

DOI: 10.19650/j.cnki.cjsi.J2412924

# 基于点云分割与主成分配准的 换电机器人位姿估计方法\*

王立辉<sup>1</sup>, 陈勇吉<sup>1</sup>, 韩华春<sup>2</sup>, 顾炜琪<sup>1</sup>, 陈良亮<sup>3</sup>

(1. 东南大学 仪器科学与工程学院 微惯性仪表与先进导航技术教育部重点实验室 南京 210096;  
2. 国网江苏省电力有限公司 南京 210024; 3. 国电南瑞科技股份有限公司 南京 211100)

**摘要:**针对新能源电动汽车换电过程中电池包加解锁精确定位问题,提出一种基于点云分割与主成分配准的电池包加解锁位姿估计方法。该方法首先使用实例分割网络分割出场景图片中的加解锁实例,将加解锁实例对应的深度数据投影成点云后,采用统计滤波和体素滤波对加解锁点云进行去噪和下采样处理;其次通过在 PointNet++ 网络的特征提取层嵌入 SGE 注意力机制模块,增强点云中点的空间语义特征并分割出锁头点云;最后采用点云主成分配准将锁头点云与目标点云配准,获取加解锁的空间位姿。实验结果表明,本文提出的位姿估计算法具有较高精度和一定抗干扰能力,其加解锁分割精度为 98.02%,位姿估计平移误差为 2.239 mm,角度误差为 1.822°,RMSE 为 1.495 mm,满足换电机器人加解锁定位精度需求。

**关键词:** 换电机器人;位姿估计;点云配准;点云分割;注意力机制

**中图分类号:** TH89 **文献标识码:** A **国家标准学科分类代码:** 460.40

## Pose estimation method for battery swapping robot based on point cloud segmentation and principal component registration

Wang Lihui<sup>1</sup>, Chen Yongji<sup>1</sup>, Han Huachun<sup>2</sup>, Gu Weiqi<sup>1</sup>, Chen Liangliang<sup>3</sup>

(1. Key Laboratory of Micro-inertial Instrument and Advanced Navigation Technology, Ministry of Education, School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China; 2. Jiangsu Provincial Electric Power Corporation, Nanjing 210024, China; 3. NARI Technology Company Ltd, Nanjing 211100, China)

**Abstract:** Aiming at the precise positioning issue of the battery pack locking mechanism during the battery swapping process of new energy electric vehicles, a pose estimation method based on point cloud segmentation and principal component registration is proposed. The method first utilizes an instance segmentation network to extract the locking mechanism instances from the scene images. The depth data corresponding to the locking mechanism instances are projected into point clouds, and statistical filtering and voxel filtering are applied to denoise and downsample the point clouds. Secondly, by embedding the SGE attention mechanism module into the feature extraction layer of PointNet++ network, the spatial semantic features of points in the point cloud are enhanced, and the lock head point cloud is segmented. Finally, the spatial pose for locking mechanism is obtained by aligning the locking mechanism head point cloud with the target point cloud through principal component registration. Experimental results show that the pose estimation algorithm proposed in this paper has high accuracy and certain anti-interference capability, achieving a locking mechanism point cloud segmentation accuracy of 98.02%, a translation error of 2.239 mm, an angular error of 1.822°, and an RMSE of 1.495 mm, meeting the positioning accuracy requirements for battery swapping robots.

**Keywords:** battery swapping robot; pose estimation; point cloud registration; point cloud segmentation; attention mechanism

## 0 引言

新能源电动汽车的及时补能是其持续运行的关键。为了克服新能源电动汽车电池包充电时间长的缺点,直接更换动力电池包是当下研究的热点。在新能源电动汽车换电过程中,给换电机器人提供一个准确的电池包加解锁位姿信息,是实现电池包拆卸和安装的重要步骤。目前成熟的换电方式为机械式,该方法通过机械装置使电池包加解锁和换电机器人之间的位姿为一个固定值,随后通过程序预定一条固定的轨迹与加解锁对接,故机械式定位具有很高的稳定性和准确性。但随着新能源汽车保有量增加,不同品牌、不同结构尺寸的电池包对换电兼容性提出了挑战。相对于机械式定位,视觉定位对汽车底盘环境具有较强的实时感知能力和适应性<sup>[1]</sup>,所以根据视觉定位结果引导换电机器人换电是目前解决换电兼容性问题的一个重要研发方向。

在换电站中,大多数零件颜色相近且纹理较弱,导致图像信息变得不够可靠<sup>[2]</sup>。点云相比于图像数据可以提供物体更准确的形状和位置信息,在空间感知任务上更有优势<sup>[3-4]</sup>。传统基于点云的位姿估计方法可以分为基于模板匹配、基于对应关系和基于投票的方法。基于模板匹配方法需要从模板库中寻找与当前场景差异最小的模板,并将其位姿作为目标位姿估计值。汪威等<sup>[5]</sup>先把位姿粗校准后的点云转换成灰度图,用得到的灰度图与标准灰度图进行模板匹配计算目标的位姿,文代洲等<sup>[6]</sup>用渲染视角采样的方法增加少量渲染模板优化位姿估计结果,此类方法受限于模板的数量和质量。基于对应点关系的位姿估计方法需要借助快速点特征直方图(fast point feature histogram, FPFH)<sup>[7]</sup>、视点特征直方图(viewpoint feature histogram, VFH)<sup>[8]</sup>等特征描述子寻找源点云和目标点云之间对应关系,根据所求对应关系用随机采样一致算法(random sample consensus, RANSAC)<sup>[9]</sup>、快速全局配准(fast global registration, FGR)<sup>[10]</sup>等方法进行点云配准得到物体的估计位姿,这类方法位姿结果受到目标几何信息和点云特征描述子鉴别能力的影响。Drost等<sup>[11]</sup>将点云点对特征存储在哈希表中,用投票的方式找到最优局部坐标系进而求解目标位姿。此外,还可以对物体关键点<sup>[12]</sup>投票获得关键点的三维坐标,建立目标关键点二维图像坐标和三维坐标的对应关系,进一步用n点透视算法(perspective n point, PnP)、最小二乘法等方法拟合目标位姿,此类方法对遮挡和堆叠的场景有较好效果。Charles等<sup>[13-14]</sup>开创性的提出了PointNet和PointNet++点云特征提取网络,为用深度学习进行位姿估计奠定了基础。借助PointNet++网络强大的点云特征提取能力,将其提取的全局特征与

FPFH点云局部特征结合,使得点云特征的表达能力更强<sup>[15]</sup>。PointNetLK<sup>[16]</sup>利用PointNet网络将点云映射成多维特征向量,将特征空间视为图像,再借助光流算法(lucas kanade, LK)进行点云配准估计目标位姿。CloudPose<sup>[17]</sup>网络用两个分离的网络分别预测物体的平移向量和方向向量,直接从点云数据中回归目标准确的位姿信息。可见,在位姿估计算法中引入深度学习可以有效提高位姿估计结果的准确性。

在换电站背景中,电池包加解锁属于小目标,为准确获取加解锁的位姿信息,本文的主要贡献为提出了一种基于点云分割与主成分配准的加解锁位姿估计方法。该方法用改进PointNet++网络对3D相机采集的加解锁点云进行精细分割,再将加解锁点云特征强的主要部分与目标点云配准获得准确的加解锁位姿信息,为用视觉的方式进行电池包加解锁定位提供一种有效途径。

## 1 加解锁点云获取

### 1.1 换电站位姿估计系统标定

新能源电动汽车驶入换电站后,停在柔性化停车平台上的固定位置,其动力电池包通过加解锁固定在汽车底盘。换电机器人位于停车平台下方,通过轨道移动。电动汽车换电时,换电机器人移动到电池包下方,在加解锁位姿信息的引导下用电池包夹具完成电池包的拆卸和安装。其中换电机器人电池包夹具和3D相机的安装方式为相机在电池包夹具外的手眼系统<sup>[18]</sup>(eye to hand),示意图如图1所示。

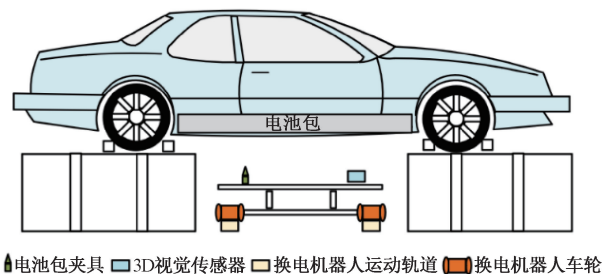


图1 新能源电动汽车换电示意图

Fig. 1 Schematic diagram of battery swapping for new energy electric vehicles

实际工作过程中,为使相机采集到准确的环境信息,首先需要标定3D相机的彩色相机获得彩色相机的内参 $K = \{f_x, f_y, c_x, c_y\}$ 和外参 $P = \{R_{3 \times 3}, T_{3 \times 1}\}$ 。其次,需要标定彩色相机和深度相机之间的对齐关系 $f$ ,为将深度图投影成点云做准备。最后,还需要对换电系统和3D相机组成的手眼系统进行标定,获得3D相机相对于换电系统的位姿信息 $T_s$ 。用位姿估计算法获得电池包加解锁位姿

信息  $T_e$  后, 最终用于引导换电机器人夹具与加解锁对接的位姿信息为:

$$T_o = T_s \cdot T_e \quad (1)$$

### 1.2 加解锁三维点云获取及预处理

电池包加解锁在庞大的汽车底盘和换电站背景中属于小目标, 这造成采集的图像数据大部分是非目标信息, 会对实际定位造成干扰。为解决这个问题, 本文采用图像实例分割技术提取彩色图像中感兴趣的加解锁区域<sup>[19]</sup>。电动汽车电池包实物图通过实例分割网络以后得到以图 2 局部放大区域为前景的掩膜图片。

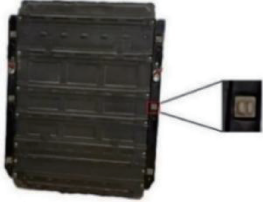


图 2 电动汽车电池包

Fig. 2 Physical picture of electric vehicle battery pack

考虑到电池包加解锁较小且需要实时处理, 本文选用对小目标检测效果突出且运行速度快的 YOLOv5 目标检测网络作为实例分割的主干网络, 结合实例分割模块完成电池包加解锁的实例分割。该网络采用 CSPDarknet53<sup>[20]</sup> 作为骨干网络提取输入图片的多层次特征, 然后通过特征金字塔结构融合不同层次的特征, 最后用目标检测头和实例分割头对目标的边界框、类别和实例掩膜进行预测, 结构如图 3 所示。

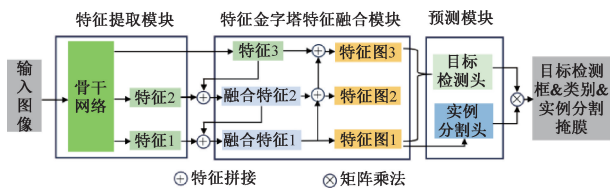


图 3 YOLOv5 实例分割网络结构

Fig. 3 YOLOv5 instance segmentation network structure

获得加解锁实例的掩膜图片后, 根据彩色相机和深度相机的对齐关系  $f$  求取彩色图掩膜前景对应的深度图区域, 然后利用彩色相机内参  $K$  将该区域深度数据逐像素投影得到初始加解锁点云, 映射关系如下:

$$\begin{cases} x = \frac{u - c_x}{f_x} z \\ y = \frac{v - c_y}{f_y} z \\ z = d \end{cases} \quad (2)$$

式中:  $u, v$  为像素坐标;  $f_x, f_y, c_x, c_y$  为彩色相机内参;  $z$  为深度值。

初始加解锁点云由深度图投影得到, 点的数量庞大。此外, 受相机工作噪声、电池包表面灰尘等因素影响, 初始点云数据不可避免会出现一些噪声点和离群点。基于上述原因, 本文首先利用统计滤波去除点云中的噪声点和离群点, 然后用体素滤波对点云进行下采样, 用以提高点云数据的质量。判断点云中点  $p_i(x_i, y_i, z_i)$  是否为离群点时, 首先计算这个点与邻域内  $n$  个点欧式距离的均值:

$$\begin{cases} \mu = \sum_{i=1}^n d_i / k \\ \sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (d_i - \mu)^2} \end{cases} \quad (3)$$

式中:  $d_i = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}$  为点  $p$  与邻域内一点的欧式距离。

如果点  $p$  邻域内一点与点  $p$  的距离在阈值范围  $(\mu - k\sigma, \mu + k\sigma)$  内, 则这个点予以保留, 否则剔除此点, 遍历点云中所有点完成点云统计滤波。然后对统计滤波后的点云建立 3D 体素方格, 计算每个非空体素内所有点的质心并将其作为该体素的代表点, 遍历所有体素方格获得经过体素滤波下采样后的点云。加解锁点云预处理后的结果如图 4 所示。

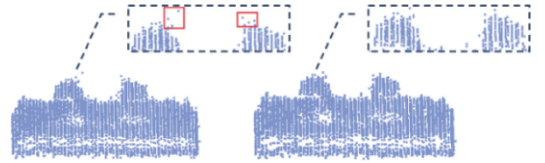


图 4 加解锁点云预处理结果

Fig. 4 Preprocessing results of locking mechanism point cloud

## 2 加解锁点云分割

初始加解锁点云预处理后, 提高了点云数据的质量和可用性。但加解锁锁头周围仍有一部分不用关注的平面点云, 这部分点云会对后续的点云配准造成影响, 所以需要进一步分割出精细的锁头点云。传统的点云聚类算法, 都要求点云有均匀的空间分布以及不同簇之间有明显的距离差异<sup>[21]</sup>。但实际情况下, 3D 相机从不同角度和距离采集到的加解锁点云数据具有不同的分辨率和分布特征, 因此传统聚类方法在换电场景中不能发挥出较好的效果<sup>[22]</sup>。针对这一问题, 本文提出了基于改进 PointNet++ 网络的点云分割方法。

相比于传统点云聚类分割方法, PointNet++ 可直接从点云数据中学习抽象特征, 更好地拟合点云的聚类边界。PointNet++ 点云分割网络整体框架采用 Encoder-Decoder 结构。其中 Encoder 部分用 3 个不同尺度的集合

采样层采集点的多层次特征,每个集合采样层由采样层、分组层和特征提取层3个部分组成。采样层通过最远点采样法从输入点云中选取确定数量的点作为关键点。分组层以采集到的关键点为中心,以球查询的方式把其周围 $k$ 个点划分成1个点云组。特征提取层则用PointNet提取每个点云组的特征。Decode部分由3个上采样层构成,每个上采样层首先通过最近邻插值的方式恢复点云的细节特征,然后通过跳跃连接的方式与集合

采样层相同尺度的特征进行融合。最后把融合特征输入多层感知机中获得点云的分割结果。PointNet++用不同尺度集合采样层级联提取点云不同层次的特征,为了增强每级特征在空间语义上的联系和弱化特征中无关的语义信息,本文在PointNet++的每个集合采样层的特征提取层后加入空间分组增强(spatial group-wise enhance, SGE)注意力机制模块<sup>[23]</sup>,并将ReLU激活函数替换为Mish激活函数以提高网络的实际分割性能,如图5所示。

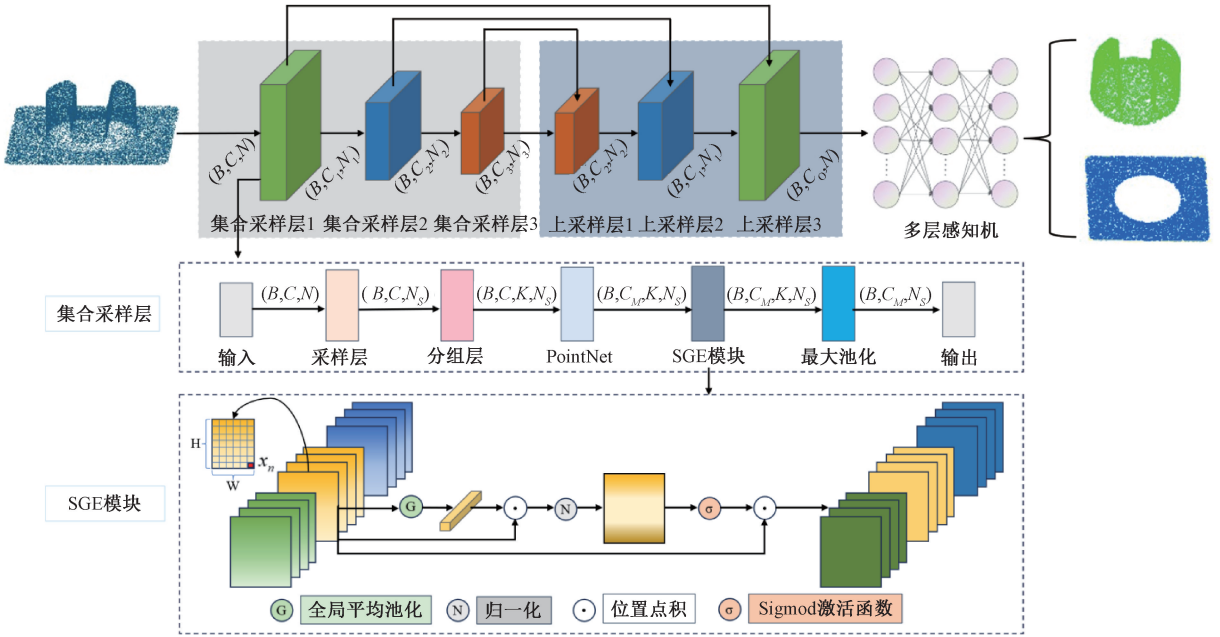


图5 改进 PointNet++点云分割网络

Fig. 5 Improved PointNet++ point cloud segmentation network

SGE注意力机制将输入特征按通道分成不同的特征组。对于每个特征组,它利用局部特征和全局特征的相似性计算出包含每个空间位置重要性的注意力掩膜,最后将输入特征和求得的注意力掩膜相乘获得增强语义特征的特征。其实现过程为:

1) 对于一个输入为 $(C, H, W)$ 的特征图,将该特征按通道数划分为 $g$ 组。针对每个特征组,特征图中每个位置可以用一个向量 $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ ,  $n = H \times W$ 表示,先提取该特征组的全局平均池化特征,再用获得的全局特征和输入特征进行点积运算,依次计算出每个位置的重要性系数:

$$\mathbf{c}_i = \left( \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \right) \cdot \mathbf{x}_i \quad (4)$$

2) 对求得的重要性系数进行归一化处理:

$$\hat{\mathbf{c}}_i = \frac{\mathbf{c}_i - \mu_c}{\sigma_c + \varepsilon}, \mu_c = \frac{1}{n} \sum_{i=1}^n \mathbf{c}_i, \sigma_c^2 = \frac{1}{n} \sum_{i=1}^n (\mathbf{c}_i - \mu_c)^2 \quad (5)$$

式中:  $\varepsilon = 1e^{-5}$  为接近0的常数。

3) 为每个归一化后的重要性系数引入一对权重和偏

置参数,求得注意力掩膜:

$$\mathbf{m}_i = \sigma(\gamma \hat{\mathbf{c}}_i + \beta) \quad (6)$$

式中:  $\gamma$  和  $\beta$  为权重和偏置参数;  $\sigma$  为 sigmoid 激活函数。

4) 用原输入特征与注意力掩膜相乘,获得增强空间语义信息的卷积特征。

$$\hat{\mathbf{x}}_i = \mathbf{x}_i \cdot \mathbf{m}_i \quad (7)$$

由于ReLU函数在负值区域输出为零,这导致了ReLU函数在负值区域存在梯度消失的问题。而Mish激活函数是一种平滑的非线性函数,在整个输入范围内都有非零梯度,这有助于避免梯度死亡问题,提高模型的训练效果和泛化性能。故用Mish函数替换了原来网络的ReLU函数,其计算公式为:

$$\text{Mish}(x) = x \cdot \tanh(\ln(1 + e^x)) \quad (8)$$

式中:  $\tanh(\cdot)$  为双曲正切函数;  $e$  为自然常数。

### 3 加解锁点云主成分配准

预处理后的加解锁点云进一步分割后,输出结构更

为精细的场景锁头点云(以下简称场景点云)。分别计算场景点云和目标点云的 FPFH 特征,通过两者特征的对对应关系求取加解锁的估计位姿。但由于加解锁具有对称的几何形状,结构相似的区域有相似的 FPFH 特征,在点云配准时易引起点云误匹配。此外,用 3D 结构光相机拍摄得到的点云具有近密远疏的特性,在某些角度下远离相机的一侧可能会出现孔洞,影响特征描述子表达。针对以上问题,本文提出了一种点云主成分配准方法,该方法利用场景主要部分点云代替整体点云完成初配准,再利用场景整体点云进行精配准以优化初配准结果。

由于 3D 相机位于汽车底盘下方,加解锁点云在世界坐标系  $XOY$  平面投影的一般情况如图 6 所示。

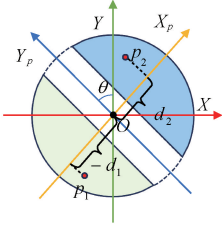


图 6 主成分坐标系示意图

Fig. 6 Schematic diagram of principal component coordinate system

本文处理的加解锁装置具有中间低,两边高的对称结构,将加解锁点云投影到世界坐标系( $O-XYZ$ )的  $XOY$  平面后,点云两边点的数量会显著多于中间低凹部分,根据此特性应用主成分分析法(principal component analysis, PCA)<sup>[24]</sup>可求出方向指向点云两边的主要向量  $X_p$ 、垂直于  $X_p$  的次要向量  $Y_p$  和同时垂直于  $X_p$ 、 $Y_p$  的向量  $Z_p$ 。为了消除  $X_p$ 、 $Y_p$ 、 $Z_p$  方向上的歧义,定义  $Z_p$  的方向与  $Z$  轴一致,  $Y_p$  与  $Y$  轴的夹角为  $\theta \in [0, \pi/2)$ ,  $X_p$  由  $Y_p$ 、 $Z_p$  两者外积得到。以  $X_p$ 、 $Y_p$ 、 $Z_p$  为坐标轴,点云质心为原点构成主成分坐标系。根据点云中每个点在主成分坐标系  $X_p$  方向上投影的正负(如图 6 点  $p_1$  和点  $p_2$  的投影  $-d_1$  和  $d_2$ ) 将点云分割成两部分,选择其中点数多且分辨率低的部分进行采样一致性初始配准(sample consensus initial alignment, SAC-IA)<sup>[7,9]</sup>,得到初步变换矩阵  $T_{SAC}$ 。利用  $T_{SAC}$  对整体场景点云进行刚体变换,使其与目标点云大致重合。最后使用迭代最近点算法(iterative closest point, ICP)<sup>[25]</sup>对变换后整体场景点云与目标点云进行精细配准,获得精配准矩阵  $T_{ICP}$ 。配准的主要过程如图 7 所示。

上述原理的详细实现过程如下:

1) 根据点云中每个点的坐标  $p_i(x_i, y_i, z_i)$ , 计算点云的质心  $p_c(x_c, y_c, z_c)$ :

$$p_c = \frac{1}{n} \sum_{i=1}^n p_i, i \in \{1, \dots, n\} \quad (9)$$

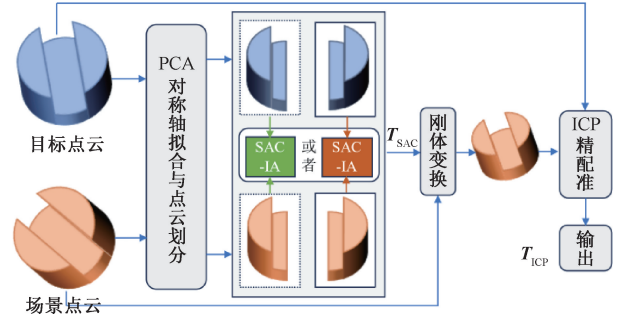


图 7 点云主成分配准过程

Fig. 7 The principal component registration process of point cloud

2) 把点云和质心投影到  $XOY$  平面,计算点云中各点  $p'_i(x_i, y_i, 0)$  与  $p'_c(x_c, y_c, 0)$  的协方差矩阵:

$$C = \frac{1}{n} \sum_{i=1}^n (p'_i - p'_c)(p'_i - p'_c)^T \quad (10)$$

3) 求解协方差矩阵的特征值,并按从小到大的顺序排列:  $\lambda_1 < \lambda_2 < \lambda_3$ 。计算每个特征值对应的特征向量  $v_1$ 、 $v_2$ 、 $v_3$ , 并求得  $X_p$ 、 $Y_p$ 、 $Z_p$ :

$$\begin{cases} Z_p = v_1 \\ Y_p = -(\|v_2\| \cos\theta) \cdot v_2 \\ X_p = Y_p \times Z_p \end{cases} \quad (11)$$

式中:  $\theta$  表示  $v_2$  与  $Y$  轴的夹角。

4) 以质心  $p_c$  为原点,  $X_p$ 、 $Y_p$ 、 $Z_p$  为坐标轴建立主成分坐标系,计算点云中每个点在  $X_p$  方向的投影,根据投影的正负把点云分割成两部分。

5) 用同样的方法完成场景点云  $P$  和目标点云  $Q$  的分割,选择 2 个点云中点数多、分辨率低的部分  $P_m$  和  $Q_m$ , 并计算 FPFH 特征  $F(P_m)$  和  $F(Q_m)$ 。点云中一点  $p_s$  的 FPFH 特征计算如下:

$$FPFH(p_s) = SPFH(p_s) + \frac{1}{k} \sum_{i=1}^k \frac{1}{\omega_i} SPFH(p_i) \quad (12)$$

式中:  $p_i$  为  $p_s$  邻域内一点;  $\omega_i$  为  $p_i$  和  $p_s$  之间的欧式距离;  $SPFH(p_s)$  为简化特征直方图。计算过程如下。

(1) 建立  $p_s$  和  $p_i$  的局部坐标系  $(u, v, w)$ , 其中  $n_s$  为  $p_s$  的法向量:

$$\begin{cases} u = n_s \\ v = u \times [(p_s - p_i) / \|p_s - p_i\|] \\ w = u \times v \end{cases} \quad (13)$$

(2) 根据局部坐标系和  $p_i$  法向量  $n_i$  计算  $p_s$  的局部特征:

$$\begin{cases} \alpha = v \cdot n_i \\ \phi = u \cdot [(p_s - p_i) / \|p_s - p_i\|] \\ \theta = \arctan(w \cdot n_i, u \cdot n_i) \end{cases} \quad (14)$$

(3) 将  $\alpha$ 、 $\phi$  和  $\theta$  的取值区间划分为 11 等份,计算  $p_s$

邻域内每一组点对 $(p_s, p_i)$ 的特征 $(\alpha, \phi, \theta)$ 并统计3个特征的直方图,然后将这3个直方图拼接成1个33维向量表示 $p_s$ 的SPFH特征。

6) 根据求得的特征 $F(P_m)$ 和 $F(Q_m)$ ,利用SAC-IA配准将 $P_m$ 和 $Q_m$ 初步配准,过程描述如下。

(1) 在 $P_m$ 中随机选择3个采样点,为保证采样点FPFH特征的差异,采样点两两之间的距离应该大于预定的距离阈值。在 $Q_m$ 中查找与 $P_m$ 中采样点FPFH特征最相似的3个对应点。

(2) 根据选出的3对匹配点,通过SVD奇异值分解计算对应点之间的刚体变换矩阵 $T_{SAC}$ ,根据 $T_{SAC}$ 求两幅点云的变换误差:

$$e = \sum_{i=1}^n h(\|q_i - T_{SAC} \cdot p_i\|) \quad (15)$$

式中: $q_i$ 为目标点云中距离 $p_i$ 最近的点; $h(\cdot)$ 为Huber loss。

(3) 重复(1)和(2),直到误差 $e$ 小于阈值或迭代次数达到上限,得到粗变换矩阵 $T_{SAC}$ 。

7) 把求得的 $T_{SAC}$ 应用到场景整体点云 $P$ 上得到点云 $P_i$ ,此时变换后的场景点云 $P_i$ 和目标点云 $Q$ 已经大致重合。为了进一步提高点云配准精度,使用ICP算法对 $P_i$ 和 $Q$ 进行精配准得到 $T_{ICP}$ ,最终求得加解锁的估计位姿为 $T_e = T_{ICP} \cdot T_{SAC}$ 。

## 4 实验结果与分析

### 4.1 实验环境参数

本文实验的操作系统为Ubuntu22.04系统,显卡为NVIDIA 4060 ti,处理器为Intel(R) Core(TM) CPU @ 2.50 GHz。深度学习框架采用PyTorch,结合开源库PCL1.10.0、OpenCV4.5.0和Open3D完成实验。

实验中使用的3D相机为国内某型号双目结构光相机。其正常工作范围为250~750 mm; $z$ 轴精度为 $\pm 0.15$  mm;彩色图最大分辨率为1920×1080,帧率为20 fps;深度图最大分辨率为1920×1200,帧率为8 fps;水平、垂直方向视场角为55°和36°。其实物如图8所示。实验时相机距离电池包0.4~0.6 m。



图8 结构光相机

Fig. 8 Structured light camera

本文实验的电池包实物图如图2所示,为了在实验室内对位姿估计算法进行试验,制作了电池包实物的3D打印模型,如图9所示。



图9 电动汽车电池包3D打印模型

Fig. 9 3D printing model of electric vehicle battery pack

### 4.2 点云分割实验

加解锁点云输入点云分割网络后,输出点云中每个点的标签信息,所有相同标签的点构成点云的一个部件。制作点云数据集时,固定电池包3D打印模型,用3D相机从不同角度、不同距离采集电池包3D打印模型的加解锁点云,并对采集到的点云进行体素滤波下采样处理。为了把加解锁点云分割成锁头和其周围背景平面两个部分,使用CloudCompare软件手动将下采样后的点云标注成锁头和背景平面两种类别的标签。此外,为提高网络的泛化能力,对采集到的点云数据进行随机小距离空间位姿变换模拟从不同角度采集加解锁点云,并对每个点云添加高斯噪声模拟3D相机在工作过程中可能出现的随机抖动,以此完成对点云数据集的扩充。将扩充后的数据集按8:2的比例随机划分为训练集和测试集,其中训练集包括624个点云,测试集包括156个点云。用相同的数据集分别对PointNet、PointNet++和改进PointNet++网络进行训练,用来验证点云分割网络对加解锁锁头点云的分割性能。

模型训练时,采用Adam优化器提高训练效率,batchsize设置为16,最远采样点个数为5500,epoch为200,初始学习率为0.001,训练过程中网络的损失和精度如图10所示。使用总体分割精度(overall segmentation accuracy, OA)来衡量点云分割网络的分割准确率,其计算公式如下:

$$OA = \sum_{i=1}^n \frac{TP_i}{W_i} \times 100\% \quad (16)$$

式中: $n$ 为测试集中点云的数量;对于测试集中的一个点云, $TP_i$ 为该点云中标签被正确预测的点数, $W_i$ 为该点云中点的总数。

模型训练完成后,经过测试集试验,PointNet和PointNet++的OA分别为97.45%和97.74%,使用SGE注意力机制改进的PointNet++网络的OA为98.02%,相比于PointNet和PointNet++分别提高了0.57和0.28个

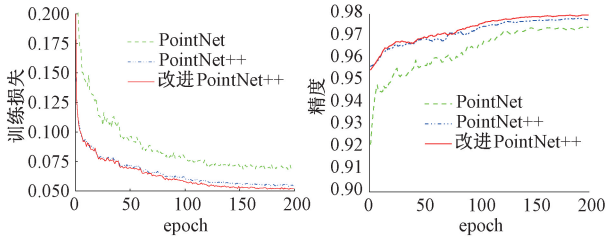


图 10 训练损失和精度

Fig. 10 Training loss and accuracy

百分点,说明改进的点云分割网络在加解锁点云分割场景中有更好的分割性能,更能满足实际使用需求,分割结果如图 11 所示。



图 11 加解锁点云分割结果

Fig. 11 Locking mechanism point cloud segmentation results

为了进一步验证同时采用 SGE 注意力机制和 Mish 激活函数的有效性,将 SGE 模块和 Mish 激活函数分别单独加入 PointNet++网络中设计 4 组消融实验,在制作的电池包加解锁数据集上进行训练和验证,结果如表 1 所示。实验结果表明,单独使用 SGE 模块和 Mish 激活函数的总体精度都没有同时使用 SGE 和 Mish 激活函数的分割精度高。

表 1 改进 PointNet++消融实验结果

Table 1 Improve PointNet++ablation experiment results

PointNet++	SGE	Mish	OA/%
✓	—	—	97.74
✓	✓	—	97.85
✓	—	✓	97.90
✓	✓	✓	98.02

#### 4.3 位姿估计实验

使用点云分割网络分割出加解锁锁头点云后,用本文所提算法求得加解锁相对于 3D 相机的估计位姿  $T_e$ 。为了验证位姿估计算法的性能,需要获取场景中加解锁点云的位姿真值。本文利用 CloudCompare 软件的配准功能,手动选场景点云和目标点云的关键点,直接实现场景点云和目标点云的精确配准,获取加解锁相对于 3D 相机的真实位姿  $T_{gt}$ 。

位姿估计的准确度用角度误差  $\Delta\theta$ 、平移误差  $\Delta t$  衡量,点云配准后的重合度用均方根误差 (root mean square

error, RMSE) 衡量,3 个指标的计算公式为:

$$\Delta\theta = \arccos\left(\frac{\text{trace}(\mathbf{R}_{gt}^T \mathbf{R}_e) - 1}{2}\right) \cdot \frac{180}{\pi} \quad (17)$$

$$\Delta t = \|\mathbf{t}_{gt} - \mathbf{t}_e\| \quad (18)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \|\mathbf{p}_i - \mathbf{q}_i\|^2}, \|\mathbf{p}_i - \mathbf{q}_i\| \leq d \quad (19)$$

式中:  $\mathbf{R}_{gt}$  和  $\mathbf{t}_{gt}$  为真实位姿的旋转矩阵和平移向量;  $\mathbf{R}_e$  和  $\mathbf{t}_e$  为算法估计位姿的旋转矩阵和平移向量;  $\mathbf{p}_i$  为场景点云中一点;  $\mathbf{q}_i$  为目标点云中距离  $\mathbf{p}_i$  距离最近的一点且该距离小于  $d$ , 本文  $d$  的值设为 20 mm。

为验证本文位姿估计算法的有效性,分别使用 RANSAC、ICP、FGR、广义迭代最近点 (generalized iterative closest point, GICP) 算法及先使用 RANSAC 粗配准再使用正态变换精配准的 SAC-NDT 算法,与本文提出的算法进行对比实验。实验时,3D 相机和电池包 3D 打印模型之间的位置保持不变,共采集 10 组场景点云,分别用所列方法估计加解锁相对 3D 相机的位姿,并计算  $\Delta\theta$ 、 $\Delta t$  和 RMSE 的平均值,结果如表 2 所示。

表 2 算法对比实验结果

Table 2 Algorithm comparison results

算法	$\Delta\theta/(\circ)$	$\Delta t/\text{mm}$	RMSE/mm
RANSAC	38.24	24.583	1.815
ICP	8.994	5.978	1.637
FGR	7.59	13.377	1.788
GICP	9.462	7.402	1.621
SAC-NDT	0.968	4.202	1.613
本文算法	1.816	1.943	1.608

电动汽车行驶过程中,会在加解锁表面积累泥沙灰尘,污渍的存在会干扰点云局部特征的表达。为验证本文所提位姿估计算法对加解锁表面污渍干扰的鲁棒性,首先把污渍建模成直径为 1~3 mm 的点,然后在上述采集的每个点云中随机选取 1/3 数量的点,在选择的中添加最小值为 1 mm,最大值为 3 mm 的均匀噪声,其示效果如图 12 所示。

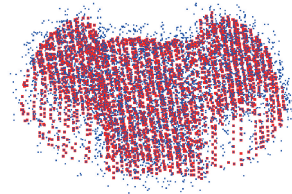


图 12 加解锁表面形貌噪声示意图

Fig. 12 Schematic diagram of locking mechanism surface morphology noise

分别用表2中所列方法对添加噪声后的点云进行位姿估计实验,并计算每种方法 $\Delta\theta$ 、 $\Delta t$ 和RMSE的平均值,实验结果如表3所示。

表3 算法对比试验结果(含噪声)

Table 3 Algorithm comparison results (with noise)

算法	$\Delta\theta/(\circ)$	$\Delta t/\text{mm}$	RMSE/mm
RANSAC	51.995	48.677	2.084
ICP	12.461	6.473	1.74
FGR	7.339	21.062	1.873
GICP	6.069	15.078	1.773
SAC-NDT	5.592	7.303	1.743
本文算法	2.385	4.451	1.718

表2得到的实验结果中,本文所提位姿估计算法 $\Delta\theta$ 为 $1.816^\circ$ , $\Delta t$ 为 $1.943\text{ mm}$ ,RMSE为 $1.608\text{ mm}$ ,整体配准精度优于表中其他方法。由表3可知,为锁头点云添加噪声后,表2所列算法的位姿误差都有明显增加,本文所提算法 $\Delta\theta$ 的增量为 $0.569^\circ$ , $\Delta t$ 的增量为 $2.508\text{ mm}$ ,RMSE的增量为 $0.135\text{ mm}$ ,相对于表中其他算法波动较小,这表明本文算法对于加解锁表面形貌噪声有较高的稳定性。

为验证本文所提位姿估计算法在实际应用中的可行性,用算法估计不同姿态下电池包3D打印模型的位姿。实验时,固定3D相机的位置,调整10次电池包3D打印模型的位姿,每个位姿采集5个场景点云并进行位姿估计。计算每个位姿下 $\Delta\theta$ 、 $\Delta t$ 和RMSE的平均值,实验结果表4所示。

表4 位姿估计实验结果

Table 4 Results of pose estimation experiment

序号	$\Delta\theta/(\circ)$	$\Delta t/\text{mm}$	RMSE/mm
1	1.644	1.962	1.559
2	1.484	0.572	1.673
3	1.165	1.740	1.574
4	1.739	1.523	1.683
5	1.980	1.894	1.713
6	1.482	1.284	1.545
7	1.489	1.889	1.558
8	1.740	1.781	1.534
9	1.226	0.600	1.120
10	0.718	1.232	1.585

由表4可知,电池包3D打印模型10个不同姿态下,估计位姿的平均角度误差达到 $1.467^\circ$ ,平均平移误差达到 $1.448\text{ mm}$ ,RMSE平均值达到 $1.554\text{ mm}$ ,配准结果的可视化图如图13所示。

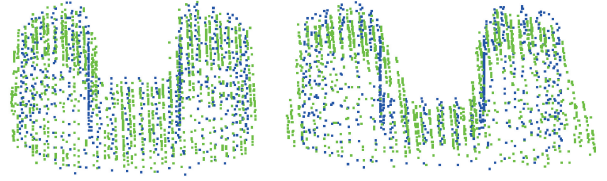


图13 加解锁锁头点云配准结果

Fig. 13 Locking mechanism head point cloud registration results

最后,验证本文所提位姿估计算法在电池包实物上的有效性。实验时,调整5次3D相机的位姿,每个位姿采集5个加解锁实物点云并进行位姿估计。计算每个位姿下 $\Delta\theta$ 、 $\Delta t$ 和RMSE的平均值,实验结果如表5所示。

表5 加解锁实物位姿估计实验结果

Table 5 Results of locking mechanism pose estimation

序号	$\Delta\theta/(\circ)$	$\Delta t/\text{mm}$	RMSE/mm
1	1.802	2.376	1.496
2	2.658	2.342	1.508
3	1.993	2.004	1.512
4	2.018	2.16	1.599
5	0.637	2.315	1.358

由表5可知,5次位姿估计结果的平均角度误差为 $1.822^\circ$ ,平均平移误差为 $2.239\text{ mm}$ ,RMSE平均值为 $1.495\text{ mm}$ 。电动汽车在换电站中换电时,为满足高效准确换电,对电池包加解锁的定位精度要求为:各方向定位精度小于 $5\text{ mm}$ ,所以本文提出的位姿估计算法满足电动汽车换电过程中的定位要求,兼具准确性和稳定性优势,可用于引导换电机器人工作。

## 5 结论

本文针对新能源电动汽车换电过程中电池包加解锁精准定位问题展开研究,提出了一种基于点云分割与主成分配准的换电机器人位姿估计方法。所提方法在PointNet++点云分割网络的集合采样层加入SGE注意力机制模块,提高了锁头点云分割的总体精度;其次,将PCA算法应用到SAC-IA粗配准方法中,结合ICP精配准算法获得加解锁的估计位姿,有效解决了因点云的对称结构和近密远疏特性引起的点云特征表达不强问题;此外,在点云预处理阶段加入统计滤波和体素滤波步骤,从而增强了算法对灰尘干扰的稳定性。实验结果表明,改进PointNet++网络的整体分割精度为 $98.02\%$ ,加解锁估计位姿的平均角度误差为 $1.822^\circ$ ,平均平移误差为



2. 239 mm, RMSE 的平均值为 1.495 mm, 满足换电过程中加解锁的定位要求, 可用于引导换电机器人进行电池包换电。后续将在算法中扩展加解锁的种类, 提高算法对不同车型电池包加解锁定位的兼容性。

## 参考文献

- [ 1 ] 张旭辉, 杨红强, 白琳娜, 等. 基于改进 RANSAC 特征提取的掘进装备视觉定位方法研究[J]. 仪器仪表学报, 2022, 43(12): 168-177.  
ZHANG X H, YANG H Q, BAI L N, et al. Research on the visual positioning method of tunneling equipment based on the improved RANSAC feature extraction[J]. Chinese Journal of Scientific Instrument, 2022, 43(12): 168-177.
- [ 2 ] 陈海永, 李龙腾, 陈鹏, 等. 复杂场景点云数据的 6D 位姿估计深度学习网络[J]. 电子与信息学报, 2022, 44(5): 1591-1601.  
CHEN H Y, LI L T, CHEN P, et al. 6D pose estimation network in complex point cloud scenes[J]. Journal of Electronics & Information Technology, 2022, 44(5): 1591-1601.
- [ 3 ] 陈慧娴, 吴一全, 张耀. 基于深度学习的三维点云分析方法研究进展[J]. 仪器仪表学报, 2023, 44(11): 130-158.  
CHEN H X, WU Y Q, ZHANG Y. Research progress of 3D point cloud analysis methods based on deep learning[J]. Chinese Journal of Scientific Instrument, 2023, 44(11): 130-158.
- [ 4 ] 龚国强, 田演, 夏鑫宇. 基于位姿参数估计的多视角点云配准方法[J]. 电子测量与仪器学报, 2024, 38(6): 241-252.  
GONG G Q, TIAN Y, XIA X Y. Multi-view point cloud registration method based on pose parameter estimation[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(6): 241-252.
- [ 5 ] 汪威, 张开颜, 刘亚川, 等. 一种新能源动力电池顶盖平面度检测方法研究[J]. 仪器仪表学报, 2020, 41(2): 218-225.  
WANG W, ZHANG K Y, LIU Y CH, et al. Research on the flatness detection method for cap of the new energy power battery[J]. Chinese Journal of Scientific Instrument, 2020, 41(2): 218-225.
- [ 6 ] 文代洲, 王晰, 任明俊. 基于渲染视角采样的轻量化模板匹配算法[J]. 激光与光电子学进展, 2024, 61(18): 404-412.  
WEN D ZH, WANG X, REN M J. A lightweight template matching algorithm based on rendering perspective sampling [J]. Laser & Optoelectronics Progress, 2024, 61(18): 404-412.
- [ 7 ] RUSU R B, BLODOW N, BEETZ M. Fast point feature histograms (FPFH) for 3D registration[C]. 2009 IEEE International Conference on Robotics and Automation. Kobe; IEEE, 2009: 3212-3217.
- [ 8 ] CHEN Z H, PAN B F. Research on key technology of water robot avoiding collision based on improved VFH algorithm[J]. Journal of Physics, 2021, 1820(1): 012064.
- [ 9 ] FISCHLER M A, BOLLES R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography [J]. Communications of the ACM, 1981, 24(6): 381-395.
- [ 10 ] ZHOU Q Y, PARK J, KOLTUN V. Fast global registration [C]. Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14. Springer International Publishing, 2016: 766-782.
- [ 11 ] DROST B, ULRICH M, NAVAB N, et al. Model globally, match locally: Efficient and robust 3D object recognition [C]. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco; IEEE, 2010: 998-1005.
- [ 12 ] 王太勇, 于恩霖. 基于三维关键点投票的物体位姿估计方法[J]. 天津大学学报(自然科学与工程技术版), 2024, 57(3): 291-300.  
WANG T Y, YU EN L. Object pose estimation method based on 3D key points voting[J]. Journal of Tianjin University (Science and Technology), 2024, 57(3): 291-300.
- [ 13 ] CHARLES R Q, SU H, KAICHUN M, et al. Pointnet: Deep learning on point sets for 3D classification and segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 77-85.
- [ 14 ] CHARLES R Q, YI L, SU H, et al. Pointnet++: Deep hierarchical feature learning on point sets in a metric space [J]. Advances in Neural Information Processing Systems, 2017, 30.
- [ 15 ] 王立辉, 苏余足威, 韩华春, 等. 智能换电站电池包锁止机构位姿视觉估计[J]. 光学精密工程, 2023, 31(21): 3135-3144.  
WANG L H, SU Y Z W, HAN H CH, et al. Visual 6D pose estimation of battery package locking mechanism in intelligent battery swapping station [J]. Optics and Precision Engineering, 2023, 31(21): 3135-3144.
- [ 16 ] AOKI Y, GOFORTH H, SRIVATSAN R A, et al. PointNetLK: Robust & efficient point cloud registration using PointNet [C]. 2019 IEEE/CV Conference on

- Computer Vision and Pattern Recognition, 2019: 7156-7165.
- [17] GAO G, LAURI M, WANG Y L, et al. 6D object pose regression via supervised learning on point clouds [C]. 2020 IEEE International Conference on Robotics and Automation(ICRA), Paris, France, 2020: 3643-3649.
- [18] ENEBUSE I, FOO M, IBRAHIM B S K K, et al. A comparative review of hand-eye calibration techniques for vision guided robots [J]. IEEE Access, 2021, 9: 113143-113155.
- [19] GU W CH, BAI SH, KONG L X. A review on 2D instance segmentation based on deep neural networks[J]. Image and Vision Computing, 2022, 120: 104401.
- [20] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. ArXiv preprint arXiv:2004.10934, 2020.
- [21] 吴文涛, 何泽洋, 杜旭, 等. 融合相机与激光雷达的目标检测与尺寸测量[J]. 电子测量与仪器学报, 2023, 37(6): 169-177.
- WU W T, HE Y Z, DU X, et al. Fusing camera and Lidar for object detection and dimensional measurement[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(6): 169-177.
- [22] 卢健, 贾旭瑞, 周健, 等. 基于深度学习的三维点云分割综述[J]. 控制与决策, 2023, 38(3): 595-611.
- LU J, JIA X R, ZHOU J, et al. A review of deep learning based on 3D point cloud segmentation [J]. Control and Decision, 2023, 38(3): 595-611.
- [23] LI Y X, LI X, YANG J. Spatial group-wise enhance: Enhancing semantic feature learning in CNN [C]. 16th Asian Conference on Computer Vision (ACCV), Macao, China, 2023: 316-332.
- [24] DUAN Y, YANG CH CH, CHEN H, et al. Low-complexity point cloud denoising for LiDAR by PCA-based dimension reduction[J]. Optics Communications, 2021, 482: 126567.
- [25] CHEN AN G, ZHUANG J L, HAN X Q. An improved ICP algorithm for 3D point cloud registration [C]. IEEE Conference on Pattern Recognition and Machine Learning (PRML), 2022: 205-210.

## 作者简介

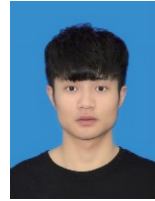


王立辉(通信作者), 2009年于哈尔滨工程大学获得博士学位, 现为东南大学仪器科学与工程学院教授, 主要研究方向为无人系统自主导航和精密传感。

E-mail: wlhseu@163.com

Wang Lihui (Corresponding author)

received his Ph. D. degree from Harbin Engineering University in 2009. He is currently a professor in the School of Instrument Science and Engineering at Southeast University. His main research interests include autonomous navigation of unmanned systems and precision sensing.



陈勇吉, 2022年于中北大学获得学士学位, 现为东南大学仪器科学与工程学院硕士研究生, 主要研究方向为三维视觉、位姿估计。

E-mail: chenyj\_11000@163.com

Chen Yongji received his B.Sc. degree from North University of China in 2022. He is currently a M. Sc. candidate in the School of Instrument Science and Engineering at Southeast University. His main research interests include three-dimensional vision and pose estimation.



韩华春, 2009年于山东大学获得学士学位, 2016年于中国科学院大学获得博士学位, 现为国网江苏省电力有限公司电力科学研究院专业主管, 主要研究方向为电动汽车充换电技术。

E-mail: 598543427@qq.com

Han Huachun received her B.Sc. degree from Shandong University in 2009 and received her Ph. D. degree from University of Chinese Academy of Sciences in 2016. She is currently a professional manager at Jiangsu Provincial Electric Power Corporation. Her main research interest is electric vehicle charging and swapping technology.



顾炜琪, 2022年于西南交通大学获得学士学位, 现为东南大学仪器科学与工程学院硕士研究生, 主要研究方向为机器人定位与规划控制。

E-mail: weiqi\_gu@seu.edu.cn

Gu Weiqi received his B.Sc. degree from Southwest Jiaotong University in 2022. He is currently a M. Sc. candidate in the School of Instrument Science and Engineering at Southeast University. His main research interests include robot localization, planning and control.



陈良亮, 1999年于南京航空航天大学获得学士学位, 2004年于南京航空航天大学获得博士学位, 现为国电南瑞科技股份有限公司业务专家, 主要研究方向为电动汽车充换电技术。

E-mail: chenliangliang@sgepri.sgcc.com.cn

Chen Liangliang received his B.Sc. degree and Ph.D. degree both from Nanjing University of Aeronautics and Astronautics in 1999 and 2004, respectively. He is currently a business expert at NARI Technology Company Ltd. His main research interest is electric vehicle charging and swapping technology.