

## ◆交通场景◆

## 基于无需校准激光雷达曲线与 RGB-D 图像的数据融合方法\*

徐正丽<sup>1</sup>, 肖素芳<sup>1\*</sup>, 杨明浩<sup>2</sup>

(1. 桂林电子科技大学, 广西桂林 541004; 2. 中国科学院自动化研究所, 北京 100190)

**摘要:**激光雷达和 RGB-D 摄像机是在各种机器人导航任务中广泛使用的两种传感器。尽管通过图像校准能够降低传感器中存在的噪声,但由于外部复杂环境的影响,融合数据中仍会有空洞和毛刺。为此,本文提出一种无需校准激光雷达曲线与 RGB-D 图像的数据融合方法。该方法利用激光雷达曲线和 RGB-D 图像数据在格式上存在显著差异但在深度信息上紧密相对的特点,通过时间配准和相关性分析,将激光雷达曲线在高度上与 RGB-D 图像配准。同时,即使在不进行校准的情况下,该方法也能将 RGB-D 图像在高度上的相应线条自动配准到激光雷达曲线在宽度上的范围。使用室内开放式机器人导航数据集 Robot@Home 中的数据对本文方法进行评估,结果表明,该方法可同时为激光雷达曲线和 RGB-D 图像的原始数据去噪。该方法已在真实的导航环境中得到了验证,可扩展应用于机器人导航的更精确的二维地图构建中。

**关键词:**激光雷达;深度图像;数据融合;二维地图重建;机器人导航

中图分类号: TP389 文献标识码: A 文章编号: 1005-9164(2024)05-1038-11

DOI: 10.13656/j.cnki.gxkx.20241127.019

移动智能体的自主漫游是指智能体在动态场景感知的基础上进行自主建图和定位,从而实现自主导航的功能<sup>[1-4]</sup>。移动智能体利用该技术,以传感器对周围环境采集的数据为基础,从中分析和感知环境,进行实时定位的同时创建地图,把自身在地图的位置与环境进行关联。因此,如何使用传感器获取的数据对环境进行感知是移动智能体实现自主漫游的关键。越来越多的研究开始将视觉传感器、激光雷

达、惯性传感单元、超宽带基线定位装置等传感器应用于移动智能体的自主漫游中。其中,使用视觉传感器的自主漫游方案最为丰富,该方案在构建可视化地图上具备优势,在室内环境中其有较好的泛用性<sup>[5]</sup>;使用激光雷达的自主漫游方案实用性高,测量精度稳定,使用场景较为广泛<sup>[6]</sup>;惯性传感单元价格经济,是测量移动智能体短时间自身运动状态的良好传感器,但由于其长时间测量误差较大,也无法采集环境信

收稿日期: 2023-11-08

修回日期: 2024-01-12

\* 国家自然科学基金项目(71463010, 22180155466), 广西科技计划项目(2021GXNSFBA220048, 桂科 AB21220038)和桂林科技计划项目(20220115-1, 20230110-1)资助。

## 【第一作者简介】

徐正丽(1982—),女,博士,副教授,主要从事数据统计分析、数字图像处理等交叉学科研究。

## 【\*\*通信作者简介】

肖素芳(1990—),女,博士,副教授,主要从事数据统计分析研究, E-mail: xiaosufang2011@163.com。

## 【引用本文】

徐正丽,肖素芳,杨明浩. 基于无需校准激光雷达曲线与 RGB-D 图像的数据融合方法[J]. 广西科学, 2024, 31(5): 1038-1048.

XU Z L, XIAO S F, YANG M H. A Fusion Method Based on Uncalibrated LiDAR Curves and RGB-D Images [J]. Guangxi Sciences, 2024, 31(5): 1038-1048.

息,因此大多和其他传感器融合使用,作为自主漫游方案的有效补充<sup>[7]</sup>;超宽带基线定位装置虽然定位精度很高,但只适合在室内特定场景使用<sup>[8]</sup>。

虽然使用多种传感器在机器人导航任务的自主定位<sup>[9-11]</sup>、自主建图<sup>[12,13]</sup>等方面取得了很好的效果,但是考虑到不同传感器的融合特性,激光雷达和 RGB-D 摄像机因其相对较高的性价比,仍然是机器人在动态场景感知的首要选择。RGB-D 摄像机的深度图像和激光雷达提供的信息视图主要有两处不同:①深度信息的视角和范围各不相同。激光雷达提供了水平方向深度信息,而 RGB-D 摄像机则提供了水平和垂直视角的深度信息<sup>[12]</sup>。②深度信息中都存在噪声,但产生噪声的原因各不相同。激光雷达深度信息中的噪声通常是由前方可能存在透明玻璃、激光投射的强光吸收和反射等原因造成的;而大部分 RGB-D 摄像机是通过红外反射来获取深度信息,会因为背光太强、远处或近处的物体超出了反射范围,以及多次红外反射造成测量误差,导致 RGB-D 图像中存在空洞和斑点毛刺<sup>[14]</sup>。

目前,研究人员将这两种传感器的深度信息结合,用于机器人的地图绘制<sup>[15]</sup>和导航定位<sup>[16,17]</sup>。然而,由于它们处理信号的原理不同,目前仍缺乏有效的融合方法或深度融合策略。本文提出一种将激光雷达曲线与 RGB-D 图像的数据融合方法,利用激光雷达曲线和 RGB-D 图像数据在格式上存在显著差异,但在深度信息上却紧密相对的特点,通过时间配准和相关性分析,发现激光雷达曲线能够在高度上与 RGB-D 图像配准。同时,本文提出的方法即使不需要校准,RGB-D 图像在高度上的相应线条也能自动配准到激光雷达曲线在宽度上的范围。

## 1 相关工作

激光雷达可提供视线前方的水平深度信息,同时得到二维地图中不可获取的导航信息,用于实时闭环<sup>[18]</sup>、位置识别<sup>[19,20]</sup>、障碍物距离估计<sup>[21,22]</sup>等任务。RGB-D 摄像机获取的深度信息与二维激光雷达不同,它提供了丰富的三维信息,可以实现全视角观测,支持障碍物分割<sup>[23]</sup>、地图绘制<sup>[9,13]</sup>、导航<sup>[9,10]</sup>、物体检测<sup>[24]</sup>等任务。然而,由于三维点云的匹配处理<sup>[10,15,17]</sup>以及 RGB-D 图像的三维重建<sup>[9,11,25-28]</sup>都非常耗时,因此利用 RGB-D 图像绘制地图和导航相对复杂。

目前,研究人员将激光雷达曲线和 RGB-D 图像

中的深度信息结合起来用于地图导航。例如,Winterhalter 等<sup>[16]</sup>采用粒子滤波技术估计平面图和真实世界数据之间的 6 个自由度姿态;Chong 等<sup>[22]</sup>基于合成激光雷达测量和里程计信息,采用蒙特卡罗定位方案进行车辆位置估计;Song 等<sup>[17]</sup>使用基于鲁棒特征的低级视觉-激光雷达融合算法,通过视觉跟踪、结构光传感器的深度信息进行相对定位。这类技术通常被称为地图扫描,主要用于将机器人眼睛看到的范围和位置与完整的激光雷达曲线地图或三维重建环境相匹配。不过,这些方法的目的是通过将激光信息与视觉映射相匹配,减少机器人导航的时间成本并保持定位精度。至于如何融合这两种深度信息,则研究较少。

尽管激光雷达和 RGB-D 摄像机都能提供深度信息,但它们的成像原理和成像形式却大相径庭。通常,激光雷达的深度信息是由激光投射和反射信号形成的,而 RGB-D 摄像机的深度信息主要来自环境中的红外反射<sup>[14]</sup>。在激光雷达和 RGB-D 摄像机的深度信息中,空洞和毛刺等噪声在位置和范围上是不同的。当激光雷达和 RGB-D 摄像机放置在相似的高度时<sup>[15,17]</sup>,有助于将两个传感器的深度信息进行匹配处理。然而,在居家服务机器人、智能送餐机器人、养老陪伴机器人等使用场景上,这两种传感器的放置高度完全不同<sup>[12]</sup>。在这种情况下,不同高度的深度信息会使它们在视图中的观测差异较大,因此难以对激光雷达和 RGB-D 摄像机进行校准。因此如何采用一种自动方法来融合激光雷达的二维数据和 RGB-D 摄像机的深度信息是研究工作中面临的难点问题。

## 2 基于无需校准激光雷达曲线与 RGB-D 图像的数据融合方法

本文提出的基于无需校准激光雷达曲线与 RGB-D 图像的数据融合方法主要包括 3 个步骤:首先,考虑到激光雷达和 RGB-D 摄像机之间工作频率的差异,采用时间维度来对齐激光雷达的曲线和 RGB-D 摄像机的深度图像,使得激光雷达曲线采样后频率向 RGB-D 图像频率对齐。其次,针对激光雷达曲线的宽度范围和 RGB-D 图像高度范围不统一的问题,采用皮尔逊相关系数、归一化余弦距离和欧拉距离来建立相似度函数,通过沿  $x$  轴坐标将 RGB-D 图像的每一行与激光雷达数据序列对齐;反之,根据它们的  $z$  值沿  $y$  轴坐标将激光雷达曲线与 RGB-D

图像对齐,对于 RGB-D 图像, $x$  轴是指图像的横坐标, $y$  轴是指图像的纵坐标, $z$  轴表示深度信息,由此得到激光雷达曲线在某段中和 RGB-D 图像在某高度处的相似度,并获得它们之间的最佳匹配度。最

后,再通过配准和去噪的方法进行水平和垂直初始对齐,消除融合数据的空洞和毛刺。具体流程如图 1 所示。

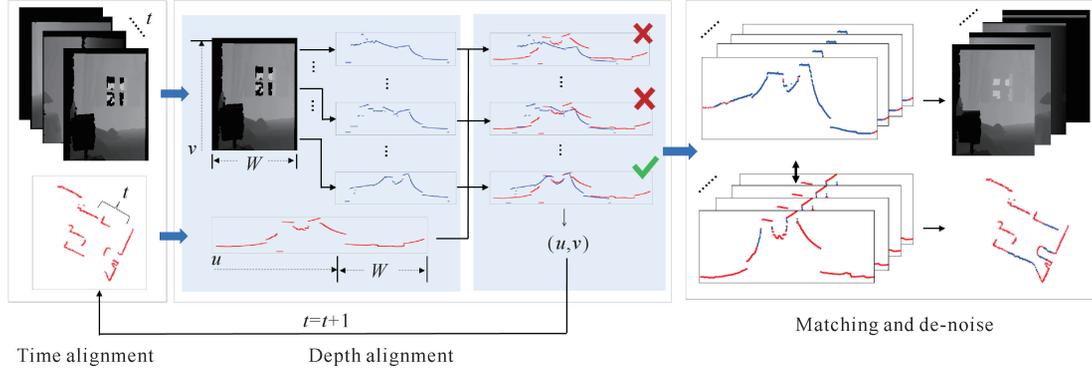


图 1 本文提出的融合方法的流程

Fig. 1 Workflow of the fusion method proposed in this study

## 2.1 时间维度对齐

假设  $D^i$  ( $1 \leq i \leq M$ ) 和  $L^j$  ( $1 \leq j \leq N$ ) 分别为 RGB-D 摄像机拍摄的第  $i$  张深度图像和激光雷达获取的第  $j$  张曲线。考虑到两个传感器之间的工作频率差异,采用  $f(i)/g(j) = t$  来对齐  $D^i$  和  $L^j$ ,其中  $f(\cdot)$  和  $g(\cdot)$  是 RGB-D 图像和激光雷达曲线的抽样函数, $f(i)$  和  $g(j)$  使得  $D^i$  和  $L^j$  在时间序列上非常相邻。本文激光雷达曲线的采样频率是 1 000 fps,而 RGB-D 图像的采样频率是 24 fps。考虑到激光雷达和 RGB-D 摄像机之间工作频率的差异,这里取  $t = 1\,000/24$ ,使得激光雷达曲线采样后的频率与深度图像的频率对齐。

## 2.2 激光雷达曲线的宽度范围与 RGB-D 图像的高度范围对齐

假设  $W$  和  $H$  分别是  $D^i$  的宽度和高度, $\Psi$  是  $L^j$  的宽度, $D_{h,w}^i(x_{h,w}^i, z_{h,w}^i)$  是  $D^i$  高度  $h$  沿  $x$  轴方向的二维点序列,其中  $1 \leq w \leq W$ ;  $L_{\omega}^j(x_{\omega}^j, z_{\omega}^j)$  是激光雷达曲线  $L^j$  的二维点序列,其中  $1 \leq \omega \leq \Psi$ 。假设  $\hat{L}_{u,\omega}^j$  ( $\omega \leq u \leq \omega + W, 1 \leq \omega \leq \Psi$ ) 是从起点  $u$  开始从  $L_{\omega}^j$  中选取的  $W$  个点。通常激光雷达曲线的宽度大于等于 RGB-D 图像的宽度,即  $\Psi \geq W$ 。深度配准的目标就是找到适当的激光雷达曲线与 RGB-D 图像中配准的目标点  $(u, v)$ 。研究人员通常采用距离来衡量一条曲线(点集)和另一条曲线(点集)的相似度。在衡量和评估两条曲线相似度的方法中,皮尔逊相关系数、归一化余弦距离和欧拉距离是最为经典和常见的 3 个函数,被广泛用于视觉计算和模式识别

中<sup>[21-24]</sup>。因此,本文采用皮尔逊相关系数、归一化余弦距离和欧拉距离来建立函数  $\text{sim}(\cdot)$  获得激光雷达曲线在某段中和 RGB-D 图像在某高度处的相似度,并获得它们间的最佳匹配度。本文通过最大化函数  $\text{sim}(\cdot)$  来获取最优的目标点  $(u, v)$ ,如公式(1)所示:

$$u, v = \max[\arg \min_{(u,v)} \sum_{v=1}^H \sum_{u=1}^{\Psi-W} \|\text{sim}(D_{v,w}^t - \hat{L}_{u,\omega}^t)\|], \quad (1)$$

其中,  $1 \leq u \leq \Psi - W, 1 \leq v \leq H, 0 \leq \text{sim}(\cdot) \leq 1$ 。

$D_{v,w}^t$  和  $\hat{L}_{u,\omega}^t$  中都存在噪声和毛刺,函数  $\text{sim}(\cdot)$  对于消除它们的噪声和毛刺,提升 RGB-D 图像和激光雷达曲线的表达有重要意义。为简化起见,将  $D_{v,w}^t$  和  $\hat{L}_{u,\omega}^t$  分别表示为  $D_v$  和  $\hat{L}_u$ , 函数  $\text{sim}(\cdot)$  的计算方式有 3 种,如公式(2)–(4)所示。

$$\begin{aligned} \text{PCC}(D_v, \hat{L}_u) &= \\ \frac{\text{Cov}(D_v, \hat{L}_u)}{\sqrt{\text{Var}(D_v) \text{Var}(\hat{L}_u)}} &= \\ \frac{E(D_v \hat{L}_u) - E(D_v) E(\hat{L}_u)}{\sqrt{E(D_v^2) - E^2(D_v)} \sqrt{E(\hat{L}_u^2) - E^2(\hat{L}_u)}}, \end{aligned} \quad (2)$$

$$\begin{aligned} \text{InvCos}(D_v, \hat{L}_u) &= \\ 1 - \text{Cos}(D_v, \hat{L}_u) &= \\ 1 - \frac{D_v \cdot \hat{L}_u}{\|D_v\| \|\hat{L}_u\|}, \end{aligned} \quad (3)$$

$$\text{InvEuler}(D_v, \hat{L}_u) = 1 - \text{Euler}(D_v, \hat{L}_u) = 1 - \frac{\|D_v - \hat{L}_u\|}{\max(\|D_v\|, \|\hat{L}_u\|)} \quad (4)$$

## 2.3 配准和去噪

### 2.3.1 配准

利用公式(1)获得  $u$  和  $v$  的值后,就能在激光雷达曲线上获得适当的  $x$  轴坐标( $u$ )的融合位置,在 RGB-D 图像上获得适当的  $y$  轴坐标( $v$ )的融合位置。尽管激光雷达曲线和 RGB-D 图像存在空洞和毛刺,但  $\hat{L}_{u,w}^t$  和  $D_{v,w}^t$  形状相似,因此它们共享相同的正面视图。常用于点集匹配和配准的方法有迭代最近点(Iterative Closest Point, ICP)<sup>[29,30]</sup>、粒子滤波<sup>[31]</sup>等。本文将随机样本共识(Random Sample Consensus, RANSAC)算法<sup>[32]</sup>和同源性<sup>[33]</sup>结合起来,以获得  $\hat{L}_{u,w}$  和  $D_{v,w}$  之间稳健的映射关系,并确定其形状上的噪声点。

在平面上相似的两个点集可通过同构法进行变换<sup>[33]</sup>,公式(5)为两个点集使用同构法得到的映射关系,其中激光雷达曲线平面上的  $m \times 3$  点集( $\mathbf{L}$ )可通过  $3 \times 3$  同构矩阵( $\mathbf{A}$ )转换为 RGB-D 图像垂直面上的  $m \times 3$  点集( $\mathbf{D}$ )。

在公式(5)中,  $(x_i^L, z_i^L)$  和  $(x_i^D, z_i^D)$  分别是激光雷达曲线和 RGB-D 图像在  $(x, z)$  二维平面上的点值,其中  $1 \leq i \leq m$ 。

$$\hat{\mathbf{L}}\mathbf{A} = \mathbf{D}, \quad (5)$$

$$\text{其中, } \mathbf{L} = \begin{bmatrix} x_1^L & z_1^L & 1 \\ \vdots & \vdots & \vdots \\ x_m^L & z_m^L & 1 \end{bmatrix}, \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \mathbf{D} =$$

$$\begin{bmatrix} x_1^D & z_1^D & 1 \\ \vdots & \vdots & \vdots \\ x_m^D & z_m^D & 1 \end{bmatrix}。$$

假设  $D_{v,w}^t$  和  $\hat{L}_{u,w}^t$  之间有  $\alpha$  个良好匹配对和  $\beta$  个噪声对,其中  $W = \alpha + \beta$ ,  $\beta$  个噪声对是由  $D_{v,w}^t$  和  $\hat{L}_{u,w}^t$  上的毛刺或空洞造成的。由于存在噪声对,理想的策略是从  $\alpha$  个良好匹配对中选出  $n$  对计算  $\mathbf{A}$  的值,其中  $n < \text{sizeof}(\alpha) < W$ 。然而,  $\alpha$  和  $\beta$  没有预先标注,因此本文从  $W$  对中随机挑选  $n$  对,当挑选出的  $n$  对都是好样本时,就可以得到  $\mathbf{A}$  的值。假设寻找到良好

匹配对的概率是  $c = \alpha / (\alpha + \beta)$ ,那么选取的  $n$  对样本都是好样本的概率为  $c^n$ ,通过找到可信度为 97.5% 的共识样本的最大尝试次数约为  $4/c^n$ <sup>[32]</sup>。

### 2.3.2 去噪

在去噪过程中,  $\beta$  个噪声对中主要包含两种噪声:空洞和毛刺。噪声区域分为 3 种类型。

①激光雷达曲线和 RGB-D 图像的深度数据中的噪声区域都是空洞。

②激光雷达曲线和 RGB-D 图像的深度数据中的噪声区域为空洞,其他区域不是空洞。

③激光雷达曲线和 RGB-D 图像的深度数据中的噪声区域为毛刺,其他区域不是毛刺。

类型①中  $\hat{L}_{u,w}^t$  和  $D_{v,w}^t$  的点为 0,很容易判断,类型②和③需要根据公式(6)进行判断使用。公式(6)中,  $\epsilon$  是阈值,用于区分配对的匹配距离。 $\hat{L}_{u,w}^t$  或  $D_{v,w}^t$  中满足公式(6)的点为空洞或毛刺。

$$\exists \beta \in W \|\hat{L}_{u,\beta} \mathbf{A} - D_{v,\beta}\| > \epsilon \quad (6)$$

对于类型②和③,在  $\hat{\mathbf{L}}$  或  $\mathbf{D}$  中发现空洞或毛刺可分别表示为  $\hat{L}_{u,\beta}$  或  $D_{v,\beta}$ 。然后,通过  $\hat{\mathbf{L}}$  和  $\mathbf{D}$  中对应的好点,使用公式(7)或(8)来检索噪声数据,其中  $\hat{L}_{u,\alpha}$  和  $D_{v,\alpha}$  分别是噪声点  $\hat{L}_{u,\beta}$  和  $D_{v,\beta}$  的修复数据。

$$\hat{L}_{u,\alpha} = D_{v,\alpha} \mathbf{A}^{-1} \quad (7)$$

$$D_{v,\alpha} = \hat{L}_{u,\alpha} \mathbf{A} \quad (8)$$

值得注意的是,在噪声区域类型①中,  $\hat{L}_{u,\beta}$  和  $D_{v,\beta}$  的点都为 0。当发现这种情况时,  $\hat{\mathbf{L}}$  和  $\mathbf{D}$  中原有的  $\hat{L}_{u,\beta}$  和  $D_{v,\beta}$  点会被移除,并使用与其相邻的好点进行替换。

## 3 实验

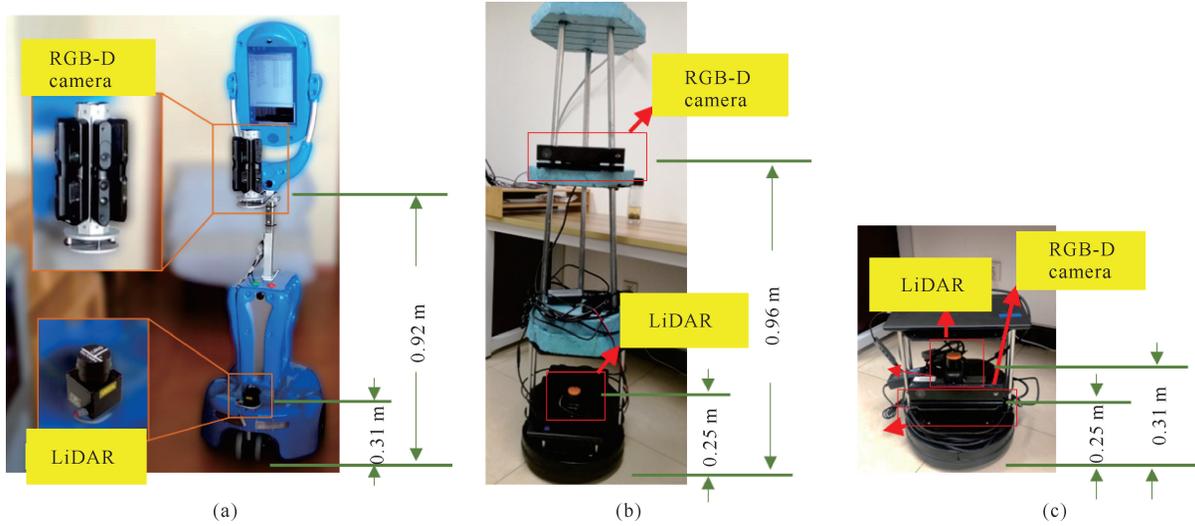
### 3.1 实验设置

使用室内开放式机器人导航数据集 Robot@Home<sup>[12]</sup>(以下简称“Robot@Home 数据集”)对本文提出的融合方法进行评估。Robot@Home 数据集收集了来自 5 个家庭环境的原始数据和经过处理的数据,旨在作为语义映射、机器人定位、二维/三维地图构建和物体分割的基准<sup>[34]</sup>。该数据集包含 1 个移动机器人收集的约 87 000 条有时间标记的观测数据,该移动机器人配备了 4 个 RGB-D 摄像机(XTion

Pro Live, 华硕电脑股份有限公司) 和 1 个激光雷达 (URG-04LX-UG01, 日本北阳电机株式会社)。在 Robot@Home 数据集中, 激光雷达安装在距离地面约 0.31 m 的位置, 4 台 RGB-D 摄像机安装在距离地面 0.92 m 的位置。图 2(a) 显示了 Robot@Home 数据集中激光雷达和 RGB-D 摄像头位置。尽管激光雷达和 RGB-D 摄像机的视角相似, 但 Robot@Home 数据集中的激光雷达曲线和 RGB-D 图像并未进行校准。

为了进一步验证本文提出的融合方法的有效性,

在真实室内导航环境中对本文提出的融合方法进行评估。此实验中, 在 Turtlebot-2 平台上放置 1 个激光雷达 (Hokuyo, 日本北阳电机株式会社) 和 1 个 RGB-D 摄像机 [Kinect 2, 微软(中国)有限公司], 它们分别位于不同高度和相似高度。图 2(b) 和 (c) 展示了这两种情况, 图 2(b) 中激光雷达和 RGB-D 摄像机分别放置在距离地面 0.25 m 和 0.96 m 的高度 (情况 1), 图 2(c) 中分别放置在距离地面 0.31 m 和 0.25 m 的高度 (情况 2)。



(a) Height setting position in the Robot@Home dataset; (b) – (c) Two height setting positions in real indoor navigation environment.

图 2 激光雷达和 RGB-D 摄像机在实验中的高度设置情况

Fig. 2 Height setting positions of LiDAR and RGB-D camera in experiment

### 3.2 Robot@Home 数据集的实验效果评估

在 Robot@Home 数据集中, 4 台 RGB-D 摄像机沿水平方向以  $180^\circ$  弧度生成  $960 \times 320$  分辨率的图像, 1 个激光雷达在  $240^\circ$  范围内生成每帧 682 像素的二维曲线。图 3 展示了经过处理后时间标记为“130986334520319765”、位置为“alma-s1 公寓”的“卧室 1”的第 4 幅 RGB 图像和 RGB-D 图像, 以及第 12 幅激光雷达曲线。

#### 3.2.1 Robot@Home 数据集上深度对齐的实验评估

在深度配准步骤中, RGB-D 图像水平方向上的 960 个像素被采样为 180 个像素, 激光雷达曲线沿水平方向的 682 个像素被采样为 240 个像素。这样, 深度配准的目标就是根据公式(1)同时计算  $u$  和  $v$  的

值。  $u$  的取值范围为  $\omega \leq u \leq \omega + W$ , 其中  $W = 240$  像素,  $1 \leq \omega \leq 180$ , 则  $u$  的取值范围为  $0 \leq u \leq 420$ 。理论上,  $v$  的取值范围为  $0 \leq v \leq 320$ 。但在移动机器人上, 激光雷达被安装在较低的位置, 因此在实验中  $v$  的最大值设定为 160 像素。通过公式(2)–(4)获得激光雷达曲线在特定区段和 RGB-D 图像在特定高度的相似度, 并确定它们之间的最佳匹配。图 4 展示了  $D_{v,\omega}^t$  和  $\hat{L}_{u,\omega}^t$  的  $\text{PCC}(D_v, \hat{L}_u)$ 、 $\text{Inv}_{\text{Cos}}(D_v, \hat{L}_u)$  和  $\text{Inv}_{\text{Euler}}(D_v, \hat{L}_u)$  中  $(u, v)$  值的分布情况, 在 [图 4: (a)–(c)] 中, 像素强度表示  $D_v$  和  $\hat{L}_u$  的匹配度, [图 4: (d)–(f)] 为  $(u, v)$  值最大时相应的激光雷达曲线和 RGB-D 曲线。

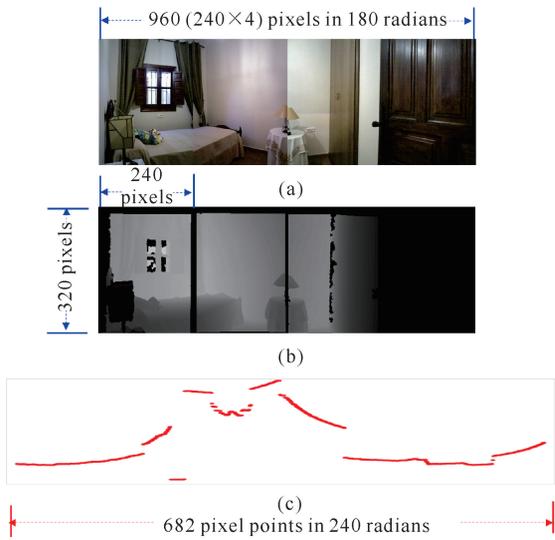


图3 在 Robot@Home 数据集中按时间维度对齐的 RGB 图像(a)、RGB-D 图像(b)和激光雷达曲线(c)

Fig. 3 RGB image (a), RGB-D image (b) and LiDAR curve (c) aligned by time dimension in the Robot@Home dataset

实验结果表明, PCC 值和  $Inv_{Euler}$  值分布相似, PCC 值和  $Inv_{Euler}$  值最大的位置在 (49, 21) 附近较为接近, 而且图 4(d) 和 (e) 中两条曲线的波动趋势比图 4(f) 中的波动趋势更加相关。这表明在深度配准任

务中, PCC 值和  $Inv_{Euler}$  值优于  $Inv_{Cos}$  值。此外, 从图 4 可以看出, PCC 值、 $Inv_{Euler}$  值和  $Inv_{Cos}$  值的范围分别为  $0.10 \leq PCC \leq 0.80$ 、 $0.70 \leq Inv_{Euler} \leq 0.93$  和  $0.10 \leq Inv_{Cos} \leq 0.35$ 。与  $Inv_{Euler}$  值和  $Inv_{Cos}$  值相比, PCC 值在 (0, 1) 范围内分布更均匀, 这有助于更好地区分  $u$  和  $v$ 。

此外, 在  $0 \leq v \leq 160$  范围内,  $0 \leq u \leq 20$  范围内的 PCC 值明显小于  $40 \leq u \leq 60$  范围内的 PCC 值。造成这种现象的原因是, 在 Robot@Home 数据集中, 激光雷达位于离地面 0.31 m 的高度, 而 RGB-D 摄像机位于离地面 0.92 m 的高度。有些物体放置在地面上, 离机器人非常近。图 3(a) 左侧床边放置了一个矮柜, 大部分矮柜没有出现在 RGB-D 图像中, 而是由激光雷达扫描到的。因此, 激光雷达的像素无法与 RGB-D 曲线同一高度的像素匹配。在图 4(d) 和 (e) 中,  $10 \leq u \leq 20$  范围内的红色点为矮柜, 而同一范围内的蓝色点为床的前部。激光雷达和 RGB-D 图像在此范围内的点的 PCC 值较低。而对于图 4(d) 和 (e) 中位于  $60 \leq w \leq 180$  右侧范围内的点, 红色的点比左侧部分的点更接近蓝色的点。这是因为图 3(a) 和 (b) 右侧部分的场景非常开阔, 在激光雷达之前没有矮小物体。

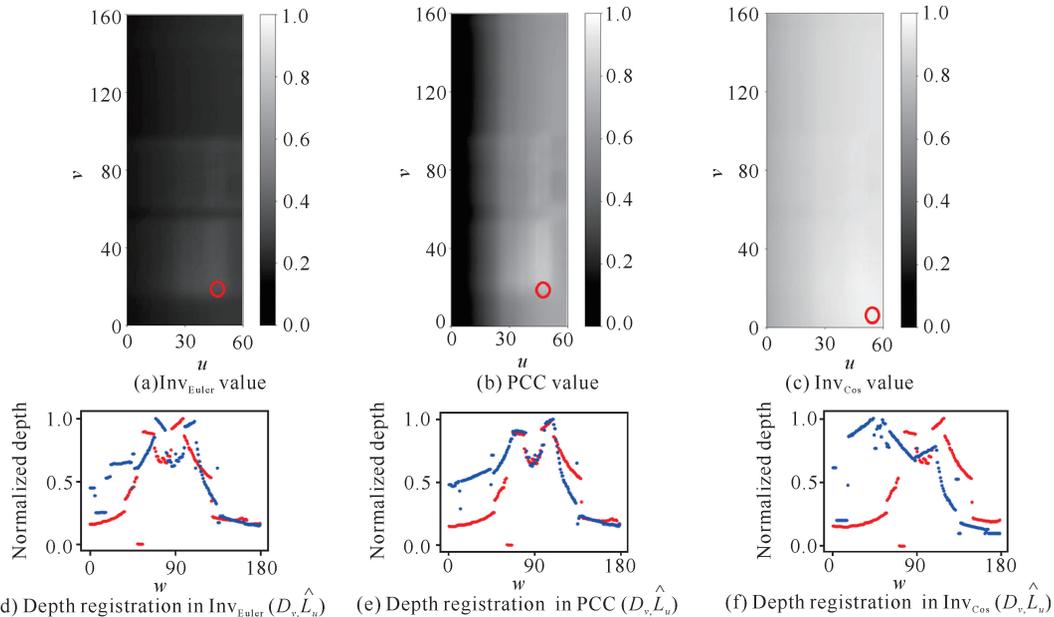


图4 激光雷达曲线(红色曲线)和 RGB-D 图像曲线(蓝色曲线)的  $Inv_{Euler}$  值、PCC 值和  $Inv_{Cos}$  值深度配准的分布情况

Fig. 4 Distribution of  $Inv_{Euler}$  value, PCC value and  $Inv_{Cos}$  value in depth registration for the LiDAR curve (red curve) and RGB-D image curve (blue curve)

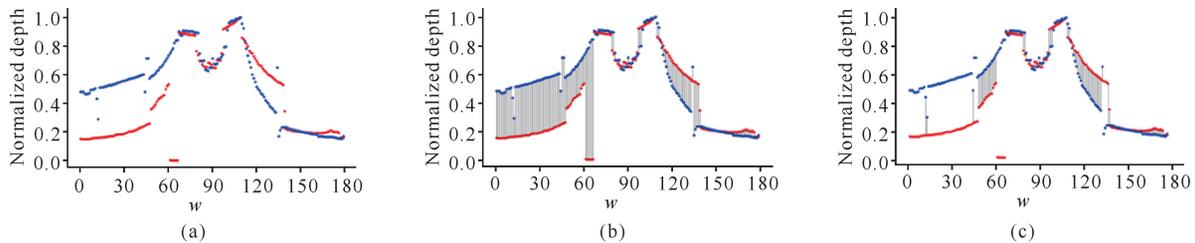
表 1 表示 Robot@Home 数据集中最大公寓“sarmis-s1”中 7 个房间的最大 PCC 值, 它们分别位于低 ( $0.0 < PCC \leq 0.3$ )、中 ( $0.3 < PCC \leq 0.6$ ) 和高 ( $0.6 < PCC \leq 1.0$ ) 范围内。从表 1 可以看

出, “corridor1”房间的高得分为 66.67%, 明显高于“bedroom3”房间 (18.48%) 和“bedroom1”房间 (20.10%)。造成这一现象的原因是在 Robot@Home 数据集中, “corridor1”房间的场景明显比其他

房间更加开放。

### 3.2.2 Robot@Home 数据集上配准的实验评估

使用 PCC ( $D_v, \hat{L}_u$ ) 对  $D_v$  和  $\hat{L}_u$  进行深度对齐后, 就可以得到  $u$  和  $v$  的值。对于 PCC 值较高的场景, 本文使用 RANSAC 算法和同构法匹配  $D_v$  和  $\hat{L}_u$ 。根据 Zlot 等<sup>[35]</sup> 的描述, 当  $c$  值大于 0.6 时, 在同源性计算中每次随机抽取  $n=6$  个样本对, 则在 84 次随机抽取中将得到 1 个共识样本, 置信度为 97.5%。在实验中, 采样次数设为 84 次, 采样对数  $n=6$ 。图 5 展示了图 3 和图 4(d) 所示场景的匹配过程, 其中图 5(a) 中的红色曲线和蓝色曲线分别是激光雷达曲线和用于匹配的 RGB-D 图像曲线。图 5(b) 和 (c) 显示了所有  $\alpha+\beta$  对和经过 84 次采样后最终  $\alpha$  个良好的配准对。



(a) LiDAR curve (red curve) points and RGB-D image (blue curve) depth points; (b) Matching line between LiDAR curve (red curve) points and RGB-D image (blue curve) depth points; (c) Matching line between LiDAR curve (red curve) points and RGB-D image (blue curve) depth points after removal of outliers.

图 5 激光雷达曲线点和 RGB-D 图像曲线对齐后的配准过程

Fig. 5 Registration process of LiDAR curve points and RGB-D image depth points after aligning

不同  $\epsilon$  值的良好匹配对和噪声对的召回率实验结果如图 6 所示。 $\xi = \xi = \alpha' \cap \alpha / \alpha$ ,  $\eta = \beta' \cap \beta / \beta$ ,  $\alpha'$  和  $\beta'$  是通过本文方法从噪声中得到的良好匹配对和修复对。 $\epsilon$  随所有线对的平均误差 ( $E$ ) 不同而变化,  $E = \sum_{u=1}^W \| \hat{L}_u A - D_u \| / W$ 。 $\xi$  和  $\eta$  的值越大, 同构矩阵  $A$  的性能就越好。当  $E \leq \epsilon \leq \sqrt{2}E$  时, 良好匹配对和噪声对都能获得较高的召回率。

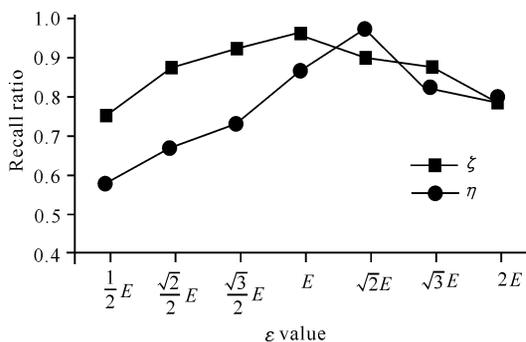


图 6 不同  $\epsilon$  值的良好匹配对和噪声对的召回率

Fig. 6 Recall ratio of good matching and noise matching under different  $\epsilon$  values

表 1 Robot@Home 数据集中最大公寓“sarmis-s1”中 7 个房间各范围内 PCC 最大值的数量

Table 1 Number of maximum PCC values for each range of seven rooms in the largest apartment ‘sarmis-s1’ in the Robot@Home dataset

房间 Room	范围 Range			高得分率/% High scoring rate/%
	低 Low	中 Middle	高 High	
bedroom1	28	316	92	21.10
bedroom2	20	212	56	19.44
bathroom2	16	84	24	19.35
bedroom3	44	256	68	18.48
corridor1	0	72	144	<b>66.67</b>
kitchroom1	4	120	20	13.89
livingroom1	28	196	76	25.33

Note: the bold data are the best data in the table.

### 3.2.3 Robot@Home 数据集上去噪的实验评估

图 7 展示了时间标记为“130986334520319765”、位置为“alma-s1 公寓”的“卧室 1”的第 4 幅 RGB-D 图像和第 12 幅激光雷达曲线的去噪结果, 图 7(a) 为原始的 RGB-D 图像和激光雷达曲线, 图 7(b) 为去噪后的 RGB-D 图像和激光雷达曲线。原始 RGB-D 图像底部的浅绿色线条表示激光雷达与 RGB-D 图像深度对齐的位置。标有“A”“B”和“C”的图块表示了 2.3.2 节中提到的 3 种噪声区域类型, 即激光雷达曲线中红色图块中的点是空心的, 其深度值由从 RGB-D 图像中提取的相应深度信息修复, 这些信息用蓝色箭头标示。同样, RGB-D 图像中位于绿线上的蓝色图块中的点是空心的, 它们的深度值由从激光雷达曲线中提取的相应深度信息修复, 该信息由红色箭头标注。这样, 原始的激光雷达曲线和 RGB-D 图像就可以相互去噪。去噪后, 激光雷达曲线中的蓝色像素点为本文提出的融合方法给出的建议修复点 [图 7(b)]。

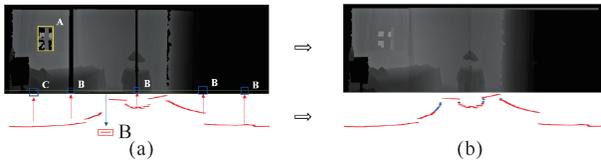


图7 Robot@Home 数据集场景的去噪结果

Fig. 7 Denoising results of the Robot@Home dataset scene

本文提出的融合方法在 Robot@Home 数据集的评估结果表明,皮尔逊相关系数在深度配准方面的效果相对优于归一化余弦距离和欧拉距离(图 5 和表 1)。此外,对于 PCC 值较高的开阔场景,激光雷达曲线能够在 RGB-D 图像的高度上进行配准,同时 RGB-D 图像也能在水平方向上与激光雷达曲线自动

配准。此外,图 6 和图 7 表明,RANSAC 算法结合同构法能够区分激光雷达曲线和 RGB-D 图像配准后的噪声。

### 3.3 真实室内导航环境中的实验分析

移动机器人在真实室内导航环境获取的数据如图 8 所示,其中上图和下图分别对应激光雷达和 RGB-D 摄像机高度设置情况 1 和情况 2 得到的图片,[图 8:(a)–(d)]分别为 RGB 图像、RGB-D 图像、激光雷达曲线(红色曲线)和 RGB-D 曲线(蓝色曲线)、经本文提出的融合方法融合后的深度图像。根据图 8(b)和(d)的对比可以看到本文提出的融合方法能够获得较好的配准融合结果。

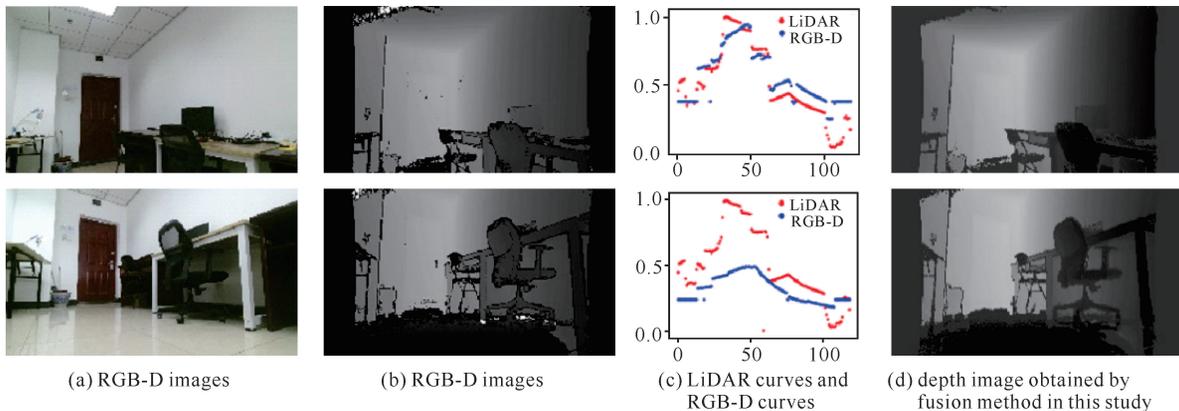


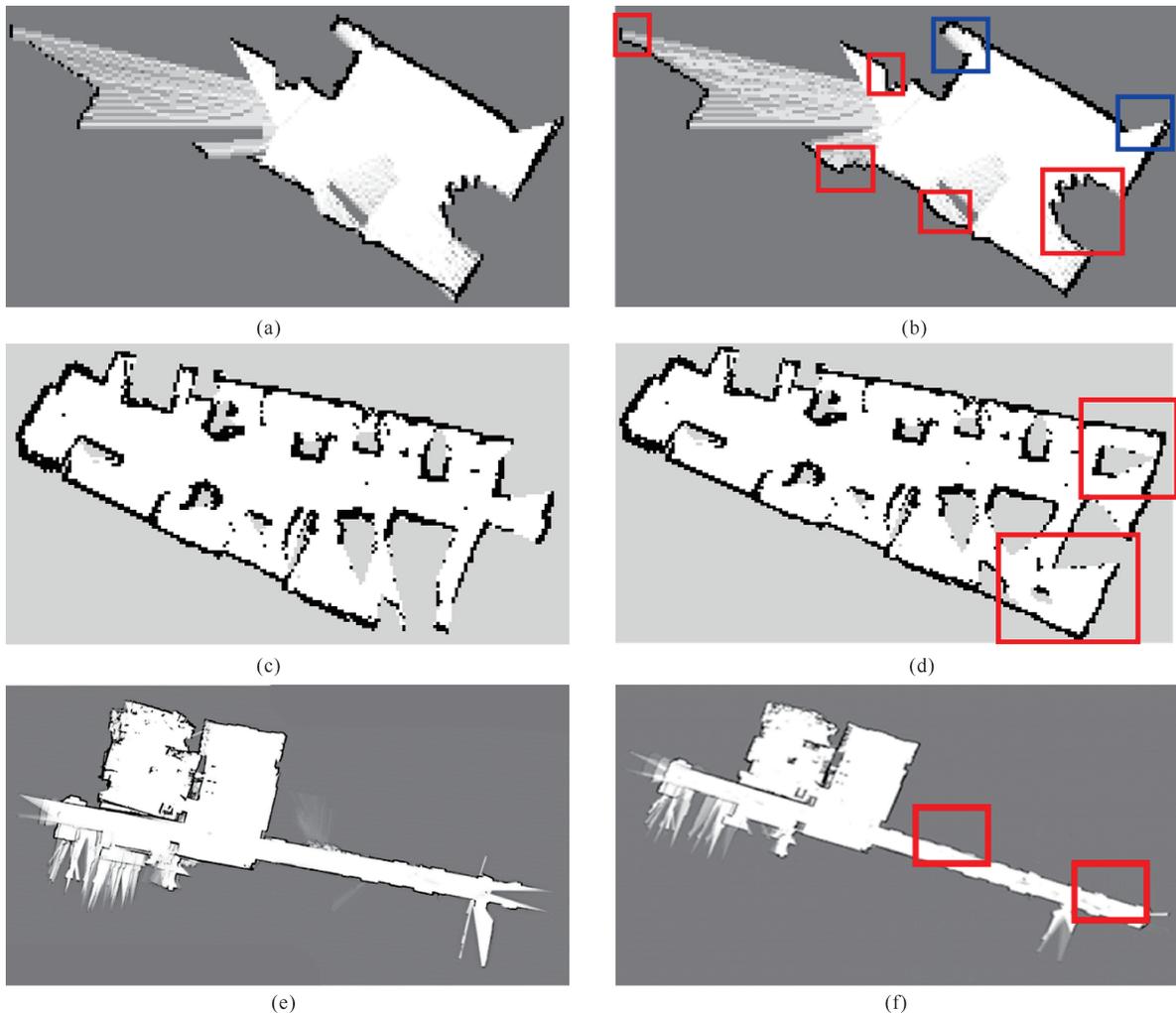
图8 移动机器人在真实室内导航环境中获得的图像

Fig. 8 Images obtained by moving robot in real indoor navigation environment

### 3.4 在 Robot@Home 数据集和真实室内导航环境数据中的融合效果对比

为了更进一步验证本文提出的融合方法,将该方法在 Robot@Home 数据集和真实导航环境数据中的融合效果进行对比。图 9(a)和(b)分别为通过 Robot@Home 数据集中“alma-s1”样本“bedroom1”房间的激光雷达数据生成的原始二维地图和经本文提出的融合方法得到的二维地图,图 9(c)和(d)分别为由真实室内导航环境数据实验房间的激光雷达数据生成的原始二维地图和经本文提出的融合方法得到的二维地图,图 9(e)和(f)分别为由真实室内导航环境数据实验房间所在楼层的激光雷达数据生成的原始二维地图和经本文提出的融合方法得到的二维地

图。从图 9(b)可以看出,红色区块中遗漏的边界在融合方法中得到了明显的改善和修复。此外,在蓝色图块的一些区域中,Robot@Home 数据集没有显示其正视图,因此这些区域的激光雷达曲线和 RGB-D 图像都没有深度信息。图 9(d)中二维地图的右下角和右上角红色区域比图 9(c)显示了更完整的信息,同时图 9(f)中二维地图的中心和右下角区域比图 9(e)显示了更为平滑的效果,并且消除了毛刺的影响。这是因为在这两个角落,激光雷达的深度信息由于桌子两侧的遮挡而部分为 0。而 RGB-D 摄像机提供的深度信息则包含了桌子和墙壁的全部深度信息,RGB-D 图像可补充激光雷达曲线的深度信息,从而得到融合结果更好的二维地图。



(a)–(b): ‘bedroom1’ in the ‘alma s1’ sample in the Robot@Home dataset; (c)–(d): the room of real indoor navigation environment experiment; (e)–(f): the floor of real indoor navigation environment data experimental room.

图9 激光雷达数据生成的原始二维地图和经本文提出的融合方法得到的二维地图

Fig. 9 Original 2D map generated by LiDAR data and 2D map generated by fusion method in this study

#### 4 结论

本文提出了一种基于无需校准的激光雷达和RGB-D图像的数据融合方法,并通过在Robot@Home数据集和真实的机器人室内环境中对本文方法进行了评估,实验结果表明,本文方法有助于去除原始数据的噪声,并在无需校准的情况下同时为激光雷达曲线和RGB-D图像提供更准确的细节信息。本文方法可以扩展到更精确的二维地图构建和三维场景重构中,用于机器人导航等应用领域。

#### 参考文献

[1] ZOU Q, SUN Q, CHEN L, et al. A comparative analysis of LiDAR SLAM-based indoor navigation for autonomous vehicles [J]. IEEE Transactions on Intelligent

Transportation Systems, 2022, 23(7): 6907-6921.

- [2] ZHOU Z, FENG X, DI S, et al. A LiDAR mapping system for robot navigation in dynamic environments [J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(7): 1-20.
- [3] SONG Y, XU F, YAO Q, et al. Navigation algorithm based on semantic segmentation in wheat fields using an RGB-D camera [J]. Information Processing in Agriculture, 2023, 10(4): 475-490.
- [4] KOLHATKAR C, WAGLE K. Review of SLAM algorithms for indoor mobile robot with LiDAR and RGB-D camera technology [C]//Innovations in Electrical and Electronic Engineering. Singapore: Springer Singapore, 2021: 397-409.
- [5] VULPI F, MARANI R, PETITTI A, et al. An RGB-D multi-view perspective for autonomous agricultural ro-

- bots [J]. *Computers and Electronics in Agriculture*, 2022, 202:107419.
- [6] MUÑOZ-BAÑÓN M Á, VELASCO-SÁNCHEZ E, CANDELAS F A, et al. OpenStreetMap-based autonomous navigation with LiDAR naive-valley-path obstacle avoidance [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(12):24428-24438.
- [7] PATOLIYA J, MEWADA H, HASSABALLAH M, et al. A robust autonomous navigation and mapping system based on GPS and LiDAR data for unconstrained environment [J]. *Earth Science Informatics*, 2022, 15(4):2703-2715.
- [8] HAO X, REN W, CHOO K R, et al. A self-trading and authenticated roaming scheme based on blockchain for smart grids [J]. *IEEE Transactions on Industrial Informatics*, 2022, 18(6):4097-4106.
- [9] TIAN B, SHIM V A, YUAN M, et al. RGB-D based cognitive map building and navigation [C]//2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ: IEEE, 2013:1562-1567.
- [10] WELLHAUSEN L, RANFTL R, HUTTER M. Safe robot navigation via multi-modal anomaly detection [J]. *IEEE Robotics and Automation Letters*, 2020, 5(2):1326-1333.
- [11] FU Y, YAN Q, YANG L, et al. Texture mapping for 3D reconstruction with RGB-D sensor [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018:4645-4653.
- [12] RUIZ-SARMIENTO J R, GALINDO C, GONZALEZ-JIMENEZ J. Robot@Home, a robotic dataset for semantic mapping of home environments [J]. *The International Journal of Robotics Research*, 2017, 36(2):131-141.
- [13] KAZEROUNI I A, FITZGERALD L, DOOLY G, et al. A survey of state-of-the-art on visual SLAM [J]. *Expert Systems with Applications*, 2022, 205:117734.
- [14] KUAN Y W, EE N O, WEI L S. Comparative study of intel R200, Kinect v2, and primesense RGB-D sensors performance outdoors [J]. *IEEE Sensors Journal*, 2019, 19(19):8741-8750.
- [15] SHIN Y S, PARK Y S, KIM A. Direct visual SLAM using sparse depth for camera-LiDAR system [C]//2018 IEEE International Conference on Robotics and Automation (ICRA). Piscataway, NJ: IEEE, 2018:5144-5151.
- [16] WINTERHALTER W, FLECKENSTEIN F, STEDER B, et al. Accurate indoor localization for RGB-D smartphones and tablets given 2D floor plans [C]//2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway, NJ: IEEE, 2015:3138-3143.
- [17] SONG H, CHOI W, KIM H. Robust vision-based relative-localization approach using an RGB-depth camera and LiDAR sensor fusion [J]. *IEEE Transactions on Industrial Electronics*, 2016, 63(6):3725-3736.
- [18] HESS W, KOHLER D, RAPP H, et al. Real-time loop closure in 2D LIDAR SLAM [C]//2016 IEEE International Conference on Robotics and Automation (ICRA). Piscataway, NJ: IEEE, 2016:1271-1278.
- [19] HIMSTEDT M, FROST J, HELLBACH S, et al. Large scale place recognition in 2D LiDAR scans using geometrical landmark relations [C]//2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ: IEEE, 2014:5030-5035.
- [20] ARRAS K O, MOZOS O M, BURGARD W. Using boosted features for the detection of people in 2D range data [C]//Proceedings 2007 IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE, 2007:3402-3407.
- [21] BRACHMANN E, ROTHER C. Learning less is more-6D camera localization via 3D surface regression [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018:4654-4662.
- [22] CHONG Z J, QIN B, BANDYOPADHYAY T, et al. Synthetic 2D LIDAR for precise vehicle localization in 3D urban environment [C]//2013 IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE, 2013:1554-1559.
- [23] DAI J, LI X, WANG K, et al. A novel STSOSLAM algorithm based on strong tracking second order central difference Kalman filter [J]. *Robotics and Autonomous Systems*, 2019, 116:114-125.
- [24] HOLZ D, HOLZER S, RUSU R B, et al. Real-time plane segmentation using RGB-D cameras [M]//RoboCup 2011: Robot Soccer World Cup XV. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012:306-317.
- [25] HOLZ D, ICHIM A E, TOMBARI F, et al. Registration with the point cloud library: a modular framework for aligning in 3-D [J]. *IEEE Robotics & Automation Magazine*, 2015, 22(4):110-124.
- [26] HOU X, SHI H, QU Y, et al. A fusion method for 2D LiDAR and RGB-D camera depth image without cali-

- bration [M]//Cognitive Computation and Systems. Singapore: Springer Nature Singapore, 2023: 89-101.
- [27] HUANG Y, ZHU M, ZHENG Z, et al. Homography-based visual servoing for underactuated VTOL UAVs tracking a 6-DOF moving ship [J]. IEEE Transactions on Vehicular Technology, 2022, 71(3): 2385-2398.
- [28] LIAN W, ZHANG L, ZHANG D. Rotation-invariant nonrigid point set matching in cluttered scenes [J]. IEEE Transactions on Image Processing, 2012, 21(5): 2786-2797.
- [29] POMERLEAU F, COLAS F, SIEGWART R, et al. Comparing ICP variants on real-world data sets [J]. Autonomous Robots, 2013, 34(3): 133-148.
- [30] WANG D Z, POSNER I, NEWMAN P. Model-free detection and tracking of dynamic objects with 2D lidar [J]. International Journal of Robotics Research, 2015, 34(7): 1039-1063.
- [31] WANG Y, WEI X, SHEN H, et al. Robust fusion for RGB-D tracking using CNN features [J]. Applied Soft Computing, 2020, 92: 106302.
- [32] YLIMÄKI M, HEIKKILÄ J, KANNALA J. Accurate 3-D reconstruction with RGB-D cameras using depth map fusion and pose refinement [C]//2018 24th International Conference on Pattern Recognition (ICPR). Piscataway, NJ: IEEE, 2018: 1977-1982.
- [33] ZHONG J, LI M, LIAO X, et al. A real-time infrared stereo matching algorithm for RGB-D cameras' indoor 3D perception [J]. ISPRS International Journal of Geo-Information, 2020, 9(8): 472.
- [34] ZHOU S K, LE H N, LUU K, et al. Deep reinforcement learning in medical imaging: a literature review [J]. Medical Image Analysis, 2021, 73: 102193.
- [35] ZLOT R, BOSSE M. Place recognition using keypoint similarities in 2D lidar maps [M]//Experimental Robotics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009: 363-372.

---

## A Fusion Method Based on Uncalibrated LiDAR Curves and RGB-D Images

XU Zhengli<sup>1</sup>, XIAO Sufang<sup>1\* \*</sup>, YANG Minghao<sup>2</sup>

(1. Guilin University of Electronic Technology, Guilin, Guangxi, 541004, China; 2. Institute of Automation of the Chinese Academy of Sciences, Beijing, 100190, China)

**Abstract:** LiDAR and RGB-D cameras are two widely used sensors in various tasks of robot navigation. Although image calibration can reduce the noise present in both sensors, there may still be holes and burrs due to the complex external environment. To address these issues, a fusion method based on uncalibrated LiDAR curves and RGB-D images is proposed. The proposed method utilizes the significant differences in format but close alignment of depth information between LiDAR curve and RGB-D image data. Through time alignment and correlation analysis, two-dimensional LiDAR curves can be matched with RGB-D images at heights. Meanwhile, the proposed method can automatically match the corresponding lines in RGB-D images to the width range of the LiDAR curve even if calibration is not necessary. The proposed method is evaluated on the public indoor robot navigation database Robot@Home. The experimental results show that the proposed method simultaneously de-noises the raw data of LiDAR curves and RGB-D images. Moreover, the proposed method is validated in real navigation environments and can be applied to the reconstruction of more accurate 2D maps for robot navigation.

**Key words:** LiDAR; depth image; data fusion; 2D map reconstruction; robot navigation

责任编辑: 梁 晓, 于子涵

---