基于二维激光雷达和多视角相机数据级融合的 3D-RGB 点云成像

田明昊[1]，李 攀[1]，阎 肃[1]，吴学良[1]，许录平[1]，阎博[1]
(1．西安电子科技大学空间科学与技术学院，陕西 710126)

摘　要：当前，基于激光和视觉的三维重建技术在地形场景的测量中具有广泛的应用。尽管目前已经开发出多种基于激光雷达和相机的三维成像方法，但它们各自存在一定的局限性。具体而言，RGB-D 相机虽然能够同时捕获颜色和深度信息，但其精度往往低于 LiDAR，而 3D 激光雷达虽然能够提供高精度的空间信息，却缺乏颜色信息，且其成本通常较高。本文提出了一种基于二维激光雷达与多视角相机数据级融合的 3D-RGB 成像方法，通过采用二维激光雷达与四组不同视角相机得到的 3D-RGB 点云数据，通过 3D-RGB 点增强、特征平面检测提取与全局一致对齐技术实现准确，稠密的 3D-RGB 成像。首先，通过 RGB 与点云信息融合提升数据质量，并利用特征平面检测优化几何结构表达。随后，结合全局一致对齐策略，降低累积误差，提高整体成像精度。实验结果表明，该方法相比多线激光雷达方案在成像密度和精度上具有优势，整体误差小于 0.15 米，验证了其在复杂环境下的三维重建与环境勘测应用潜力。

关键词：3D-RGB 成像设备、低成本、3D 激光雷达、点云着色、室内场景重建
中图分类号：TN958.98

3D-RGB Point Cloud Imaging Based on Data-level Fusion of 2D LiDAR and Multi-view Camera

Minghao Tian[1] , Pan Li[1] , Su Yan[1] , Xueliang Wu[1] , Luping Xu[1] , Bo Yan[1] , Xiaojun Li[1]

([1] School of Space Science and Technology, Xi'an University of Electronics and Technology, Shaanxi 710126, China)

Abstract: Currently, LiDAR- and vision-based 3D reconstruction technologies are widely used in terrain scene measurement. Although various 3D imaging methods based on LiDAR and cameras have been developed, each has limitations. RGB-D cameras can capture both color and depth information but often have lower accuracy than LiDAR, while 3D LiDAR provides high-precision spatial data but lacks color information and is typically expensive. This paper proposes a 3D-RGB imaging method based on data-level fusion of a 2D LiDAR and multi-view cameras, integrating 3D-RGB point cloud data from a 2D LiDAR and four cameras from different viewpoints. The method achieves accurate and dense 3D-RGB imaging through 3D-RGB point enhancement, feature plane detection and extraction, and global consistency alignment. First, fusing RGB and point cloud data enhances data quality, while feature plane detection optimizes geometric structure representation. Then, a global consistency alignment strategy reduces accumulated errors and improves overall imaging accuracy. Experimental results show that compared with multi-line LiDAR solutions, the proposed method offers advantages in imaging density and accuracy, with an overall error of less than 0.15 meters, demonstrating its potential for 3D reconstruction and environmental surveying in complex environments.

Key words:3D-RGB imaging device; low-cost; 3D LiDAR; Point cloud coloring; indoor scenario reconstruction.

1. 引言

近年来，随着智能手机、平板电脑和可穿戴设备的普及，以及物联网设备性能的提升，位置感知技术在室内外应用越来越广泛。室外定位依赖于 GPS、GLONASS、北斗等全球导航卫星系统，已经取得显著进展。但在室内，卫星信号因建筑物遮挡和反射，导致定位精度和稳定性下降，室内定位因此成为挑战。

目前，蓝牙、超声波和超宽带（UWB）等基于距离的室内定位方法具有较好的实时性和精度，但要实现精确的路径规划和导航，还需要依赖高精度的室内场景地图。室内建图不仅能为路径规划提供支持，还能提升系统的智能化水平，在自动导航等应用中发挥关键作用。

Newcomer 等人验证了获取 3D-RGB 点云可以通过 RGB-D 相机[1]实现，这种相机能够同时提供彩色图像和深度信息，并基于这两种信息实现 3D 地图重建，从而简化了 3D 重建过程**错误!未找到引用源。**，成为目前最常用的 3D 重建方法。但直接为彩色图像中的像素点提供深度信息[3]虽然实现简单，但其在点云的稠密度和精度存在局限，造成 3D 地图重建效果不佳。在最近的研究中，文献**错误!未找到引用源。**中提出了一种基于彩色图像引导的 RGB-D 相机追踪与三维重建方法，通过在相邻的图像中获取特征提高相机追踪的精度，并通过传统 icp 匹配方法进一步对相邻帧点云进行配准，有效提升了所构建点云的成像精度；文献**错误!未找到引用源。**中则通过基于事件信息和深度学习算法的高动态场景三维测量方法，增强弱激光条纹三维图像质量。但由于在低纹理的室内环境中缺乏足够的视觉特征，基于 RGB-D 相机获取 3D 点云依然会产生较大的误差，并且测量视野较小，难以应用到大范围室内环境。

为了实现获取更大的视野，进行更大范围的地图构建，FEHRMAN Brian 等人**错误!未找到引用源。**通过结合多个相机来创建室内环境的密集实时深度图。然而，对于距离相机较远的物体，其在图像之间的移动会显得相对较小，从而导致深度估计严重下降，依然难以应用到大范围室内场景。

相较于以上方法，激光雷达具备更广的探测距离与视野，并且具备更高的探测精度，且不易受环境的影响，并且可以通过与图像数据的融合实现为点云获取颜色信息。所以各种基于激光雷达和相机的测绘和检测设备与方法正在不断开发[6]-[7]。激光雷达与相机的融合可以分为基于体素的方法[6]，基于投影的方法**错误!未找到引用源。**，基于检测结果的融合方法[7]三类。基于体素的方法即将激光雷达点云体素化为 3D 网格，并将网格的特征投影到 RGB 图像上。基于投影的方法通过将激光雷达点云投影到 RGB 图像上，并嵌入到原始 RGB 图像中以生成新的 3D-RGB 图像。基于检测结果的融合则通过将激光雷达和图像的检测结果与感兴趣目标的空间和语义信息关联起来。然而，基于体素的方法缺乏语义信息，基于投影的方法在投影到不同视图时会损失大量空间信息，而基于检测结果的融合只关注感兴趣的目标，难以还原完整室内场景。

此外，通常用于 3D 重建的 3D 激光雷达成本昂贵，难以广泛应用于室内三维重建中，因此一些 2D 激光雷达已经应用于不需要高密度的精确测量且需要严格成本控制的场景，如室内跟踪或室内测绘**错误!未找到引用源。**。然而，2D 激光雷达提供的空间信息对于视野范围来说相当有限，所以为了实现高性价比，许多研究正在开发基于 2D 激光雷达实现低成本 3D 激光雷达，例如文献**错误!未找到引用源。**中师瑞卓等人就通过基于云台控制单线激光雷达旋转，并通过数据解算实现获取 3D 点云。在我们之前的工作中[9]**错误!未找到引用源。**，也验证了此类 3D 扫描设备具有简单高效的结构，无需校准即可实现 3D 建图，并应用到了室内场景目标检测跟踪场景中。然而其获取的三维点云密度仍无法满足室内高精度三维重建的需求，依然具有对点云密度进行增强处理的必要性。

基于上述当前 3D 成像研究存在的问题，本文研究了一种基于二维激光雷达和多视角相机数据级融合的低成本 3D-RGB 成像方法。具体地，本文基于二维激光雷达与相机搭建了低成本的 3D-RGB 点云采集设备，

并通过设计实现基于超像素改进的雷视融合算法有效提升了获取点云图像的密度与精度，最后通过基于特征平面提取实现点云的全局一致对齐，实现大范围室内 3D-RGB 点云地图构建。

所用采集设备包括一种基于 2D 激光雷达与云台组合的 3D 激光点云采集装置用于采集 3D 激光点云，四个不同视角相机用于采集图像信息，搭载小车用于搭载采集设备移动。

在室内 3D-RGB 点云地图构建处理中，本文提出了以下五个步骤：

(1) 用低成本扫描设备获取 3D 激光点云和四个方向的相机图像，并标记激光点云中的特征点。(2) 对相机图像进行超像素分割，将图像划分为若干相似区域。(3) 根据相机参数，将 3D 激光点与超像素关联，映射特征点及普通激光点，生成 3D-RGB 点云图。(4) 检测并标记 3D-RGB 点云中的特征线和特征平面。(5) 通过不同相机和帧间特征点云的匹配，关联相似的 3D-RGB 点，实现完整 3D-RGB 点云地图构建。

本文方法的优势体现在：(1)相较于 RGB-D 相机，3D 激光雷达等地图重建设备，本文通过 2D 激光雷达与云台组合，更低成本采集 3D 点云。(2) 通过基于超像素改进方法实现雷视融合，获得更密集的 3D-RGB 点云，有效减少了因点云稀疏造成的测量误差，实现更大的探测范围。(3) 利用特征平面检测后的 3D-RGB 点云信息进行帧间点云配准，有效提升了整体建图精度。

本文 3D-RGB 点云采集设备结构及处理数据结构流程如图 1 所示，蓝色部分为本文 3D-RGB 点云采集及传输设备结构，黄色部分为点云建图数据处理流程。

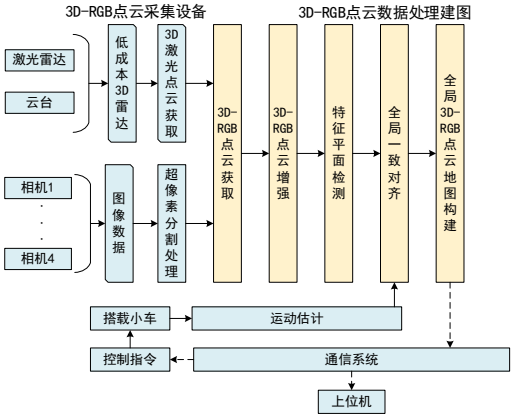

图 1 本文结构

Fig.1 Structure of this paper

2. 设备结构

本文搭建了一种低成本的基于二维激光雷达与相机的高精度 3D-RGB 点云成像设备采集 3D-RGB 点云数据，其结构包括一种基于 2D 激光雷达与云台组合的 3D 激光点云采集装置，四个不同方向的相机，装置搭载小车与 5g 通信系统四部分。

2.1.低成本 3D 点云采集装置

本文低成本 3D 点云采集设备由一个 2D 激光雷达和一个可以水平旋转的云台组成，雷达与云台夹角为 90 度，通过云台带动雷达旋转实现 3D 点云采集。

2.2.多视角相机

将四个相机安置在 3D 点云采集装置的周围，每个相机与相邻相机夹角为 90°，从而实现全方位的 360° 覆盖。通过这四个方向的相机，可以获取周围环境的 RGB 图像，从而获取颜色信息。

2.3.搭载小车

搭载成像装置的小车配备了四个全向轮，车辆可以在不偏转的情况下向四个方向移动（前进、后退、右移、左移）。通过向四个全向轮发送固定脉冲，可以完成四个方向的精确移动。这意味着在第 k 次数据扫描期间，可以直接获得车辆的精确移动位置 $x_{\delta}^{k,k+1}$，$y_{\delta}^{k,k+1}$，$z_{\delta}^{k,k+1}$ 与旋转角度 $\theta_{\delta}^{k,k+1}$，$\varphi_{\delta}^{k,k+1}$。在室内场景中，通常小车只涉及在 x,y 平面内的平移与旋转，故 $z_{\delta}^{k,k+1}$ 和 $\varphi_{\delta}^{k,k+1}$ 通常为 0。

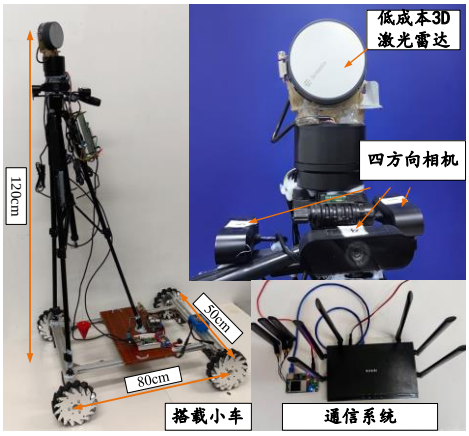

图 2 本文设备示意图

Fig. 2 Schematic diagram of the device in this paper

2.4.通信系统

本文建立了基于5G和WiFi的通信系统用于小车与上位机通信。在此系统中，车辆可以通过远程接收用户发出的运动指令，并且将采集获取的 3D 地图通过互联网传输并发送到上位机用户。此节所提设备如图 2 所示，其中标出了设备尺寸和所介绍的四个组成部分。

3. 数据处理

3.1.三维 RGB 点云获取

高分辨率图像的像素点数量通常大于10^6，远超激光雷达采集的点云点数。为了生成稠密的 3D-RGB 点云，可采用超像素分割方法对图像进行聚类，将颜色相近的像素归为一类，并结合 3D 点云与图像的映射关系，将图像中的像素点映射到三维空间，从而增强点云的密度。

步骤一：超像素分割处理

对图像进行超像素分割处理[10]，可将颜色相近的像素归入同一超像素聚类，从而将图像划分为多个相似颜色的区域。在光学图像处理中，简单线性迭代聚类（Simple Linear Iterative Clustering, SLIC）是一种常用且有效的超像素分割方法，其主要包括以下四个步骤：(1)在图像上以规则网格采样像素，初始化聚类中心，并设定网格间距；(2)使用标准的 K 均值聚类(K-means clustering,k-means)迭代[15]算法，将每个像素分配给相似度最高的聚类中心；(3)当所有像素都归属到对应的聚类中心后，计算每个聚类中心的平均坐标并进行更新；(4)重复 K-means 聚类过程，直至聚类中心位置稳定，完成 SLIC 超像素分割。

步骤二：在基础映射中，激光雷达的每个 3D 激光点与图像的一个像素相关联。如图 3 所示，$\breve{A}$表示由激光雷达获取的 3D 激光点，其位置为(x_A, y_A, z_A)。在融合过程中，将 3D 点所在坐标系的坐标原点设为标记为$\breve{O}$的焦点。点$\breve{A}$与点$\breve{O}$的距离为z_A。像素$A(u_A, v_A)$是是与 3D 激光点$\breve{A}$对应的图像像素，根据相机中的比例关系，有

$$\frac{f}{z_A}=\frac{\delta(u_A-u_o)}{x_A}=\frac{\delta(v_A-v_o)}{y_A} \tag{1}$$

然后，与点$\breve{A}$对应的像素A的像素坐标系坐标可以通过以下方式计算：

$$u_A=\frac{x_A f}{z_A\delta}+u_o$$

$$v_A=\frac{y_A f}{z_A\delta}+v_o \tag{2}$$

然后，3D 点$\breve{A}$即可从对应像素获取颜色信息，可以表示为一个 3D-RGB 点$\dot{A}=(x_A, y_A, z_A, r_A, g_A, b_A)$，其中 3D-RGB 点中的$(r_A, g_A, b_A)$分别表示红色、绿色和蓝色的颜色分量。

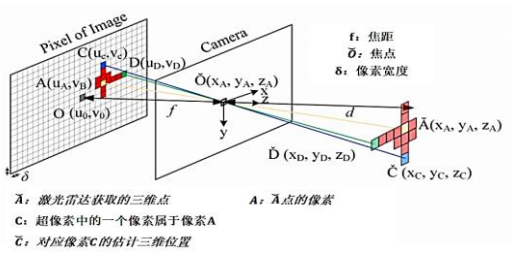

图 3 点云与超像素对应关系

Fig. 3 Correspondence between point cloud and superpixels

步骤三：估计属于像素 A 的超级像素的像素在 3D 激光点坐标系中的位置。在图 3 中，根据超像素分割结果，像素 C 和 D 与像素 A 处于同一超像素区域，属于同一超级像素。基于同一超级像素中像素的一致性，我们认为焦点$\breve{O}$到三维点$\breve{A}$，$\breve{C}$，$\breve{D}$的距离是相等的，即$z_A=z_C=z_D$。所以以像素点 C 为例，有

$$\frac{f}{z_A}=\frac{\delta(u_C-u_A)}{x_C-x_A}=\frac{\delta(v_C-v_A)}{y_C-y_A} \tag{3}$$

所以像素点 C 在 3D 点坐标系中对应的三维位置可以通过以下方式计算：

$$x_C=\frac{(x_C-x_A)f}{z_A\delta}+u_A$$

$$y_C=\frac{(y_C-x_A)f}{z_A\delta}+v_A$$

$$z_C=z_A \tag{4}$$

因为同一超像素聚类中的像素 RGB 颜

色近似，所以像素点 D 及同一超像素聚类中所有其他像素的 3D 位置也可以通过上述公式进行估计。

在此步骤结束时，可以计算出四幅图像中所有像素点在 3D 点云空间中的位置，可以获得一个包含 $4\times N_X \times N_Y$ 个 3D-RGB 点的稠密 3D-RGB 点云集合，表示为

$$\left\{P_i=(x,y,z,r,g,b)\middle| i=1,\cdots 4\times N_X \times N_Y\right\} \tag{5}$$

其中 N_X 与 N_Y 分别表示图像的宽度和高度的像素范围，即图像的分辨率。基于此方法获取到的 3D-RGB 点数量相对激光雷达采集的 3D 点云有明显提升，获取到更稠密的 3D-RGB 点云。

3.2.3D RGB 点云增强

在对环境三维重建应用中，更连续、密集的点云可以实现对周围环境中目标更好的纹理还原效果。因此，本文采用了一种包含膨胀和腐蚀操作的闭合操作对从四幅图像中获得的 3D-RGB 点云 $4\times N_X \times N_Y$ 进行增强处理以获得更密集的点云，可以将一个 3D 线条或平面中的散乱 3D-RGB 点转换为连续的点。

根据文献[11]，通过结构元素（或内核 B 对矩阵 A 进行闭合操作，得到一个精细化的矩阵 $\ddot{A}$。定义为：

$$\begin{aligned}\dot{A}&=A+B_1\\ \ddot{A}&=\dot{A}-B_2\end{aligned} \tag{6}$$

在对点云 A 的闭合操作中，$A+B_1$ 表示通过结构元素 B_1 对 A 进行的膨胀操作。而 $\dot{A}-B_2$ 表示通过 B_2 对 $\dot{A}$ 进行的腐蚀操作。此阶段包括以下三个步骤。

步骤 1：使用 3D 结构元素 B_1 膨胀 3D-RGB 点云 A 得到膨胀后的点云 $\dot{A}$。3D 结构元素 B_1 由 18 个单元组成，其中一个中心单元作为原点（0,0,0），对应于 (x,y,z) 在 B_1 的 3×3×3 单元中，欧式距离小于等于 2 的单元格为 1，共有 18 个。

步骤 2：基于 3D 结构元素 B_2 腐蚀膨胀后的点云 $\dot{A}$，得到腐蚀后的点云 $\ddot{A}$。在所使用的 3D 结构元素 B_2 中，在 3×3×3 的单元格中欧式距离小于等于 2 的单元格为 1，共 6 个。

步骤 3：对点云闭合操作后新创建的 3D 点进行颜色增强。在闭合操作中，在连续的直线与平面内填充了一些新的 3D 点，此步骤将此 3D 点的颜色设置为与其距离最近的具有 RGB 颜色信息的点相同。

基于上述步骤，通过补充 3D 点实现了对点云密度的进一步增强。

3.3.特征平面检测

在这一阶段，将检测并提取场景点云中的地板、屋顶和墙壁，这些平面区域具有显著的几何形态特征，并且在不同采集角度下能够保持形状的一致性，不易受到视角变化的影响。通过在后续阶段直接对提取出的特征平面进行对齐，可以更高效的将所有点云关联，生成全局 3D-RGB 地图。

在本文的设备中，激光雷达水平放置，而相机的图像垂直于地面和屋顶。因此，墙壁部分的 3D- RGB 点云在 x-y 位置上是相同的，而地面和屋顶的点在 z 轴位置上是相同的。对于墙壁部分，有

$$\begin{matrix}(x_p,y_p,z_p,r_p,g_p,b_p)\in P_{w,i}\\ (x_q,y_q,z_q,r_q,g_q,b_q)\in P_{w,i}\end{matrix}\Rightarrow\begin{matrix}(x_p=x_q)\\ (y_p=y_q)\end{matrix} \tag{7}$$

其中，用 $P_{w,i}$ 表示第 i 面墙的点云。同样地，用 P_r 和 P_f 分别表示地面与屋顶的点云。对于地面部分点云，有

$$\begin{matrix}(x_p,y_p,z_p,r_p,g_p,b_p)\in P_f\\ (x_q,y_q,z_q,r_q,g_q,b_q)\in P_f\end{matrix}\Rightarrow\begin{matrix}(z_p=z_q)\\ (z_p<0)\end{matrix} \tag{8}$$

对于屋顶部分点云

$$\begin{matrix}(x_p,y_p,z_p,r_p,g_p,b_p)\in P_r\\ (x_q,y_q,z_q,r_q,g_q,b_q)\in P_r\end{matrix}\Rightarrow\begin{matrix}(z_p=z_q)\\ (z_p>0)\end{matrix} \tag{9}$$

基于上述特征，可以通过以下三个步骤检测地板、屋顶和墙壁。

步骤 1：检测地板和屋顶部分点云。首先统计 3D-RGB 点云在 z 轴上的位置分布，其在 z<0 与 z>0 处会有两个峰值，分别对应地板和屋顶点云。这两个峰值分别记作 z_f 和 z_r。然后，将 z 轴位置接近 z_f 的 3D-RGB

点视为地板的潜在测量点，接近 z_r 的点则视步骤 2：通过二维区域生长法对地板和屋顶进行分割。首先对上一步地板的潜在测量点集{ z_f }进行处理，对其在 x-y 平面上的位置进行二维区域生长，得到的最大的二维区域即为地板，并表示为点集 P_f 。屋顶也可以通过相同方法基于点集{ z_r }获取，屋顶的测量点表示为 P_r 。

步骤 2：在对 3D-RGB 点云地面，屋顶部分检测提取，即将地面部分点云 P_f 与屋顶部分点云 P_r 去除后，对剩余点云进行墙面检测。首先创建一个 x-y 网格地图 M_w 来定位墙壁，将三维空间中的点云根据其 x,y 值投射到网格地图 M_w 中，这样 M_w 中的每个网格内包含了点云中全部 x,y 值位于此区间的点。在 M_w 中，墙壁所在的网格单元会包含更多的 3D-RGB 点。因此，如果某个网格单元的点数超过设定的阈值，则该网格单元可以被认为是可能的墙壁。在对网格单元进行投票后，可以在 M_w 地图中获得潜在墙壁的网格单元。接下来对 M_w 执行二维区域生长，将孤立的网格单元丢弃，保留与墙壁轮廓相符的区域。每个区域对应一面墙，区域内的点表示第 i 面墙的尺寸记作 $P_{w,i}$ ，由此可提取出每次扫描中获取的点云中的墙面部分 $\left\{P_{w,i}\middle|i=1,...,N_{w,k}\right\}$。

3.4.全局一致对齐

通过对各帧点云中具有丰富特征的平面部分进行关联对齐，实现各帧点云的全局一致对齐，完成整体建图。

在 3.3 特征平面检测中，记在第 k 次扫描中获取的点云其地板、屋顶和 N_w^k 面墙壁的点云部分表示为 P_f^k ， P_r^k 与 $\left\{P_{w,i}^k\middle|i=1,...,N_w^k\right\}$，第 $k+1$ 次扫描中的地板、屋顶和墙壁表示为 P_f^{k+1} ， P_r^{k+1} 与 $\left\{P_{w,i}^{k+1}\middle|i=1,...,N_w^{k+1}\right\}$。

在对点云进行关联匹配时，由于移动搭载小车在室内场景中运动，所以在考虑小车姿态变化时只需考虑 x-y 平面内的偏转角，记为 θ 。车辆在第 k 次扫描和第 $k+1$ 次扫描期间的大致位移与姿态变化可以通过车辆本身里程计提供的信息获取，并且其中沿 x,y,z 轴的位移分别表示为 $x_\delta^{k,k+1}$， $y_\delta^{k,k+1}$， $z_\delta^{k,k+1}$ 在 x-y 平面内的偏转角表示为 $\theta_\delta^{k,k+1}$。但在实际应用中，车辆本身返回的位姿信息存在较大误差，对应包括沿 x,y,z 轴的位移误差 $\Delta x^{k,k+1}$， $\Delta y^{k,k+1}$， $\Delta z^{k,k+1}$ 与 x-y 平面内的偏转角度误差 $\Delta\theta^{k,k+1}$，所以单纯基于里程计信息难以构建高精度的点云地图。

所以在这一阶段，本文开发了一种基于粒子滤波的点云配准优化方法，应用于对相邻两帧点云进行关联匹配，通过对当前帧点云相对上一帧点云的位姿修正，减小匹配误差。该方法分为以下四个步骤：

步骤 1：建立一个包含 N_p 个粒子的初始集合。初始粒子是随机选择的，每个粒子表示一个可能的估计误差值。

$$e=\left\{x_\Delta^{k,k+1},y_\Delta^{k,k+1},z_\Delta^{k,k+1},\theta_\Delta^{k,k+1}\right\} \tag{10}$$

给定一个粒子和第 k 次扫描中获取的 3D-RGB 点云，可以通过公式获得第 k+1 次扫描的估计 3D-RGB 点云。

$$\begin{bmatrix}x_p\\y_p\\z_p\end{bmatrix}\begin{bmatrix}\cos\theta_{\delta\Delta}&-\sin\theta_{\delta\Delta}&0\\\sin\theta_{\delta\Delta}&\cos\theta_{\delta\Delta}&0\\0&0&1\end{bmatrix}+\begin{bmatrix}x_{\delta\Delta}\\y_{\delta\Delta}\\z_{\delta\Delta}\end{bmatrix}=\begin{bmatrix}x_q\\y_q\\z_q\end{bmatrix} \tag{11}$$

$$\begin{bmatrix}r_p\\g_p\\b_p\end{bmatrix}=\begin{bmatrix}r_q\\g_q\\b_q\end{bmatrix} \tag{12}$$

其中，公式中的 $\theta_{\delta\Delta}^{k,k+1}$， $x_{\delta\Delta}^{k,k+1}$， $y_{\delta\Delta}^{k,k+1}$， $z_{\delta\Delta}^{k,k+1}$ 为

$$\begin{aligned}\theta_{\delta\Delta}&=\alpha_\delta^{k,k+1}+\alpha_\Delta^{k,k+1}\\x_{\delta\Delta}&=x_\delta^{k,k+1}+x_\Delta^{k,k+1}\\y_{\delta\Delta}&=y_\delta^{k,k+1}+y_\Delta^{k,k+1}\\z_{\delta\Delta}&=z_\delta^{k,k+1}+z_\Delta^{k,k+1}\end{aligned} \tag{13}$$

基于此粒子估计的第 k+1 次扫描的 3D-RGB 点基于特征平面提取表示为 $\hat{P}_f^{k+1}$， $\hat{P}_r^{k+1}$ 和 $\left\{\hat{P}_w^{i,k+1}\middle|i=1,...,N_w^{k+1}\right\}$。

步骤 2：定义一个代价函数，用于评估第 k+1 次扫描中获取的 3D-RGB 点云 $P_f^{i,k+1},P_r^{i,k+1},\left\{P_w^{i,k+1}\middle|i=1,...,N_w^{k+1}\right\}$ 与基于粒子的估计 3D-RGB 点云 $\hat{P}_f^{i,k+1}$， $\hat{P}_r^{i,k+1}$，

$\left\{\hat{P}_w^{i,k+1} \middle| i=1,...,N_w^{k+1}\right\}$的一致性，定义的代价函数为

$$f(p_n)=D(P_f^k,\hat{P}_f^{k+1})+D(P_r^k,\hat{P}_r^{k+1})+D(\{P_w^{i,k}\},\{\hat{P}_w^{i,k+1}\})$$

(14)

其中，p_n表示第 n 个粒子$\{x_\Delta^{k,k+1},y_\Delta^{k,k+1},z_\Delta^{k,k+1},\theta_\Delta^{k,k+1}\}$。函数$D(P^k,P^{k+1})$表示点云区域$P^k$与$P^{k+1}$之间未重叠部分的体积，因此当点云$P^k$与$P^{k+1}$区域没有很好地重叠，$D(P^k,P^{k+1})$会有较大值。因此，如果两次扫描中的对象重叠良好，代价函数$f(p_n)$的值将非常小。接下来计算每个粒子的代价函数，并比较N_p个粒子的代价函数，以找到最优的粒子。

步骤 3：按照优化算法迭代更新粒子。通过不断更新粒子，最终可以得到最优粒子$(\hat{\theta}_{\delta\Delta},\hat{x}_{\delta\Delta},\hat{y}_{\delta\Delta},\hat{z}_{\delta\Delta})$，从而获得最小的代价函数值。

步骤 4：通过最优变换粒子矩阵$(\hat{\theta}_{\delta\Delta},\hat{x}_{\delta\Delta},\hat{y}_{\delta\Delta},\hat{z}_{\delta\Delta})$对第 k+1 次扫描中获取的 3D-RGB 点$P^{k+1}$进行位姿修正，从而计算出精确的 3D-RGB 点云$\hat{P}^{k+1}$。

$$\begin{bmatrix}\hat{x}_q\\ \hat{y}_q\\ \hat{z}_q\end{bmatrix}==\begin{bmatrix}x_q-\hat{x}_{\delta\Delta}\\ y_q-\hat{y}_{\delta\Delta}\\ z_q-\hat{z}_{\delta\Delta}\end{bmatrix}\begin{bmatrix}\cos\hat{\theta}_{\delta\Delta} & -\sin\hat{\theta}_{\delta\Delta} & 0\\ \sin\hat{\theta}_{\delta\Delta} & \cos\hat{\theta}_{\delta\Delta} & 0\\ 0 & 0 & 1\end{bmatrix}^{-1}$$

(15)

$$(x_q,y_q,z_q,r_q,g_q,b_q)\in P^{k+1}$$
$$(\hat{x}_q,\hat{y}_q,\hat{z}_q,\hat{r}_q,\hat{g}_q,\hat{b}_q)\in \hat{P}^{k+1}$$

(16)

经测试，修正后的 3D-RGB 点集与之前扫描中获取的点集可以很好地重叠。

3.5.全局 3D-RGB 地图构建

将 3D-RGB 地图的原点设为设备在第一次扫描时的位置，并且可以获得第一次扫描的 3D-RGB 点云A_1。然后，将设备移动到第二个位置，通过 3.2，3.3，3.4 中的处理，可以获得位姿校准后的 3D-RGB 点云$\hat{A}_2$。随着设备的移动，依次获得校准后的 3D-RGB 点云集合$\hat{A}_3$到$\hat{A}_k$。通过叠加所有的 3D-RGB 点云A_1、$\hat{A}_2 ... \hat{A}_k$，可以完成构建一个覆盖大范围的全局 3D-RGB 地图。

4. 实验与结果分析

为了全方面评估本文提出的 3D-RGB 成像方法 3D-RGB 建图的性能与成本优势，与其它三维地图构建设备应用成本进行对比，并对 3D-RGB 点云建图密度，精度进行实验验证。

4.1.低成本成像设备

为了证明本文采用的 3D-RGB 点云成像设备在构建成本上的优势，将本文设备的构建成本与 RGB-D 相机成像[1]，多相机组合成像[6]与基于 3D 激光雷达与相机组合[9]共三种 3D-RGB 成像设备进行成本与性能的对比，如表 1 所示。

表 1 本文设备与其它三维成像设备对比

Table 1 Comparison of This Device with Other 3D Imaging Devices

	RGB-D	多相机组合	3D 雷达-相机	本文设备
探测范围（m）	0.5-4	0.5-50	0.4-150	0.4-10
成本（$）	≥210	≥440	≈4040	≈610
低纹理环境	较差[1]	较差	较好[2]	较好
低光照环境	较差[3]	较差	较好[4]	较好

[1]：低纹理环境下基于 RGB-D 和多相机组合的方法可能导致数据丢失或匹配失败，整体误差可能达到 5-20%的测量距离。

[2]：低纹理环境下基于激光雷达进行点云成像完全不受影响，测量依赖激光返回信号。

[3]：低光照环境下，影响 RGB 传感器对特征提取的效果，基于多相机组合的方法在黑暗环境中可能完全失效。

[4]：激光本身是主动光源，不受低光照环境影响。

对比发现，本文采用的 3D-RGB 点云成像设备在低纹理和低光照环境下的成像效果优于 RGB-D 相机和多相机组合成像设备，展现了更强的适应性。并且相比于 3D 激光雷达与相机组合的成像设备，本设备成本显著降低，仅为其应用成本的约 15%，具备更高的性价比。尽管本设备的有效探测范围相较于多相机组合成像设备和 3D 激光雷达与相机组合设备稍近，但 0.4-10m 的探测范围已能满足室内环境的成像需求。

4.2.点云密度增强

为了验证本文基于超像素分割改进的雷视融合方法在点云密度增强中的有效性，我们将本文中基于超像素分割的雷视融合方法与应用文献[7]中基于映射的雷视融合方法对成像效果进行对比。在设置云台转速为高速模式和低速模式下进行了点获取实验。在低速模式（7 转每分钟）下，转盘转一圈需要 8.40 秒。而在高速模式（140 转每分钟）下，仅需 0.51 秒。更高的速度模式会导致点云更稀疏，但成像帧率更高。通过地图融合方法和超像素改进方法，将激光雷达获取的三维点与点的颜色信息进行融合。随后，获

表 2 本文基于超像素分割改进的雷视融合方法与基于映射的方法在不同云台转速成像效果对比

Table 2 Comparison of imaging results between the improved LiDAR-vision fusion method based on superpixel segmentation and the mapping-based method at different gimbal rotational speeds

方法	云台转速	$Num(k)$	$\bar{d}(cm)$	$E_r(cm)$	$E_p(cm)$
本文基于超像素分割的雷视融合方法	低速	8356.5	0.192	1.628	1.950
	高速	7086.5	0.188	1.750	1.979
基于映射的雷视融合方法	低速	62.1	0.982	3.542	3.567
	高速	3.4	5.605	10.479	12.256

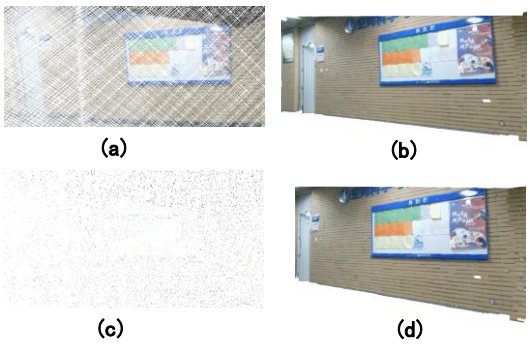

图 4 两方法在不同转速下成像对比效果图

Fig. 4 Comparison of imaging results of two methods at different rotational speeds

得了两种方法在不同旋转速度下的对比效果图，如图 4 所示。

如图 4 所示，在高速模式下，使用基于映射的融合方法可视化结果如（c）所示，在云台高速旋转时成像效果并不理想，因为获取的点云数量有限，导致点云密度低，整体质量较差。当云台在低速模式下运行时，如（a）所示，可实现更好的可视化效果，但帧率会降低至每帧 8.40 秒，这种帧率的降低削弱了构建地图的效率。

然而，通过本文 3.1 超像素改进雷视融合方法，使所有图像像素的三维位置都被计算出来补充到点云当中，这意味着点云增加的数量接近四台相机捕获的像素数量，点云密度和可视化效果对比都得到了增强，如图 4 中（b），（d）所示。因此，本文的方法在低速和高速模式下都能实现较高的点云密度，从而同时实现高成像分辨率和高帧率。此外，表 2 中展示了在对室内场景采集的一帧 3D-RGB 点云图像基于两种方法在两种模式下效果的定量统计对比数据，其中第三列 $Num(k)$ 为这两种方法在不同模式下获取的点云数量。可见超像素改进方法在低速模式下获取的点数是地图融合方法的 100 多倍，而在高速模式下获取的点数则超过了 2000 倍。这意味着在高速模式下，超像素改进方法也可以实现高成像分辨率。表中第四列 $\bar{d}$ 为点之间的平均距离，其是通过计算所有点的最近邻之间的平均欧氏距离得出的，较小的平均距离密度意味着点云更密集，对应更高的分辨率，表明本方法可以获得更密集的点云。第五和第六列的 E_r 与 E_p 分别显示了对室内场景内距离采集装置 5 米处物体的平均测量误差。随着物体与设备之间距离的增加，点云的平均点间距离逐渐增加，这表明点云的分辨率降低，点云变得更加稀疏。第四列的点云间平均距离表明超像素改进方法在两种模式下都能实现更密集的点云。第五和第六列的结果表明，超像素改进方法由于在测量过程中物体生成的点数更多，从而实现了更低的测量误差。

4.3.点云成像效果分析

为了验证本文设备结合基于超像素分割改进的雷视融合方法可以获取到更稠密的点云图像，有效减少由于点云稀疏、点间隔较大造成的测量误差，获取到更丰富的纹理信息，设计点云成像效果对比实验进行验

证。

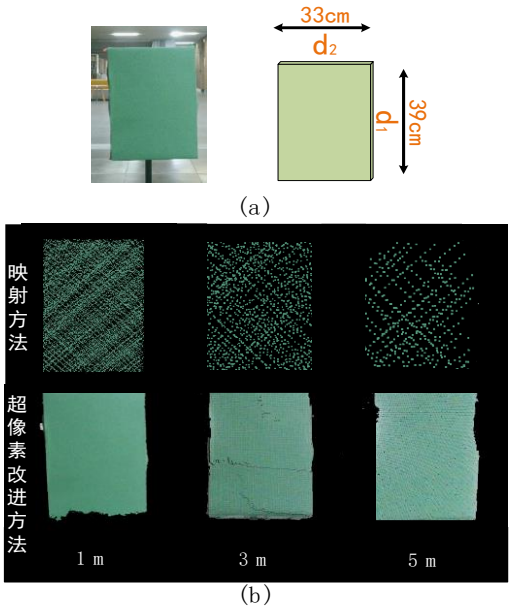

图 5 两种融合方法对平面点云成像效果

Fig. 5 Imaging results of two fusion methods on planar point clouds

首先对比本文给予超像素分割的雷视融合方法与基于映射的雷视融合方法成像效果，如图 5(a)所示，尺寸为 39cm×33cm 的平面被放置在设备前方。实验在距离范围从 1 米到 10 米的不同位置进行测量，每个距离共测量 50 次。图 5(b)展示了在 1,3,5m 的距离下采集到此平面的 3D-RGB 点云成像效果。

测量结果如图 6 所示，图 6(a)为使用两种方法在两种模式下从平面获取的点数量。显然，随着平面与设备之间距离的增加，点的数量会随之减少。然而绿色柱状图表明，即使平面距离设备较远，使用超像素改进方法获取的点数也显著更多，并且在获取的点数量始终大于10^4，能够充分的还原测量平面的纹理特征。

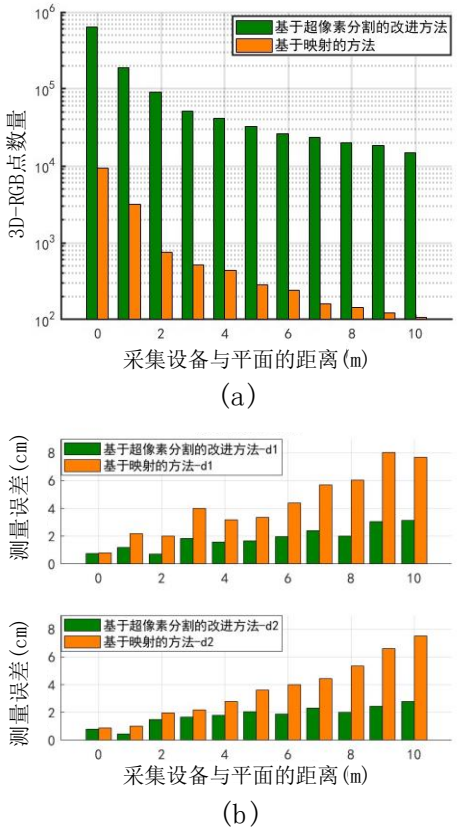

图 6 两种方法平面点数量及测量误差对比

Fig. 6 Comparison of the number of planar points and measurement errors between the two methods

图 6(b)则分别显示了对物体平面两边长度d_1，d_2的测量误差，可以看出随着平面与设备之间距离的增加，误差也随之增加。但对两种方法对比可以看出随着距离的增加，超像素改进方法的误差上升速度比地图融合方法更慢。这是由于超像素改进方法用来表示检测对象的点数量更多，具有更高的目标分辨率，所以在目标距离设备较远时，检测误差仍然较小。

此外，为了验证本文所提出设备与方法在室内区域建图中的优势，与目前对环境进行三维重建精度较高的多线激光雷达的成像效果进行对比。选择的多线激光雷达为 Velodyne VLP16，图 7 为其在平面上成像的效果，可见由于多线激光雷达激光束在俯仰方向上的发射角度保持不变，这导致其点云在平面上的分布呈现出离散的行状特征，这一特性限制了其在俯仰方向上的点云分辨率，随着探测距离的增加，平面上的激光点云数逐渐减少，从而导致难以精确获取平面的边缘和纹理信息。

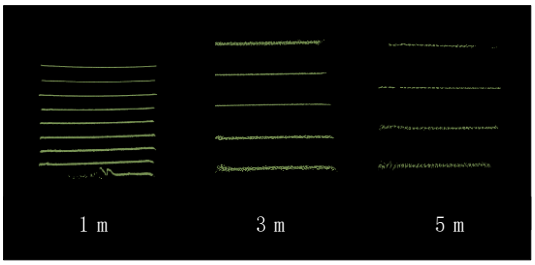

图 7 多线激光雷达平面点云成像效果
Fig. 7 Imaging results of planar point clouds from a multi-line LiDAR

表 3 本文方法与多线激光雷达对平面成像测量面积对比
Table 3 Comparison of planar imaging measurement area between the proposed method and multi-line LiDAR

	本文设备及方法		多线激光雷达	
	$S(cm^2)$	E_S	$S(cm^2)$	E_S
1m	1252.08	2.71%	890.16	30.83%
3m	1245.04	3.26%	1166.58	9.36%
5m	1203.28	6.51%	1011.60	21.40%

表3统计了本文设备及方法多线激光雷达基于不同距离对平面多次成像后的点云信息计算测量面积误差，其中 S 为基于点云边缘信息计算的平面面积，E_S 为测量面积与平面真实面积的相对误差。可见多线激光雷达对平面的成像面积误差受分辨率限制相对本文方法在 3m 与 5m 距离分别增加了 6.10%与 14.89%。此外，在距离平面仅 1m 时，由于激光雷达在俯仰方向存在较大的采集盲区，缺失平面上方部分信息，使得平面面积的测量误差达到了 30.83%。由此可见，多线激光雷达虽然具备较高的成像精度，但其在室内环境中对近距离的物体目标成像分辨率较低，证明了本文设备及基于超像素改进的雷视融合方法有效提升了在室内场景中的建图精度与视觉效果。

4.4.3D-RGB 地图高精度构建

对本文基于特征平面提取后进行点云帧间匹配构建的 3D-RGB 点云地图的精度进行验证。在此实验中，使用本文提出的设备即处理方法获取了大规模室内场景的 3D-RGB 点云地图，该室内场景为一个大约 100 米×12 米的矩形区域。建图效果如图 8 所示，第一行的子图展示了本文所用低成本 3D 激光雷达获取的不带颜色信息的 3D 点云地图，清晰地显示了地板和墙壁的结构。第二行的子图展示了 3D-RGB 点云地图。两张子图均准确捕捉到了建筑物的整体结构和颜色信息，验证了 3D 图像在全局一致对齐中的有效性。

为了进一步证明本文基于特征平面提取后基于粒子滤波的点云帧间匹配构建 3D-RGB 点云地图的精度优势，图 8 第一行中建筑物的长度和宽度被标注为四条线段：a、b、c 和 d，将这些线段的估计长度 d_r 与其实际长度 d_m 进行比较，结果汇总在表 4 中，可见地图构建的误差 E_{a} 始终在 0.15 米以内，且相对误差 E_{r} 小于 0.4%，这表明地图构建具有较高的准确性。

表 4 本文设备生成的 3D-RGB 地图测量误差
Table 4 Measurement errors of 3D-RGB maps generated by the device in this paper

S_d	$d_r(m)$	$d_m(m)$	$E_{\mathrm{a}}(m)$	E_{r}
a	8.37	8.3792	0.0092	0.110%
b	35.53	35.6612	0.1312	0.370%
c	9.86	9.8418	-0.0182	0.185%
d	21.08	21.0253	-0.0547	0.259%

并选择与改进的 LeGO-LOAM 方法[16]的精度进行比较。表 5 展示了经其处理后图 8 地图中 a、b、c 和 d 线段的精度。比较结果显示，我们的地图的平均相对误差为 0.231%，远低于改进的 LeGO-LOAM 映射方法[16]的 0.503%。

表 5 使用改进的 LEGO-LOAM 映射方法[11]测量 3D-RGB 地图的误差 Table 5 Measurement errors of 3D-RGB maps using the improved LEGO-LOAM mapping method [17]

S_d	$d_r(m)$	$d_m(m)$	$E_{\mathrm{a}}(m)$	E_{r}
a	8.33	8.3792	0.046	0.55%
b	35.44	35.6612	0.218	0.61%
c	9.86	9.8418	-0.034	0.35%
d	20.92	21.0253	0.105	0.50%

第三行子图对 3D-RGB 点云地图两个局部场景的点云进行放大展示，可见这些点的颜色信息为物体提供了丰富的纹理信息，这对地图构建、定位、导航以及目标识别有重要的应用价值。例如，墙上的画可以被识别出来，通过物体的颜色有助于区分和识别不同的目标，此外，特定的墙壁或物体也可以被识别，从而用于对采集车辆定位，获得

车辆的精确位置。

5. 结论

基于激光和视觉的室内三维重建技术应用需求，本文搭建了一种 3D-RGB 建图设备，用于获取大规模的 3D-RGB 地图。该设备由一个低成本的组建 3D 激光雷达、四个不同视角的相机、一个车辆和一个通信系统组成。并在对所采集点云数据在数据处理阶段通过基于超像素改进的雷视融合方法有效提升了点云密度、减小了三维点云地图的测量误差，再通过基于特征平面检测实现了高精度的点云全局一致性对齐建图。经实际实验验证，本文所提方法能够实现使用低成本的数据采集设备提供准确且稠密的 3D-RGB 点云地图，能够充分还原大范围室内场景及其内部的各类物体，具备较高的实际应用价值。

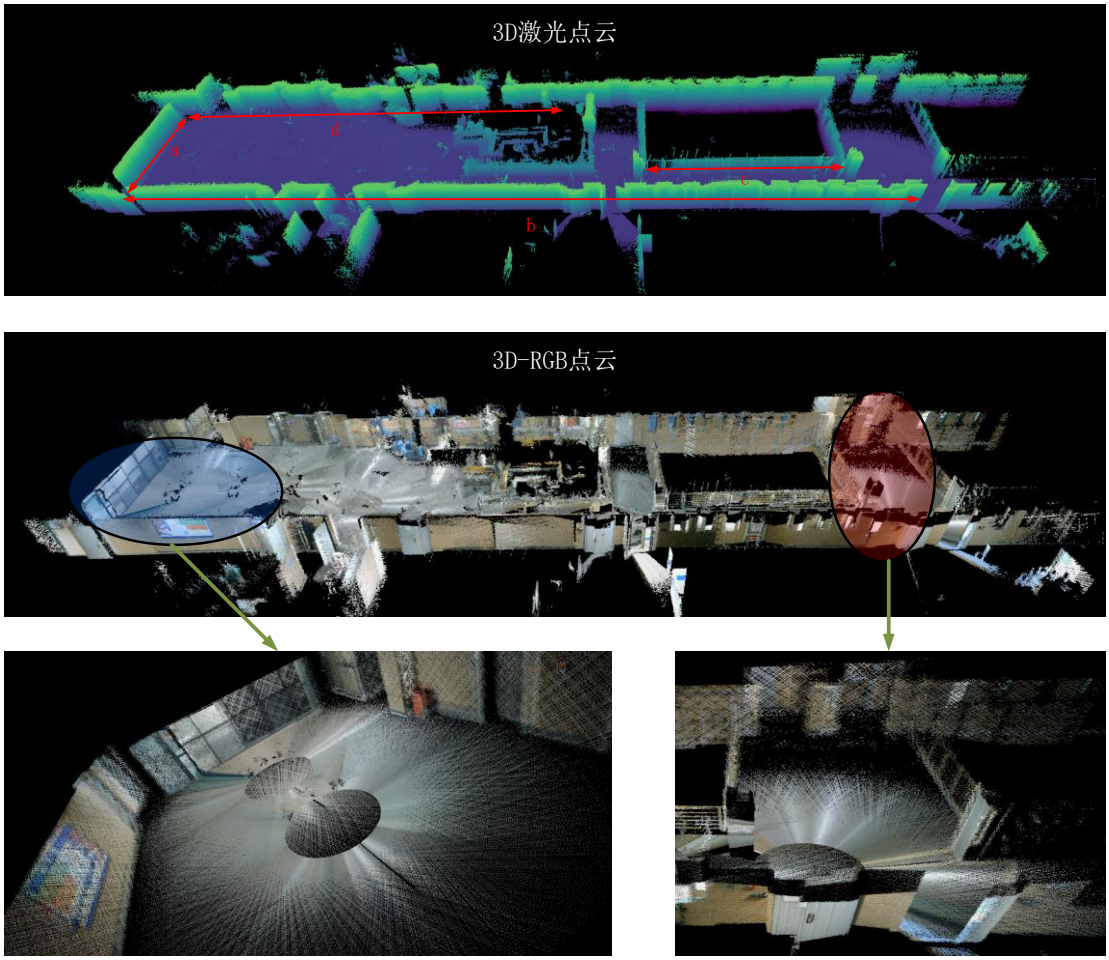

图 8 室内场景 3D-RGB 点云地图构建

Fig. 8 Construction of 3D-RGB point cloud map for indoor scene

参考文献：

[1] NEWCOMER, Robert, HARTLEY, Richard, and ROTH, S. "Real-Time Depth Camera Motion Tracking for 3D Reconstruction" [J]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012, pp. 123–130.

[2] 李嘉辉,范馨月,张干,张阔.基于背景修复的动态 SLAM [J].数据采集与处理,2024,(5):1204-1213

[3] ENGEL Jakob, KOLTUN Vladlen, and CREMERS Daniel, "Direct Sparse Odometry" [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(3): 611-625.

[4] 付燕平,严庆安,廖杰,等.基于彩色图像引导的 RGB-D 相机追踪与三维重建[J].武汉大学学报(工学版),2022,55(01):92-100.DOI:10.14188/j.1671-8844.2022-01-012.

[5] 王杰,魏振东,王启江,等.基于事件信息与深度学习的高动态范围三维重建[J].数据采集与处理,2024,39(02):337-347.DOI:10.16337/j.1004-9037.2024.02.007.

[6] FEHRMAN Brian, and McGOUGH Jason, "Depth Mapping Using a Low-Cost Camera Array" [C]. 2014 Southwest Symposium on Image Analysis and Interpretation, 2014: 101-104.

[7] WEN Li-Hong, and JO Kyung-Hwan, "Fast and Accurate 3D Object Detection for LiDAR-Camera-Based Autonomous Vehicles Using One Shared Voxel-Based Backbone" [J]. IEEE Access, 2021, 9: 22080-22089.

[8] 张成,杨耿,李钦,等.雷视融合自动联合标定方法研究[J].中国交通信息化,2024,(04):117-120+124.DOI:10.13439/j.cnki.itsc.2024.04.018.

[9] ZHAO Xiang, SUN Peng, XU Zhi, et al. "Fusion of 3D LiDAR and Camera Data for Object Detection in Autonomous Vehicle Applications" [J]. IEEE Sensors Journal, 2020, 20(9): 4901-4913.

[10] L. Mu, P. Yao, Y. Zheng, et al,“Research on slam algorithm of mobile robot based on the fusion of 2d lidar and depth camera,” IEEE Access, vol. 8, pp. 157 628–157 642, 2020.

[11] 师瑞卓,张小俊,孙凌宇,等.基于单线激光雷达的三维形貌重建方法研究[J].激光与红外,2022,52(02):188-195.

[12] YAN Bo, WANG Wei, YAN Yong, et al. "A Low-Cost 3D Imaging Device Using 2D LiDAR and Reflectors" [J]. IEEE Sensors Journal, 2023, 23(8): 8797-8809.

[13] 秦庆国,许录平,孙学荣,等.基于激光雷达的多帧联合目标跟踪系统[J].雷达科学与技术,2022,20(01):98-108+118.

[14] YU Wei, WANG Yan, LIU Hui, et al. "Superpixel-Based CFAR Target Detection for High-Resolution SAR Images" [J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(5): 730-734.

[15] ZHANG, Y., SONG, Q. "Image segmentation based on K-means clustering and multi-scale feature fusion" [J]. International Journal of Computer Applications, 2019, vol. 178, pp. 1–5.

[16] YAN Bo, GIORGETTI Andrea, and PAOLINI Enrico, "A Track-Before-Detect Algorithm for UWB Radar Sensor Networks" [J]. Signal Processing, 2021, 189: 108257.

[17] 崔洋,顾恒之,徐震.基于 LeGO-LOAM 的实际场景下的同步定位与建图方法[J].汽车实用技术,2023,48(01):44-47.DOI:10.16638/j.cnki.1671-7988.2023.001.009.

作者简介

田明昊(1999-)，男，硕士研究生，研究方向：雷达遥感探测及相关数据处理等，E-mail：22131214201@stu.xidian.edu.cn。

李攀（1991-），男，博士研究生，研究方向：目标探测，室内外定位导航等。

闫肃（2000-），男，硕士研究生，研究方向：目标探测，室内外定位导航等。

吴学良（2002-），男，硕士研究生，研究方向：雷达遥感探测及相关数据处理等。

许录平（1961-），男，教授，博士生导师，研究方向：先进导航技术及应用，精确制导与智能控制等，E-mail：xd203@aliyun.com。

阎博（1991-），**通信作者**，男，博士，硕士生导师，研究方向：雷达成像与目标识别，目标跟踪等，E-mail：boyan@xidian.edu.cn。