的感知 语义特征 s_i :





其中, $I(s_i, D_t)$ 表示地图元素 x^l 是否与感知语义特征 s_i 匹配。

如果得到最优的匹配感知语义特征的加权置信度远大干次优的匹配感知语义元素,则最优的匹配感 知元素被认为是地图特征 x^l 的匹配。否则,地图元素 x^l 被认为具有不确定的匹配,并给出与每 个感知语义特征的匹配置信度。该过程构建了确定匹配和不确定匹配的概率表达。即能够解决强匹 配的误匹配问题,又能解决普通概率数据关联方法的计算效率问题。

优化层:基于关联结果和里程计观测,优化自车位姿序列

本文通过里程计信息 $\mathbf{z}_{i,i+1}^{o}$ 和语义数据关联结果,构建滑动窗口非线性最小二乘估计器来估计最近 T个时刻的位姿

▲ 已赞同 80











首发于

语义定位与建图



其中,里程计误差定义为:



语义观测误差定义为重投影横向误差,来消除对地图精确高度的依赖



知乎 首发于 语义定位与建图



地图误差因子表示为:



该非线性优化问题能够直接通过迭代算法求解。滑窗外的状态直接被忽略,这是因为边缘化方法会 累计线性化误差,使得系统矩阵稠密,造成deadlock。边缘化方法基于过去数据来约束位姿,但 是使用地图特征作为先验足够约束自车位姿。

实验结果:

仿真实验验证本文算法的特征:

▲ 已赞同 80



■ 2条评论

7 分享 ♥ 喜欢

★ 收藏

语义定位与建图









首发于

语义定位与建图





分享

实际城市道路场景验证定位精度与数据关联方法效果:

▲ 已赞同 80











语义定位与建图







首发于

语义定位与建图









首发于

语义定位与建图

- 1. 本文的核心贡献就是提出了一套完整的语义数据关联算法流程来解决语义数据关联的奇异性;
- 2. 将来为该框架添加更加有效的实时定位效果评估算法;
- 3. 定位初始化是否可以通过纯语义特征在普通GPS基础上内实现?

文章2: A Light-Weight Semantic Map for Visual Localization towards Autonomous Driving



作者: Tong Qin, Yuxin Zheng, Tongqing Chen, Yilun Chen, and Qing Su



单位:华为IAS BU 智能驾驶产品部

视频效果:

知乎 | 首发于 语义定位与建图

视觉语义建图与定位

摘要:





精确定位对于自动驾驶任务是至关重要的。如今,我们可以见到许多sensor-rich车辆自动驾驶在街道上,这些车辆依赖于高精度的传感器(例如激光和RTK)和高分辨率地图。然后,低成本的量产车不能负担如此昂贵的传感器和地图。如何降低成本?如何低成本车辆受益于sensor-rich车辆?本文中,我们提出了一种轻量级定位解决方案,该方案依赖于低成本的相机和视觉语义地图。地图是很容易通过sensor-rich车辆以众包的形式进行构建和更新的。具体来说,地图包含一些语义元素,例如道路表面的车道线、人行横道、路面标志和停止线等。本文提出了车端建图、云端维护和user-end定位的框架。地图数据在车端进行收集和预处理。然后,众包数据被上传到云端服务器。来自多个车辆的大量数据在云端进行聚合,使得语义地图实现实时更新。最后,语义地图被压缩和分发到量产车上用来定位。我们在实际环境中验证了本文提出的地图的性能,并和其它算法进行了比较。该语义地图的平均大小为36kb/km。我们强调该框架是一个可靠且可实用的自动驾驶定位解决方案。

创新点:

- 1. 提出了一个应用于自动驾驶任务的**轻量级定位框架**,包含**车端建图、云端维护和user-end定位** 三个部分;
- 2. 提出了一个新的想法使得sensor-rich**车辆有益于低广** 据并每天自动更新地图;





• 1 4 4 7 . . 11 44 . 141







知乎 首发于 语义定位与建图

系统框架:





本文系统由三个部分组成,**车端建图、云端众包建图和use-end定位**。

车端建图:

车辆装有**前视相机、RTK-GPS、IMU和论速记**。语义特征通过**语义分割网络**从前视相机图像中提 取。然后语义特征基于IPM变换到自车坐标下,再基于自车位姿变换到世界坐标下。然后局部地图 就能够被构建,并上传到云端服务器。



首发于

语义定位与建图









语义定位与建图















语义定位与建图











通过图像语义分割,每个点包含一个类别标签。每个点表征世界坐标系下的一个小区域。当车辆运动时,一个区域会被观测到很多次。然后,由于语义分割噪声,该区域可能被分类为错误的类别。为了解决该问题,本文**使用统计方法来滤除噪声**。将地图分为分别率为**0.1 x 0.1 x 0.1 m的栅格**,每个栅格包含位置、类别和每个类别的数量。*初始每个类别标签的score都为0,当一个语义点进入该栅格,那么对应语义标签的score增加1*。具有最高分数的语义标签作为该栅格的语义标签。通过该方法,语义地图变得精确且对语义分割噪声非常鲁棒。

语义定位与建图







语义定位与建图

云端服务器收集来自多个车辆的局部地图,并将这些局部地图融合为一个全局地图。融合方法与 车端局部构图一样,将云端服务器上的语义地图也分为分辨率为0.1 x 0.1 x 0.1 m的栅格,然后统计 每个栅格类别数量确定栅格类别标签。然后将该全局地图通过**轮廓提取进行压缩**。最后,压缩的语 义地图被分发到自动驾驶车辆。

End-user定位:



7 分享

低成本量产车辆装配有低沉本传感器,例如相机、低精度GPS、IMU和论速记。使用端从云端服务 器下载语义地图并解码。与车端建图一样,语义特征通过语义分割从前视相机提取,并通过与语义 地图进行语义特征匹配来实现自车定位,采用ICP方法。最后通过EKF框架来融合语义定位结果和 里程计结果来提高定位结果的鲁棒性和平滑性。







实验结果:

语义定位与建图







语义定位与建图







知乎 | 首发于 语义定位与建图

结论:





- 1. 本文主要提出了一个车端建图、云端众包构图和user-end定位的框架;
- 2. 实现紧凑地图表达,语义地图数据36KB/KM;
- 3. 将来会扩展更多的语义元素,例如红绿灯、标牌和灯杆等;
- 4. 在结果方面,本文声称得到的定位精度非常高,达到了4cm,比激光还要高?

文章3: Lightweight Semantic Mesh Mapping for Autonomous Vehicles

作者: Markus Herb, Tobias Weiherer, Nassir Navab and Federico Tombari

单位: AUDI AG, Ingolstadt, Germany

该文章是奥迪IROS2019众包语义构图文章Crowd-sourced Semantic Edge Mapping for Autonomous Vehicles的后续工作,扩展了更多语义物体,例如杆、交通标志等,作者也增加了增量地图更新功能。参考:

首发于

语义定位与建图



视频效果:





分享

自动驾驶轻量级语义mesh建图

摘要:













首发于

语义定位与建图

征为3D mesh。本文系统继承了已有的基于特征的视觉里程计、基于学习的深度预测和语义图像分割来识别和重建语义环境结构。我们引入了一个概率融合框架来增量式修正和扩展带有语义标签的3D mesh。为了验证算法的有效性,我们在室内场景结合单目深度预测和双目验证了我们的系统,并且给出了与真值之间定性和定量的比较结果。实验结果表明本文方法能够实现与当前最先进的方法相当的重建质量,但是需要更小的存储和计算资源。









语义定位与建图





分享

主要贡献:





★ 收藏

知乎 | 首发于 语义定位与建图

3. 引入了一个**概率融合方案**,在没有基于体素的中间融合的情况下,逐步细化和扩展带有语义标签的3D网格。

系统框架:





- 1. **视觉里程计:** 采用ORB-SLAM2的前端里程计,将关键帧给到后面建图系统;
- 2. **关键帧语义和深度**:使用**PSPNet**进行语义分割,深度预测采用双目或者单目深度预测。深度预测的不确定性采用**参考文献[9]**的方法;
- 3. 从单帧图像中提取Triangulated 2D Mesh;
- 4. Mesh地图初始化:通过深度估计,并给出不确定从2D传播到3D的结果;
- 5. **语义mesh匹配**:采用了概率匹配,如下图所示

▲ 已赞同 80











首发于 语义定位与建图





分享

6. **语义mesh融合:** 匹配的mesh投影、冲突检测、更新和扩展;--该步骤都是比较细节的东西,详 细参考原文。

实验结果:







★ 收藏

语义定位与建图







语义定位与建图













分享

结论:

1. 对于未来的工作,结合全景分割和3D 四面体网格是一个有趣的改进途径来增加遮挡物体的性能。

文章4: Road Mapping and Localization using Sparse Semantic Visual Features

作者: Wentao Cheng, Sheng Yang, Maomin Zhou, Ziyuan Liu, Yiming Chen, Mingyang Li

单位: 阿里巴巴

▲ 已赞同 80











知乎 | 首发于 语义定位与建图





使用稀疏视觉语义特征进行建图和定位

摘要:

本文呈现了一个新的**自动驾驶视觉建图和定位方法**,该方法提取、建模和优化道路语义元素。具体来说,本文方法集成深度学习模型来检测**标准道路元素**而不是传统视觉特征点,来寻求提高的位姿精度和紧凑地图表达。为了使用这些结构特征,本文通过骨架和边界的代表性点来建模红绿灯和标牌,并通过分段三次曲线来参数化线。基于道路语义特征,我们**构建了一个完成的定位建图框架**,

包括**a)图像处理前端,b)传感器融合策略和c)优化** 验证明了本文方法的有效性和相比于传统方法的优势。













知乎 | 首发于 语义定位与建图

- 1. 一个基于**卷积神经网络**的图像处理前端来提取语义特征;
- 2. **参数化道路元素**方法和设计的loss函数;
- 3. 语义优化模块能够被同时用于离线建图和在线定位。

系统框架:





分享

本文建图与定位系统的核心框架式一个紧耦合的batch和滑动窗口状态优化框架。具体上,本文算法首先构建基本道路语义元素的语义地图,然后使用该地图进行在线定位。语义元素主要包含三种类型:垂直物体、路面物体和线,如图2所示。











首发于 语义定位与建图





在离线建图过程,感知模型(采用DLA-34 和 DCN模块)对每个关键帧执行来提取语义元素。然后 进行连续关键帧之间**语义元素跟踪**,来建立多视角关联,用来估计相机轨迹和路标位置。然后,重 新观测场景通过**闭环检测**发现,并进行融合。最后该地图用于定位。

在线基于地图定位过程中,感知模块运行频率低,因此,语义特征通过一个**混合检测和跟踪策略**来 获得。这些特征与地图进行匹配,并通过滑动窗口优化来减小全局漂移。

首发于

语义定位与建图





分享

实验结果:









语义定位与建图







首发于

语义定位与建图









语义定位与建图





语义定位与建图





首发于

语义定位与建图





分享

结论:

- 1. 本文呈现了一个完整的语义建图与定位框架;
- 2. 本文的建图方法得到的建图精度和认知中的视觉精度











知乎 | 首发于 语义定位与建图

综合总结:

- 1. 自动驾驶场景视觉语义建图与定位的文章大多来自于业界,学术界对该问题关注较少;
- 2. 视觉定位的鲁棒性还需进一步研究,尤其是如何从2D图像信息获得鲁棒的3D结构信息;
- 3. 这几篇文章的精度高低差别比较大,可能是使用数据不同或评估方法不同,急需相关数据集提供统一的标准来验证各个算法的效果。

编辑于 06-19

高精度地图

自动驾驶

同时定位和地图构建(SLAM)

文章被以下专栏收录



语义定位与建图

专注于自动驾驶场景语义定位与建图、语义数据关联

推荐阅读

▲ 已赞同 80



■ 2条评论

✔ 分享

● 喜欢

★ 收藏

首发于

语义定位与建图



很高,就算没做回环和后端优化定位精度也很高。高精度地图中,里程计也是考当前帧和地图参…

蒋成

TOOK (OCALIZATION

测试数据:

https://pan.baidu 提取码: n9ys本篇2

任乾











