

文章编号 1004-924X(2023)21-3135-10

智能换电站电池包锁止机构位姿视觉估计

王立辉^{1*}, 苏余足威¹, 韩华春², 陈良亮³, 张浩³

- (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,
江苏 南京 210096;
2. 国网江苏省电力有限公司, 江苏 南京 210024;
3. 国电南瑞科技股份有限公司, 江苏 南京 211100)

摘要:面向电动汽车的电池换电需求,对换电站电池包对接中的锁止机构定位问题,提出了一种基于点云分割的电池包锁止机构 6D 位姿估计方法。该方法使用 YOLOv5 网络从场景中分割出锁止机构的点云,并使用体素滤波与移动最小二乘拟合进行点云的滤波与平滑;通过引入点云分割网络预测点云标签,为快速点特征直方图特征加入全局语义特征,弥补快速点特征直方图只有点云局部特征的缺陷,并根据该特征进行随机抽样一致性刚体点云配准,估计锁止机构点云的 6D 位姿,最后使用迭代最近点方法算法校正位姿估计结果。实验结果表明,基于点云分割的锁止机构 6D 位姿估计算法精度较高,可以克服环境噪声导致的误匹配,精确获取锁止机构位姿,其位姿估计的角度误差可以达到 1.90° ,位移误差可以达到 1.4 mm, RMSE 可以达到 1.5 mm,为换电站电池对接定位提供了有效的解决途径。

关键词:换电机器人;电池包;位姿估计;点云配准;点云分割

中图分类号:U666.1 **文献标识码:**A **doi:**10.37188/OPE.20233121.3135

Visual 6D pose estimation of battery package locking mechanism in intelligent battery swapping station

WANG Lihui^{1*}, SU Yuzuwei¹, HAN Chunhua², CHEN Liangliang³, ZHANG Hao³

- (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*)
* *Corresponding author, E-mail: wangyi_sy1256@163.com*

Abstract: In order to meet the battery replacement demand of electric vehicles, a 6D pose estimation method of battery package locking mechanism based on point cloud segmentation is proposed to solve the positioning problem of locking mechanism during battery package docking in battery swapping station. This method uses YOLOv5 network to segment the point cloud of locking mechanism from the scene, and uses voxel filtering and moving least square fitting to filter and smooth the point cloud. The point cloud labels are predicted by the point cloud segmentation network, and the global semantic feature is added to the

收稿日期:2023-02-26;修订日期:2023-04-27.

基金项目:国家重点研发计划资助项目(No. 2021YFB2501603)

Fast Point Feature Histograms (FPFH) feature to make up for the defect that the FPFH has only the local feature of the point cloud. According to this feature, the Random Sample Consensus (RANSAC) rigid point cloud registration is leveraged, and the 6D pose of the locking mechanism point cloud is estimated. Finally, the Iterative Closest Point (ICP) algorithm is used to correct the pose estimation results. The experimental results show that the 6D pose estimation algorithm of locking mechanism based on point cloud segmentation has high accuracy, and can overcome the mismatching caused by environmental noise, and accurately obtain the position and attitude of locking mechanism. The angle error of position and attitude estimation can reach 1.90° , the displacement error can reach 1.4 mm, and the RMSE can reach 1.5 mm, which provides an effective solution for battery docking positioning in battery swapping station.

Key words: battery swapping robot; battery pack; pose estimation; point cloud registration; point cloud segmentation

1 引言

电动汽车具有节能环保、能量转换效率高优点,在我国得到了大力推广。更换电池包作为一种补电方式,可让用户在较短的时间内充满电能,弥补了充电方式补电缓慢的缺点,具有更好的用户体验。在电动汽车换电过程中,引导换电机构与电池包精准对接是高效快速换电的关键。该过程可通过机械浮动对接装置或 3D 视觉定位系统实现。而相对于机械式定位,3D 视觉定位更具兼容性与灵活性,因此是目前换电站技术的研发重点。

根据定位目标的不同,3D 视觉定位方法可以大致分为标识物定位与无标识物定位^[1]。标识物定位提前在物体上设置定位标识,使用相机对标识物进行定位,从而推算出物体的坐标与姿态。杨宁等^[2]以航天器本体和星箭对接环作为识别特征,在不依靠额外测量信息的情况下,利用视觉系统求解特征顶点的约束关系,从而计算出目标物的相对姿态。李莹等^[3]提出了一种用于火星探测的人工标识定位方法,通过边缘检测、椭圆拟合等手段识别定位着陆器,并对标志中心进行直线交点检测获得精确定位结果。尽管标识物定位相对步骤较少,精度相对较高,但并不是所有情况下都能提前安装标识,灰尘、污渍等环境因素也会干扰标识的定位。无标识物定位则直接识别物体的特征,以物体特征或特征匹配的方式进行定位,相对于换电场景具有更强的灵活性。Drost 等^[4]提出了一种基于点对特征(Point Pair Feature, PPF)投票的 6D 位姿估计方法,可

基于模板的全局 PPF 哈希表估计场景点云中物体的位姿,该方法可适应多种场景下的位姿估计,但是计算量过大。Hutchison 等^[5]提出 linemod 算法,采用模板匹配的方法计算 3D 物体 RGB-D 图像的彩色梯度与深度图梯度,与模板进行比对得到估计的位姿,可在较短时间内得到位姿估计结果,但由于存在多个匹配结果,需要对结果进行筛选。Verikas^[6]通过将构件的 CAD 模型与获取的场景点云数据进行特征描述子匹配,再通过点云配准实现对构件位姿的识别和估计,但是由于缺乏全局信息,特征描述子容易误匹配,影响位姿估计结果。

针对电动车换电对接中的锁止机构视觉定位问题,本文使用 3D 相机拍摄电池包对接部位的锁止装置,获得包含锁止装置的 RGB 图与点云,并针对锁止装置的位姿估计问题提出了基于点云分割的 6D 位姿估计方法,结合深度学习点云分割网络的全局感知能力与点云特征描述子的局部描述能力,实现更为精准、稳性的点云配准,从而获得满足电动车换电要求的锁止机构位姿估计结果。

2 原理

电动车换电示意图如图 1 所示。在换电站中,换电平台的作业空间处于车辆停靠区域的正下方,换电机器人置于该空间底部的柔性作业平台上,通过导轨沿相应的轨迹移动。电池包通过锁止机构固定在车辆底盘之上,换电机器人则经由电池包定位信息的引导,利用搭载的加解锁装置对电池包进行对接解锁并取下电池包。

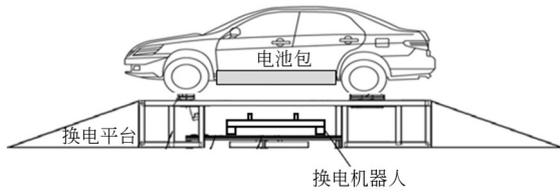


图 1 电动车换电示意图

Fig. 1 Schematic diagram of electric vehicle battery swapping

在换电过程中,电池包定位信息来源于电池包锁止机构的定位结果。对于电池包锁止机构的定位,有机械式定位与视觉定位两种方法。机械式定位方法根据已知的车辆尺寸、车辆停靠位置、电池包尺寸和换电站结构尺寸推算出电池包锁止机构的空间位置,通过加解锁装置的柔性机构弥补定位估计的误差。视觉定位方法则通过 3D 视觉传感器,如双目相机、RGBD 相机、激光扫描仪等,拍摄车辆底部电池包的图像,通过相应算法处理得到锁止机构的具体空间位姿。

相较于机械式定位方法,视觉定位方法更加精确,可获得旋转角度信息,并能兼容更多的车辆型号与电池包型号,同时能更好地应对环境干扰与意外情况,因此选用 3D 视觉定位方式。在 3D 视觉定位方案中,视觉传感器选用高精度双目结构光相机,可获得 RGBD 图像。相机置于换电平台底部 0.5 m 处的固定位置,图 1 中换电机器人下方,设置方式为眼到手(Eye to hand)型,相机使用前要经过标定,获得彩色摄像头的内参参数 $K = \{f_x, f_y, c_x, c_y\}$ 、彩色图片与深度图片的对齐关系 f 以及相机相对于换电站坐标系的外参参数,即视觉传感器坐标系相对换电坐标系的相对位姿 T_c 。视觉传感器拍摄上方的汽车底盘电池包,获得包括车底盘电池包、部分换电站结构以及锁止机构的 RGBD 图像,其中锁止机构为位姿估计的目标。使用 K 将深度图投影为点云,即:

$$\begin{cases} x = \frac{u - c_x}{f_x} z \\ y = \frac{v - c_y}{f_y} z \\ z = d \end{cases}, \quad (1)$$

其中: x, y, z 为点云三维坐标, d 为像素的深度值, u, v 为图像像素坐标。

对点云使用位姿估计算法继续处理,得到锁止机构相对于相机的位姿 T_r 。最后经过相机外参的转换,计算锁止机构在换电站空间中的绝对位姿:

$$T_o = T_c T_r. \quad (2)$$

T_o 用于引导换电机器人与电池包的对接。

但使用 RGBD 相机拍摄生成的点云不可避免会有噪声,其值约为 $\pm 1 \sim 3$ mm; 车辆经过较长时间的行驶后,其电池包锁止机构也可能不均匀地黏附泥泞和污渍,产生新的表面形貌噪声,其值约为 $\pm 2 \sim 4$ mm。而锁止机构尺寸较小,约为 $40 \text{ mm} \times 40 \text{ mm} \times 30 \text{ mm}$,上述噪声会对位姿估计产生较大的干扰,导致位姿估计算法的准确度下降,甚至可能得出错误的结果。为了解决上述问题,本文提出一种稳定的位姿估计方法。

3 基于点云分割的 6D 位姿估计算法

针对换电视觉定位稳定性的需求,提出基于点云分割的锁止机构 6D 位姿估计方法,通过引入点云分割神经网络解决点云配准算法中源点云与目标点云之间对应关系难以估计的问题。该锁止机构 6D 位姿估计方法由锁止机构目标检测,锁止机构点云预处理,锁止机构点云部分分割,锁止机构源点云与目标模型点云配准四部分组成。

3.1 RGB 图像目标检测

拍摄得到的底部电池包图像包含车底盘、换电站背景等信息,其中多余信息会严重干扰算法的处理结果。为顺利进行锁止机构的位姿估计,首先需要从图像中找出锁止机构的位置。实例分割是一种基于深度学习的图像分割技术,可在存在如灰尘污渍等干扰的图像中较为准确地得到目标物的分割结果。由于需要分割的锁止机构在图像中相对较小,且分割任务对于速度有一定的要求,因此,本文使用小目标检测性能良好、运行速度较快的 YOLOv5 网络^[7],并配合实例分割模块^[8]进行目标分割。

该网络使用多层残差卷积神经网络进行图像的特征提取,利用金字塔结构得到多尺度的图像特征,并经过包围框预测层与分割掩码预测层,

输出目标检测结果与实例分割掩膜。网络结构如图 2 所示。输出掩膜为每个图像像素对前景与背景的隶属概率 p 。当 $p < 0.5$ 时,像素隶属于背景;当 $p > 0.5$ 时,像素隶属于前景。选取所有 $p > 0.5$ 的像素,作为锁止机构目标的分割模板,同时根据网络的目标检测框确定目标所属类别。

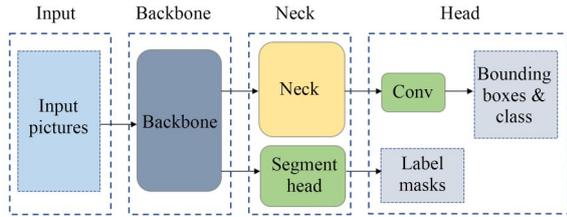


图 2 YOLOv5 网络结构

Fig. 2 Structure of YOLOv5 network

3.2 点云预处理

由神经网络分割得到电池包锁止机构的 RGBD 图像,通过相机的内参矩阵 K 将该部分深度图像素投影至三维空间而转换为点云 P_0 。 P_0 点云数目巨大,且由于相机精度的限制,灰尘、物体表面材料特性的影响,点云坐标 z 轴会包含一定的噪声,因此,后续步骤首先通过体素滤波^[9]进行点云降采样,再利用移动最小二乘算法^[10]尽可能去除点云的表面噪声。设计体素滤波器的体素方格边长为 $f_z = 2.0 \text{ mm}$,在锁止机构输入点云 P_0 中建立 3D 体素方格,计算每个体素内的点云的质心并作为输出点云,得到滤波后的点云 P_1 。 P_1 点云数量更少并最大限度地保留了点云精细结构,但仍然包含一定的 z 轴噪声并且还有部分偏差较大的离群点,因此,通过移动最小二乘方法进行点云 z 轴方向的平滑处理。在待拟合点的一个邻域上,拟合函数表示为:

$$f(x) = \mathbf{p}^T(x) \mathbf{a}(x), \quad (3)$$

其中: $\mathbf{a}(x) = [a_1(x), a_2(x), a_3(x)]$ 为待求的拟合系数, $\mathbf{p}(x) = [p_1(x), p_2(x), p_3(x)]$ 为基函数,对于所要拟合的点云 z 轴坐标,基函数 $\mathbf{p}(x) = [1, p_{1x}, p_{1y}]$, p_{1x}, p_{1y} 分别为点的 x, y 轴坐标。利用移动最小二乘方法得到:

$$\mathbf{a}(x) = \mathbf{A}^{-1}(x) \mathbf{B}(x) \mathbf{y}, \quad (4)$$

$$\mathbf{A}(x) = \sum_{i=1}^m \omega(x - x_i) \mathbf{p}^T(x_i) \mathbf{p}(x_i), \quad (5)$$

$$\mathbf{B}(x) = [\omega(x_1 - x_i) \mathbf{p}(x_1), \dots, \omega(x_m - x_i) \mathbf{p}(x_m)], \quad (6)$$

$$\mathbf{y} = [y(x_1), y(x_2), y(x_3)], \quad (7)$$

其中: $\omega(x - x_i)$ 为权函数,用来控制拟合的光滑度, y 为函数原值。经过体素滤波与移动最小二乘平滑的点云 P_2 如图 3 所示,可以看到,处理后点云更加精简,且曲面表面更加光滑,有益于后续步骤的处理。

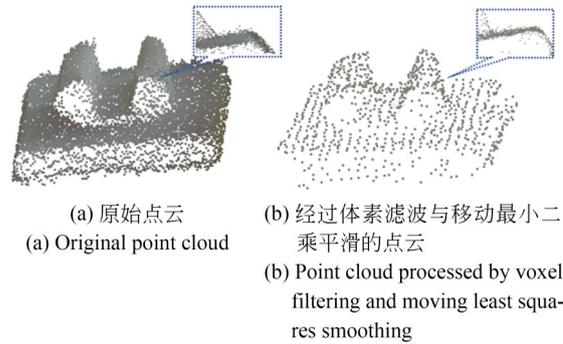


图 3 锁止机构点云预处理结果

Fig. 3 Pretreatment result of locking mechanism point cloud

3.3 点云分割

传统的全局点云配准方法如快速全局配准 (FGR)^[11] 使用快速点特征直方图 (Fast Point Feature Histogram, FPFH) 局部 3D 描述子^[12-13] 匹配的方法来寻找目标点云与模板点云之间的对应关系,从而由对应关系计算得到目标点云的位姿估计。但由于 FPFH 较为依赖点邻域的局部形貌特征,而噪声会对 FPFH 的计算造成较大的干扰,导致误匹配;本文的锁止机构中存在着具有相同特征的不同表面,其 FPFH 描述子相似,该部分描述子后续步骤中也会发生误匹配。以上因素会较大程度地减小描述子匹配的准确率,降低后续随机抽样一致性 (Random Sample Consensus, RANSAC) 步骤的效率和精度。针对这一问题,本文提出为 FPFH 描述子引入全局的语义特征。锁止机构点云具有明显的形状特征和分明的几何轮廓,其全局几何语义特征相较于局部的几何特征更为可靠稳定,可改善 FPFH 的稳定性。因此,利用深度学习网络较为强大的特征感知能力,提取点云的全局语义特征,弥补 FPFH 仅考虑局部特征的缺点。

首先,利用点云分割网络对锁止机构点云进行语义分割,对锁止机构点云的不同部位进行分类,赋予点云全局的语义信息。由于锁止机构点云包含噪声和离群点,需要网络能够关联全局特征信息,具备较好的稳定性,因此,采用包含 Encoder-Decoder 结构的点云分割^[14]网络,网络结构如图 4 所示。其中,Encoder 层分为 3 个采样-聚合层,每层使用 fps 采样算法进行下采样,并聚合采样点的领域特征,由 PointNet 进行运算^[15];Decoder 层分为两个插值层,每层都对 Encoder 层的输出特征进行邻域插值;最后,通过 Shared-MLP 层与 LogSoftmax 函数输出点云的部件分割标签。通过逐层的特征聚合和逐层的插值上采样,该网络可以更好地捕捉上下文信息,在存在干扰的情况下输出准确的结果。

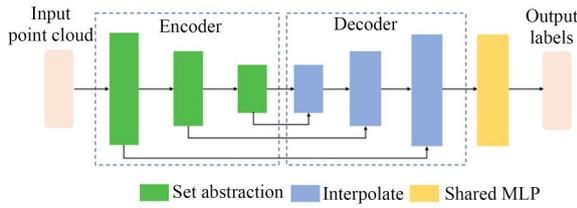


图 4 点云分割网络结构

Fig. 4 Structure of point cloud segmentation network

选择锁止装置几何特征明显的部分作为分割对象,如图 5 所示(彩图见期刊电子版),设置锁头顶部(黄)、锁头中部(红),锁头底部(绿)和锁头支撑面(蓝)共 4 个分割部分,对每个部分使用 One-Hot 标签进行标记。

将 P_2 随机采样至 4 096 个点,并进行归一化,输入点云分割网络,令网络输出 One-Hot 标签最大值对应位为输出标签。

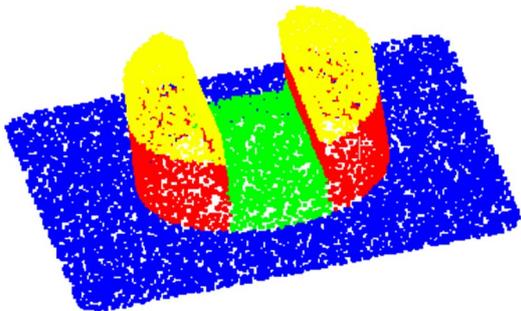


图 5 锁止机构分割示意图

Fig. 5 Schematic diagram of locking mechanism division segmentation

3.4 点云配准

采用 FPFH 3D 描述子寻找目标点云与模板点云的对应关系。FPFH 是一种姿态不变的实时 3D 描述子,主要描述点云的局部几何特征。某一点的 FPFH 为特征直方图的加权和,即:

$$FPFH(p) = SPFH(p) + \frac{1}{k} \sum_{i=1}^k \frac{1}{\omega_i} * SPFH(p_i), \quad (8)$$

其中:SPFH 为简化的特征直方图, p_i 为 p 的邻域点, ω_i 为 p_i 到 p 的欧式距离。SPFH 的计算流程如下:

(1) 建立 p 与各邻域点 p_i 的局部坐标系正交坐标轴 (u, v, ω) , 其中 n_s 为 p 处的法向量:

$$u = n_s, \quad (9)$$

$$v = u \times \frac{(p - p_i)}{\|p - p_i\|}, \quad (10)$$

$$\omega = u \times v. \quad (11)$$

(2) 根据局部坐标系与法向量建立特征:

$$\alpha = v \cdot n_i, \quad (12)$$

$$\phi = u \cdot \frac{(p - p_i)}{\|p - p_i\|}, \quad (13)$$

$$\theta = \arctan(\omega \cdot n_i, u \cdot n_i), \quad (14)$$

$$f = \langle \alpha, \phi, \theta \rangle. \quad (15)$$

(3) 对特征 f 进行统计,计算其对应的直方图。具体而言,将特征向量中的每一个角度化为 11 等分,共有 33 维,将特征放入对应的维度中,即为相应的 SPFH 特征。

然后计算得到 FPFH 特征,结合上一小节的内容,将点云语义分割结果融合进 FPFH 特征直方图中。改进后的 FPFH 特征为:

$$FPFH^* = \langle FPFH, \mu L \rangle, \quad (16)$$

其中: $L \in \{1, 2, 3, 4\}$ 为点云语义分割的标签, μ 为超参数。该特征描述子可确保点云在后续的匹配过程中能够同时结合局部特征与全局特征,使具有相同语义标签的点匹配,从而减少误匹配,提升后续点云配准步骤的精度与速度。

获得 FPFH 特征后,记由 3D 相机拍摄且经过前述步骤处理的目标点云为 P_s , 其 FPFH 特征记为 $F(P_s)$, 由 CAD 模型采样的模板点云为 Q_m , 其 FPFH 特征为 $F(Q_m)$ 。点云配准的目的就是求得 Q_m 至 P_s 的姿态变换矩阵 T_{sm} 。

首先使用 KD 树进行特征点的匹配,对于每

一个 $p_s \in P_s$, 寻找 $F(Q_m)$ 中离 $F(p_s)$ 最近的点 $F(q_m)$, 将 q_m 与 p_s 作为一对匹配点。同理对于每一个 $q_m \in Q_m$, 可获得另一组匹配点。上述得到的匹配点, 再通过互换性测试与三角测试, 筛选得到匹配关系。

使用 RANSAC^[12,16] 方法获得位姿估计 T_{sm} , 其流程如下:

(1) 从上述的匹配关系中, 随机选择 3 对匹配点, 进行 SVD 奇值分解, 计算变换位姿 T ;

(2) 根据 T 计算两幅点云之间的误差, $l = \sum_{i=1}^n h(\|p_i - Tq_i\|)$, 其中 $h(\bullet)$ 为 Huber Loss^[17];

(3) 重复步骤(1)和(2), 直到 l 小于阈值 ξ 或迭代次数 $n \geq 100$;

(4) 取所有迭代中 l 最小时的位姿 T 作为输出 T_{sm} 。

最后, 使用迭代最近点方法 (Iteration Closest Point, ICP)^[18] 进行估计位姿 T_{sm} 的精配准, 得到更为精准的 T_{sm} 。

4 实验与结果分析

电池包实物图如图 6 所示, 其底面的锁止机构位于电池包的两边。



图 6 电池包实物图

Fig. 6 Photo of battery pack

3D 相机选用知像光电的 SURFACE HD50 双目结构光相机, 其有效工作距离为 250~750 mm, RGB 相机的分辨率可达 $1\,920 \times 1\,080$, 深度图的分辨率为 960×680 , 精度最高可达 ± 0.15 mm。相机置于换电作业空间的底部, 用于拍摄

上方车辆底部的电池包锁止机构, 相机与电池包的距离为 0.55 m。

为了进行室内实验, 并更好地模拟各种姿态下的电池包, 建立电池包的 CAD 模型, 对电池包中包含锁止机构的部分进行了 3D 打印, 如图 7 所示。3D 打印模型摆放在距离相机 0.4~0.7 m 处, 实验中在一定范围内调整它相对于相机的姿态。



图 7 电池包的 3D 打印模型

Fig. 7 Three dimensional printing model of battery pack

实验首先需要采集电池包照片, 制作数据集以训练 YOLOv5 网络, 并根据锁止机构标签制作锁止机构分割数据集, 训练点云分割网络, 然后使用训练后的网络进行位姿估计。实验流程如图 8 所示。

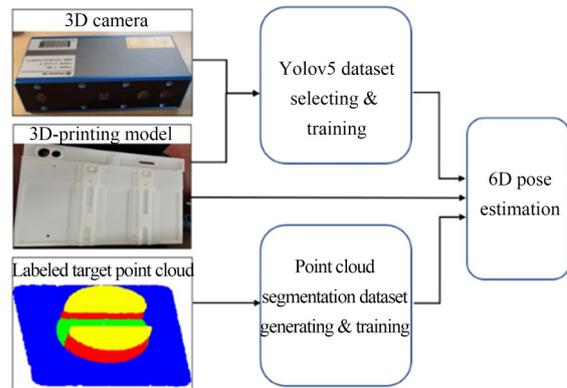


图 8 六维位姿估计实验流程

Fig. 8 Flow chart of 6D pose estimation experiment

4.1 网络训练

点云分割网络输出标签包含锁头顶部、锁头中部, 锁头底部和锁头支撑面 4 类。由于难以直

接标注 3D 相机拍摄得到的点云,因此采用由 CAD 模型仿真的方式来生成用于网络训练的点云数据集。首先,由锁止机构 CAD 模型均匀采样得到模型点云,然后采用 CloudCompare 软件进行点云的标签标注,按照 3.3 节所述将锁止机构点云分为锁头顶部、锁头中部、锁头底部和锁头支撑面 4 个部分进行标注,其标签分别为 $\{1, 2, 3, 4\}$ 。之后,为点云添加 $\sigma = 1.5 \text{ mm}$ 的随机高斯噪声,同时进行随机位姿变换,共生产 600 个具有不同姿态的点云。为了尽可能模拟从某一视角拍摄而得的点云,对每个数据集中的点云,根据相机内参 K 将点云投影至深度图并消去重叠部分中的远点,保留近点,再根据 K 将深度图重新投影至三维空间,获得具有视角遮挡的点云。生成数据集后,将数据集按照 5:1 的比例分为训练集与测试集。

为了尽可能提升模型训练的效果,采用迁移训练。首先,使用 ShapeNet 数据对模型进行预训练,再使用自制的数据集对模型进行微调。使用交叉熵作为损失函数,其学习率设置为 0.01,采用 Adam 优化器,共训练 50 个 epoch。训练结束后锁止机构的点云分割结果如图 9 所示。

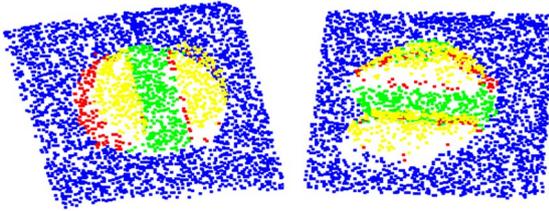


图 9 锁止机构的点云分割结果

Fig. 9 Point cloud segmentation result of locking mechanism

模型训练完成后,通过检验,其标签预测准确率可达 95%,分割交并比 (Intersection Over Union, IoU) 达 90%,满足使用要求。其结果如图 10 所示。

4.2 位姿估计实验

实验中,使用相机拍摄多组电池包 RGBD 照片,用于验证基于点云分割的 6D 位姿估计方法。在位姿估计时,首先使用 YOLOv5 网络对 RGB 图进行目标检测与分割,并通过对齐关系获得相应的深度图分割结果,将它投影为点云,对点云进

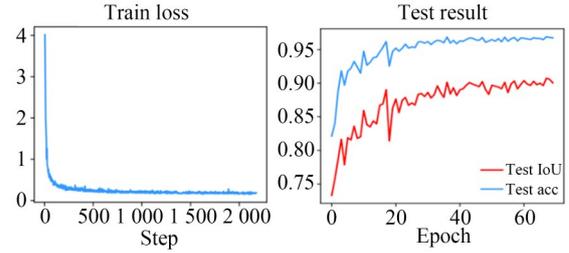


图 10 网络训练损失与测试结果

Fig. 10 Network training loss and test results

行预处理。然后,利用点云分割结果计算改进的 FPFH 描述子,进行点云初配准,最后使用 ICP 方法进行点云精配准,得到估计位姿。

为了获取拍摄得到的点云位姿的真值,首先使用 YOLOv5 获取分割后的锁止机构点云,再使用 CloudCompare 软件的手动配准功能,利用人工放置匹配点的方式实现拍摄的目标点云与模型点云之间的精准配准,获得准确的位姿变换矩阵 T_{gt} 。

配准精度用角度误差 $\Delta\theta$ 、距离误差 Δt 和均方根误差 (Root Mean Square Error, RMSE) 衡量,即:

$$\Delta\theta = \frac{1}{2} \arccos [\text{trace}(\mathbf{R}_{gt}^T \mathbf{R}) - 1], \quad (17)$$

$$\Delta t = \|\mathbf{t}_{gt} - \mathbf{t}\|_2, \quad (18)$$

$$RMSE_{\Delta d} = \left(\frac{1}{n} \sum_{i=1}^n \|p_i - q_i\|_2^2 \right)^{\frac{1}{2}}, \quad (19)$$

其中: \mathbf{R}_{gt} , \mathbf{t}_{gt} 为真实的旋转矩阵与位移向量, \mathbf{R} , \mathbf{t} 为旋转矩阵与位移向量的估计值, q_i 为模板点云中的一点, p_i 为目标点云中距离 q_i 最近的点。

为验证基于点云分割的改进 FPFH 点云配准算法 (记为 D-SAC-IA) 的有效性,分别使用 ICP、正态分布变换配准 (NDT)、FGR、采样一致性初始配准算法 (SAC-IA) 等算法与提出的 D-SAC-IA 算法对同一姿态下相机拍摄的锁止机构进行对比实验,共拍摄 10 组并计算相应的 $\Delta\theta$, Δt 和 RMSE 平均值,结果如表 1 所示。

为验证算法对由泥渍等因素引起的表面形貌误差的稳定性,在上述的锁止机构点云中额外添加标准差为 4 mm 的高斯噪声和大小为 2 mm 的表面形貌噪声,实验结果如表 2 所示。

表 1 中, D-SAC-IA 的 $\Delta\theta = 2.86^\circ$, $\Delta t = 1.4 \text{ mm}$, $RMSE = 1.6 \text{ mm}$; 表 2 中, D-SAC-IA 的

表 1 算法对比实验结果

Tab. 1 Algorithm comparison results

算 法	$\Delta\theta/(\circ)$	$\Delta t/\text{mm}$	RMSE/mm
ICP	40.23	2.8	5.5
NDT	45.54	2.5	6.1
FGR	16.59	2.1	2.2
SAC-IA	5.67	1.5	1.8
D-SAC-IA	2.86	1.4	1.6

$\Delta\theta = 2.51^\circ$, $\Delta t = 1.9 \text{ mm}$, $\text{RSME} = 1.8 \text{ mm}$ 。在两次实验中, D-SAC-IA 均精度最高, 证明了该方法的有效性。由表 2 可知, 即使添加较大噪声, D-SAC-IA 仍能较为精确地估计锁止机构的位姿, 表明其点云分割网络可更好地避免误匹配, 使得算法对噪声有更好的稳定性, 因此精度更高。

表 2 算法对比实验结果(含噪声)

Tab. 2 Algorithm comparison results(with noise)

算 法	$\Delta\theta/(\circ)$	$\Delta t/\text{mm}$	RMSE/mm
ICP	49.24	4.6	6.2
NDT	44.85	5.5	5.9
FGR	25.50	4.7	3.1
SAC-IA	10.67	2.5	2.3
D-SAC-IA	5.51	1.9	1.8

最后进行包含 ICP 精配准步骤的完整位姿估计实验, 以验证整个方法的有效性。将 3D 打印模型置于 10 个不同的姿态下, 每个姿态下进行 5 次位姿估计实验并计算平均值, 结果如表 3 和图 11 所示。

表 3 位姿估计实验结果

Tab. 3 Results of pose estimation experiment

序号	$\Delta\theta/(\circ)$	$\Delta t/\text{mm}$	RMSE/mm
1	1.221	1.4	1.29
2	1.364	1.4	1.34
3	2.358	1.6	1.65
4	0.782	0.8	1.13
5	1.266	1.4	1.31
6	1.590	1.3	1.32
7	0.962	0.9	1.19
8	0.895	1.0	1.15
9	1.365	1.1	1.26
10	1.244	1.0	1.25

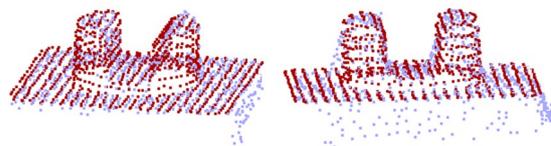


图 11 位姿估计匹配结果

Fig. 11 Pose estimation matching result

由表 3 可知, 位姿估计的平均角度误差达到 1.30° , 平均位移误差达到 1.2 mm , RMSE 平均值达到 1.3 mm 。

为了更进一步验证所提方法的有效性, 采用电池包实物进行包含 ICP 精配准的完整位姿估计实验。从 5 个不同的视角拍摄电池包的锁止机构, 在每个视角下进行 5 次位姿估计实验并计算实验结果的平均值, 结果如表 4 所示。

表 4 位姿估计实验结果

Tab. 4 Results of pose estimation experiment

序号	$\Delta\theta/(\circ)$	$\Delta t/\text{mm}$	RMSE/mm
1	1.813	1.5	1.44
2	2.364	1.3	1.59
3	2.018	1.3	1.48
4	1.797	1.3	1.38
5	1.510	1.4	1.31

由表 4 可知, 位姿估计的平均角度误差达到 1.90° , 平均位移误差达到 1.4 mm , RMSE 平均值达到 1.5 mm 。换电站换电作业对于对锁止机构定位精度的要求为: 重合误差 $\leq 3 \text{ mm}$ 。因此, D-SAC-IA 算法对于电池包锁止机构的位姿估计具有较高的精度和稳定性, 可以满足电动车换电过程中的定位要求。

5 结 论

针对换电电池包的对接定位问题, 本文提出了一种基于点云分割的电池包锁止机构 6D 位姿估计方法。该方法使用 YOLOv5 从场景中分割除锁止机构的点云, 并使用体素滤波与移动最小二乘拟合进行点云的滤波与平滑; 通过引入点云分割网络为 FPFH 特征加入全局语义特征, 弥补 FPFH 只有点云局部特征的缺陷, 并根据该特征进行 RANSAC 点云配准, 估计锁止机构点云的

位姿;最后使用ICP算法校正位姿估计结果。实验结果表明,基于点云分割的锁止机构6D位姿估计算法精度较高,其位姿估计的平均角度误差

为 1.90° ,平均位移误差为1.4 mm, RMSE平均值为1.5 mm,能够实现换电站电动车的快速稳定换电。

参考文献:

- [1] 周芮,刘延芳,齐乃明,等. 面向空间应用的视觉位姿估计技术综述[J]. 光学精密工程, 2022, 30(20): 2538-2553.
ZHOU R, LIU Y F, QI N M, *et al.* Overview of visual pose estimation methods for space missions [J]. *Optics and Precision Engineering*, 2022, 30(20): 2538-2553. (in Chinese)
- [2] 杨宁,申景诗,张建德,等. 基于立体视觉的空间非合作航天器相对位姿自主测量[J]. 光学精密工程, 2017, 25(5):1331-1339.
YANG N, SHEN J SH, ZHANG J D, *et al.* Autonomous measurement of relative attitude and position for spatial non-cooperative spacecraft based on stereo vision[J]. *Optics and Precision Engineering*, 2017, 25(5):1331-1339. (in Chinese)
- [3] 李莹,叶培建,彭兢,等. 基于火星探测的人工标志识别定位[J]. 光学精密工程, 2015, 23(2): 566-572.
LI Y, YE P J, PENG J, *et al.* Artificial target recognition and location based on Mars exploration[J]. *Optics and Precision Engineering*, 2015, 23(2): 566-572. (in Chinese)
- [4] 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*. June 13-18, 2010, San Francisco, CA, USA. IEEE, 2010: 998-1005.
- [5] HINTERSTOISSER S, LEPETIT V, ILIC S, *et al.* Model Based Training, Detection and Pose Estimation of Texture-less 3D Objects in Heavily Cluttered Scenes [M]. *Computer Vision - ACCV 2012*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013: 548-562.
- [6] OGUN P S, USMAN Z, DHARMARAJ K, *et al.* 3D vision assisted flexible robotic assembly of machine components [C]. *Proc SPIE 9875, Eighth International Conference on Machine Vision (ICMV 2015)*, 2015, 9875: 342-348.
- [7] BOCHKOVSKIY A, WANG C Y, LIAO H. YO-LOv4: optimal speed and accuracy of object detection [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [8] WU D, LIAO M W, ZHANG W T, *et al.* YO-LOP: You only look once for panoptic driving perception [J]. *Machine Intelligence Research*, 2022, 19(6): 550-562.
- [9] 艾卫涛. 三维激光扫描点云去噪及在建模中的应用 [D]. 西安: 西安科技大学, 2017.
AI W T. *Denoising of 3D Laser Scanning Point Cloud and Its Application In Modeling* [D]. Xi'an: Xi'an University of Science and Technology, 2017. (in Chinese)
- [10] 曾清红,卢德唐. 基于移动最小二乘法的曲线曲面拟合 [J]. 工程图学学报, 2004(1): 84-89.
ZENG Q H, LU D. Curve and surface fitting based on moving least-squares methods [J]. *Journal of Engineering Graphics*, 2004(1): 84-89. (in Chinese)
- [11] ZHOU Q Y, PARK J, KOLTUN V. *Fast Global Registration* [M]. *Computer Vision - ECCV 2016*. Cham: Springer International Publishing, 2016: 766-782.
- [12] RUSU R B, BLODOW N, MARTON Z C, *et al.* Aligning point cloud views using persistent feature histograms [C]. 2008 *IEEE/RSJ International Conference on Intelligent Robots and Systems*. September 22-26, 2008, Nice, France. IEEE, 2008: 3384-3391.
- [13] RUSU R B, BLODOW N, BEETZ M. Fast point feature histograms (FPFH) for 3D registration [C]. 2009 *IEEE International Conference on Robotics and Automation*. May 12-17, 2009, Kobe, Japan. IEEE, 2009: 3212-3217.
- [14] QIC R, YIL, SUH, *et al.* PointNet++: Deep Hierarchical Feature Learning on Point Sets in A Metric Space [EB/OL]. 2017: *arXiv*: 1706.02413. <https://arxiv.org/abs/1706.02413.pdf>.
- [15] CHARLES R Q, HAO S, MO K C, *et al.* PointNet: deep learning on point sets for 3D classifica-

- tion and segmentation[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 77-85.
- [16] FISCHLER M A, BOLLES R C. *Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography* [M]. Readings in Computer Vision. Amsterdam: Elsevier, 1987: 726-740.
- [17] HUBER P J. Robust estimation of a location parameter [J]. *The Annals of Mathematical Statistics*, 1964, 35(1): 73-101.
- [18] BESL P J, MCKAY N D. A method for registration of 3-D shapes[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, 14(2): 239-256.

作者简介:



王立辉(1979—),男,教授,博导,主要从事无人系统自主导航、智能算法等方面的应用研究。E-mail: wlhseu@163.com



苏余足威(1998—),男,硕士研究生,主要研究方向为三维视觉。E-mail: wangyi_sy1256@163.com