

# 基于三维变径螺旋线的路侧激光雷达和相机联合标定的方法

高 博, 吴 霞\*, 王 平

同济大学电子与信息工程学院, 上海

收稿日期: 2024年6月12日; 录用日期: 2024年7月16日; 发布日期: 2024年7月24日

## 摘要

车路协同技术是车联网发展新阶段迫切需求的一个方向, 进行路侧激光雷达和相机设备的联合标定成为实现智慧道路的必要工作。鉴于路侧与单车智能传感器部署场景的不同, 本文设计了一种适用于路侧联合标定的特殊标记物, 主体结构为三维变径螺旋线。首先, 在激光雷达和相机的共视区分别采集标记物的点云数据和图像数据; 对于点云, 采用改进的随机采样一致性算法提取螺旋线结构并进行点云补全; 对于图像, 采用语义分割技术获取标记物的像素信息, 进行距离变换形成激励掩码。我们将标定计算流程分为两个阶段。在粗标定阶段, 构建pnp问题的模型获得联合标定参数旋转矩阵R和平移向量t的初始值; 在细标定阶段, 构建目标函数, 最大化标记物点云在图像掩码上的投影覆盖率来优化标定参数。实验结果验证了所提方法相比传统方法在标定效率、鲁棒性和精度上具有显著优势, 为车路协同系统中传感器的空间校准提供了有效的解决思路。

## 关键词

车路协同, 联合标定, 激光雷达, 相机, 路侧感知

# Roadside LiDAR-Camera Calibration Based on 3D Variable Diameter Spiral

Bo Gao, Xia Wu\*, Ping Wang

School of Electronics and Information Engineering, Tongji University, Shanghai

Received: Jun. 12<sup>th</sup>, 2024; accepted: Jul. 16<sup>th</sup>, 2024; published: Jul. 24<sup>th</sup>, 2024

## Abstract

**Cooperative Vehicle Infrastructure System (CVIS) is a crucial direction in the development of Intelligent Vehicle-Highway Systems (IVHS).** The joint calibration of roadside LiDAR and cameras is a key component of CVIS. This paper proposes a novel calibration method based on a 3D variable-diameter spiral marker. The marker is composed of a series of concentric spirals with varying diameters. The calibration process involves two main steps: 1) Data collection: point cloud data from LiDAR and image data from cameras are collected in the common field of view. 2) Calibration: the spiral structure is extracted from the point cloud using an improved Random Sample Consensus (RANSAC) algorithm, and the point cloud is then completed. The image data is processed using semantic segmentation to obtain pixel-level information, which is then used to generate an incentive mask via distance transformation. The calibration parameters, including the rotation matrix R and translation vector t, are estimated using the PnP problem model. The optimization process involves maximizing the projection coverage rate of the marker point cloud on the image mask to refine the calibration parameters. Experimental results demonstrate that the proposed method significantly outperforms traditional methods in terms of calibration efficiency, robustness, and accuracy, providing an effective solution for the spatial calibration of sensors in the vehicle-road协同 system.

ternet of Vehicles, making the spatial calibration of roadside LiDAR and camera essential for implementing intelligent roadside systems. Given the differences in deployment scenarios between roadside and onboard intelligent sensors, a unique marker is designed specifically for roadside spatial calibration, featuring a three-dimensional variable-diameter helical structure. Initially, point cloud and image of the marker are collected respectively in the common field of view of the LiDAR and camera. For the point cloud, an improved Random Sample Consensus (RANSAC) algorithm is used to extract the marker and complete the point cloud. For the images, semantic segmentation technology is employed to obtain pixel information of the marker and perform distance transformation to create an excitation mask. The calibration process is divided into two stages. In the coarse calibration stage, a model based on the perspective-n-point (PNP) problem is constructed to obtain initial values for the spatial calibration parameters, including the rotation matrix  $R$  and translation vector  $t$ . In the fine calibration stage, an objective function is constructed to optimize the calibration parameters by maximizing the projection coverage of the marker's point cloud on the image mask. Experimental results demonstrate that the proposed method significantly outperforms traditional methods in terms of calibration efficiency, robustness, and accuracy, providing effective technical support for spatial calibration of sensors in Cooperative Vehicle Infrastructure System.

## Keywords

Cooperative Vehicle Infrastructure System, Joint Calibration, LiDAR, Camera, Roadside Perception

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着自动驾驶技术的快速发展，国家不断出台相关政策和指导意见，以适应我国智能网联汽车发展的新趋势、新特征和新需求。《国家车联网产业标准体系建设指南》明确指出以智能网联汽车为核心载体和应用载体，“车-路-云”协同发展。智能的道路可以提供更加丰富、准确和完整的信息，包括交通基础设施、车辆和行人等环境信息，为自动驾驶系统提供关键的感知能力。

激光雷达探测距离精度高，范围广[1]。相机颜色信息和纹理特征丰富，在计算机视觉领域已有诸多成熟的算法，因此激光雷达和相机被广泛应用于智能交通系统中。然而，单独使用其中一种传感器都存在一定的局限性。激光雷达对目标的识别和分类能力相对较弱；相机在低光照、强反光等特殊条件下容易受到影响，且对距离的感知能力较差。目前诸多研究聚焦于点云和图像的异构数据融合[2]，充分发挥两者的优势，实现更全面准确的环境感知。特别地，融合阶段发生在原始数据进行特征提取前，联合标定的工作至关重要，直接影响感知的精度。

多模态传感器融合正成为自动驾驶实现目标感知的主流技术方案，联合标定是实现融合感知的先决工作。激光雷达和相机的联合标定是指对两种传感器在空间上进行校准，确定激光雷达点云和相机图像之间的几何变换关系，使其数据能够在统一的坐标系下融合。

目前联合标定的方法大致可以分为两类，一是基于标靶标定，二是无标靶标定[3]。基于标靶标定的方法通常需要的一个特殊的目标，利用传感器中标靶对应特征的几何约束对外在参数进行优化。传统方法采用棋盘格[4]，特征为黑白相交的角点[5]；采用实心圆或圆环标定板[6]，同时使用多个多边形标定板[7]，在棋盘格上增加了圆标靶[8]，将 ArUco 码与圆标靶相结合，进一步提升特征点的提取精度。但是这

些标靶都是一个二维平面, 采集数据时需要控制标靶位置及角度。无标靶标定通常利用共视区内自然场景边缘信息或图像灰度值与点云强度的联系[9], 可分为基于特征信息[10] [11]和基于互信息[12] [13]标定。无标靶不需要人为设计特殊目标, 可实现在线标定。基于标靶标定的方法由于位置和姿态是已知的, 精度高, 易于操作, 可控性强。由于无标靶标定使用的是环境中的特征, 其准确性受到环境的影响, 存在一定误差, 精度相对较低。

上述联合标定算法, 主要针对机器人或者单车智能平台的传感器部署场景, 目前的研究对于路侧场景仍较为欠缺。车载端和路侧端使用标靶在实施标定的过程中, 主要存在传感器视角朝向不同的区别。本质上, 激光雷达和相机都是光学传感器, 两者均是接受光信息进行数据处理, 最终成像。无论是转子式机械激光雷达或扫描式固态激光雷达[1], 在单车智能平台上往往激光雷达和摄像头的视场角朝向一致, 传感器接收光线的传播方向相同, 控制角度可以使共视区内的标定板在两种传感器都有清晰可辨的成像, 为联合标定可以准确求解奠定基础。而路侧智能平台的部署则会出现两个传感器朝向相反的情况。此时二维的标定板无法满足需求, 因为当其正面朝向其中一个传感器时, 极大概率背面朝向另一个传感器。另外, 由于路侧场景空旷, 无标靶方法在很多情况下难以提取到丰富的特征信息。

基于上述路侧场景的特点, 针对激光雷达和相机的联合标定, 本文设计了一种三维标记物, 力求解决现有标定板无法适用于路侧场景的问题。该标记物的主体结构为经参数方程严格定义的三维变径螺旋线, 在三维笛卡尔坐标系下螺旋线上各点的相对位置是已知的, 因此可以利用这一约束关系, 采用改进的 RANSAC 算法进行点云的特征提取和补全。重要的是, 螺旋线结构允许传感器在不同位置进行观测, 只要螺旋线的厚度足够薄, 在误差允许范围内, 激光雷达和相机总能在共视区内观测到螺旋线上的同一坐标点。

## 2. 基于三维变径螺旋线的联合标定



Figure 1. Architecture of the calibration algorithm

图 1. 联合标定算法流程

本文提出的算法流程如图 1 所示。基于标靶进行联合标定的主要思想是在共视区内激光雷达和相机识别提取标记物特征, 通过特征的对应关系建立约束和优化方程, 以此获得标定参数旋转矩阵  $\mathbf{R}$  和平移向量  $\mathbf{t}$ 。同时路侧的传感器部署场景不同于车载端, 要考虑传感器朝向相对的情况。本文重点关注的是外参数的求解, 因此假设相机的内参数已知。点  $\mathbf{P}$  在激光雷达坐标系表示为  $\mathbf{P}_L = [X_L, Y_L, Z_L]^T$ , 在像素平面上的投影坐标为  $\mathbf{p}_C = [u, v]^T$ , 那么满足:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K}_c [\mathbf{R} \quad \mathbf{t}] \begin{bmatrix} X_L \\ Y_L \\ Z_L \\ 1 \end{bmatrix} \quad (1)$$

式中  $\mathbf{K}_c$  为相机的内参数, 旋转矩阵  $\mathbf{R} \in SO(3)$ , 平移向量  $\mathbf{t} \in \mathbb{R}^3$ 。

## 2.1. 三维螺旋线结构

### 2.1.1. 标靶条件

根据前文所述, 路侧场景下, 标靶应满足以下三个特性:

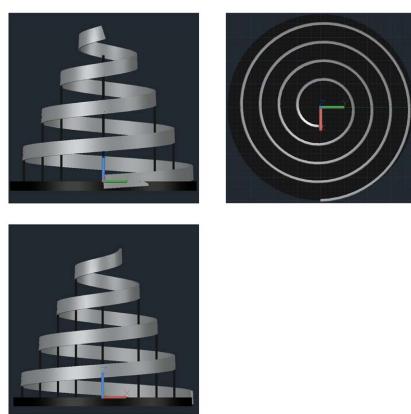
1) 视角稳定性: 对于不同视角, 激光雷达和相机都能观测到标靶。视角稳定性换言之, 即固定标靶位置, 改变观测者观察角度, 依然可以观测到有效的特征信息, 这一特性使得标靶在多样化的多传感器部署场景可以发挥关键作用。二维平面标定板总是存在观察盲区, 即使在背面存在一致的设计, 从侧面观察时也无法得到有效信息, 因此标靶理应是三维结构, 并且对旋转不敏感。

2) 特征统一性: 激光雷达和相机在同一时刻观测到的特征是相对应的, 即同时观测到同一特征。假设同时从不同角度观测一固定的实心圆球, 那么符合统一性的特征为该球的球心, 因为无论从何角度, 球心的绝对位置不变且唯一。然而球心的坐标并不是先验信息, 采用估计的方法会引入偏心误差[5]。至于利用直接观测到的球面, 若存在交集, 要确定交集各元素的对应关系则是一个难题, 因为点、线(边缘)比面的信息更有助于寻找准确的标定参数[14]; 而交集之外的数据, 或者不存在交集的情况, 则为无效观测。因此考虑标靶为线型结构, 最大化传感器原始观测数据的利用效率。

3) 特征可检测性: 对于激光雷达和相机观测到的特征, 要能够采用相关算法在点云数据和图像数据中进行提取。这一特性使得标定可以自动化进行, 简化标定流程提高标定效率, 同时为实现在线标定奠定基础。

上述三个特性, 视角稳定性是对路侧部署场景提出的新要求; 特征统一性是自动驾驶中联合标定的基本条件, 在路侧新场景下包括了更丰富的含义; 特征可检测性扩展了标靶的应用场景。

### 2.1.2. 结构设计



**Figure 2.** Multiview orthographic projections of the structure  
**图 2.** 标靶结构三视图

三维变径螺旋线的结构三视图如图 2 所示。支架底座为一圆盘, 为了突出螺旋线的特征以及减少支架对螺旋线的干扰, 支架部分颜色为黑色, 螺旋线部分为白色。对于激光雷达发射的激光, 白色物体的

反射率要远远高于黑色物体，因此在整个结构的点云内部可以利用反射强度过滤支架的点云，保留螺旋线；在图像中，黑色和白色的对比度差异大，有利于分割出螺旋线的像素信息。

以底座圆盘中心为原点，Z轴方向垂直向上，构建三维直角坐标系，则螺旋线的参数方程  $\mathbf{r}(t)$  如式(2)所示，参数  $t \in (0, 8\pi)$ 。假设坐标系的单位为 mm，本文式中常数系数  $a$  取 200， $b$  取  $-20/\pi$ ， $c$  取  $40/\pi$ 。螺旋线的曲率半径沿 Z 轴方向逐渐减小，最大半径为 200 mm，最小半径为 40 mm。事实上，参数  $t$  代表绕 Z 轴旋转的角度，与极坐标的表达形式相似。

$$\begin{aligned} x(t) &= (a + bt) \cos(t) \\ y(t) &= (a + bt) \sin(t) \\ z(t) &= ct \end{aligned} \quad (2)$$

### 2.1.3. 结构特性

三维变径螺旋线的良好特性是其能在路侧场景下充当标靶的关键。

#### 1) 旋转不敏感

标靶旋转一定角度后其特征保持一致或相似，称其对旋转不敏感，这一特性满足了视角稳定性。

假设该螺旋线绕 Z 轴旋转角度  $\theta$ ，旋转矩阵  $\mathbf{R}_z(\theta)$  表示为：

$$\mathbf{R}_z(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

螺旋线结构  $\mathbf{r}(t) = (x(t), y(t), z(t))$  旋转变换后得到  $\mathbf{r}'(t) = (x'(t), y'(t), z'(t))$  为：

$$\begin{bmatrix} x'(t) \\ y'(t) \\ z'(t) \end{bmatrix} = \mathbf{R}_z(\theta) \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} (a + bt) \cos(t) \\ (a + bt) \sin(t) \\ ct \end{bmatrix} \quad (4)$$

展开并化简得到：

$$\begin{aligned} x'(t) &= (a + bt) \cos(t + \theta) \\ y'(t) &= (a + bt) \sin(t + \theta) \\ z'(t) &= ct \end{aligned} \quad (5)$$

螺旋线在旋转  $\theta$  角后， $x$  和  $y$  分量在形式上保持不变，而  $z$  分量的值不变。这表明螺旋线绕 Z 轴旋转其特征保持一致。

#### 2) 深度不消失

假设三维曲线  $\mathbf{r}(t)$  在某平面  $\mathbf{l}$  的投影为  $\mathbf{P}(t)$ ， $(a_1, b_1)$  是定义区间内的非空子区间，若当  $t_1 \in (a_1, b_1)$  时总存在另一非空子区间  $(a_2, b_2)$  且  $t_2 \in (a_2, b_2)$ ，使得  $\mathbf{P}(t_1) = \mathbf{P}(t_2)$ ，那么称在平面  $\mathbf{l}$  的投影发生深度消失现象，即  $\mathbf{r}(t)$  上的两段曲线在投影重合。本文定义的螺旋线在任一投影平面均不存在深度消失现象，保证了观测信息的有效性。

假设平面  $\mathbf{l}$  的单位法向量  $\mathbf{n} = (n_x, n_y, n_z)$ ，那么螺旋线  $\mathbf{r}(t)$  在平面  $\mathbf{l}$  上的投影  $\mathbf{P}(t)$  可以通过移除曲线在法向量方向的分量得到：

$$\mathbf{P}(t) = \mathbf{r}(t) - (\mathbf{r}(t) \cdot \mathbf{n}) \mathbf{n} \quad (6)$$

其中  $\mathbf{r}(t) \cdot \mathbf{n}$  是曲线与法向量的点积，给出曲线在法向量方向上的长度。根据定义，深度不发生消失的必要条件为投影  $\mathbf{P}(t)$  不存在交点，或若有  $\mathbf{P}(t_1) = \mathbf{P}(t_2)$  且  $t_1 \neq t_2$ ，投影  $\mathbf{P}(t)$  在  $t_1$  和  $t_2$  的导数不相等。 $\mathbf{P}(t)$  的

导数为:

$$\mathbf{P}'(t) = \mathbf{r}'(t) - (\mathbf{r}'(t) \cdot \mathbf{n}) \mathbf{n} \quad (7)$$

式中,  $\mathbf{r}'(t)$  为螺旋线的导数, 由式(2)求导有:

$$\mathbf{r}'(t) = (-(a+bt)\sin(t) + b\cos(t), (a+bt)\cos(t) + b\sin(t), c) \quad (8)$$

由于绕  $Z$  轴旋转特征不变, 因此对于投影平面的法向量  $\mathbf{n}$ , 可以简化为证明以下两种条件满足定义的情况, 一是  $n_x = 0$  或  $n_y = 0$ , 二是  $n_z = 0$ 。

当  $n_y = 0$  时,  $\mathbf{P}'(t)$  各分量为:

$$\begin{aligned} \mathbf{P}'_x(t) &= n_z^2 b \cos t - n_z^2 (a + bt) \sin t - n_x n_z c \\ \mathbf{P}'_y(t) &= (a + bt) \cos t + b \sin t \\ \mathbf{P}'_z(t) &= n_x n_z b \cos t - n_x n_z (a + bt) \sin t + c - cn_z^2 \end{aligned} \quad (9)$$

整理式(9), 考虑消除常数项和相同系数, 发现  $\mathbf{P}'(t)$  各分量的值分别由以下表达式决定:

$$\begin{aligned} \mathbf{P}'_x(t) &= \rho_x(f_x(t)), f_x(t) = b \cos t - (a + bt) \sin t \\ \mathbf{P}'_y(t) &= \rho_y(f_y(t)), f_y(t) = (a + bt) \cos t + b \sin t \\ \mathbf{P}'_z(t) &= \rho_z(f_z(t)), f_z(t) = b \cos t - (a + bt) \sin t \end{aligned} \quad (10)$$

注意  $f_x(t) = f_z(t)$ , 且有:

$$h(t) = f_x^2(t) + f_y^2(t) = (a + bt)^2 + b^2 \quad (11)$$

在定义域内  $h(t)$  为单调递减函数, 若有  $f_x(t_1) = f_x(t_2)$ , 则必有  $f_y(t_1) \neq f_y(t_2)$ , 那么  $\mathbf{P}'_x(t) \neq \mathbf{P}'_y(t)$ , 命题情况一得证。而当  $n_z = 0$  时,  $\mathbf{P}(t)$  的  $z$  分量  $\mathbf{P}_z(t) = ct$ , 为单调递增函数, 因此投影中不存在  $\mathbf{P}(t_1) = \mathbf{P}(t_2)$ , 命题情况二得证。

## 2.2. 点云数据处理

由于模型经过严格的数学方程定义, 因此可以采用 RANSAC (Random Sample Consensus, 随机采样一致性) [15] 进行螺旋线点云的提取。但是点云数量庞大, 直接使用 RANSAC 算法效率低下, 计算成本高。因此设计一种三维滑动窗口采样策略, 以局部而精确地搜索总体样本, 下面介绍改进的 RANSAC 算法。

一帧点云的点数量为  $N_L$ , 其中螺旋线的点数量为  $N_h$ , 假设在一帧点云中随机采样一点是螺旋线点群(内点群)的概率为  $w$ , 则:

$$w = \frac{N_h}{N_L} \quad (12)$$

假设每次采样  $n$  个点, 由于点云样本数量远远大于螺旋线点群数量, 因此不放回采样的概率与放回采样的概率近似相等, 则所有点均为内点群的概率为  $w^n$ , 至少一个点不是内群点的概率为  $1 - w^n$ 。重复  $k$  次采样, 每次采样均存在离群点(即没有一次采样的  $n$  个点均为内群点)的概率为  $(1 - w^n)^k$ 。假设重复  $k$  次后采样的所有点均为内群点的概率为  $p$ , 则:

$$p = 1 - (1 - w^n)^k \quad (13)$$

得到采样次数  $k$  为:

$$k = \frac{\log(1 - p)}{\log(1 - w^n)} \quad (14)$$

所以, 如果希望采样符合要求的概率  $p$  越高, 可以增加采样次数  $k$ , 减少每次采样的点数  $n$ , 提高内点群在所有样本中的概率  $w$ 。以速腾聚创 RS-Ruby Lite 80 线激光雷达为例, 一帧点云数量为 144,000 个, 而 ouster OS1 128 线激光雷达每帧点云数量高达 260,000 个, 螺旋线模型的点云数量依据模型尺寸和激光雷达的分辨率估计,  $w$  的值将会非常小。

针对上述问题, 算法改进的目标是提高内群点占样本总量的比例  $w$ 。具体来说, 在三维空间中定义一个三维滑动窗口, 窗口大小  $W = (W_x, W_y, W_z)$ , 在各坐标轴上的滑动步长为  $\lambda$ 。窗口按照预设的步长在点云样本集中逐步移动, 每个窗口内的点云数据作为独立的采样子集, 用于后续的 RANSAC 算法处理。合理的设置窗口尺寸和滑动步长, 可以极大提高子集内群点占比  $w$ 。通过多次迭代, 选择符合预设模型内群点数占比最高的结果作为螺旋线的点云数据。

在滑动窗口中建立三维直角坐标系, 螺旋线模型在坐标系中不可避免会发生旋转和平移变换, 设  $(x', y', z')$  为变换后的坐标, 则:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \mathbf{R}_w \cdot \begin{bmatrix} (a+bt)\cos(t) \\ (a+bt)\sin(t) \\ ct \end{bmatrix} + \mathbf{t}_w \quad (15)$$

式中  $\mathbf{R}_w$  为滑动窗口内的旋转矩阵,  $\mathbf{t}_w$  为滑动窗口内的平移向量。当点云数量较为稀疏时, 可以利用拟合的映射参数生成点云进行补充, 以增加点云的信息量。

### 2.3. 图像数据处理

为了从图像中精确地识别出螺旋线结构, 采用图像语义分割技术。语义分割是一种基于深度学习的图像分析方法, 能够将图像中的每个像素分类到预定的类别。本文采用 yolo v9 [16] 的语义分割部分实现。

将语义分割的结果作为优化标定参数过程的激励掩码, 用于指导激光雷达螺旋线的点云在相机图像上投影位置。具体来说, 定义一种局部距离变换, 以在特定阈值范围内增强目标区域周围像素的灰度值。这种变换依据像素到最近目标区域的距离来调整其灰度值, 使得距离目标区域越近的像素灰度值越高。

设图像语义分割的结果为  $\mathbf{I}$ 。 $\mathbf{I}$  为二值图像, 目标区域(即螺旋线)  $\Omega$  为白色, 背景区域  $\mathbf{B}$  为黑色。定义一个距离函数  $d(p, \Omega)$  表示像素  $p$  到最近目标区域  $\Omega$  的距离,  $t$  为参与变换的像素与目标区域的最大距离。变换后的图像  $\mathbf{I}'$  可以通过以下公式定义:

$$\mathbf{I}'(p) = \begin{cases} \alpha \left(1 - \frac{d(p, \Omega)}{t}\right), & \text{if } 0 \leq d(p, \Omega) \leq t \\ \mathbf{I}(p), & \text{otherwise} \end{cases} \quad (16)$$

其中,  $\alpha$  为一个常数, 用于调节灰度值的增幅, 确保目标区域的灰度值不变。通常取  $\alpha = 255$ , 即图像灰度的最大值。

### 2.4. 目标函数

利用基于语义分割的掩码来优化激光雷达和摄像头的联合标定过程。通过语义分割, 我们能够准确地识别图像中的关键区域, 并生成相应的激励掩码。这一掩码不仅定义了感兴趣的目标区域, 而且还指导了激光雷达点云数据的投影, 以确保点云主要集中在感兴趣区域内。

定义激励函数  $U$  作为点云投影到目标掩码区域内的度量:

$$U = \sum_{i=1}^N M(\mathbf{p}_i) \quad (17)$$

其中  $N$  为参与投影的点云总数,  $M$  是式(16)的归一化表达,  $\mathbf{p}_i$  是投影到图像平面的点, 满足:

$$\mathbf{p}_i = \mathbf{K}_c (\mathbf{R} \mathbf{p}_i^L + \mathbf{t}) \quad (18)$$

式中  $\mathbf{K}_c$  为相机内参,  $\mathbf{p}_i^L$  表示激光雷达坐标系的坐标,  $\mathbf{R}$  为旋转矩阵,  $\mathbf{t}$  为平移向量。根据罗德里格斯公式将含有九个自由度的旋转矩阵  $\mathbf{R}$  用只有 3 个自由度的旋转向量  $\mathbf{r}$  表示  $\mathbf{R} = h(\mathbf{r})$ , 最终的目标函数为:

$$\varphi(\mathbf{R}, \mathbf{t}) = U(h(\mathbf{r}), \mathbf{t}) \quad (19)$$

求解的联合标定参数为:

$$\hat{\mathbf{R}}, \hat{\mathbf{t}} = \arg \max_{\mathbf{R}, \mathbf{t}} \varphi(\mathbf{R}, \mathbf{t}) \quad (20)$$

由于螺旋线模型尺寸有限, 在图像中的目标区域比较小, 因此初始值的选择十分重要, 否则可能出现参与投影的目标点云全部处于激励掩码的零空间中, 将难以找到最佳搜索方向。本文分别从 3 帧对应的图像和点云中提取螺旋线的特征点, 为方便计算可以选择螺旋线结构的最高点, 采用 P3P [17] 先求得粗标定结果, 再此基础上进行进一步的细标定。由于包含有离散函数, 目标函数非凸且不可微, 因此采用子微分策略[18], 在子微分锥中选择更新方向。

### 3. 实验设计与结果

#### 3.1. 场景部署

我们使用 ouster OS1 128 线激光雷达和元戎启行 Deeproute Vision Camera V2 相机部署在真实道路上进行实验如图 3, 部署示意图如图 4 所示。相机安装在高约 6.5 米的横杆上, 激光雷达使用抱箍固定在路灯柱上, 安装高度距离地面约 2 米, 横杆与路灯的水平距离约为 11 米。图像数据和点云数据通过千兆网线传输到边缘服务器, 边缘服务器为 Nuvo-8208GC 工控机。传感器与服务器经过时间硬同步, 确保获取同一时刻的图像数据和点云数据。



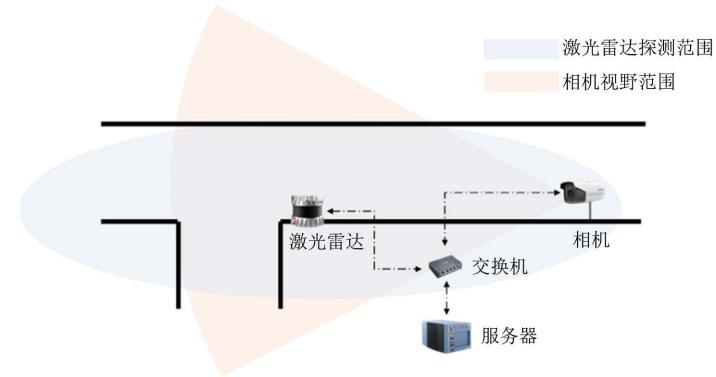
Figure 3. Installed LiDAR and camera  
图 3. 激光雷达和相机

#### 3.2. 评价指标

在激光雷达和相机联合标定中, 常用的评价指标是重投影误差。这一指标衡量的是在三维空间中的点经过投影变换到二维图像平面后与观测像素的误差大小, 用以评估标定的准确性和精度。重投影误差可以用像素坐标系下的欧氏距离来表示:

$$e = \frac{1}{N} \sum_i^N \sqrt{(\mathbf{p}_x^i - \hat{\mathbf{p}}_x^i)^2 + (\mathbf{p}_y^i - \hat{\mathbf{p}}_y^i)^2} \quad (21)$$

其中,  $(p_x^i, p_y^i)$ 是观测的像素坐标,  $(\hat{p}_x^i, \hat{p}_y^i)$ 是由激光雷达坐标投影变换得到的像素坐标。



**Figure 4.** Sketch map of devices installed  
**图 4.** 设备部署示意图

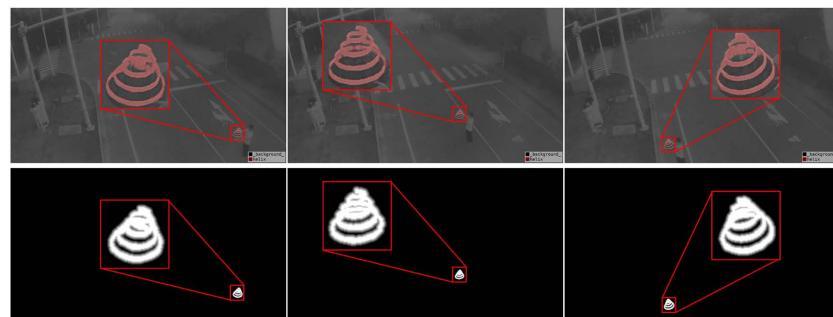
### 3.3. 实验结果

采集同一时刻螺旋线模型的图像和点云数据, 如图 5 所示。为了验证模型可以适用于路侧场景, 特将模型放置于激光雷达和相机之间的位置。粗标定阶段选择的对应点在图中使用红点标注。



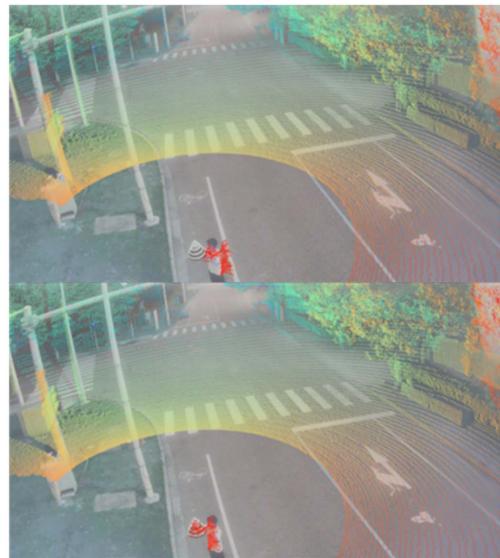
**Figure 5.** Images and point cloud of the target  
**图 5.** 标靶的图像和点云数据

图像数据采用 yolo v9 进行训练识别, 平均 IoU(交并比)为 97.3%, 图像语义分割和激励掩码的距离变换的结果如图 6 所示。激光雷达检测的三维滑动窗口大小  $W = (60, 60, 60)$ , 单位为 cm; 滑动步长  $\lambda$  为 20 cm。激光雷达识别结果使用螺旋线参数方程将点云数量补齐至 1000 个, 参与投影至激励掩码。



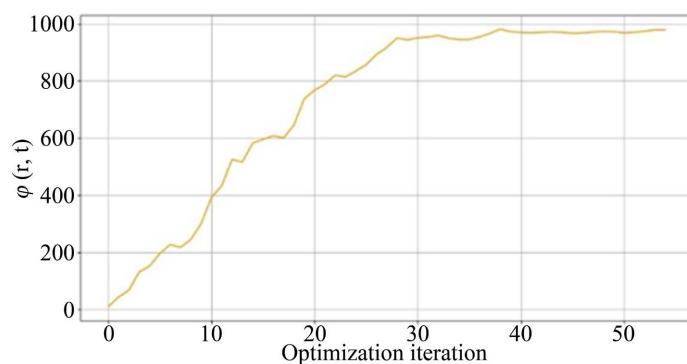
**Figure 6.** Images segmentation and objective mask  
**图 6.** 语义分割和激励掩码

粗标定的结果如图 7 上所示, 重投影误差平均为 38.14 个像素点。由于粗标定阶段只使用了三组对应点, 且对应点的选择为模型估计的结果, 因此误差较大, 但其为细标定阶段提供了一个可靠的初始值, 使细标定阶段的优化获得准确的参数。细标定的结果如图 7 下所示, 重投影误差平均为 1.91 个像素点。将点云直接投影到图像上可以很直观地分辨参数求解的准确性, 在工程上也常常以此定性分析联合标定的结果。粗标定阶段, 模型的点云投影到图像上有明显的偏移, 包括背景里的路灯杆等易于分辨的明显特征。经过细标定的优化, 点云准确地投影到相应位置上。



**Figure 7.** Calibration result: Coarse calibration (above), fine calibration (below)  
**图 7.** 标定结果: 粗标定(上), 细标定(下)

我们对比了可以应用于路侧场景的标定方法, 包括基于标靶的棋盘格法[4]和无标靶的边缘特征匹配法[14], 结果如表 1 所示, 单位 px 代表像素, 迭代优化过程如图 8。对比可知, 我们的方法精度比棋盘格法[4]高 0.6 px, 比边缘检测法高 1.41 px。我们的方法更具有鲁棒性, 避免了陷入局部最优解; 同时效率更高, 并且能够在路侧场景中满足标定需求。



**Figure 8.** Objective function value changes with iterations  
**图 8.** 目标函数值随迭代轮次变化情况

位于激光雷达和相机之间的共视区内, 我们的方法表现良好。不同于采用棋盘格法, 在采集数据时无需特别注意标靶的角度, 同时由于标靶设计时采用高反射率材料, 颜色对比度大, 在点云和图像中特

征明显,识别率高。我们测试了螺旋线标靶可以被检测识别的距离,以 ouster OS1 128 线激光雷达为例,受限于模型尺寸大小,标靶在距离激光雷达 30 米的区域就因采样点稀少而难以识别,而距离相机 50 米时图像特征依然明显,语义分割效果良好。因此本文所述方法应用于路侧标定,最远距离的性能取决于激光雷达的分辨率和标靶的尺寸。

**Table 1.** Comparison of different calibration methods

**表 1. 不同标定方法的对比**

方法	重投影误差	稳定性	效率
Zhang [4]	2.51 px	稳定	需要控制棋盘格角度,且采集数据区域有限
Levinson [14]	3.32 px	存在陷入局部最优解	无需标靶
本文方法	1.91 px	稳定	无需严格控制标靶角度,且采集范围大

## 4. 结束语

为了实现路侧场景下激光雷达和相机的联合标定问题,本文提出了一种三维变径螺旋线的结构,该结构从不同的观测角度都能获得相似的特征,以此满足激光雷达和相机视野方向相对场景下观测标靶的标定需求。该方法充分结合了基于标靶和基于无标靶两类方法的优点,先进行粗标定获得距离全局最优解较近的初始值,再进行细标定对初始值进行优化,避免陷入局部最优,同时保证较高的精度。通过实验表明,该方法具有较好的性能。后续工作可以基于此结构加以改进,例如固定在装有 GPS 的可移动装置上,搭建一站式解决路侧传感器设备的自动标定系统;另外螺旋线模型可以选择更高反射率和更加不易发生形变的材料,以进一步提高标定性能。

## 基金项目

这项工作得到了上海市科委重点课题“空港智慧货运通道构建关键技术”(项目编号: 22dz1203400)的资助。

## 参考文献

- [1] Roriz, R., Cabral, J. and Gomes, T. (2021) Automotive LiDAR Technology: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, **23**, 6282-6297. <https://doi.org/10.1109/TITS.2021.3086804>
- [2] Wang, X., Li, K. and Chehri, A. (2023) Multi-Sensor Fusion Technology for 3D Object Detection in Autonomous Driving: A Review. *IEEE Transactions on Intelligent Transportation Systems*, **25**, 1148-1165. <https://doi.org/10.1109/TITS.2023.3317372>
- [3] Li, X., Xiao, Y., Wang, B., et al. (2023) Automatic Targetless LiDAR—Camera Calibration: A Survey. *Artificial Intelligence Review*, **56**, 9949-9987. <https://doi.org/10.1007/s10462-022-10317-y>
- [4] Zhang, Q. and Pless, R. (2004) Extrinsic Calibration of a Camera and Laser Range Finder (Improves Camera Calibration). 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, 28 September-2 October 2004, 2301-2306.
- [5] Datta, A., Kim, J.S. and Kanade, T. (2009) Accurate Camera Calibration Using Iterative Refinement of Control Points. 2009 IEEE 12th International Conference on Computer Vision Workshops, Kyoto, 27 September-4 October 2009, 1201-1208. <https://doi.org/10.1109/ICCVW.2009.5457474>
- [6] Park, Y., Yun, S., Won, C.S., et al. (2014) Calibration Between Color Camera and 3D LIDAR Instruments with a Polygonal Planar Board. *Sensors*, **14**, 5333-5353. <https://doi.org/10.3390/s14030533>
- [7] Yan, G., He, F., Shi, C., et al. (2023) Joint Camera Intrinsic and Lidar-Camera Extrinsic Calibration. 2023 IEEE International Conference on Robotics and Automation, London, 29 May-2 June 2023, 11446-11452. <https://doi.org/10.1109/ICRA48891.2023.10160542>

- 
- [8] Beltrán, J., Guindel, C., de la Escalera, A., *et al.* (2022) Automatic Extrinsic Calibration Method for Lidar and Camera Sensor Setups. *IEEE Transactions on Intelligent Transportation Systems*, **23**, 17677-17689.  
<https://doi.org/10.1109/TITS.2022.3155228>
  - [9] 王世强, 孟召宗, 高楠, 等. 激光雷达与相机融合标定技术研究进展[J]. 红外与激光工程, 2023, **52**(8): 111-124.
  - [10] Yuan, C., Liu, X., Hong, X., *et al.* (2021) Pixel-Level Extrinsic Self Calibration of High Resolution Lidar and Camera in Targetless Environments. *IEEE Robotics and Automation Letters*, **6**, 7517-7524.  
<https://doi.org/10.1109/LRA.2021.3098923>
  - [11] Liu, X., Yuan, C. and Zhang, F. (2022) Targetless Extrinsic Calibration of Multiple Small FoV LiDARs and Cameras Using Adaptive Voxelization. *IEEE Transactions on Instrumentation and Measurement*, **71**, 1-12.  
<https://doi.org/10.1109/TIM.2022.3176889>
  - [12] Song, W.S., Zhang, Z.H. and Gao, N. (2022) Spatial Pose Calibration Method for Lidar and Camera Based on Intensity Information. *Laser and Optoelectronics Progress*, **59**, Article 0215003. <https://doi.org/10.3788/LOP202259.0215003>
  - [13] Pandey, G., McBride, J.R., Savarese, S., *et al.* (2015) Automatic Extrinsic Calibration of Vision and Lidar by Maximizing Mutual Information. *Journal of Field Robotics*, **32**, 696-722. <https://doi.org/10.1002/rob.21542>
  - [14] Levinson, J. and Thrun, S. (2013) Automatic Online Calibration of Cameras and Lasers.  
<https://doi.org/10.15607/RSS.2013.IX.029>
  - [15] Fischler, M.A. and Bolles, R.C. (1981) Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, **24**, 381-395.  
<https://doi.org/10.1145/358669.358692>
  - [16] Wang, C.Y., Yeh, I.H. and Liao, H.Y.M. (2024) YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information. arXiv: 2402.13616.
  - [17] Gao, X.S., Hou, X.R., Tang, J., *et al.* (2003) Complete Solution Classification for the Perspective-Three-Point Problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **25**, 930-943.  
<https://doi.org/10.1109/TPAMI.2003.1217599>
  - [18] Zhu, Y., Li, C. and Zhang, Y. (2020) Online Camera-Lidar Calibration with Sensor Semantic Information. 2020 *IEEE International Conference on Robotics and Automation*, Paris, 31 May-31 August 2020, 4970-4976.  
<https://doi.org/10.1109/ICRA40945.2020.9196627>