

基于多模态感知融合的自然环境精准三维重建

王鸿鹏^{1,2}, 韩霄^{1,2}, 曹仲之^{1,2}, 陈新伟^{3*}

(1. 南开大学人工智能学院, 天津 300350;

2. 南开大学智能技术与机器人研究所, 深圳 518000;

3. 闽江学院信息处理与智能控制重点实验室, 福建 350004)

摘要: 本文研究解决了自然环境高精度三维重建及效果优化问题。在基于图像单模数据三维重建的基础上提出了基于激光外点拒绝策略和视觉-惯性测量单元的轨迹修正的多模态融合三维重建效果优化算法。其中, 激光外点拒绝策略通过设置阈值去除图像的误匹配特征点; 视觉-惯性测量单元融合轨迹修正在无旋转增量场景中纠正机器人系统的运动轨迹, 解决了重建效果分层交叉问题。在增量式重建环节中的每次迭代优化中加入上述两种方法优化了重建效果。最后, 将多模态扫描仪采集到的异种多源信息, 包含激光雷达点云, 深度相机散斑点云, 彩色相机 RGB 图像, 惯性测量单元的加速度计数据和陀螺仪数据。对自然环境进行三维重建实验, 证明了算法的有效性, 鲁棒性和精准性。

关键词: 多传感器融合; 三维重建; 轨迹规划; 多尺度

中图分类号: TP242

文献标识码: A

Multi-modal Sensing Fusion based Accurate 3D Reconstruction of Natural Environment

Wang Hongpeng^{1,2}, Han Xiao^{1,2}, Cao Zhongzhi^{1,2}, Chen Xinwei^{3*}

(1. College of Artificial Intelligence, Nankai University, Tianjin, 300353;

2. Institute of Intelligence Technology and Robotic Systems, Shenzhen Research Institute of Nankai University, Shenzhen, China;

3. Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University), 350108

Abstract: This thesis studies and solves the problem of high-precision 3D reconstruction and effect optimization of natural environment. Based on the 3D reconstruction of single-mode image data, a multi-mode fusion 3D reconstruction effect optimization algorithm based on laser external point rejection strategy and vision inertial measurement unit trajectory correction is proposed. The laser outlier rejection strategy removes the mismatched feature points by setting a threshold; Vision inertial measurement unit fusion trajectory correction corrects the trajectory of robot system in the scene without rotation increment, and solves the problem of hierarchical cross reconstruction effect. These two methods are added to each iteration of incremental reconstruction to optimize the reconstruction effect. Finally, the heterogeneous multi-source information collected by the multi-mode scanner includes LIDAR point cloud, depth camera speckle cloud, color camera RGB image, inertial measurement unit accelerometer data and gyroscope data. Experiments on 3D reconstruction of natural environment show that the algorithm is effective, robust and accurate.

Keywords: multi sensor fusion; calibration of external parameters; 3D reconstruction; trajectory planning; Multi-scale

1 引言 (Introduction)

多传感器融合及三维重建是机器人领域中关键技术点。对于机器人多传感器融合及三维重建的相关探索, 涵盖了计算机视觉, 计算机图像学, 计算机通信学, 车辆工程学, 控制最优化论等相关学科。同时, 由于计算机图形学和深度学习等在近年来的

突破。使得机器人及无人系统在感知领域也获得了很大的进步, 例如在识别和分类行人, 车辆, 车道旁垃圾桶等障碍物, 以及对于可行道路和车道线的检测, 红绿灯检测等相关关键性的检测有着待研究方向的提高和突破, 采用的感知传感器主要包括激光测距雷达, RGB 彩色相机, 深度相机, 惯性测量

基金项目: 国家自然科学基金(61973173, 91848203), 天津市科技重大专项(20YFZCSY00830), 天津市科技重点研发计划(18ZXZNGX00340)

通信作者: 陈新伟, chenxinwei@126.com

单元等。

深度相机与激光测距雷达是采用主动测距方式，深度相机常用于室内体感游戏以及人脸识别，获得小尺度环境的散斑点云。激光测距雷达常用于室外，获得大尺度环境的点云。常用的激光测距雷达能获得 360 度的环境信息，常用在自动驾驶中。激光雷达与深度相机都能获得点云信息，其中激光雷达的点云信息利用 TOF 原理，而深度相机的点云信息采用红外发射器投射散斑点，由红外接收器接收。二者都会因环境中灰尘，雾霾而导致测量距离失准。RGB 彩色相机是最常用的图像传感器，价格低廉，技术成熟，基于纯视觉的算法琳琅满目。但由于单目相机难以直接获得深度信息，因此搭配主动测距的激光雷达或深度相机来弥补距离信息，或使用立体视觉，单目多视角等计算机图形学知识，通过三角测量来估计深度信息。

机器人系统技术研究的核心是多种传感器的联合使用，将不同种类传感器采集来的数据，即多模态信息进行融合，从而实现感知任务或定位任务。如激光雷达，惯性测量单元，彩色摄像机，深度摄像机，融合不同多源传感器信息并进行实时检测，定位，以及真实场景的识别。然而，为了便于科学研究，本文将重点放在核心部分和关键技术。因此，本文提出了一种多模态感知融合系统是研究无人系统的核心技术。不同类型的传感器，即不同的模式，将提供不同类型的信息。同时，不同的传感器在不同的范围内工作。因此，这是非常重要的多模态信息融合的实用性和必要性传感器。例如，RGB 相机和深度相机高分辨率，可达到像素级。深度照相机可以弥补 RGB 显示的深度距离信息相机没有。然而，深度相机采用结构光原理，工作范围较小因此，为了弥补距离远处的信息。

基于异种多源传感器模态数据信息融合及三维重建任务的技术研究是各个领域建立无人系统所不可或缺的重要组成部分。

移动多模态感知融合系统的领域是结合了机器人学，车辆工程学，计算机通信，计算机视觉，计算机图形学，最优化等领域的交叉结合领域。在移动无人系统中（包括无人机和无人车），需要考虑所工作的环境因素，续航，以及对光照等因素的鲁棒性。在军事领域，工厂加工，物流快递等各个领域都需要无人机或无人车来代替人们执行基本的搬运工作。

三维重建是计算机视觉，机器人领域，医疗领域的十分重要的研究方向，人们生活在三维空间中，接触到的也是三维的各种物体。相对于二维图像信息，三维模型的真实感更加强烈，能够呈现更多的信息。尤其是，随着虚拟现实和增强现实技术的发展，可以明显的感到，三维重建带给我们生活的便利。此外，在近十年来，最先兴起的是语音处理，然后是二维图像的识别，检测。并且这些领域都取得了显著的成果。然而，相关研究使用相机所获得的二维图像，是由真实的三维物体投影到相机平面的数据，损失的深度信息。直接处理三维数据，减

少信息在源头的损失，是最直接也是最有效的解决问题的方法。虽然近些年来，硬件的快速发展，使得内存，中央处理单元等硬件功能强大，但对于三维的数据来说仍然是一个挑战。三维环境的数据更加复杂，数量级更大。对于三维环境的数据处理既需要满足实时性，又需要满足准确性。因此，对三维重建的算法研究和改进还有非常多亟待解决的痛点需要解决。

重建三维环境的数字模型已经成为机器人领域，旅游导航和地图绘制等很多领域应用中重要的任务，且与人民的生活息息相关。重建三维模型的关键问题是有效的融合来自异种多源传感器模态信息数据。相机通常提供颜色和纹理信息，但是不能直接提供深度信息，即景深。而深度信息可以由测距传感器进行直接采集，例如，激光雷达和深度相机等主动测量仪器，以及为了与三维的点云信息融合。一些现有的研究工作使用多个相机来满足环境的 360 度场景全覆盖[20]。

本文研究工作基于移动机器人为载体的环境下，对室外非结构化环境进行传感器多模态信息的感知，识别，三维重建以及在三维重建的基础上进行与现实环境的虚实结合。

2. 多模态数据融合

(Multi-model reconstruction)

自然环境中的实景重建不仅要求重建算法鲁棒性高，并且由于非结构化环境的复杂性，对三维重建算法的实时性和重建效果提出了较高的要求。

单目增量式从运动恢复结构是对无序的图像集合进行三维重建的一种流程式的策略，其本质是单目多视角。至今，经典的从运动恢复结构在各种场景下都有应用，且有巨大的进步。但是它的鲁棒性，完整性，实时性和可扩展性仍然是增量式从运动恢复结构的三维重建面临的关键问题。具体来说，经典的从运动恢复场景的结构是对无序的 RGB 图像集合进行暴力搜索，特征匹配。可以减少人工的干预，但是，不能够满足实时性，因此，增量式从运动恢复结构的三维重建常用于离线的三维重建。

2.1 多模态数据融合

2.1.1 惯性测量单元数据

惯性测量单元的数据可表示为：

$$\omega_t = \omega_m - b_a - \eta_\omega \quad (1)$$

公式（1）中，下标 t 和 m 分别表示真实值和测量值，g 表示重力的矢量， η_ω 表示惯性测量单元中的陀螺仪的噪声， b_a 表示惯性测量单元中的加速度计的噪声。在第三部分中使用 IMU 与视觉图像进行融合。其基本思想是在多视角运动恢复结构增量式重建环节中增加 VIO 融合轨迹修正环节。惯性测量单元高频采样和精准位姿估计的优点，来抑制图像模糊和矫正图像之间错误的相对运动

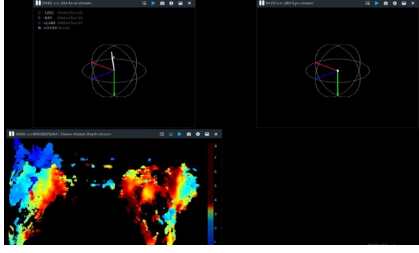


图 7 深度相机数据信息

Fig.7 depth camera data information

2.1.2 RGB 数据

$$\begin{bmatrix} U \\ V \\ 1 \end{bmatrix} = \begin{bmatrix} K_u & 0 & U_0 \\ 0 & K_v & V_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2)$$

公式(2)表示本研究使用的 BGR 数据信息所在的针孔模型,其数据表示为二维的 UV 平面。其中, U 和 V 是三维空间中点投影在图像平面中的坐标, X , Y 和 Z 是 3D 世界坐标系中的三个互相垂直的坐标系轴, k 是像素研制像素坐标的比例因子, k 常取值为 1。 U_0 和 V_0 是相机中心的像素坐标。 f 是相机的焦距。



图 8 RGB 相机数据信息

Fig.8 RGB camera data information

2.1.3 激光雷达数据表示

激光雷达是一种光学遥感技术,测量物体到激光雷达的距离和角度,通过激光雷达发射的激光脉冲和接受的反射脉冲之间的时间间隔来计算激光雷达与物体之间的距离。即 TOF 原理,应用于创建高分辨率的三维重建效果图,以及同时建图与定位领域中的建图。

其数据表示为:

$$x = d \cos(\phi) \quad (3)$$

$$y = d \sin(\phi) \quad (4)$$

$$Z = \theta \quad (5)$$

其中, d 为激光雷达到被扫描点之间的距离, ϕ 是光束角,注意, x 和 y 是激光雷达局部帧中扫描点的坐标,而不是相对于全局参考帧的坐标,想要算出全局坐标,必须讲局部坐标转换为全局帧。 θ

表示多线激光中,不同线数与水平面的夹角。

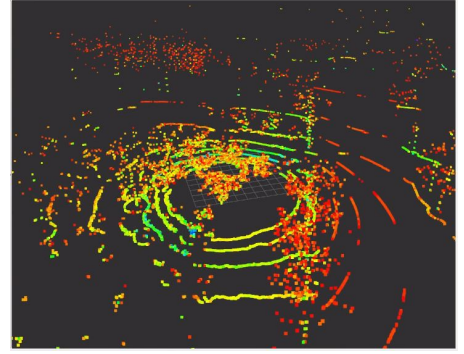


图 9 激光雷达数据信息

Fig.9 LiDAR data information

3. 多源点云数据融合 (Multi source point cloud data fusion)

3.1 多源点云自适应多尺度切换算法

如图 10 所示,为多源点云多尺度重建切换原理示意图。以半径为 10m 的圆锥形。由深度相机放置于底部,因其工作距离为 0.2m-10m。可弥补激光雷达盲区的感知信息,实现自然环境的全场景感知和重建。

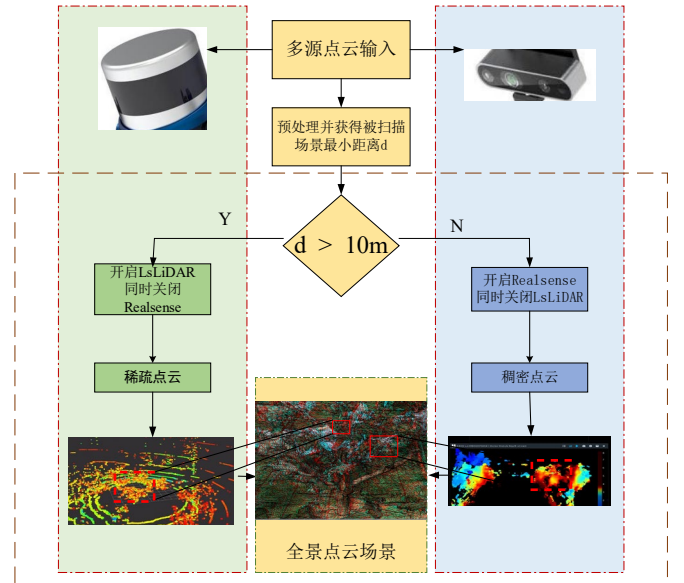


图 10 点云测距组合:多线激光+深度相机

Fig 10 point cloud ranging combination: multiline laser + depth camera

$$D_{obj} = w_1 D_L + w_2 D_D \quad (6)$$

公式(6)中所示,表示自适应多尺度切换算法。其中, w_1 和 w_2 表示激光雷达测距项与深度相机测距项的权重。当满足式(7)时, $w_1=0, w_2=1$ 。自动切换只有深度相机进行测距,关闭激光雷达测距功能,获取小尺度环境测距信息。同理当满足式(8)时, $w_1=1, w_2=0$ 。自动切换为只有激光雷达进行测距,关闭深度相机测距,获取大尺度环境测距信息。

$$w_1 = \begin{cases} 0, & \text{if } (x^2 + y^2 < 100) \\ 1, & \text{otherwise} \end{cases} \quad (7)$$

$$w_2 = \begin{cases} 1, & \text{if } (x^2 + y^2 \geq 100) \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

图（11）所示为点云与影像融合的三维重建算法流程图。包括数据关联环节和重建结果环节。图像经过三角测量提取到特征点，进行特征点之间的匹配关系。多源点云（来自多线激光雷达和深度相机）通过特征提取和匹配，经过注册配准，使用点面对之间的数据关联来匹配点云之间的关系。最后视觉匹配的特征点和点云匹配好的特征点一同作为重建输入。

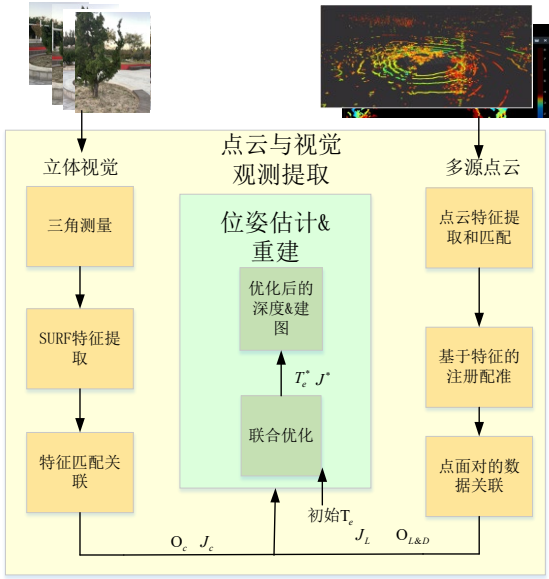


图 11 点云影像融合的三维重建算法流程

Fig 11 3D reconstruction algorithm flow of point cloud image fusion

3.2 基于点云异常值拒绝策略重建优化

使用激光雷达来拒绝异常点（即，拒绝策略）：如图所示，由于系统运动速度过快，光照等因素，导致视觉模糊，特征点检测匹配不正确。图中标识出来的两部分区域应为同一个对象，但是显示为不正确的对齐。因此，针对这种情况，我们提出了一致性比率 r 来量化源点云和目标点云的一致性测量，如式（9）：

$$r = \sum_{c \in \{image\}} I(g) - \sum_{c \in \{LiDAR\}} I(g) \quad (9)$$

其中， $I(*)$ 表示标识函数，如果其中的运算为真，则 $I(*)$ 赋值为 1，否则为 0。本文通过在同一个尺寸大小的区域内，将点云特征与图像特征进行匹配，在匹配的点云图像特征为中心，对多线点云设定阈值为 5m。点云特征在 5m 外的图像特征点标记为外点，增量式重建环节中的每次循环迭代都使用激光点云的特征对图像提取到的特征进行外点拒绝，即异常值拒绝策略。定义测量指标 r 作为定量分析指标。并且为其设定阈值，

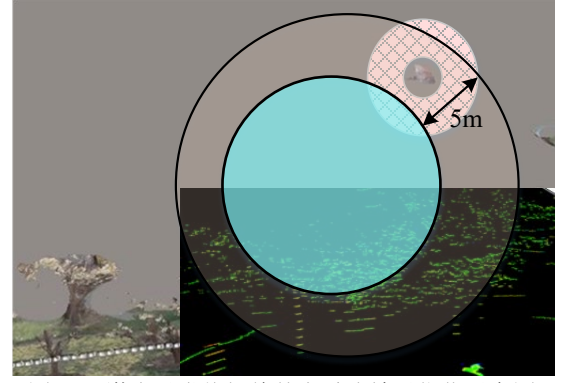


图 12 激光异常值拒绝外点重建效果优化示意图

Fig 12 Optimization of reconstruction effect of outliers rejected by laser

当 r 小于 5m 时，则视点云图像匹配的特征点云为内点，即有效点。否则，图像中的点云视为无效点，及时去除。不参与下一次迭代的增量式注册。

在提出的方法中，激光雷达加强增量式的 SFM 体现在以下两方面：

1、使用激光雷达点云去除图像点云中的特征匹配外点，即图像重建点云中与特征点距离大于阈值的杂点。激光雷达的精确点云数据信息可以检测和排除无效的图像匹配，使得多模态系统鲁棒性更强，并且降低三维重建失败的概率。

2、使用 3D 激光扫描仪，进行回环检测。纠正纯图像增量式 SFM 在大环境下由于累计误差产生的轨迹估计漂移。正如，第二章所提到的外参联合标定算法中，对相机和激光的重投影误差的定义。这里本文简化的表述激光点云和 RGB 图像两种传感器数据信息的重投影误差：

相机的重投影误差定义：

$$E_c = Q(x; T_i^{i+1}, P) - u \quad (10)$$

图像的观测误差损失函数定义为图像上特征点坐标在图像平面上的投影点即通过在三角测量得到的投影点与观测点之间的差值。其中， E_c 表示相机观测到特征点的重投影误差， $Q(*)$ 表示估计空间三维点映射在图像平面的投影， u 表示已经观测特征点集到的图像平面的投影。

$$x = d \cos(\phi) \quad (11)$$

$$\Delta T = 0 \text{ \& } \Delta R \neq 0, \forall M_i, i=1,2,3...N \quad (12)$$

$$Z = \theta \quad (13)$$

激光雷达的观测误差定义为点到面片的距离。具体阐述如下：首先，激光雷达的观测数据定义为匹配的关键点 P ，三维点。该关键点选取在局部面

片中，而该局部面片是由多模态系统的两个被筛选过的相对运动中生成的。这里的 Y 是 p 在两个相对运动中的邻 n 是局部面片的法向量。激光雷达的观测误差定义为点到面片的距离。

$$E_l = L^T (\tau(p; T) - y) \quad (14)$$

$$E_j = n^T (\tau(x; T_e^{-1} T_i^{-1}) - Y) \quad (15)$$

其中，函数 $\tau(\bullet)$ 为变换函数，将一个激光点使用给定的变换矩阵进行变换。给定一个初始化位姿 T_e 随机值，在联合优化的迭代优化环节中会进行迭代优化得到最优估计。具体而言，本文从每个系统位置下的点云中取出 3000 个左右的特征点，并从 5m 范围内的位置中进行搜索最近邻的位置，形成局部面片。

联合误差定义：

$$\arg \min \lambda_e \Sigma e_e^2 + \lambda_l \Sigma e_l^2 + \lambda_j \Sigma e_j^2 + \lambda_j \Sigma e_j^2 \quad (16)$$

本文将联合误差定义类似于激光雷达的检测定义，将点面对的距离作为其误差如式 (16) 所示。显然，激光雷达的误差表示不用改变，而对于在相机中检测到的特征点，但不属于激光雷达的点云的特征点，以匹配的激光雷达点云为中心，设置阈值。阈值以外的图像特征点视为外点，即异常值；小于阈值的图像特征点视为内点，即为激光-图像配准的特征点。将激光-图像配准的特征点作为过滤后的点云进行点云匹配，参与面片生成，纹理贴图等环节。

当优化过程收敛的时候，利用误差作为指标过滤观测值，并观测的集合中去除掉误差大于阈值的观测值。在多模态感知融合系统中，本文使用的阈值为：相机是 4 像素；激光雷达是 0.1m；联合观测的阈值：0.1m。当去除了异常值之后，再次进行损失函数的求解

由于是使用最近邻搜索算法找到激光雷达和联合观测的下一帧，并增量式扩展，因此，初始化十分重要。一旦获得更新，就应立即进行解算，重新计算激光雷达和联合观测的观测集合。使用列文-麦夸特算法对泰勒公式进行二次项展开并进行求导迭代优化，经实验验证，最终经过 5-8 次的迭代，优化结果会区域收敛，得到相应的联合优化结果。

4. 基于惯性测量单元与多线激光重建效果优化 (Optimization of reconstruction effect based on inertial measurement unit and multi-line laser)

在机器人和智能系统中，需要得到高精地图，

地标性建筑的三维重建模型，为避障，导航和识别等后续环节提供坚实的基础。其中，室外和大型工厂的重建中，使用大视角，高分辨率的相机可以提供清晰地数据和丰富的纹理。因此，基于图像的三维重建成为热门的研究点，其中以单目多视角为代表，即增量 SFM。虽然具有重建精度高，对特征匹配外点鲁棒性高和捆绑调整能全局优化三维场景结构等优点。然而，仍然存在着很多缺点。其中主要有三方面问题如下：

1、对于大型的工厂或野外环境的三维重建中，由于快速的运动或光照等因素，导致产生累计误差导致场景漂移，出现重建结果错位等问题。

2、大量的进行捆绑调整，效率低，时间复杂度高，远不能满足实时性。

3、最终重建结果与初始化图像视点对的选取，视点对的选取与添加顺序十分敏感。在增量式的从运动恢复结构的三维重建过程中，顺序添加视点要比乱序添加视点低一个数量级的时间复杂度。



图 10 纯视觉快速运动产生重建效果分层

Fig.10 pure visual fast motion produces layered reconstruction effect

如图 10 所示，在进行三维重建的实验中，在转角处，由于只存在旋转增量而没有平移增量。使得纯视觉的面片生成产生分层，重建失败。

$$\Delta T = 0 \text{ \& } \Delta R \neq 0, \forall M_i, i = 1, 2, 3 \dots N \quad (17)$$

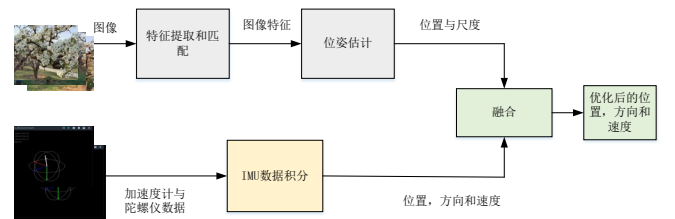


图 11 视觉-惯性测量单元融合运动估计

Fig 11 motion estimation based on vision inertial measurement unit fusion

如图 11 所示为视觉与惯性测量仪融合流程，对于视觉传感器的处理，先根据相机获得的图片进行

特征提取和匹配，得到图像特征。进而根据图像的特征进行粗略运动估计，由于图像的运动估计受到光度一致性不连续影响，例如光照，水潭，湖面等自然环境，以及在使用基于图像的增量式运动恢复结构必须有平移的增量时才会得到较为准确的运动估计，而对于如图所示的拐角处，由于没有平移增量而只有旋转增量，使得图像的特征匹配出现误差，导致最终的重建效果出现分层和交叉，三维重建的模型生成失败。因此，基于视觉-惯性测量单元的融合，正是结合了惯性测量单元 300HZ 的高频采样特性，快速的获取陀螺仪和加速度计数据，通过二次积分得到平移运动增量和角度增量。弥补图像重建失灵的情况。

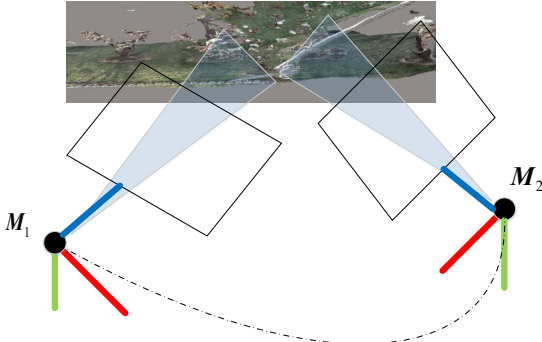


图 12 惯性测量单元+视觉运动估计

Fig.12 IMU + visual motion estimation

如图 12 所示，在梨树园进行实验时，当多模态系统运动轨迹从由北向南移动突然转换为由西向东时，由于没有平移上的增量，在只有旋转增量没有平移增量时，并且由于树底的土坡松软，导致多模态系统在运动时发生抖动。此时，使用多模态系统中的视觉-惯性测量单元融合，得到融合优化后的位置，方向和速度。

$$\min_{\mathbf{N}} f_1 + f_2 + f_3, \text{其中:}$$

$$f_1 = \rho(\|r_p - J_p \mathbf{N}\|^2)$$

$$f_2 = \omega_{IMU} \sum_{i \in B} (\|r_b(z_{bi, bi+1}, \mathbf{N})\|^2) \quad (18)$$

$$f_3 = \omega_{image} \sum_{(i,j) \in F} \rho(\|r_f(z_{ij}^c, \mathbf{N})\|^2)$$

如式 (18) 所示为使用基于滑动窗口的视觉-惯性测量单元的融合优化目标函数，其中，第一项表示惯性测量单元的预积分量，第二项表示惯性测量单元的误差项，第三项表示图像的重投影误差项。将三项之和的最小化作为损失函数。其中， ω_{IMU} 代表惯性测量单元的运动估计权重， ω_{image} 代表图像数据进行运动估计时的权重。当机器人系统执行向前或向后等具有平移增量的移动时，惯性测量单元与图像运动估计的权重各位 0.5。当机器人系统在拐角

处执行扫描任务时， ω_{IMU} 取值为 1， ω_{image} 取值为 0。此时，只有惯性测量单元进行运动估计，关闭图像数据的运动估计。

使用惯性测量单元时，需要进行预积分。如式 (18) 第一项所示，该预积分量仅仅与惯性测量单元的测量值有关，预积分量的获取时将一段时间内的惯性测量单元数据直接积分获得。

在计算时，为了减少计算量，尽可能达到实时性，采用滑动窗口形式的捆绑调整。其中滑动窗口内需要优化的系统状态量定义如下：

$$\dot{\mathbf{A}} = [\mathbf{X}_n, \mathbf{X}_{n+1}, \mathbf{X}_{n+2}, \dots, \mathbf{X}_{n+N}, \boldsymbol{\lambda}_m, \boldsymbol{\lambda}_{m+1}, \dots, \boldsymbol{\lambda}_{m+M}] \quad (19)$$

$$\mathbf{X}_i = [p_{wb_i}, q_{wb_i}, V_i^w, b_a^{b_i}, b_g^{b_i}]^T, i \in [n, n+N] \quad (20)$$

式 (19) 和式 (20) 中， \mathbf{X}_i 包含 i 时刻惯性测量单元机体的位置，速度，姿态以及惯性测量单元中的加速度和角速度的偏置量估计。 n, m 表示视觉与惯性测量单元的状态量， landmark 在滑动窗口内的起始时刻。 N 表示滑动窗口中的关键帧的数量。 M 表示所有关键帧观测到的 landmark 的数量。

因此，在三维重建的拐角处，使用如上的视觉-惯性测量单元融合轨迹估计，矫正运动。使得三维重建系统更具有鲁棒性。通过实验，验证了算法的有效性。

基于多线激光与惯性测量单元增强影像单模数据重建的流程如图 13。整个流程分为三部分，分为多模态数据的采集和预处理，相关性搜索匹配和增量式运动恢复重建。最后输出三维重建结果。

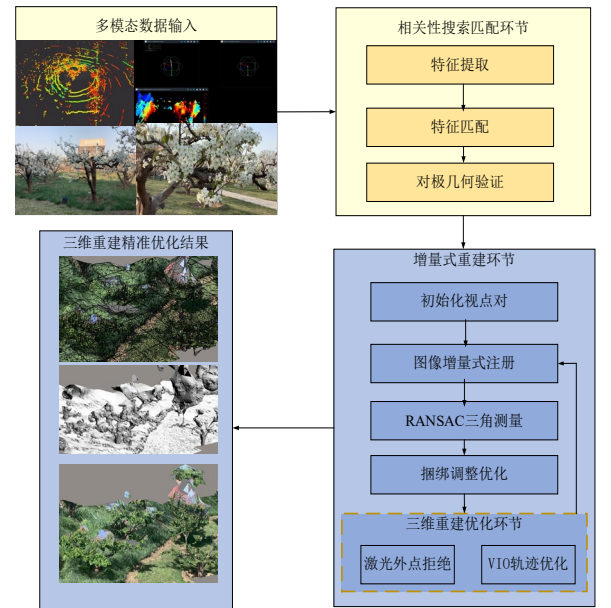


图 13 多模态系统 SFM 重建效果优化流程

Fig.13 SFM reconstruction effect optimization process of multimodal system

第一阶段是预处理。在多模态数据采集集中，使用时间戳对齐方法物理对齐多传感器数据信息。将

图像数据输入相关性搜索匹配模块，使用 SIFT 算法提取特征，进行特征匹配，使用多传感器模态系统中的立体视觉进行极线约束，求解本征矩阵，获得图像传感器内参。

第二阶段是改进增量式运动恢复结构重建。首先，与单目多视角运动结构估计相同，用五帧图像进行初始化三角测量，得到结构的初始值。然后进行图像的增量式注册，使用基于 RANSAC 算法的三角测量。最后进行基于重投影误差的捆绑调整，在计算迭代优化之后，由于纯图像的重建时，传感器模态信息单一，并且有轨迹估计偏移，图像模糊的问题。采用激光雷达进行特征匹配外点的筛选，以及使用惯性测量单元在运动中，光度突发不一致时使用惯性测量单元进行代替三角测量轨迹估计。

5 实验 (Experiments)

基于本文中提出的外参联合标定算法和面向自然环境的基于多模态增量式 SFM 精准三维重建算法，分别设计了外参联合标定实验和三维重建效果优化实验。后续可空地协同更加优化三维重建。

5.1 三维重建实验 (Experiments)

为了验证将 VIO 融合定位用于重建效果优化环节，实验选取了梨树园作为实验场景，针对拐角处重建生成的模型分层问题，进行实验。采用的机器人系统是本文设计的移动机械臂多模态感知融合机器人系统。装配了 16 线 360° 激光扫描测距仪，Realsense D435i 深度相机，远景立体视觉和近景立体视觉，其中惯性测量单元采用 Realsense 深度相机中自带的惯性测量单元。激光的数据获取频率为 20HZ，远近景立体视觉采用的相机数据获取频率为 30HZ，惯性测量单元的数据采集频率为 300HZ。上位机处理器是 16GB 的 RAM，Intel (R) core (TM) i7-8700kcpu，gtx1070。测试环境选用 Ubuntu LTS16.04，编程语言选用 C/C++语言。

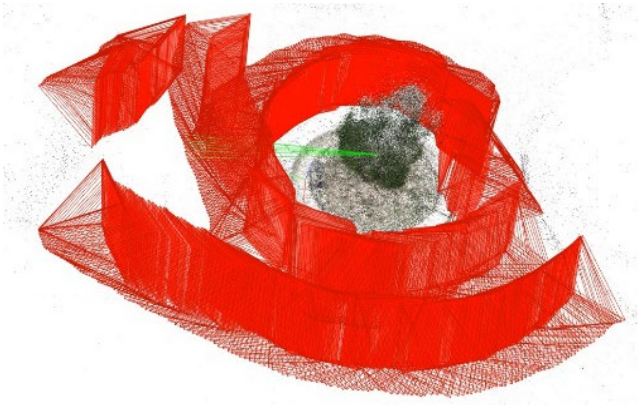


图 14 使用 COLMAP^[16]对梨园松树三维重建
Fig.14 3D reconstruction of pear trees using COLMAP^[16]

如图 14 所示，使用图像单模数据对梨树园进

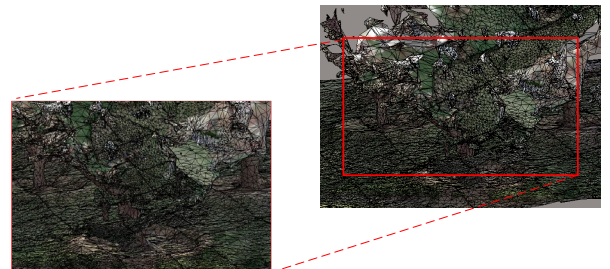
行三维重建效果展示。

由图 15 可得，机器人系统在拐角处，旋转时，由于没有平移增量，导致重建模型出现分层，使得三维重建实验失败。因此，采用视觉-惯性测量单元融合方法，得到精确的机器人系统运动估计。最终修正了三维重建的效果，满足实验需求。将优化前的轨迹与优化后的轨迹可视化如图 15 所示，验证了算法的有效性和实用性。



图 15 图像单模态数据拐角处三维重建交叉分层
Fig.15 Cross slicing of 3D reconstruction at corner of image single mode data

运动恢复结构重建需要注册环节中机器人系统运动具有平移增量以供运动估计。然而，机器人系统在拐角处，由于机器人系统此时的运动没有平移增量，只有旋转增量，导致重建模型出现交叉分层，重建模型无法使用，如图 16 所示。因此，采用视觉-惯性测量单元融合方法，得到精确的机器人系统运动估计。最终修正了三维重建的效果，满足实验需求。将优化前的轨迹与优化后的轨迹可视化如图 17 所示，验证了算法的有效性和实用性。



(a) 图像单模重建效果分层示意图



(b) VIO 轨迹矫正重建效果优化示意图

图 16 积水形成的朗博面，形成重建空洞
Fig.16 Motion path correction based on VIO fusion

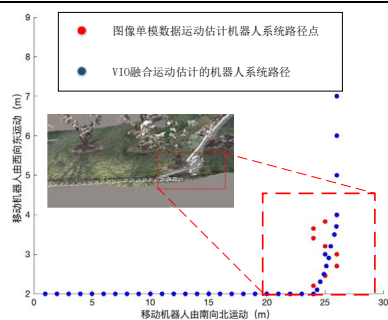


图 17 视觉-惯性测量单元融合轨迹矫正
Fig 17 Vision inertial measurement unit fusion trajectory correction.



图 16 使用激光雷达点云补点生成面片
Fig 16 using LiDAR point cloud patch to generate patch

6 结论 (Conclusion)

本研究中, 融合多传感器数据模态信息进行了自然环境的精准重建。首先, 提出了多模态数据融合的方法。然后, 提出了激光拒绝图像异常值策略。最后, 提出基于激光与 IMU 融合增强的三维重建效果优化算法, 并设计实验验证其有效性。

参考文献 (References)

- [1] Chen C Y and Chien H J. "On-Site Sensor Recalibration of a Spinning Multi-Beam LiDAR System Using Automatically-Detected Planar Targets," *Sens.*, vol. 12, no. 10, pp. 13736-13752, 2012.
- [2] LI J, HE X and LI J. "2D LiDAR and Camera Fusion in 3D Modeling of Indoor Environment," in *IEEE National Aerosp. Electron. Conf.*, 2015, pp. 379-383.
- [3] Chai Z, Sun Y and Xiong Z. "A Novel Method for LiDAR Camera Calibration by Plane Fitting," in *Proc. Int. Conf. Adv. Intell Mechatron.*, 2018, pp. 286-291.
- [4] Yohannan B and Chandy D A. "A Novel Approach for Fusing LIDAR and Visual Camera Images in Unstructured Environment," in *Proc. Int. Conf. Adv. Comput. Commun. Syst.*, 2017, pp. 1-5.
- [5] Mirzaei F M, Kottas D G, Roumeliotis S I, et al. "3D LIDAR - Camera Intrinsic and Extrinsic Calibration: Identifiability and Analytical Least-Squares-Based Initialization," *Int. J. Rob. Res.*, vol. 31, no. 4, pp. 452-467, 2012.
- [6] Guindel C, Beltrón J, Martín D, et al. "Automatic Extrinsic

Calibration for Lidar-Stereo Vehicle Sensor Setups," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.* pp. 1-6, 2017.

[7] Pandey G, McBride J R, Savarese S, et al. "Automatic Extrinsic Calibration of Vision and Lidar by Maximizing Mutual Information," *J. Field Rob.*, vol. 32, no.5, pp. 696-722, 2015.

[8] Schneider N, Piewak F, Stiller C, et al. "RegNet: Multimodal Sensor Registration Using Deep Neural Networks," in *Proc. IEEE Intell. Veh. Symp.*, pp. 1803-1810, 2017.

[9] Willis A R, Zapata M J and Conrad J M. "A Linear Method for Calibrating LIDAR-and-Camera Systems," in *Proc. IEEE Int. Symp. Model. Anal. Simul. Comput. and Telecommun. Syst.*, pp. 1-3, 2009.

[10] Kasper R and Schmidt S. Sensor-Data-Fusion for an Autonomous Vehicle Using a Kalman-Filter," in *Proc. IEEE Int. Symp. Intell. Syst. Inf.*, pp. 1-5, 2008.

[11] Man Q, Dong P and Guo H. Pixel and Feature-Level Fusion of Hyperspectral and Lidar Data for Urban Land-Use Classification," *Int. J. Remote Sens.*, vol. 36, no. 6, pp. 1618-1644, 2015.

[12] Fankhauser P, Bloesch M, Rodriguez D, et al. "Kinect v2 for Mobile Robot Navigation: Evaluation and Modeling," in *Proc. IEEE Int. Conf. Adv. Rob.*, pp. 388-394, 2015.

[13] Guindel C, Beltrón J, Martín D, et al. "Automatic Extrinsic Calibration for Lidar-Stereo Vehicle Sensor Setups," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.*, pp. 1-6, 2017.

[14] Ishikawa R, Oishi T and Ikeuchi K. "Lidar and Camera Calibration Using Motions Estimated by Sensor Fusion Odometry," in *Proc. IEEE/RSJ Int. Conf. Intell. Rob. Syst.*, pp. 7342-7349, 2018.

[15] Nagy B, Kovács L and Benedek C. "Online Targetless End-to-End Camera-LiDAR Self-Calibration," in *Proc. Int. Conf. Mach. Vision Appl.*, pp. 1-6, 2019.

[16] Schonberger J L, Frahm J M. Structure-from-motion revisited[in *Proc. IEEE Int. Conf. (CVPR)*. pp.4104-4113,2016:

作者简介:

王鸿鹏 (1979--), 男, 博士, 副教授。研究领域: 机器人与智能系统, 虚拟现实与智能仿真。

韩 霄 (1995--), 男, 硕士, 研究领域: 机器人与智能系统。

曹仲之 (1998--), 男, 硕士, 研究领域: 机器人与智能系统。

陈新伟 (1984--), 男, 博士, 副教授。研究领域: 机器人与智能系统, 工业视觉系统
