

基于Catmull-Rom样条的在线单目车道映射

Zhijian Qiao, Zehuan Yu, Huan Yin 和 Shaojie Shen

摘要——在本研究中，我们提出了一种在线单目车道映射方法，该方法仅依赖于单个摄像头和里程计来生成基于样条的地图。我们提出的技术将车道关联过程建模为利用二分图的分配问题，并通过结合Chamfer距离、姿态不确定性和横向序列一致性为边分配权重。此外，我们精心设计了控制点初始化、样条参数化和优化，以逐步创建、扩展和细化样条。与先前使用自建数据集评估性能的研究不同，我们的实验是在公开可访问的OpenLane数据集上进行的。实验结果表明，我们提出的方法提高了车道关联和里程计的精度，以及整体车道地图的质量。我们已将本项目的代码^{开源}。

1. 引言

高精度（HD）地图对于自动驾驶至关重要。作为HD地图中的关键要素，车道标线在车辆姿态估计、轨迹规划以及高级任务决策中发挥着至关重要的作用。然而，创建和维护包含车道的HD地图通常涉及多个流程[1]、[2]，例如，通过专用车辆进行高质量数据采集、三维重建和地标标注。这种昂贵的映射系统阻碍了自动驾驶在大规模环境中的推广。一个有前景的方向是利用标准车辆上的廉价传感器实现在线车道映射。有人可能会建议通过直接累积使用车辆里程计信息获取的车道检测结果来获得车道地图，而无需设计额外的模块。这种方法将面临两个重大问题。首先，从累积的语义点云中获得的地图不包含实例级信息。其次，里程计漂移和检测错误会使车道标线变得模糊和不准确。因此，如果使用累积的车道检测结果，仍然需要离线手动标注和矢量化。

在本研究中，我们提出使用同时定位与建图（SLAM）技术来实现在线车道映射。然而，在线车道映射仍然是一项具有挑战性的任务，因为车道标线通常是弯曲的，比SLAM领域中的传统点、线和面表示更难表示和优化到

本研究部分由香港科技大学研究生奖学金资助，部分由香港科技大学-大疆联合创新实验室资助，部分由香港建筑机器人中心（InnoHK中心，由香港创新科技署支持）资助。

作者就职于中国香港香港科技大学电子与计算机工程系。邮箱：zqiaoac@connect.ust.hk, zyuay@connect.ust.hk, eehyin@ust.hk, eeshaojie@ust.hk

通讯作者：Huan Yin

¹<https://github.com/HKUST-Aerial-Robotics/MonoLaneMapping>

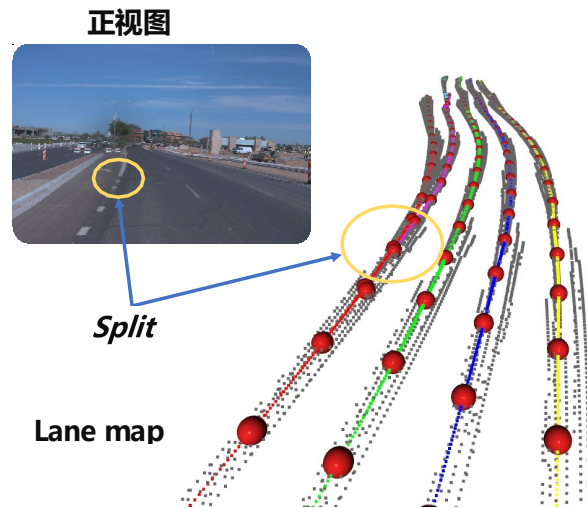


图 1. OpenLane 数据集上的实验结果。灰色点表示通过里程计累积的多帧检测结果。彩色曲线表示地图中不同实例的样条采样点。红色球体表示样条的控制点。

为了应对这些挑战，我们提出使用Catmull-Rom样条[3]来对车道标线进行建模（下文中为简洁起见，也称为车道）。我们选择Catmull-Rom样条主要有三个原因。首先，它自然具有 C^1 连续性，并且足够平滑，适用于车道标线建模。其次，与点集表示法相比，它是一种轻量级的表示方法；因此，我们只需存储控制点即可构建连续的车道标线模型。第三，所有控制点都位于样条曲线上，这有助于获取最优的控制点初始化，并能够开发出从粗到细的样条参数化技术。

基于Catmull-Rom样条表示法，设计了一个完整的在线车道映射系统，如图1所示。该系统允许自动驾驶车辆使用时间图像和里程计数据实时构建局部车道地图，这些地图可用于自定位、规划和众包更新[4]。总体而言，具体贡献可概括如下：

- 我们提出了一种在线单目车道映射系统，包括车道跟踪和地图优化子系统。该系统可以直接输出由Catmull-Rom样条表示的轻量级实例级车道地图，无需离线矢量化。
- 系统的每个部分都经过精心设计，结合了车道标记和样条的特性，包括车道关联、姿态估计、样条初始化、扩展和优化。

- 在公开数据集 OpenLane 上的实验表明，我们提出的方法能够提高车道关联性、里程计精度和地图质量。

II. 相关研究

A. 3D 车道检测

近年来，在线视觉车道检测领域在准确性和速度方面取得了显著进展 [5]、[6]。这些方法大多强调在图像空间中检测 2D 车道。与点不同的是，当通过三角测量将 2D 车道投影到世界坐标系时，会面临挑战。相反，大多数研究利用逆透视映射 (IPM) [7] 将车道投影到路面上。然而，这种方法严重依赖于地面平坦的假设和摄像头俯仰角的估计，这些因素在车辆振动或道路高度变化时可能导致 3D 车道的扭曲 [8]。

一种独特的方法是在鸟瞰图 (BEV) 空间中处理图像特征并检测车道标记，如 HDMapNet [9] 和 MapTR [10] 所示，或直接从 BEV 图像中识别车道图 [11]。然而，这些方法并不回归高度。相比之下，Persformer [12] 是一种开源的 3D 车道检测前沿方法，它通过在 BEV 空间上建立 3D 车道锚点来回归偏移量和高度，从而直接预测 3D 车道。

B. 车道映射

由于车道映射对自动驾驶汽车的重要性，过去十年中已对其进行了广泛研究。文献 [13]、[14] 中的工作使用了语义分割结合深度或平面地面信息来重建三维车道点云。然而，将点云与姿态信息累积会使车道地图尺寸变得庞大。文献 [15]、[16] 中的研究人员提出了创建栅格以减小地图尺寸，但栅格也会引入量化误差。更好的方法是使用折线来近似车道，并估计每个节点的位置和不确定性 [17]、[8]。然而，这种表示方法在某些弯曲车道场景中显然不够充分。另一种选择是样条 [18]、[19]，它紧凑且自然连续，但缺点是控制点初始化、扩展和优化具有挑战性。尽管已有样条辅助的激光雷达映射方法 [20]、[21] 被提出，但 these 方法不能直接用于单目车道映射，因为它们依赖于高精度的激光雷达测量。

此外，一些基于深度学习的在线地图构建方法 [9]、[10] 取得了令人瞩目的成果。然而，很少有研究实施时间融合来提高地图的一致性和准确性。

与以往的研究相比，我们无需使用激光雷达 (LiDAR) 或平面地面假设。相反，我们直接通过单目摄像头预测三维车道标线。为实现这一目标，我们的方法采用了

Catmull-Rom 样条，能够高效地建模和优化车道标线复杂且常呈弯曲的特性。

C. 基于边缘的视觉里程计 (VO)

我们的方法在一定程度上也受到了基于边缘的视觉里程计的启发。边缘最初被引入视觉里程计流程中，以避免陷入局部最优解，并增强对光照的鲁棒性 [22]、[23]。对边缘进行最近邻搜索可以忽略基本亮度恒定性的假设。早期的基于边缘的视觉里程计将图像边缘表示为一组像素，并为这些像素创建距离场，以获取从地图投影点到边缘的残差，并优化姿态 [22]、[24]。然而，这种方法存在三个问题。首先，距离场可能导致优化过程中的不可微性。因此，Ling 等人 [23] 提出了使用次梯度法。Zhou 等人 [25] 则提出使用最近邻场代替距离场。其次，研究人员在文献 [26] 中指出了建立残差的两种方式：数据到模型和模型到数据。同时，后者在部分遮挡的情况下可能会导致部分观测。在这种情况下，Zhou 等人 [25] 仍然使用了前者，因为他们声称所提出的点到切线残差可以减少这种影响。此外，如果模型指的是样条和数据点，部分观测还源于离散表示和连续表示之间的差异。第三，3D 世界中并不存在自然的 3D 边缘，这意味着在图像上检测到的边缘是错误的，如明显的轮廓或遮挡 [27]。

在本研究中，考虑到 CatmullRom 样条的特性，我们设计了一种从粗到细的参数化方法，以寻找点与曲线之间的最近邻（即脚点），计算残差，并采用数据到模型的方法。此外，检测到的车道作为语义边缘，既自然又独特。

III. 方法

为了更好地描述我们提出的系统，我们首先提供系统概述，然后在以下小节中对子系统进行全面说明。

A. 系统概述

所提出的单目车道映射系统的结构如图2所示。该系统仅使用单目摄像头和里程计（如 VIO、LIO）作为输入，输出以样条表示的紧凑车道标记图，无需先验导航图或航拍照片。具体而言，所提出的框架由两个子系统组成：车道跟踪和地图优化。在车道跟踪中，神经网络直接基于输入图像预测 3D 车道标记。然后，对预测结果进行进一步处理以满足后续需求，详见第III-B.1节。随后，将处理后的车道标记与里程计提供的姿态相结合，与地图中的车道标记相关联（第III-B.2节）。最后，基于关联结果更新姿态（见第IIIB.3节）。在地图优化中，首先从头开始

初始化样条，或者基于新获得的检测结果扩展原始样条 (第III-C.1节)。最后，应用增量优化框架 iSAM2[28] 来添

加新的观测值，以增量方式更新地图中的样条，同时不丢失过去观测值的信息 (第III-C.2节)。

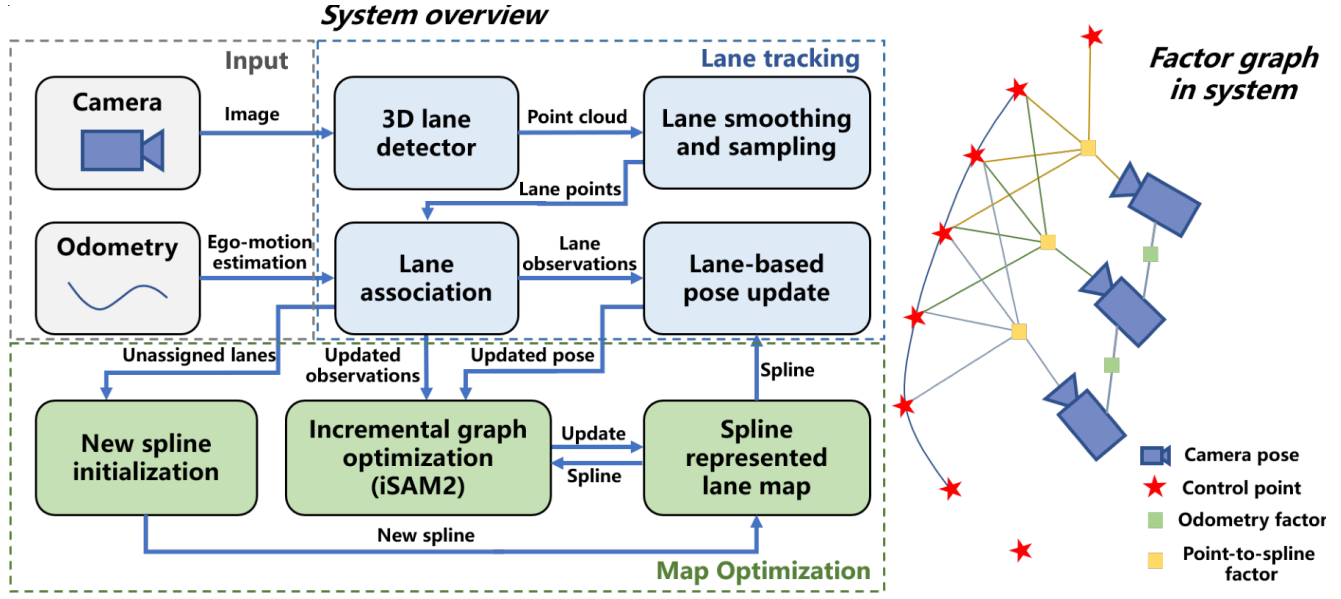


图 2. 框图展示了所提单目车道映射系统的整个流程。该系统分为两部分：车道跟踪和地图优化。前者用于车道标记关联和姿态更新，后者用于样条初始化、扩展和优化。右侧显示了因子图。与传统的二元视觉因子不同，在优化四个控制点地标时，引入了点到样条的因子。

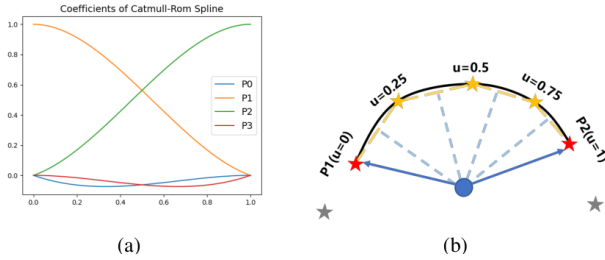


图 3. (a) $\tau = \frac{1}{5}$ 的四个控制点的系数。(b) 从粗到细的参数化。红色星形表示控制点，黄色表示采样航点。我们首先找到两个最近的控制点，然后通过折线找到脚点来确定参数。请注意，实际车道标线的曲率较小，并不像图中那么大。

实线)。 $d_{\sigma_{1:M}}$ 是检测噪声的标准差，可以设置为与点的 2 范数成正比。

$$D_i = \{d_{p_{1:M}}, f_{xy}, f_{xz}, c, d_{\sigma_{1:M}}\} \quad (1)$$

地图中的车道地标随后用 Catmull-Rom 样条表示，记为 $L_j = \{P_{0:N+1}, c\}$ ，其中 $P_{0:N+1}$ 是样条中的控制点， c 是类别。分段样条曲线 L_j 有 N 个段（每两个段共享三个控制点），每个段 l 有四个相邻的控制点，这些控制点确定了曲线的点 $l_p(u)$ ：

$$l_p(u) = \begin{bmatrix} 1 \\ u \\ u^2 \\ u^3 \end{bmatrix}^T \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\tau & 0 & \tau & 0 \\ 2\tau & \tau-3 & 3-2\tau & -\tau \\ -\tau & 2-\tau & \tau-2 & \tau \end{bmatrix} \begin{bmatrix} lP_0^T \\ lP_1^T \\ lP_2^T \\ lP_3^T \end{bmatrix} \quad (2)$$

B. 车道跟踪

1) 车道表示：在本研究中，我们利用 Persformer[12] 来获取车道检测结果，这些结果包括无序的车道标记点及其相关的实例级标签。此外，我们区分了关于观测（检测）和地标（地图）的车道表示。特别是对于网络预测，考虑到其稀疏性和噪声，我们首先将其转换到局部参考系（LRF），其中车道的主要方向与 X 轴对齐。随后，对 X-Y 和 X-Z 坐标进行三次多项式拟合，并以特定分辨率（在我们的实验中设置为 0.5m）进行采样。因此，车道观测可以表示为，其中 $d_{p_{1:M}} \in \mathbb{R}^3$ 是检测车道上的采样点， f_{xy} 和 f_{xz} 分别是多项式系数， c 是类别（例如，双黄线、白

其中 $u \in [0, 1]$ 被称为参数。寻找与曲线上某点对应的 u 的过程称为参数化； τ 控制曲线的过渡平滑程度，通常设置为 $\frac{1}{2}$ 。在不失一般性的情况下，我们将这四个控制点称为 $[lP_0, lP_1, lP_2, lP_3]$ 。在后续章节中，为简洁起见，可省略 l 和 d 。在这种情况下，曲线上的某点可视为四个控制点的加权。系数如图3(a)所示。

2) 车道关联：给定一组检测结果 D 和一组地标 L ，车道关联的目标是将检测结果 D_i 与已有的地标 L_j 进行匹配，或识别出新的车道。为此，我们使用二分图将该问题形式化为指派问题，并使用 KM 算法 [29]、[30] 进行求解，而关

键点在于如何确定边及其权重。首先，关联的车道应属于同一类别。另一个自然的想法是，可以先在样条上采样一些点，然后计算两个点云之间的距离，如 Chamfer 距离，以确定 D_i 与 L_j 的相似性。然而，Chamfer 距离总是可以计算的，因此在二分图中的每个顶点都会有一条边，导致不会生成新的车道标记。为此，受近期工作 KISS-ICP[31] 和 Kim 等人 [32] 工作的启发，我们限制了搜索范围以及 Chamfer 距离的上限。

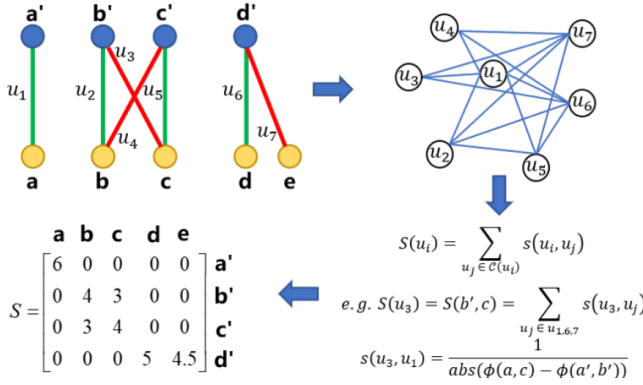


图 4. (a) 车道标线检测以黄色点表示，地图中的车道标线以蓝色点表示。 u 表示从欧几里得距离获得的关联，其中红色表示错误。(b) 创建一个无向图，顶点表示关联。如果两个关联具有横向序列一致性，则它们之间有一条边。(c) 给出了每条边的权重的定义。一个顶点的度是所有边的总和。(d) 顶点的度将作为指派问题中二分图的边。

我们计算了里程计位姿与真实位姿之间的差异 (R, t)。与李群切空间中噪声的传统假设不同，我们仅使用两个参数，即旋转和平移的标准差，分别表示为 σ_θ 和 σ_t ，其中：

$$\theta = \arccos\left(\frac{\text{tr}(\Delta R) - 1}{2}\right), t = \|\Delta t\|_2 \quad (3)$$

因此，对于点 p_k 在 D_i 中，其与真实匹配点距离的上界 δ_k (95% 的双西格玛规则) 可表示为，

$$\delta_k = 2\|p_k\|_2 \sin \frac{2\sigma_\theta}{2} + 2\sigma_t + 2^d \sigma_k \quad (4)$$

因此，给定样条上的采样点 $1:Q$ 、检测点 $p_{1:M}$ 以及里程计位姿 T ，我们将 D_i 和 L_j 之间的距离重新定义为

$$d_{ji} = \sqrt{\frac{M}{n_a}} \frac{1}{n_a} \sum_{k=1}^M \mathbb{I}(d_k < \delta_k) d_k, d_k = \|Tp_k - q_k'\|_2 \quad (5)$$

其中， n_a 表示满足距离阈值的点数， \mathbb{I} 是指标函数， q_k' 是 p_k 的最近采样点， $\sqrt{\frac{M}{n_a}}$ 用于惩罚低匹配率。此外，我们为 d_{ji} 设定了一个上限 $\sqrt{2\text{mean}(\delta_k)}$ ，以确定是否出现了新的车道。 $\sqrt{2}$ 表示至少有一半的点被匹配。

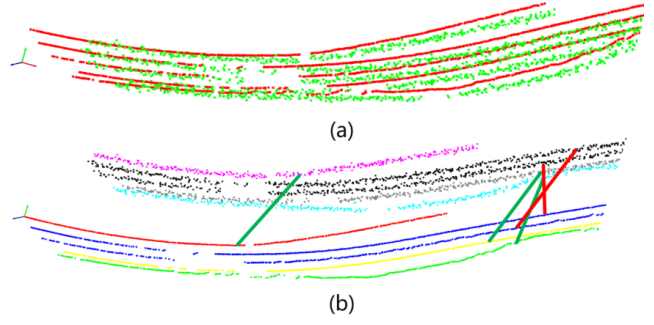


图 5. (a) 需要关联的两帧车道标线。(b) 这两帧的关联结果可视化。在每帧中，颜色代表车道标线类别。在关联中，红色表示错误，绿色表示正确。由于姿态不确定性，仅基于欧几里得距离的关联可能会导致错误结果。

尽管如此，仅依赖距离进行数据关联可能会产生歧义，尤其是在存在 R 的情况下，如图5所示。在欧几里得空间中，检测结果和地标并未很好地分离，甚至存在交叉的情况。为此，我们进一步使用横向顺序一致性对边进行加权。与大多数图匹配方法 [33] 一样，我们定义了两个关联对（二分图中的边）之间的一致性。为了更好地说明数据关联，我们在图4中配置了一个示例。该示例中有四条车道、五个地标和七个基于欧几里得距离的关联对，其中三个是错误的（但仍处于距离上限范围内）。每个关联对在图中作为一个顶点。两个顶点之间是否存在边取决于它们产生的关联结果的横向顺序一致性。例如，对于 BEV 中的 u_1 和 u_3 ，我们在车道 b_i 上采样两个点（例如，起点和终点）以创建一条直线，并确定车道 a' 的采样点（例如，中点）是位于直线上方还是下方，同样针对 a 和 c 。如果它们具有相同的相对关系，则 u_1 和 u_3 之间存在一致性。此外， u_2 和 u_3 也不一致，因为它们共享地标 b' 。进一步地，在不失一般性的前提下，我们以 u_3 为例来计算其一致性得分 $S(u_3)$ 。

$$S(u_3) = \sum_{u_j \in C(u_3)} s(u_3, u_j) \quad (6)$$

其中 C 表示与 u_3 相连的顶点集。对于 u_3 和 u_1 ，我们有，

$$s(u_3, u_1) = \frac{1}{\text{abs}(\phi(a, c) - \phi(a', b'))} \quad (7)$$

其中， $\text{abs}(\cdot)$ 表示绝对值， ϕ 表示前述点到线距离的最小值。

最后，二分图的边可以通过将两个分数相乘得出，其中一个分数是前文提到的 Chamfer 距离的倒数，另一个分数是 S 中相应的横向序列一致性。

3) 姿态更新：问题可以表述如下。设 T_t 表示在时间 t 时相机到世界坐标系的姿态变换。结合第III-B.2节的关

联结果, 使用 T_t 将 p_k 投影到世界坐标系, 通过寻找脚点 $p(u_k)$ 来确定关联样条上的参数 u_k 。给定 $p(u_k)$ 的局部方向 d_k , 由于车道标记仅提供姿态估计的横向约束, 因此定义了点到切线的残差。配准的总体目标是找到

$$\hat{T}_t = \underset{T_t}{\operatorname{argmin}} \sum_{k=1}^K \rho(\|(I - d^T d)(T_t p_k - p(u_k))\|_2^2 + \|(T_{O_t})^{-1} T_t\|_2^2) \quad (8)$$

其中, K 为所有检测到的车道中的点数, d_k 可通过 $p(u_k)$ 的一阶导数轻松获得, ρ 为鲁棒核函数, T_{O_t} 为里程计得到的姿态。

另一个关键问题是参数化。由于检测的不确定性和里程计漂移, 常用的参数化方法, 如弦长法和向心法, 可能无法有效工作。为此, 我们设计了一种从粗到细的参数化方法, 该方法考虑了 Catmull-Rom 样条的特性: 控制点位于曲线上, 且线段被限制在 P_1 和 P_2 之间。具体而言, 如图3(b)所示, 我们首先找到两个最近的控制点, 并给出三个条件:

- 1) 这两个控制点必须顺序相邻。2) 它们既不是样条曲线的第一个控制点, 也不是最后一个控制点。
- 3) 从 p_k 到两个控制点的距离必须小于这两个控制点之间的距离。

请注意, 由于出于舒适性考虑, 车道的曲率通常较小, 基于这一观察, 我们假设脚点必须位于由两个最近的控制点所创建的线段上。然后, 我们根据 u 值在该线段上均匀采样三个航点, 并用折线近似曲线。在这种情况下, 曲线上的脚点被转换为折线上的脚点。我们根据脚点来确定 u 值。

C. 地图优化

1) 车道初始化: 在优化地图中的样条之前, 需要对控制点 P 进行良好的初始化。为了避免自交, 期望 P 在弦长为 r 的样条上均匀分布。为此, 该问题可以描述为在给定检测点 $p_{1:M}$ 的情况下扩展 P 。

算法 1 描述了初始化过程的主要步骤。首先, 如果 P 为空 (新车道标线), 我们随机选择 p_0 作为控制点。然后, 我们使用 $p_{1:M}$ 来初始化候选集 V 并保持其更新。为了确定 V , 我们取第一个和最后一个控制点来建立法线平面, 法线向量是样条曲线上控制点的切线。如图6所示, 我们只取这两个法线平面之外的点, 并将它们添加到 V 中。否则, 算法终止。

Algorithm 1: Control Point Initialization

Input : control points P and detected $p_{1:M}, f_{xy}, f_{xz}$
Output : extended P

```

1 if  $P$  is  $\emptyset$  then  $P.append(p_0)$ ;
2  $V \leftarrow p_{1:M}$ ;
3 while true do
4    $N_h, N_t \leftarrow \text{ComputeNormalPlane}(P_{0:N+1})$ ;
5    $V \leftarrow \text{BeyondNormalPlane}(V, N_h, N_t)$ ;
6   if  $V$  is  $\emptyset$  then break;
7   if  $\|V_0 - P.head\|_2 \leq \|V_0 - P.head\|_2$  then
8     |  $\text{Expand}(P.head.next, P.head, f_{xy}, f_{xz})$ ;
9   else
10    |  $\text{Expand}(P.tail.prev, P.tail, f_{xy}, f_{xz})$ ;
11  end
12 end
13 def  $\text{Expand}(P_i, P_j, f_{xy}, f_{xz})$ :
14    $P_i, P_j \leftarrow \text{LRF.transform}(P_i, P_j)$ ;
15   if  $P_j.x \geq P_i.x$  then
16     |  $P_s = P_j + [3r, 0, 0]$ ;
17   else
18     |  $P_s = P_j + [-3r, 0, 0]$ ;
19   end
20   for  $i \leftarrow 1$  to 10 do
21     |  $P_f = [P_s.x, f_{xz}(P_s.x), f_{xz}(P_s.z)]$ ;
22     |  $P_s = \text{ProjectToSphere}(P_f)$ ;
23     | if  $P_s$  changes very little then break;
24   end
25   // Add  $P_s$  to  $P$ ;
26    $P_s \leftarrow \text{LRF.inv.transform}(P_s)$ ;
27   return  $P_s$ ;

```

在获得 V 后, 我们采用 V_0 来确定是延长样条的头部还是尾部。在不失一般性的情况下, 我们以尾部延伸为例。一种简单的方法是以 $P.tail$ 为中心, r 为半径绘制一个球体, 并将 V_0 在球体上的投影作为下一个控制点。该投影有一个解析解, 且易于计算。然而, 由于 $p_{1:M}$ 可能非常嘈杂且稀疏, 因此这种方法并不稳健。相反, 我们使用第三节B.1中介绍的 LRF 以及 f_{xy} 和 f_{xz} 。该问题转化为寻找球体与三次多项式曲线交点的问题。然而, 求解五次方程并不容易, 因此我们设计了一种启发式算法, 如算法1中的 Expand 所示。更直观的描述见图6。通常, 此过程在三到四次迭代后结束。如果不是这样, 则球体和曲线没有交点, 算法将 $P.tail$ 更新为 P_s 。它继续迭代, 直到满足条件并退出。

2) 样条优化: 本节将继续描述如何优化控制点。我们根据关联结果 (D, L) 和优化的姿态 T_t 对每个检测点 p_k 进行参数化, 以找到其参数 u_k 。我们将点对样条的残

差定义为

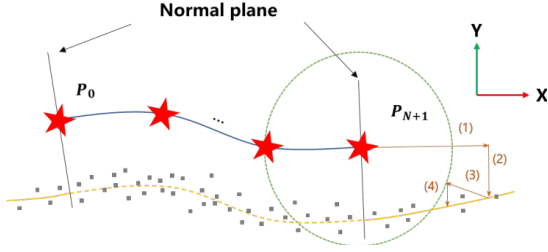


图 6. 该图解释了样条的扩展方式。首先，所有数据将在局部参考系下进行转换。然后，法线平面将空间进行划分，并选择用于扩展的观测点。最后，扩展问题转化为寻找球面与沿样条方向的曲线交点的问题。

$$r_k = \hat{T}_t p_k - p(u_k) \quad (9)$$

其中 $p(u_k)$ 是控制点的线性函数，且其雅可比矩阵已求得。

此外，我们需要在优化过程中增加另一个正则化项。尽管不同的样条段共享控制点，但仍存在一些点始终扮演着 P_0 和 P_3 的角色。如图 3(a) 所示， P_0 和 P_3 的系数始终较小，导致点对样条的残差无法很好地约束这些变量。因此，我们对这些点与其相邻点之间的相对关系施加了常数约束。

整体因子图如图 2 所示。为了保留历史信息，我们使用 GTSAM[34] 中的 iSAM2[28] 进行基于因子图的增量优化。这样，我们可以在新帧到来时仅更新部分变量，而不丢弃历史信息，从而在速度和准确性之间取得平衡。

IV. 实验

我们在基于 Waymo 数据集 [35] 构建的 OpenLane[12] 车道基准上评估了我们提出的方法。我们选择这个数据集是因为它具有 3D 车道标记和实例级跟踪标注。OpenLane 包含 1000 个标注的路段，其中 798 个用于训练车道检测，202 个用于评估映射质量和姿态估计。每个路段平均长约 135 米，以 10 FPS 的速度包含 198 帧。整个数据集包含 14 个车道标记的标注类别。大约 25% 的帧包含超过 6 个车道标记，这使得车道关联非常具有挑战性。此外，该数据集涵盖了各种天气条件和复杂的道路场景，如大弯道、上下坡和交叉口，能够有效测试车道映射方法在实际场景中的性能。

A. 车道关联评估

得益于 OpenLane 数据集为车道跟踪提供的标注，我们能够获取一段视频中所有车道标记的实例 ID。我们每 10 帧抽取一对用于关联。为了增加数据关联的难度，我

们施加了随机的 3 自由度刚体姿态变换。其中，X 和 Y 服从正态分布 $N(0 \text{ m}, 3 \text{ m})$ ，偏航角服从 $N(0, 2)$ 。在两帧中具有相同车道标记实例 ID 的配对集合记为 S_{true} 。对于关联方法给出的匹配对，如果它在 S_{true} 中，则为真阳性；否则为假阳性。

表 I

COMPARISON OF LANE ASSOCIATION PERFORMANCE
(↓: LOWER BETTER, ↑: HIGHER BETTER)

Method	F1↑	Precision(%)↑	Recall(%)↑	Time(ms)↓
Shell[32]	0.763	80.01	74.55	15.64
CLIPPER[37]	0.795	86.77	75.66	9.027
CD-XYZ	0.840	84.30	83.84	1.317
CD-LMR[38]	0.848	85.15	84.70	27.05
Ours-LMR[38]	0.946	94.89	94.62	28.85
Ours	0.931	93.39	93.07	2.225

为了进行比较，我们复现了文献 [32] 中提出的基于壳层的方法，并使用默认的壳层间隙 1.5 米 [36]。此外，我们基于图匹配进行数据关联，并使用 CLIPPER[37] 进行求解。在 CLIPPER 中，关键问题是定义点到点距离和噪声边界。我们的工作使用中间车道标记点来代替自身，并对边界进行微调。此外，我们对自己的方法进行了消融实验。CD+XYZ 表示我们省略了横向序列一致性。LMR 表示将 XYZ 特征转换为文献 [38] 中提出的 LMR 特征。

表 I 表明，我们提出的解决方案在 F1 分数方面是最优的，并且保持了较高的效率水平。值得注意的是，尽管 LMR 特征优于仅使用 XYZ，但它们也引入了更高的计算成本。基于壳层的方法的初衷是考虑欧几里得距离和形状相似性。在我们的实验设置中，由于显著的姿态误差，这种方法效果不佳。在纯基于点的距离方法中，CLIPPER 由于其严格的最大加权团约束而更为准确，但这也降低了召回率。

B. 姿态更新评估

由于 OpenLane 不提供 IMU 测量或轮式编码器测量，这使我们无法运行 VIO 算法来获取实际的里程计姿态。为了模拟里程计引起的漂移，我们在每两帧之间添加了 3 自由度的随机噪声。所添加的噪声远大于在 KITTI 里程计基准 [39] 上进行的大多数里程计工作。

表二

Noise	10m		30m		50m	
	Rot($^{\circ}$) \downarrow	Trans(m) \downarrow	Rot($^{\circ}$) \downarrow	Trans(m) \downarrow	Rot($^{\circ}$) \downarrow	Trans(m) \downarrow
(0.1, 0.1)	0.415/0.460	0.498/0.534	0.620/0.654	0.794/0.802	0.753/0.804	0.997/1.022
(0.2, 0.2)	0.687/0.921	0.866/1.068	0.994/1.308	1.358/1.604	1.215/1.609	1.657/2.044
(0.3, 0.3)	0.944/1.381	1.224/1.603	1.331/1.962	1.893/2.407	1.568/2.413	2.278/3.067
(0.4, 0.4)	1.178/1.841	1.632/2.137	1.601/2.616	2.460/3.210	1.884/3.217	2.928/4.091
(0.5, 0.5)	1.464/2.302	2.088/2.672	1.980/3.271	3.127/4.013	2.291/4.022	3.677/5.116

表II展示了车道标线在位姿估计中的作用。第一列说明了添加的噪声水平，例如，(0.1, 0.1) 表示旋转噪声标准差为 0.1 $^{\circ}$ ，平移噪声标准差为 0.1 米。我们分别计算了

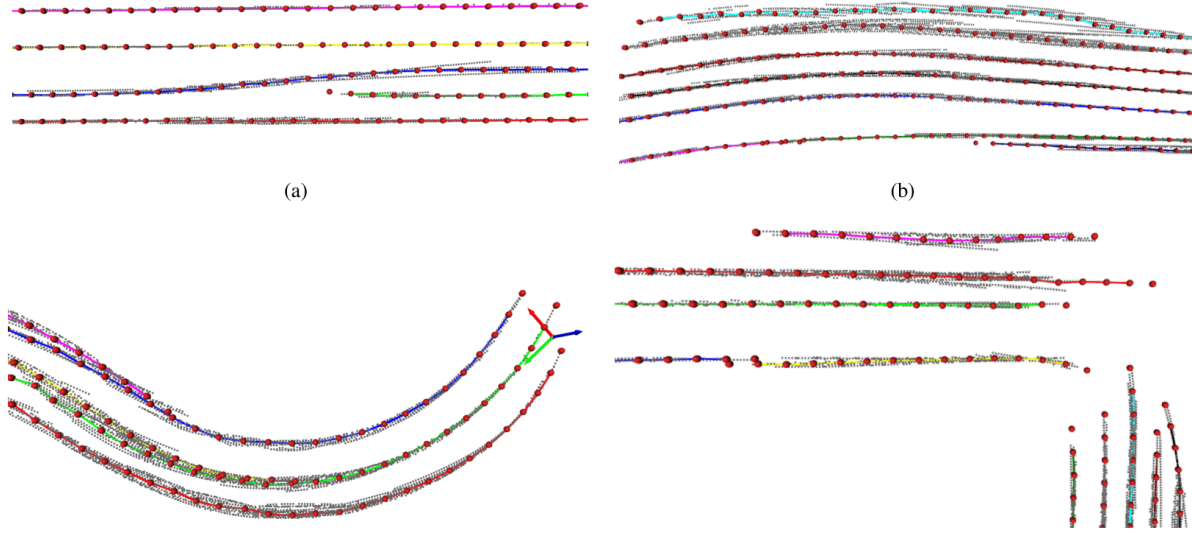


图 7. 车道地图可视化。红色球体代表控制点，彩色点表示不同样条实例上的采样点。灰色点云是下采样后的原始检测结果。该图展示了车道的各种场景：分叉、直行、弯曲和交叉。

C. 地图质量评估

OpenLane 并不提供全局高清地图的真实数据，而仅提供每帧的 3D 车道标记标注。因此，我们将样条从地图投影到每帧并进行点采样。为了进行公平比较，我们不将全局地图投影到每帧，而仅投影使用历史信息构建的在线局部地图。与 OpenLane 中的评估方法 [12] 类似，我们报告了前方 50 米局部区域内每帧的召回率和精确率。¹如果车道标记上的点与真实数据上的点之间的距离小于 0.5 米，则该点有效。²真阳性定义为有效点数量超过相应真实数据 (GT) 75% 的车道标记。

为了更精确地评估所提出方法的性能，我们将验证集划分为不同的场景。表III中的结果显示，我们提出的方法通过利用所有场景中的过往检测结果，提高了车道标记的召回率和精确率。然而，我们注意到在曲线场景中，改进幅度最小。通过比较原始检测结果与优化后的样条曲线，我们发现这种情况在大曲率下会产生许多错误，尤其是在

距离为 10 米、30 米和 50 米时的相对位姿误差 [39]。斜线右侧为添加噪声后的原始里程计结果。相比之下，左侧表示更新后的位姿。表中的结果表明，整体车道标线有助于提高仅基于里程计的位姿估计。然而，当噪声较小时，这种提高效果并不明显。这揭示了车道标线在位姿估计中可能存在的上限。这可能是由于两个原因：网络检测到的 3D 车道标线不够准确，以及地图中样条的不确定性建模不够完善。

远距离时。这可能与我们的检测器基于直线锚点有关。在这种情况下，大量的异常值会对优化器产生干扰。

表 III

C^O 比较 L_{ANE} 与 M_{APPING} $F1SCORE \uparrow$							
Method	Up & Down	Curve	Extreme Weather	Intersection	Merge & Split	Night	All
PersFormer[12]	0.407	0.595	0.514	0.474	0.575	0.518	0.559
Ours	0.459	0.609	0.552	0.523	0.631	0.566	0.609

表 iv

COMPARISON OF LANE MAPPING ON ALL VALIDATION SEGMENTS

Method	F1 \uparrow	Recall \uparrow	Precision \uparrow	F1 increase \uparrow
PersFormer (0.8)	0.497	0.418	0.614	-
PersFormer (0.6)	0.515	0.445	0.613	-
PersFormer (0.4)	0.530	0.467	0.611	-
PersFormer (0.0)	0.559	0.517	0.609	-
Ours (0.8)	0.539	0.450	0.672	8.45%
Ours (0.6)	0.559	0.479	0.670	8.54%
Ours (0.4)	0.576	0.505	0.670	8.68%
Ours (0.0)	0.609	0.559	0.668	8.94%

此外，为了评估我们的系统对于不同性能检测器的改

¹在 OpenLane 中，该距离为 100 米。我们选择 50 米是因为我们仅在 50 米范围内使用车道检测进行地图构建。

²在 OpenLane 中，该值默认为 1.5 米。此外，我们使用 KD 树来查找最近点，而不是沿固定的 Y 轴间隔进行采样，因为我们发现某些 Y 轴间隔没有车道标记点。

进程度, 我们以 p 的概率从当前的检测中随机移除一个车道标记。如表IV所示, 最大 p 值为 0.8。结果表明, 我们提出的方法能够在不同水平的干扰下稳定提升地图中车道标记的质量。

V. 讨论与结论

与单帧在线映射相比, 所提出的系统能够利用历史检测信息。保留历史信息有助于提高车道标记的召回率。此外, 将历史观测值纳入优化过程, 可以更精确地表示车道标记。另外, 该系统直接生成由样条曲线表示的矢量化地图。然而, 该方法的性能仍然取决于车道检测性能。此外, 与物理传感器不同, 对网络预测的不确定性进行建模需要付出努力。最后, 环路闭合对于构建全局一致的车道地图至关重要。在后续研究中, 我们将继续探索在线车道映射, 以应对上述挑战。

参考文献

- [1] N. Homayounfar, W.-C. Ma, J. Liang, X. Wu, J. Fan, and R. Urtasun, “Dagmapper: Learning to map by discovering lane topology,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 2911-2920.
- [2] K. Tang, X. Cao, Z. Cao, T. Zhou, E. Li, A. Liu, S. Zou, C. Liu, S. Mei, E. Sizikova *et al.*, “Thma: Tencent hd map ai system for creating hd map annotations,” *arXiv preprint arXiv:2212.11123*, 2022.
- [3] E. Catmull and R. Rom, “A class of local interpolating splines,” in *Computer aided geometric design*. Elsevier, 1974, pp. 317-326.
- [4] D. Pannen, M. Liebner, W. Hempel, and W. Burgard, “How to keep hd maps for automated driving up to date,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 2288-2294.
- [5] Z. Qin, P. Zhang, and X. Li, “Ultra fast deep lane detection with hybrid anchor driven ordinal classification,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
- [6] A. Bar Hillel, R. Lerner, D. Levi, and G. Raz, “Recent progress in road and lane detection: a survey,” *Machine vision and applications*, vol. 25, no. 3, pp. 727-745, 2014.
- [7] J. Jeong and A. Kim, “Adaptive inverse perspective mapping for lane map generation with slam,” in *2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*. IEEE, 2016, pp. 38-41.
- [8] Y. Zhou, X. Li, S. Li, and X. Wang, “Visual mapping and localization system based on compact instance-level road markings with spatial uncertainty,” *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10 802-10 809, 2022.
- [9] Q. Li, Y. Wang, Y. Wang, and H. Zhao, “Hdmapnet: An online hd map construction and evaluation framework,” in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 4628-4634.
- [10] B. Liao, S. Chen, X. Wang, T. Cheng, Q. Zhang, W. Liu, and C. Huang, “Maptr: Structured modeling and learning for online vectorized hd map construction,” *arXiv preprint arXiv:2208.14437*, 2022.
- [11] J. Zlurn, J. Vertens, and W. Burgard, “Lane graph estimation for scene understanding in urban driving,” *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 8615-8622, 2021.
- [12] L. Chen, C. Sima, Y. Li, Z. Zheng, J. Xu, X. Geng, H. Li, C. He, J. Shi, Y. Qiao *et al.*, “Persformer: 3d lane detection via perspective transformer and the openlane benchmark,” in *Computer Vision-ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23-27, 2022, Proceedings, Part XXXVIII*. Springer, 2022, pp. 550-567.
- [13] Q. Cheng, N. Zeller, and D. Cremers, “Vision-based large-scale 3d semantic mapping for autonomous driving applications,” in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 9235-9242.
- [14] J. Jeong, Y. Cho, and A. Kim, “Road-slam: Road marking based slam with lane-level accuracy,” in *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2017, pp. 1736-1473.
- [15] D. Paz, H. Zhang, Q. Li, H. Xiang, and H. I. Christensen, “Probabilistic semantic mapping for urban autonomous driving applications,” in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 2059-2064.

- [16] T. Qin, Y. Zheng, T. Chen, Y. Chen, and Q. Su, "A light-weight semantic map for visual localization towards autonomous driving," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 11 248-11 254.
- [17] T. Heidenreich, J. Spehr, and C. Stiller, "Laneslam-simultaneous pose and lane estimation using maps with lane-level accuracy," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*. IEEE, 2015, pp. 2512-2517.
- [18] W. Cheng, S. Yang, M. Zhou, Z. Liu, Y. Chen, and M. Li, "Road mapping and localization using sparse semantic visual features," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 8118-8125, 2021.
- [19] K. Meier, S.-J. Chung, and S. Hutchinson, "Visual-inertial curve simultaneous localization and mapping: Creating a sparse structured world without feature points," *Journal of Field Robotics*, vol. 35, no. 4, pp. 516-544, 2018.
- [20] M. Usman, A. M. Khan, A. Ali, S. Yaqub, K. M. Zuhaib, J. Y. Lee, and C.-S. Han, "An extensive approach to features detection and description for 2-d range data using active b-splines," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2934-2941, 2019.
- [21] L. Pedraza, D. Rodriguez-Losada, F. Matia, G. Disanayake, and J. V. Mirto, "Extending the limits of feature-based slam with b-splines," *IEEE Transactions on Robotics*, vol. 25, no. 2, pp. 353-366, 2009.
- [22] L. Kneip, Z. Yi, and H. Li, "Sdicp: Semi-dense tracking based on iterative closest points." in *Bmvc*, 2015, pp. 100-1.
- [23] Y. Ling, M. Kuse, and S. Shen, "Edge alignment-based visual-inertial fusion for tracking of aggressive motions," *Autonomous Robots*, vol. 42, pp. 513-528, 2018.
- [24] J. J. Tarrio and S. Pedre, "Realtime edge-based visual odometry for a monocular camera," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 702-710.
- [25] Y. Zhou, H. Li, and L. Kneip, "Canny-vo: Visual odometry with rgb-d cameras based on geometric 3-d-2-d edge alignment," *IEEE Transactions on Robotics*, vol. 35, no. 1, pp. 184-199, 2018.
- [26] I. Nurutdinova and A. Fitzgibbon, "Towards point-less structure from motion: 3d reconstruction and camera parameters from general 3d curves," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 2363-2371.
- [27] Z. Wang and L. Kneip, "Fully automatic structure from motion with a spline-based environment representation," *arXiv preprint arXiv:1810.12532*, 2018.
- [28] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert, "isam2: Incremental smoothing and mapping using the bayes tree," *The International Journal of Robotics Research*, vol. 31, no. 2, pp. 216-235, 2012.
- [29] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval research logistics quarterly*, vol. 2, no. 1-2, pp. 83-97, 1955.
- [30] J. Munkres, "Algorithms for the assignment and transportation problems," *Journal of the society for industrial and applied mathematics*, vol. 5, no. 1, pp. 32-38, 1957.
- [31] I. Vizzo, T. Guadagnino, B. Mersch, L. Wiesmann, J. Behley, and C. Stachniss, "Kiss-icp: In defense of point-to-point icp simple, accurate, and robust registration if done the right way," *IEEE Robotics and Automation Letters*, 2023.
- [32] K. Kim, S. Cho, and W. Chung, "Hd map update for autonomous driving with crowdsourced data," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1895-1901, 2021.
- [33] M. Leordeanu and M. Hebert, "A spectral technique for correspondence problems using pairwise constraints," in *Tenth IEEE International Conference on Computer Vision (ICCV' 05) Volume 1*, vol. 2. IEEE, 2005, pp. 1482-1489.
- [34] F. Dellaert and G. Contributors, *borglab/gtsam*, May 2022. [Online]. Available: <https://github.com/borglab/gtsam>

- [35] P. Sun, H. Kretzschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. Caine *et al.*, “Scalability in perception for autonomous driving: Waymo open dataset,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2446-2454.
- [36] D. Pannen, M. Liebner, and W. Burgard, “Lane marking learning based on crowdsourced data,” in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 7040-7046.
- [37] P. C. Lusk, K. Fathian, and J. P. How, “Clipper: A graph-theoretic framework for robust data association,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 13 828-13 834.
- [38] M. tA. Muñ oz-Bañ ton, J.-H. Pauls, H. Hu, and C. Stiller, “Da-lmr: A robust lane marking representation for data association,” in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 2193-2199.
- [39] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for autonomous driving? the kitti vision benchmark suite,” in *2012 IEEE conference on computer vision and pattern recognition*. IEEE, 2012, pp. 3354-3361.