TrajMatch：通过轨迹匹配实现路侧激光雷达自动时空标定

任浩杰、张莎、李素刚、李瑶、李昕辰、纪建民、张宇、张岩勇，*IEEE* 会士

arXiv:2302.02157v1 [cs.RO] 4 Feb 2023

摘要—近年来，在路侧部署激光雷达等传感器，以监测过往交通并辅助自动驾驶车辆感知，已日益普及。与自动驾驶车辆系统不同，路侧传感器通常隶属于不同的子系统，且在时间和空间上缺乏同步。标定是一项关键技术，可使中央服务器融合不同位置基础设施生成的数据，从而扩大感知范围并提升检测鲁棒性。遗憾的是，现有标定算法往往假设激光雷达之间存在显著重叠，或已完成时间标定。由于这些假设在现实世界中并非总是成立，现有算法的标定结果往往不尽如人意，且始终需要人工参与，从而带来高昂的人力成本。

本文提出 TrajMatch ——首个可自动对路侧激光雷达进行时间和空间标定的系统。其核心思想是基于检测/跟踪任务的结果自动标定传感器，而非提取特征。进一步而言，我们提出了一种与算法相一致的标定参数评估机制，并通过实验验证了该方案的有效性；该机制还可用于指导多传感器标定中的参数迭代。最后，为评估 TrajMatch 的性能，我们收集了两个数据集：一个仿真数据集LiDARnet-sim1.0 和一个真实世界数据集。实验结果表明，TrajMatch 可实现小于 $10cm$ 的空间标定误差和小于$1.5ms$的时间标定误差。

索引词—**路侧交通监测、时空标定、轨迹匹配**

I. 引言

IT 让智能交通系统在重要位置，尤其是繁忙路口，使用摄像机/激光雷达进行近距离交通监控，以改善交通流、驾驶安全和能源效率，已成为近期趋势。具体而言，激光雷达越来越多地应用于此类场景，因为它能够提供精确的测距信息，这对交通监控尤为重要。随着交通监控对精度的要求不断提高，在交叉路口部署多个传感器以协同获得更好的覆盖范围 [1]也已成为一种常见做法。

H. Ren、S. Zhang、Y. Li、X. Li、Y. Zhang、J. Ji 和 Y. Zhang 均就职于中国合肥中国科学技术大学计算机科学与技术学院（电子邮箱：rhj@mail.ustc.edu.cn、zhsh1@mail.ustc.edu.cn、zkdly@mail.ustc.edu.cn、yuzhang@ustc.edu.cn、jianmin@ustc.edu.cn、yanyongz@ustc.edu.cn）。S. Li 就职于美国加利福尼亚州桑尼维尔的 Google Cloud（电子邮箱：sugan-gli@google.com）。通讯作者：张岩勇。

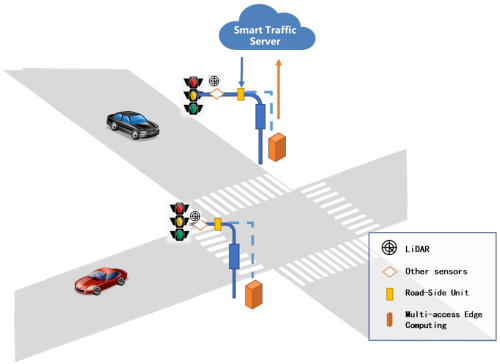


图 1. 路侧交通监控系统的一种部署示例。我们展示了交叉路口的经典部署场景，其中激光雷达以对角方式安装，以实现协同感知。

图 1 展示了一个交叉路口场景：配备计算与感知单元的智能路侧基础设施会采集实时环境数据并执行检测和跟踪等必要计算，随后将数据传输至智能交通服务器（STS），在那里融合这些数据。近年来，激光雷达技术在检测和跟踪方面已逐渐成熟，也很容易部署到路侧基础设施上。在这项工作中，我们关注基础设施上传感器的另一个关键问题——时空标定，因为它能让交通服务器准确融合来自不同位置激光雷达的数据，从而提升感知能力、扩大感知范围、实现更稳健的目标检测/跟踪，并生成更全面的交通状况视图。

我们注意到，关于多激光雷达标定的研究很少。 [2]尽管已有一些相关方法，但由于自身局限性和算法性能较低，它们并不适用于路侧多激光雷达。大多数研究聚焦于空间中两个传感器的标定，并假设传感器之间的时间已经同步 [2], [3], [4]。在时间维度上， [5]使用随机采样一致性（RANSAC）算法对齐激光雷达的轨迹，但假设最多只能看到一个人。同样的假设也出现在 [6]中。显然，这一假设在路侧交通场景中并非总是成立。在空间维度上，主要有两类方法：基于目标的方法和无目标方法。基于目标的方法需要可识别物体

（例如棋盘格、多边形标定板和 AprilTag 标记），并通过对齐各传感器观测到的目标位置来估计传感器之间的相对位姿 [7], [8], [3]。这些方法需要耗费大量成本和时间进行人工操作，无法应用于大规模部署。至于无目标方法，通常会提取环境中的特征。这类方法在交通场景中无法实现高精度，因为其中存在大量重复结构，且缺乏显著特征 [9]。

鉴于上述挑战，本文提出 TrajMatch——首个能够自动对路侧激光雷达进行时间和空间标定，以支持各种智能交通应用的系统。其核心思路是利用基于点云的目标检测与跟踪模块提供的运动物体轨迹信息，搜索不同激光雷达轨迹之间的对应关系。考虑到算法的精度和鲁棒性，我们采用由局部到全局的三阶段框架。第一阶段，我们提取物体出现的各个位置和时刻上的旋转与平移不变运动特征，并根据特征距离建立位置层面的对应关系。第二阶段，我们考虑更具判别性的物体语义信息及其邻域分布（我们称之为中层信息），以筛除错误对应关系。最后，我们基于全部数据进行迭代优化，实现高效标定。

为评估 TrajMatch 的性能，我们收集了一个仿真数据集（LiDARnet-sim 1.0）和一个真实世界交通数据集，两者均针对典型的四岔路口。对于空间标定，结果表明，TrajMatch 的成功率超过 90%，且可达到厘米级精度。对于时间标定，TrajMatch 的精度可达到 1.5 毫秒以内。

总之，本文的主要贡献如下：

1）我们提出了 TrajMatch，这是首个能够自动实现路侧激光雷达精确时空标定的系统。TrajMatch 具有以下主要优势。首先，该算法利用目标检测与跟踪模块的结果进行激光雷达标定，无需为标定任务单独提取特征。其次，该算法利用物体运动中蕴含的时空信息，无需依赖环境中的任何特征。这意味着算法效果不受特定环境影响，具有良好的可迁移性。第三，该系统不对任何初始状态作出假设，包括各激光雷达的位置和旋转，以及它们之间的空间/时间距离。因此，我们的系统提供了一种轻量、自动、高效且稳健的激光雷达标定方法。

2）我们提出了一种连续标定方法，可在汽车驶过时持续优化标定。具体而言，我们的系统能够自行评估每次标定会话的质量，并将表现良好的

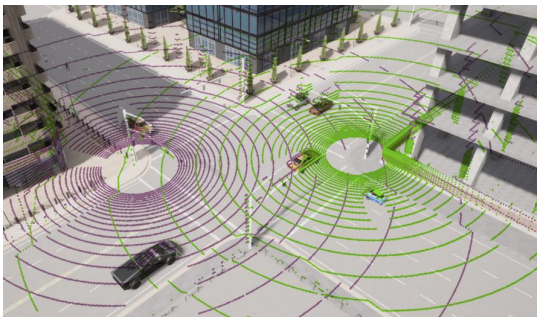


图 2. **四岔路口中对角部署的两台激光雷达生成的点云。**

会话整合起来，以持续提高标定精度。3）我们的系统无需依赖任何有关部署配置和环境的外部信息，即可实现厘米级空间标定精度和毫秒级时间标定精度。

II. 背景与动机

A. 路侧交通监测场景

本文聚焦于图 1 所示的场景。激光雷达传感器安装在路侧，协同监测经过的交通流。在此类多激光雷达协同感知系统中，各传感器的检测区域存在一定程度的重叠。图 3 列举了三种部署场景：(a) 四岔路口；(b) 三岔路口；(c) 人行道。这些激光雷达并未部署在同一位置，因此彼此之间存在显著的变换。此外，它们通常隶属于不同的子系统，导致时钟不同步。有时，时间偏差可达数秒[10][11]。值得一提的是，我们在这些场景中评估了该算法的性能，并在第 V-C 节取得了良好结果。

协同交通感知系统的工作流程通常包括以下步骤：(1) 数据采集：激光雷达传感器获取点云数据；(2) 目标检测：通过检测模块获取目标的边界框；(3) 多目标跟踪：通过跟踪模块为目标分配轨迹标识符，以区分不同目标的轨迹，并生成目标轨迹；(4) 将数据上传至中央服务器；(5) 合并来自不同子系统的数据，这需要了解子系统之间的相对位置。本文设计了 TrajMatch，通过匹配不同激光雷达检测到的同一目标的轨迹，完成时空标定。自然，这有助于中央服务器融合数据。

B. 问题表述

接下来，我们正式定义激光雷达时空标定问题。假设 $P = \{p_1, p_2, p_3, ...\}$ 和 $Q = \{q_1, q_2, q_3, ...\}$ 是两组包含时间维度的四维点云。P 和 Q 中的点具有相似的形式：

$$p_i = \{x_i^p, y_i^p, z_i^p, t_i^p\},\ q_j = \{x_j^q, y_j^q, z_j^q, t_j^q\}, \qquad (1)$$

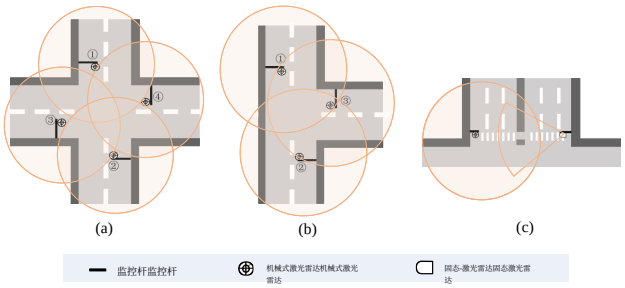


图 3. **路侧交通监测的典型部署场景：(a) 四岔路口，(b) 三岔路口，以及 (c) 人行道。**

其中，x、y和z分别表示各自坐标系中的三个维度坐标，t表示时间戳。特别地，在标定之前，P和Q所处的坐标系不同。例如，P的x轴可能对应于Q的y轴。此外，同一时刻采集的点云在时间维度t上也可能存在一定偏差。

在这项工作中，我们在四维空间中定义该问题。对于一个点$p = [x, y, z, t]^T$，坐标变换可定义为：

$$p' = \begin{bmatrix} R_3 & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \\ d_t \end{bmatrix} = R \times p + T, \tag{2}$$

其中，d_t表示时间偏移，$\mathbf{R}$和$\mathbf{T}$分别表示激光雷达之间的旋转矩阵和平移向量。

两个四维点云P和Q的时空标定问题可描述为：

$$R, T = \arg\min_{R,T} \frac{1}{|P|} \sum_{i=1}^{|P|} \|p_i - (R \times q_i + T)\|^2, \tag{3}$$

其中，p_i和q_i分别是P和Q中的对应点。

通过这种方式，我们将激光雷达标定问题转化为一个优化问题，旨在寻找旋转参数R和T所表示的平移向量，使两个激光雷达的数据在时间和空间上的匹配误差最小化。我们将R和T统称为系统的变换参数。

III. 系统概述

路侧激光雷达之间的时空标定旨在获得激光雷达之间的变换参数，并对齐它们的时间线。根据点云中的检测/跟踪结果，我们的任务是找到合适的变换，以便连贯地融合来自不同激光雷达的数据。我们的系统支持两种数据融合模式。首先，完成标定后，可以直接合并原始点云，形成统一点云。其次，我们的算法计算出的变换也可用于融合高层信息，例如深度学习模型提取的特征，甚至是目标列表及其轨迹。融合原始点云能够最大限度地保留信息，可能进一步实现点云分割等细粒度增强。然而，这种融合方法需要昂贵的硬件支持，尤其是在网络通信方面。以 RS-Ruby Lite [12] 为例，其每秒输出约 1，440，000 个点，每秒生成约 $30M$ 的点云数据。因此，在本研究中，我们采用第二种方案，融合目标级信息。鉴于当今基础设施的能力，这是一种务实且有效的方法。

为实现这一目标，我们提出了 TrajMatch，其框架如图4所示。我们的标定方法不假设系统具有任何初始状态，无论是变换参数还是激光雷达之间的时间偏差。作为一个无状态系统，我们通过高效的迭代优化过程，同时对激光雷达进行空间和时间对齐：匹配物体的位置，据此计算变换参数，然后使用这些参数对齐轨迹。我们力求在每个步骤中取得进展，并通过迭代逐步收敛到最优状态。下面简要概述各个步骤：

1) 运动特征提取。 获得各个激光雷达的轨迹信息（由一系列位置组成）后，我们首先提取每个位置处的运动特征，包括速度和曲率。选取这两种特征，是因为它们具有旋转和平移不变性。2) 语义特征匹配。基于运动特征的相似性，我们可以获得一组初始位置匹配。然而，在当前场景中，运动特征的区分能力较弱，因此会出现许多错误匹配。为过滤错误匹配，我们采用目标检测/跟踪阶段生成的一系列区分度更高的语义信息。3) 迭代优化。接下来，我们采用一种迭代方法（类似于迭代最近点 [4]），同时优化位置匹配并计算变换参数。4) 连续标定。最后，利用路侧激光雷达的位置相对固定这一特点，我们提出一种连续标定方法。该方法的核心思想是结合多次通行的标定结果，持续减小误差。

IV. 时空同步方法

如图4所示，TrajMatch 包含三个主要步骤：(i) 基于点云的轨迹生成，(ii) 轨迹运动特征提取，(iii) 语义感知位置匹配，(iv) 基于迭代优化的标定，以及一种通过持续标定进一步提高成功率的方法。下面我们以两台激光雷达的标定为例，逐一讨论这些步骤。

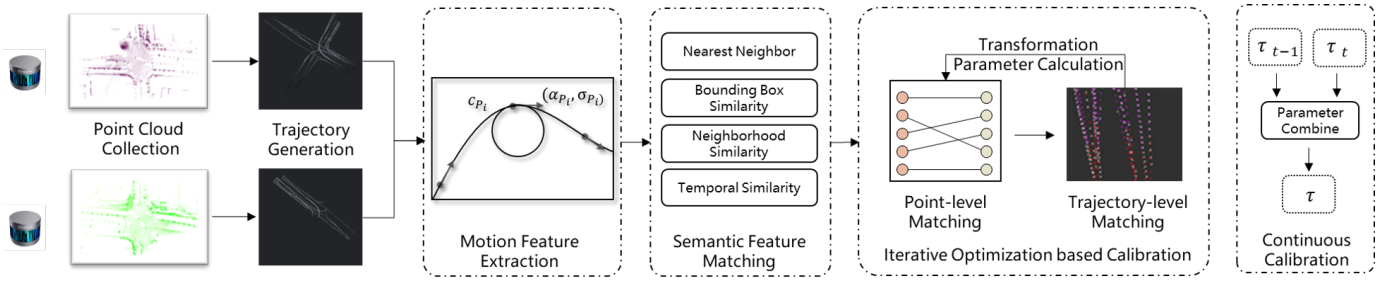
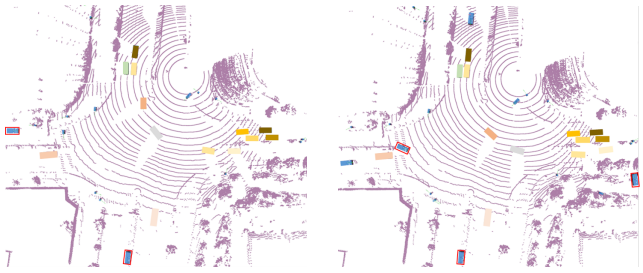


图 4. **我们提出的TrajMatch 系统**框架。本文利用一个系统对路侧激光雷达进行时空标定。该系统以激光雷达的检测和跟踪结果为输入，包含以下主要模块：(1) 运动特征提取（**局部**）、(2) 语义特征匹配（**中间**）、(3) 迭代优化（**全局**）和 (4) 连续标定。最后，我们提出一种策略来调整连续标定过程中的误差。



(a) (b)

图 5. **来自两个相邻点云帧的检测和跟踪结果示例**。检测算法负责输出目标的边界框，跟踪算法负责关联不同帧中的目标。我们用相同颜色表示同一目标，并用红框标注误检目标。

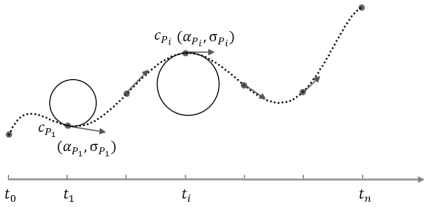


图 6. **轨迹运动特征示意图**。

A. 基于点云的轨迹生成

不同于以原始点云为输入的标定算法，我们使用一段时间内的点云检测/跟踪结果 T。我们注意到，选择 T 时应考虑两台激光雷达之间的时间偏移。

激光雷达生成点云帧后，便会执行三维目标检测。近年来，研究者提出了许多基于点云的检测算法，例如 [13]中的基于点的方法和 [14]中的基于体素的算法。这些检测算法会识别每个目标，将其归类为汽车、卡车、行人、自行车等交通参与者类别之一，并估计其三维边界框。图 5 展示了一个点云帧示例，以及使用 PointPillars 算法 [15]从该帧中检测出的目标。

在对每一帧执行目标检测后，我们接着执行多目标跟踪（MOT）任务，将各帧中的物体位置关联起来，从而获得运动轨迹。在这项工作中，我们采用 AB3DMOT 算法 [16][17]计算轨迹。图 5 展示了相邻两帧点云帧之间的一个关联结果示例，其中我们用相同的颜色表示这两帧中的同一辆汽车。

在这里，检测和跟踪任务构成了 TrajMatch 系统的预处理步骤。

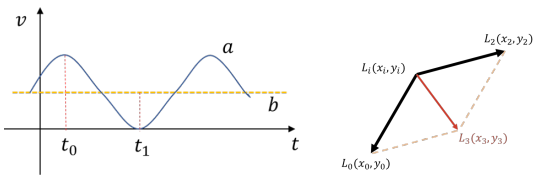
B. 轨迹运动特征提取

首先，我们遵循 [18]中的讨论，给出轨迹的数学定义。轨迹（TR）是按时间顺序排列的一系列物体在一段时间内的三维位置坐标，用于表示物体的运动信息，记作 $TR=\{P_0,...,P_i,...,P_n\}(1\le i\le n)$。这里，$P_i$表示物体在时间 t_i的位置，每个位置都由一个空间坐标集和一个时间戳组成，例如 $P_i=(x_i,\ y_i,\ z_i,\ t_i)$。更具体地说，位置 P_i 对应于从在时间 t_i采集的点云帧 f_i 中检测到的物体中心坐标（$(x_i,\ y_i,\ z_i)$）。此外，我们将激光雷达检测到的一组轨迹称为轨迹数据库（TD），其中 $TD=\{TR_1,TR_2,TR_3,...,TR_m\}$，$m$ 表示检测到的物体数量。

接下来，我们为轨迹提取合适的特征，这些特征应能准确描述物体在轨迹上的运动模式。考虑到一条轨迹由一系列位置构成，我们选择为每个位置提取运动特征，以捕捉更细粒度的运动模式。选择特征时，我们需要考虑以下因素。首先，存在未知的坐标变换，因此特征需要具有旋转不变性和平移不变性。也就是说，无论坐标如何旋转和平移，算法提取的特征都应保持一致。

因此，我们使用每个位置处的速度和曲率信息（两者均为旋转不变），来描述该位置处的运动。假设一条轨迹由 n 个位置组成。那么，该轨迹的特征向量由 n 个速度特征和 n 个曲率特征组成（如图6所示）。

速度： 速度是区分移动模式的重要特征。首先，我们注意到，仅使用平均速度（如 [19]）不足以实现这一目的。如图7(a)所示，在 t_0 至 t_1这段时间内，尽管轨迹 a 和 b 的平均速度相同，但它们的移动模式却截然不同。因此，我们在此同时使用速度的均值和方差。



(a) (b)

图 7。（a）平均速度相同但运动模式截然不同的轨迹；（b）曲率计算。

对于一条轨迹 $TR = \{P_0, P_1, P2, ..., P_n\}$，我们可以获得该轨迹位置对应的速度序列 $\{v_0, v_1, ..., v_{n-1}\}$，其中 $v_i = dis(P_i, P_i + 1)/(t_{i+1} - t_i)$。然后，位置 P_i 的速度特征 $V(P_i) = \{\alpha_i, \sigma_i\}$ 可以计算为：

$$\alpha(P_i) = \bar{v} = \frac{1}{2m}\sum_{j=i-m}^{j=i+m-1} v_j,$$

$$\sigma^2(P_i) = \frac{1}{2m}\sum_{j=i-m}^{j=i+m-1} v_j - \bar{v}^2,$$

其中，m 是用于平衡检测/跟踪噪声的超参数。例如，如果从跟踪器获得的轨迹存在显著噪声，选择较小的 m 可能会导致特征估计误差较大。

曲率：曲率用于描述轨迹在某一位置处的角度。假设有三个连续位置：P_{i-1}、P_i、P_{i+1}。位置 P_i 处的曲率计算如下：

$$c(P_i) = cos\theta = \frac{\overrightarrow{P_iP_{i-1}} \cdot \overrightarrow{P_iP_{i+1}}}{\left|\overrightarrow{P_iP_{i-1}}\right| \left|\overrightarrow{P_iP_{i+1}}\right|}, \ \theta \in (0, \pi), \quad (4)$$

其中，θ 表示线段 $\overrightarrow{P_{i-1}P_i}$ 和 $\overrightarrow{P_iP_{i+1}}$之间的角度。需要指出的是，轨迹在位置 P_i处越平滑，θ 就越接近 π，且 $c(P_i)$ 越小。

C. 语义感知位置匹配

在此步骤中，我们建立来自不同轨迹的两个位置之间的对应关系。首先，根据位置自身的信息，获得一系列位置的匹配关系；随后，我们将考虑其邻居的分布，以剔除错误匹配。

基于运动的位置匹配：我们使用欧氏距离来衡量两个位置的运动特征之间的距离。特征 $A = \{c_A, \alpha_A, \sigma_A\}$与 $B\{c_B, \alpha_B, \sigma_B\}$ 之间的距离计算如下：

$$d(A,B) = \lambda_c |c_A - c_B| + \lambda_\alpha |\alpha_A - \alpha_B| + \lambda_\sigma |\sigma_A - \sigma_B|, \quad (5)$$

其中，λ_c、λ_α、λ_σ 的值可以调整，以为各个特征属性赋予不同权重。例如，在速度的均值和方差之间，方差 σ 对噪声更为敏感。当系统中的检测/跟踪噪声高于某个阈值时，我们可以减小 λ_σ，以降低其权重。

基于上述距离计算，接下来我们执行基于运动的位置匹配。假设交叉路口有两台 LiDAR，一台的轨迹数据库为 P，另一台为 Q。对于每个位置 $p_i \in P$，我们在 Q 中找到距离最近的位置 q_j。如果 $d(p_i, q_j) < d_{th}$，则将 (p_i, q_j) 对加入匹配集合。

通过位置匹配算法，我们获得了大量位置对，其中一个位置可能有多个匹配。然而，我们发现其中许多匹配实际上是“错误”匹配。错误匹配可能由以下原因造成。首先，受交通规则限制，车辆的行驶速度和方向大多相似。其次，受车辆运动学模型限制，车辆的速度和曲率通常变化缓慢，[20],尤其是与激光雷达采样率相比时。第三，点云检测/跟踪算法存在误差。所有这些原因共同导致了较高的错误匹配率。

例如，在一个典型场景中，2台激光雷达部署在一个四岔路口（见图2），我们获得了875对匹配位置，其中最多只有175对实际上有效。

语义感知位置匹配：我们认为，如果考虑在早期目标检测和跟踪任务中生成的目标语义信息，就能过滤掉大量错误匹配。在这项工作中，我们提出探索添加语义信息，以增强位置匹配：

- 互为最近邻。这是一个简单的观察。如果位置 $P_a \in TD_a$ 和 $P_b \in TD_b$ 是真实匹配，那么在 TD_a中，P_a 是距离 P_b 最近的位置；在 TD_b中，P_b 是距离 P_a 最近的位置。

- 边界框相似度。如果位置 $P_a \in TD_a$ 和$P_b \in TD_b$ 是真实匹配，那么它们对应的目标边界框（由检测器生成）应该彼此相似，包括边界框的长、宽和高。

- 邻居数量相似度。如果位置 $P_a \in TD_a$和 $P_b \in TD_b$ 是真实匹配，那么一定区域内邻近目标的数量相同。

- 邻域分布一致性。如果位置$P_a \in TD_a$ 和 $P_b \in TD_b$ 是真实匹配，那么周围位置的分布相似，并且在相邻帧中保持不变。基于这一点，我们可以构建时域直方图来捕捉这种分布。

考虑这些语义信息后，我们可以有效地过滤掉错误匹配。在上文讨论的同一示例中，过滤后真实匹配的比例从 22% 增加到 61%。

D. 基于迭代优化的标定

接下来，我们考虑数据的整体分布，并采用一种迭代方法（类似于 ICP [21]）来优化位置匹配，同时实现精确标定。该优化方法包含三个主要操作：位置匹配、变换参数计算和轨迹匹配。具体而言，我们的优化过程如下：

1) $S1$：变换参数计算。如第 II 节所述，将一组匹配位置对作为几何约束后，我们可以通过求解公式 3 来计算变换参数。因此，获得正确变换参数的关键在于找到一组精确匹配的位置对。

2) $S2$ ：轨迹匹配。 根据变换参数，我们可以将每个位置的坐标转换到公共坐标系中。在公共坐标系下表示各个位置后，我们便可以计算两条轨迹之间的距离。具体而言，我们可以利用匹配位置对以及跟踪器提供的轨迹信息，获得匹配轨迹对。同时，我们还可以计算匹配轨迹之间的距离。如果该距离低于预设阈值，则说明标定结果令人满意，可以终止优化过程。

3) $S3$ ：位置匹配。借助匹配轨迹对以及每条轨迹的位置序列，我们可以通过补充漏匹配项并移除错误匹配项，进一步优化匹配位置对。然后返回 $S1$，持续进行优化，直到匹配轨迹之间的距离小于阈值，或迭代次数超过最大次数。

E. 连续标定

讨论完标定过程后，接下来我们将介绍系统如何评估标定效果，以及如何融合多次标定的效果。由此，我们可以实现路侧激光雷达的连续标定。

假设我们的标定结果给出了变换参数 $\tau = (R,\ T)$。假设 n_{pp} 是匹配位置对的数量， n_{po} 是两个激光雷达重叠区域内位置的数量。那么，我们使用比率 $s = 2 * n_{pp}/n_{po}$ 粗略评估标定操作的得分。完美标定的得分为 $s = 1$。在实际系统中，我们以一定频率执行激光雷达标定。例如，可以在检测到一个目标时立即执行标定，或者每检测到 k 个目标执行一次。在此过程中，我们保留当前的变换参数 τ_t 及其得分 s_t。此外，假设最新一次标定过程得到 τ_{t-1} 和 s_{t-1}。

然后，我们调整变换参数以实现连续标定：

$$\tau = \frac{s_{t-1}}{s_{t-1} + s_t} \times \tau_{t-1} + \frac{s_t}{s_{t-1} + s_t} \times \tau_t. \tag{6}$$

V. 性能评估

为了全面评估 TrajMatch，我们收集了一个仿真数据集 LiDARnet-sim1.0 和一个真实世界数据集。数据集的详细信息见第 V-A 节。我们在第 V-B 节定义了评估指标，并在第 V-C 节展示评估结果。

A. 评估数据集

为了验证所提出的方法，我们构建了两个数据集，一个真实世界数据集和一个仿真数据集。

1) LiDARnet-sim 1.0：由于目前没有公开数据集支持路侧激光雷达标定研究，我们利用 CARLA[22] 和 SUMO[23] 联合仿真平台构建了 LiDARnet-sim 1.0。使用仿真数据集有几项优势。首先，这类数据集可以提供真实标注信息，包括激光雷达的平移，以及目标检测和跟踪信息。其次，生成各种交通场景要容易得多，因此可以进行更全面的评估。

具体而言，我们基于 SUMO 模拟交通流，然后基于 CARLA 记录传感器数据。构建该数据集时，我们考虑了以下因素。第一个因素是交通流量：通过向地图中添加不同数量的汽车，形成不同的交通流量。第二个因素是激光雷达类型，包括感知范围为 50 米的 32线激光雷达，以及感知范围为 150 米的 80线激光雷达。

2) 真实世界数据集：对于真实世界数据集，我们在两个场景中部署两台激光雷达并采集数据：(1) 使用 RS-Ruby Lite（ 80 束）部署在交叉路口的斜对角位置；(2) 使用 RS-LiDAR- 32（32 束）部署在人行道两侧。我们使用 PointPillars[15]作为检测器，使用 AB3DMOT[16] 作为跟踪器。

B. 指标和基准算法

1) 指标*::* 为了将性能与其他标定算法进行比较，我们采用了 [24], 中定义、且 [25], [9] 中也使用的指标。首先，相对旋转误差（**相对旋转误差**）为：

$$\begin{aligned} angle &= F\left(R_T^{-1} \cdot R_E\right), \\ RRE &= \sum_{i=1}^{3} |angle(i)|, \end{aligned} \tag{7}$$

其中， R_T 和 R_E 分别是真实变换和估计变换的旋转矩阵。$F(\cdot)$将旋转矩阵转换为三个欧拉角。其次，相对平移误差（**RTE**）定义为：

$$RTE = ||t_T - t_E||, \tag{8}$$

其中， t_T 和 t_E 分别是真实变换和估计变换的平移向量。

此外，当 RRE 和 RTE 均低于预定义阈值 [26] 时，我们将一次标定会话定义为“成功”。本文选择阈值为 $1m$, 这足以满足自动驾驶导航 [27]的要求。**成功率**定义为会话成功的平均比率。

为测试算法的时间同步能力，我们采用时间偏移误差（ **TOE**）。TOE 定义为：

$$TOE = |T_T - T_E|, \tag{9}$$

其中， T_T 和 T_E 分别为真实时间偏移量和估计时间偏移量。

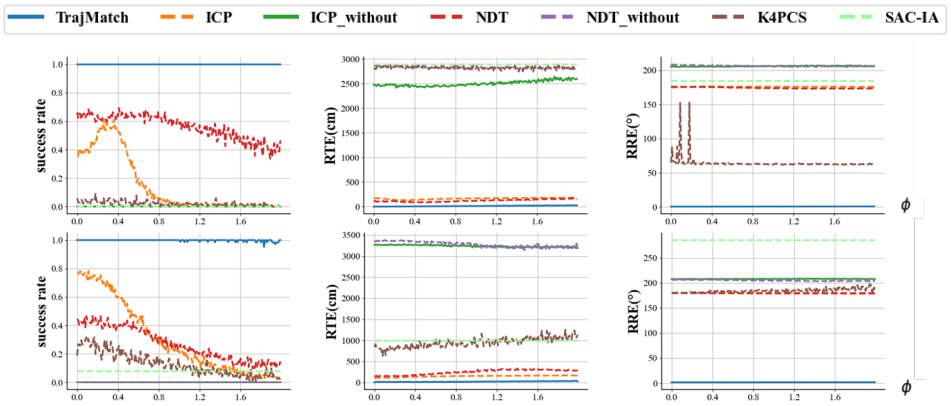


图 8. 此处展示了该算法在两种不同场景下的效果。这里使用平均值，并使用我们的评估工具对算法结果进行筛选。第一行中，我们将激光雷达放置在交叉路口的对角线上；第二行中，我们将激光雷达放置在三岔路口。

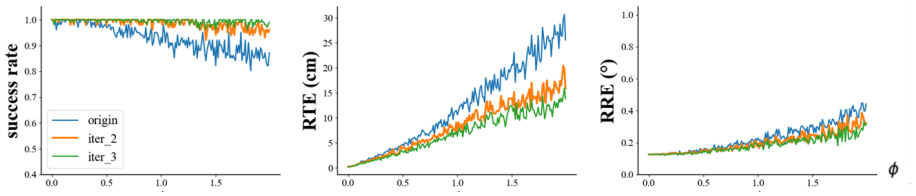


图 9. 经过 1、2 和 3 次遍历后的空间标定结果。包括 RTE 和 RRE。

2）基准算法：我们将 TrajMatch 与四种激光雷达标定基准算法进行比较：(i) ICP [21]，通过迭代优化获得标定结果；(ii) NDT [28]，根据点云概率密度函数（PDF）之间的相似性确定标定结果；(iii) K-4PCS[29]，通过提取关键点提高优化过程的效率；(iv) SAC-IA[30]，根据点周围的局部几何特征进行标定，这些特征称为快速点特征直方图（FPFH）。这些算法均基于点云库（PCL）[31]实现。我们仔细调整了各算法的参数值，以确保其在我们的数据集上达到最佳性能。在这些算法中，只有 SAC-IA 基于原始点云，其余算法均基于检测/跟踪结果。

C. 标定结果

1）仿真数据集上的结果：我们首先展示仿真数据集上的结果。

空间标定与部署设置： 在这组实验中，我们从空间标定结果方面比较 TrajMatch 与基准算法。为了模拟不同的检测误差，我们在真实轨迹上添加符合高斯分布的噪声，噪声范围为 (0，ϕ)。我们将 ϕ 从 0 变化到 2m。根据 nuScenes[32]，排行榜，该值远大于基于激光雷达检测产生的测距误差，后者约为 0.2m。选择较大的值，可以确保我们的评估涵盖最坏情况。此外，由于迭代最近点和正态分布变换的性能很大程度上取决于初始值的选择，我们在两种情况下对其进行测试：一种使用人工选取的初始值，另一种不使用初始值。

表 I 我们的参数评估方法的有效性。

	RTE（厘米）	相对旋转误差(°)	分数
1	4.117	1.550	0.848
2	5.373	1.573	0.841
3	14.759	1.568	0.824
4	17.087	1.520	0.823
5	3569.470	183.878	0.00

图8第一行展示了四岔路口处的成功率、相对平移误差和相对旋转误差。激光雷达沿对角线安装在交叉路口，间距为 28.8m，如图3(a)所示。激光雷达线束数量为 32，检测距离为 50m。在这三项指标上，我们的结果始终最佳。即使当 ϕ 大于 1.2m 时，TrajMatch 仍能实现超过 95% 的成功率。我们还测试了 80 线激光雷达，结果与 32 线激光雷达相似，因此没有在此处放置结果图。第二行展示了三岔路口处的结果，我们得出了类似的结论。

此外，我们展示了评估机制的有效性。借助该机制，我们可以轻松识别并剔除这些离群值。表 I 汇总了不同案例的分数，我们可以轻松区分离群值（分数为 0）和其他正常案例（分数高于 0.8）。

空间标定与交通流：在第二组实验中，我们测试了交通流的影响。具体而言，我们通过在城市中随机放置 50、100 和 200 辆车辆来模拟不同的交通流量。该参数会影响经过我们交叉路口的汽车数量，并检验我们同时处理多辆汽车的能力。在此，我们设置 $\phi = 0.2m$ 和时间偏移 $\Delta_t = 0.5s$。结果汇总于表 II。我们观察到，交通流量

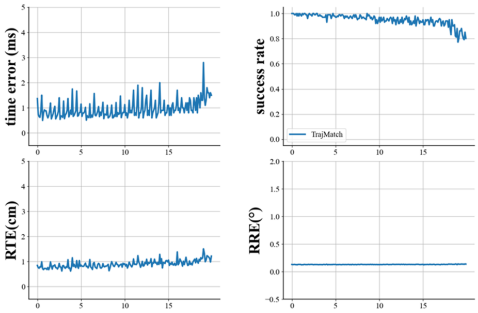


图 10. 四岔路口在不同时间偏移值下的时空标定结果。横轴为时间偏差，范围为 0–20 秒。

对性能没有明显影响。在这三种情况下，相对平移误差均小于 $5cm$ ，相对旋转误差均小于 3 度。这表明 TrajMatch 能够有效应对不同的交通流量。

空间标定与旋转角度：在第三组实验中，我们评估了系统对激光雷达旋转角度的鲁棒性。在固定其余参数的情况下改变旋转角度：时间偏移为 $0.5s$，高斯噪声的标准差为 $\phi = 0.2m$。标定结果汇总于表 III，充分表明，由于我们选用了旋转不变特征，系统对不同的激光雷达旋转角度具有鲁棒性。

连续标定：在第四组实验中，我们评估了连续标定能力的有效性。也就是说，进行多轮标定后，能否提高标定精度? 我们在图 9 中展示了四岔路口场景下的结果。结果表明，连续标定能够稳步提升性能。具体而言，进行两轮标定时，相较于只进行一轮，精度有显著提升。然而，进行三轮或更多轮标定时，提升幅度就不那么明显了。事实上，经过三轮标定后，RTE 从 30cm 降至 10cm，RRE 从 0.4° 降至 0.2°。这使 TrajMatch 成为一种非常可行的方法，能够进一步优化其性能

表 II 不同交通流下的空间标定结果。此处报告相对旋转误差和相对平移误差的中位数值。

场景	数量汽车	相对旋转误差(°)	相对平移误差(cm)	成功率（%）
四岔路口交叉路口	50	2.55	3.58	99
	100	0.13	2.67	99
	200	2.53	2.08	98

不同旋转角度下的空间标定结果。此处报告 RRE 和 RTE 的中位数。表 III

场景	旋转角度(°)	相对旋转误差(°)	相对平移误差(cm)	成功率（%）
四岔路口交叉路口	0	0.04	2.65	97
	30	0.14	1.92	99
	60	0.05	3.08	98.1
	90	0.04	2.61	99
	120	0.05	3.02	99

随着标定会话次数的增加。

时间标定：最后，我们报告时间标定结果。在这组实验中，我们有 $\phi = 0.2m$。我们将两台激光雷达的时间偏差从 0 调整到 20s。图 10 展示了四岔路口场景下的结果。如图 10(d) 所示，当时间偏移小于 18s 时，我们能够实现帧间同步，且时间同步误差小于 2ms。我们认为，这主要是因为车辆的相对位置能够形成空间约束，在存在多条轨迹时辅助时间标定。我们再次认为，这种时间标定精度对于智能交通应用而言相当有竞争力。

2）真实世界数据集的结果：除了仿真环境外，我们还在真实世界中采集了数据集进行测试。我们主要测试了两种场景：（1）在大型交叉路口对角部署两台激光雷达，如图3(a)所示；（2）在人行道两侧各部署一台激光雷达，如图3(c)所示。

图11展示了 TrajMatch 与其他 4 种算法生成的融合点云。在该数据集上，所有 4 种基线算法都未能标定激光雷达并对齐点云。TrajMatch 则很好地对齐了两组点云。在这种情况下，不同激光雷达采集的同一物体点云能够很好地匹配。此外，我们还在表IV中展示了定量标定结果。我们发现，该算法的平移误差小于 $15cm$。

VI. 相关工作

由于系统部署的特性，路侧激光雷达的时空标定面临若干重要的技术挑战。下面我们将详细说明这些挑战。

A. 空间标定：

我们面临的第一个挑战是两组点云之间的重叠率相对较低。具体而言，重叠率是交集大小与并集大小之比。来看一个真实世界部署案例（如图2所示），其参数如下：（1）四岔路口的宽度为 $60m$；（2）安装在车辆上的激光雷达传感器的探测范围为 80–100 米（根据 nuScenes [32]，并考虑安全因素）。在这种情况下，重叠率约为 26%。通常，部署在路侧的激光雷达彼此相距数十米，以最大限度地扩大整体覆盖范围，这在实际应用中可能导致更低的重叠率。

然而，现有的激光雷达标定方法通常假设较大的重叠率才能获得准确结果。例如，正如 [33], 中所讨论的，最先进的标定方法要求两个点云在真实变换下至少有 30% 的重叠率。一些算法要求更高的重叠率；例如，迭代最近点更倾向于重叠率达到 80% 或以上 [34]。

此外，交通监控系统通常会在不同高度安装激光雷达（见图2）。由于激光雷达的安装位置，生成的

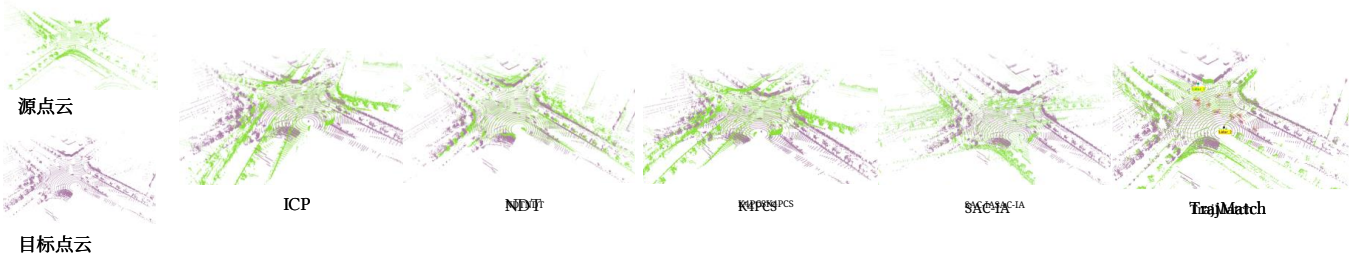


图 11. **不同标定算法在场景1中的可视化结果。TrajMatch 可实现最佳点云配准。**

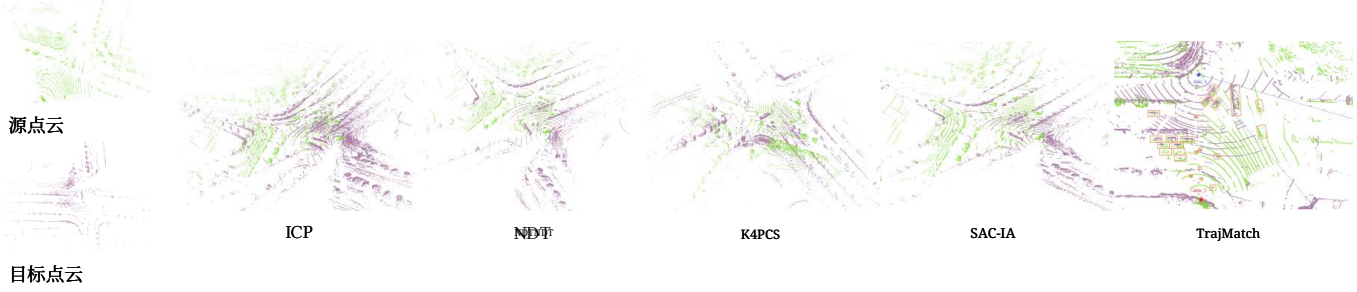


图 12. **不同标定算法在场景2中的可视化结果。** TrajMatch **可实现最佳点云配准。**表IV 场景1中真实轨迹的标定结果。

	角度误差(°)			变换误差					
	R_x	R_y	R_z	T_x		T_y		T_z	
				数量（厘米）	比率（%）	数量（厘米）	比率（%）	数量（厘米）	比率（%）
1	-0.499	1.787	-0.148	11.7	0.2	13.5	3.9	0.0	0.0
2	-0.5	1.785	-0.187	9.9	0.2	7.1	2.0	0.0	0.0
3	-0.499	1.787	-0.238	1.2	0.0	19.1	5.5	0.0	0.0
4	-0.5	1.78	0.195	19.3	0.3	9.5	2.7	0.0	0.0

两个点云之间的三维重叠几乎缩减为沿路面的二维区域，高度非常低 [35]。图2展示了两个点云之间的重叠。许多现有的标定方法依赖环境中具有区分性的手工设计特征，例如曲线[36], 交线[37],自适应协方差[38], , 而在我们的场景中很难找到这些特征。

B. 时间标定：

对于时间标定，我们需要确定传感器之间的恒定时间偏移，并对齐它们的时间戳 [39]。时间标定方法主要分为两类：基于自运动的方法和基于轨迹的方法。

第一类方法（例如 [40],[41], 中讨论的方法）尝试根据传感器的运动轨迹同步时间戳。具体而言，这类算法通常包括两个步骤：估计每个传感器的自运动，然后通过对齐轨迹（例如速度[42],方向变化[43]）来计算时间偏移。显然，这种方法不适用于我们的系统，因为我们的激光雷达安装在固定位置。

第二类方法（例如 [44], 中讨论的方法）在让传感器跟踪同一目标的同时，进行时间和空间标定。这类方法适用于所有传感器都已知该目标的场景。然而，在我们的系统设置中，激光雷达通常部署在繁忙的道路上，传感器能够检测到许多目标，但这些目标之间的对应关系未知。此外，还有一些方法利用额外线索进行标定，例如 LiDARTag [45]。这类方法通常需要人工参与，因此不适用于我们。

VII. 结论

本文提出了 TrajMatch，这是一种可自动完成路侧激光雷达时空标定的系统。我们的方法不对系统的初始状态作任何假设。作为一种无状态系统，我们能够同时对激光雷达进行空间对齐，达到厘米级精度，并进行时间对齐，达到毫秒级精度。该标定通过高效的迭代优化过程实现。此外，我们还提出了一种连续标定机制，可准确评估每次标定会话的质量，并在评估结果高于某个阈值时更新整体标定。

我们认为，利用路侧传感器进行交通监控将成为未来城市的重要组成部分，而我们的工作迈出了构建稳健、准确的路侧感知系统的第一步。今后，我们将继续努力实现这一愿景。具体而言，我们将研究适用于不同部署设置和不同传感器的标定流程。我们还将充分利用这些传感器融合后的数据，以改善交通通行状况并提高驾驶安全。

参考文献

[1] E. Arnold、M. Dianati、R. de Temple 和 S. Fallah，“利用基础设施传感器在驾驶场景中进行三维目标检测的协同感知”，《IEEE智能交通系统汇刊》，2020年。[2] J. Kim、C. Kim、Y. Han 和 H. J. Kim，“利用任意平面标定板进行距离偏移校正的三维激光雷达自动外参标定”，载于《2021年IEEE国际机器人与自动化会议（ICRA）》。IEEE，2021年，第5082–5088页。[3] Y. Xie、R. Shao、P. Guli、B. Li 和 L. Wang，“利用AprilTag标记对多相机多激光雷达系统进行基于基础设施的标定”，载于《2018年IEEE智能车辆研讨会（IV）》，第605–610页。[4] J. Quenzel、N. Papenberg 和 S. Behnke，“多个固定式激光测距仪的稳健外参标定”，载于《2016年IEEE国际自动化科学与工程会议（CASE）》。IEEE，2016年，第1332–1339页。[5] D. F. Glas、D. Brˇsˇciˇc、T. Miyashita 和 N. Hagita，“Snapcat-3d：用于行人跟踪的三维测距传感器网络标定”，载于《2015年IEEE国际机器人与自动化会议（ICRA）》。IEEE，2015年，第712–719页。[6] K. Schenk、A. Kolarow、M. Eisenbach、K. Debes 和 H.-M. Gross，“通过匹配运动轨迹自动标定固定式激光测距仪网络”，载于《2012年IEEE/RSJ智能机器人与系统国际会议》，2012年，第431–437页。[7] Z. Pusztai 和 L. Hajder，“利用普通箱子实现激光雷达-相机系统的精确标定”，载于《IEEE计算机视觉国际会议研讨会论文集》，2017年，第394–402页。[8] B. Xue、J. Jiao、Y. Zhu、L. Zhen、D. Han、M. Liu 和 R. Fan，“利用贴有逆反射胶带的两根杆自动标定双激光雷达”，载于《2019年IEEE成像系统与技术国际会议（IST）》。IEEE，2019年，第1–6页。[9] Y. He、L. Ma、Z. Jiang、Y. Tang 和 G. Xing，“Vi-eye：面向基础设施辅助自动驾驶的基于语义的三维点云配准”，载于《第27届移动计算与网络年度国际会议论文集》，2021年，第573–586页。[10] Y. Du、B. Qin、C. Zhao、Y. Zhu、J. Cao 和 Y. Ji，“一种用于传感器融合的路侧异步毫米波雷达与相机时空同步新方法”，《IEEE智能交通系统汇刊》，第1–12页，2021年。[11] M. Wu 和 B. Coifman，“量化仪器化车辆和自动驾驶车辆感知传感器数据中不可见的内容——案例研究”，《交通研究C部分：新兴技术》，第107卷，第105–119页，2019年。[12] “Rs-lidar lite”，https://www.robosense.ai/rslidar/RS-Ruby Lite， [无日期。]。[13] C. R. Qi 和 Su，“PointNet：面向点集三维分类与分割的深度学习”，载于《计算机视觉与模式识别会议》，2017年，第652–660页。[14] J. Deng，“Voxel R-CNN：迈向高性能基于体素的三维目标检测”，载于《AAAI人工智能会议论文集》，第35卷，第2期，2021年，第1201–1209页。[15] A. H. Lang、S. Vora、H. Cāesar、L. Zhou、J. Yang 和 O. Beijbom，“PointPillars：用于点云目标检测的快速编码器”，载于《计算机视觉与模式识别会议》，2019年，第12 697–12 705页。[16] X. Weng、J. Wang、D. Held 和 K. Kitani，“AB3DMOT：一种三维多目标跟踪基线及新的评估指标”，2020年8月。[17] Y. Li、J. Deng、Y. Zhang、J. Ji、H. Li 和 Y. Zhang，“ezfusion：深入探究激光雷达、毫米波雷达与相机的融合，以实现准确的三维目标检测与跟踪”，《IEEE机器人与自动化快报》，第7卷，第4期，第11 182–11 189页，2022年。[18] G. Yuan、P. Sun、J. Zhao、D. Li 和 C. Wang，“移动目标轨迹聚类算法综述”，《人工智能评论》，第47卷，第1期，第123–144页，2017年1月。[19] R. S. D. Sousa、A. Boukerche 和 A. A. F. Loureiro，“车辆轨迹相似性：模型、方法与应用”，《ACM计算调查》，第53卷，第5期，2020年。[20] J. Kong、M. Pfeiffer、G. Schildbach 和 F. Borrelli，“面向自动驾驶控制设计的车辆运动学与动力学模型”，载于《2015年IEEE智能车辆研讨会（IV）》。IEEE，2015年，第1094–1099页。[21] P. Besl 和 N. D. McKay，“一种三维形状配准方法”，《IEEE模式分析与机器智能汇刊》，第14卷，第2期，第239–256页，1992年。[22] A. Dosovitskiy、G. Ros、F. Codevilla、A. Lopez 和 V. Koltun，“CARLA：一个开放式城市驾驶模拟器”，载于《机器人学习会议》。PMLR，2017年，第1–16页。[23] D. Krajzewicz、J. Erdmann、M. Behrisch 和 L. Bieker，“近期发展

evelopment and applications of sumo-simulation of urban mo能力y,”

,国际 *journal on advances in s* 艺术 J. Lenz 和 R. Urtasun，第5卷，第3&4期，2012年。[24] A. Geiger、P. Lenz 和 R. Urtasun，《我们准备好迎接自动驾驶了吗？KITTI 视觉基准套件》，载于2012年IEEE计算机视觉与模式识别会议。IEEE，2012年，第3354–3361页。[25] C. Choy、W. Dong 和 V. Koltun，《深度全局配准》，载于IEEE/CVF计算机视觉与模式识别会议论文集，2020年，第2514–2523页。[26] G. Elbaz、T. Avraham 和 A. Fischer，《使用深度神经网络自编码器进行定位的三维点云配准》，载于 IEEE计算机视觉与模式识别会议论文集，2017年，第4631–4640页。[27] J. Liu、H. Wu、C. Guo、H. Zhang、W. Zuo 和 C. Yang，《高精度道路导航地图的发展与思考》，《中国工程院战略研究》，第20卷第2期，第99–105页，2018年。[28] P. Biber 和 W. Straßer，《正态分布变换：一种激光扫描匹配新方法》，载于2003年 IEEE/RSJ智能机器人与系统国际会议论文集（IROS 2003）（目录号03CH37453），第3卷。IEEE，2003年，第2743–2748页。[29] P. W. Theiler、J. D. Wegner 和 K. Schindler，《基于关键点的四点全等集——激光扫描的自动无标记配准》，《ISPRS摄影测量与遥感期刊》，第96卷，163，2014年，第149页。–。[30] R. B. Rusu、N. Blodow 和 M. Beetz，《用于三维配准的快速点特征直方图（FPFH）》，载于2009年 IEEE机器人与自动化国际会议。IEEE，2009年，第3212–3217页。[31] R. B. Rusu 和 S. Cousins，《三维时代已至：点云库（PCL）》，载于2011年 IEEE机器人与自动化国际会议。IEEE，2011年，第1–4页。[32] 《nuScenes上的检测排行榜》，https://www.nuscenes.org/object-detection?externalData=all&mapData=all&modalities=Any，[无日期]。[33] Z. Gojcic、C. Zhou、J. D. Wegner 和 A. Wieser，《完美匹配：基于平滑密度的三维点云匹配》，载于 IEEE/CVF计算机视觉与模式识别会议论文集，2019年，第5545–5554页。[34] P. Kim、J. Chen 和 Y. K. Cho，《在建筑环境中利用视觉和平面特征自动进行点云配准》，《土木工程计算机应用》，第32卷第2期，第04017076页，2018年。[35] R. Yue、H. Xu、J. Wu、R. Sun 和 C. Yuan，《利用地面点对路侧激光雷达传感器进行数据配准》，《遥感》，第11卷第11期，2019年。 [在线]。网址：https://www.mdpi.com/2072-4292/11/11/1354[36] B. Yang 和 Y. Zang，《利用曲线自动配准密集地面激光扫描点云》，《ISPRS摄影测量与遥感期刊》，第95卷，第109–121页，2014年。[37] I. Stamos 和 M. Leordeanu，《大规模城市场景的自动特征配准》，载于2003年计算机视觉与模式识别会议，第2卷。IEEE，2003年，第II–Ii页。[38] D. Zai 和 J. Li，《利用协方差描述子和非合作博弈对地面激光扫描点云进行成对配准》，《ISPRS摄影测量与遥感期刊》，第134卷，第15–29页，2017年。[39] J. Rehder、R. Siegwart 和 P. Furgale，《多传感器系统时空标定的一般方法》，《IEEE机器人学汇刊》，第32卷第2期，第383–398页，2016年。[40] Z. Taylor 和 J. Nieto，《多模态传感器外参的基于运动的标定及时间偏移估计》，《IEEE机器人学汇刊》，第32卷第5期，第1215–1229页，2016年。[41] M. Li 和 A. I. Mourikis，《相机–IMU系统的在线时间标定：理论与算法》，《国际机器人研究期刊》，第33卷第7期，第947–964页，2014年。[42] T. Qin 和 S. Shen，《单目视觉惯性系统的在线时间标定》，载于2018年 IEEE/RSJ智能机器人与系统国际会议（IROS）。IEEE，2018年，第3662–3669页。[43] J. Kelly 和 G. S. Sukhatme，《多种本体感知与外感知传感器时间标定的通用框架》，载于《实验机器人学》。Springer，2014年，第195–209页。[44] J. Perˇsi´c、L. Petrovi´c、I. Markovi´c 和 I. Petrovi´c，《基于高斯过程运动目标跟踪的时空多传感器标定》，《IEEE机器人学汇刊》，第37卷第5期，第1401–1415页，2021年。[45] J.-K. Huang 和 Wang，《LiDARTag：面向点云的实时基准标记系统》，第6卷第3期，第4875–4882页，2021年。