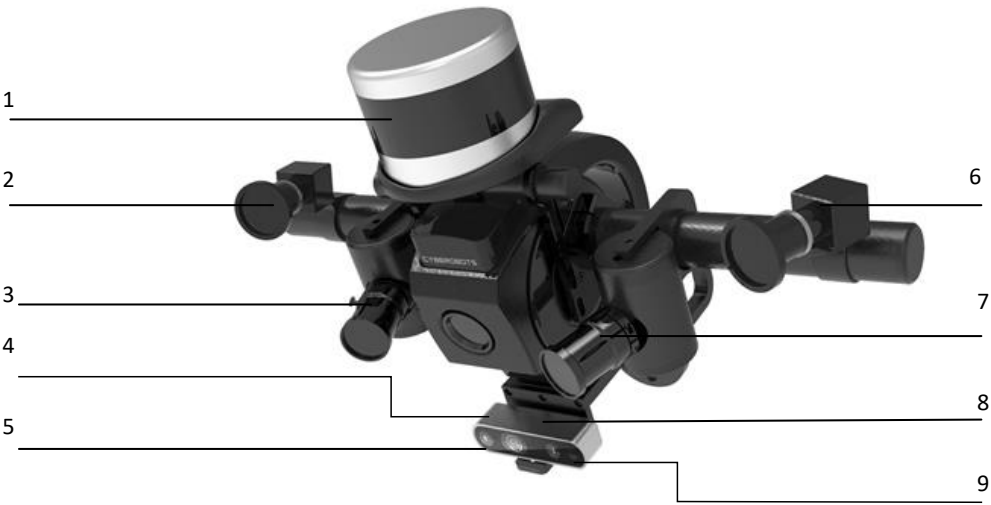


## 摘 要 说 明

---

本发明提供一种用于全场景的多模态感知融合系统，所述多模态感知融合系统包括上位机、激光雷达、多目相机、IMU、红外深度相机、电源，所述多目相机包括两个 FLIR 工业网口相机和两个 USB3.0 相机，其组成多模态感知融合系统的步骤为：安装各硬件、软件的安装以及数据的获取、模型的构建。本发明使用建模辨识中的麦夸特算法对外参数进行迭代优化，得到最优估计，从而得到最精确地模型和效果图，使得融合更加精准，能够达到实时的感知环境，并且该多模态感知融合系统小巧，重量较轻，可用于无人车车载，无人机机载，医疗行业，军事无人环境的建模，也可用于室内室外等各种复杂环境，为规划导航奠定基础

摘要附图



1. 一种用于全场景的多模态感知融合系统，其特征在于，所述多模态感知融合系统包括上位机、激光雷达、多目相机、IMU、红外深度相机、电源，所述多目相机包括两个 FLIR 工业网口相机和两个 USB3.0 相机，其组成多模态感知融合系统的步骤为：

S1：安装硬件：将激光雷达以以太网接口连接的方式连接到上位机，将两个 FLIR 工业网口相机以以太网接口方式连接到上位机，将两个 USB3.0 相机、IMU 以及红外深度相机分别连接到上位机的 usb3.0 接口，将各部分连接好后通过数据线与电源相连接；

S2：软件的安装和数据的获取：打开 Linux Ubuntu 系统，安装配置好各个模块的驱动和软件，使用 Robot Operating System 启动各个模态的节点，并且使用 RVIZ 将获取到的激光雷达的点云、多目相机的 RGB 图像、IMU 的加速度计以及陀螺仪信息以及红外深度相机的景深图的这些数据都显示出来；

S3：模型构建：接着使用 slam 理论体系将获取到的数据进行处理，该处理流程分为两步，分别是前端和后端，前端负责各个模块的特征提取和特征之间的相关性的表示，后端负责参数的优化和三维重建以及定位，最后得出融合的最终模型和效果图。

2. 根据权利要求 1 所述的一种用于全场景的多模态感知融合系统，其特征在于，所述多模态感知融合系统采用的操作系统为 Linux Ubuntu 系统，采用的中间件为 Robot Operating System，使用的编程语言为 c++和 python。

3. 根据权利要求 1 所述的一种用于全场景的多模态感知融合系统，其特征在于，所述激光雷达采用镭神智能 c16-151B。

4. 根据权利要求 1 所述的一种用于全场景的多模态感知融合系统，其特征在于，所述红外深度相机的数目为两个，所述红外深度相机和 IMU 采用 Intel Real Sense D435i。

5. 根据权利要求 1 所述的一种用于全场景的多模态感知融合系统，其特征在于，所述最终模型和效果图的得到是使用建模辨识中的麦夸特算法对外参数进行迭代优化，得到最优估计，从而得到最精确地模型和效果图。

## 权 利 要 求 书

---

6. 根据权利要求 1 所述的一种用于全场景的多模态感知融合系统，其特征在于，所述激光雷达投影到地面的距离为 10m，投影后下方会有锥形盲区，所述红外深度相机工作距离为 0.2-10m，可以弥补激光雷达投不到的盲区。

7. 根据权利要求 1 所述的一种用于全场景的多模态感知融合系统，其特征在于，所述激光雷达、多目相机、IMU、红外深度相机都分别具有独立的传感器。

## 一种多模态感知融合系统

### 技术领域

本发明属于多模态感知融合系统领域，尤其涉及一种用于全场景的多模态感知融合系统。

### 背景技术

随着传感器技术和互联网的迅速发展，各种不同模态的大数据正在以前所未有的发展速度迅速涌现。对于一个待描述事物（目标、场景等），通过不同的方法或视角收集到的耦合的数据样本就是多模态数据，通常把收集这些数据的每一个方法或视角称之为一个模态。

狭义的多模态信息通常关注感知特性不同的模态，而广义的多模态融合则通常还包括同一模态信息中的多特征融合，以及多个同类型传感器的数据融合等，因此，多模态感知与学习这一问题与信号处理领域的“多源融合”、“多传感器融合”，以及机器学习领域的“多视学习”或“多视融合”等有密切的联系；多模态数据可以获得更加全面准确的信息，增强系统的可靠性和容错性。

在多模态感知与学习问题中，由于不同模态之间具有完全不同的描述形式和复杂的耦合对应关系，因此需要统一地解决关于多模态的感知表示和认知融合的问题。多模态感知与融合就是要通过适当的变换或投影，使得两个看似完全无关、不同格式的数据样本，可以相互比较融合，这种异构数据的融合往往能取得意想不到的效果。

目前，多模态数据已经在互联网信息搜索、人机交互、工业环境故障诊断和机器人等领域发挥了巨大的作用，视觉与语言之间的多模态学习是目前多模态融合方面研究成果较为集中的领域，在机器人领域目前仍面临很多需要进一步探索的挑战性问题；由此，我们研制了一套多模态感知系统，将多目视觉，激光，双目红外，深度，IMU等多模态，这些硬件按照不同的方位进行安装。实现对大型场景，小型工件的自动化感知，扫描与建模，能够实现对全场景

的感知，适用于室内和室外，对环境的 RGB 图像信息赋予深度信息和距离信息，但其中最主要的难点在于：异种多源传感器，特征的提取，以及特征之间相关性的求解使得融合更加精准，能够达到实时的感知环境。

### 发明内容

为了解决上述技术问题，本发明提供一种用于全场景的多模态感知融合系统，以实现大型场景，小型工件的自动化感知，扫描与建模。所述多模态感知融合系统包括上位机、激光雷达、多目相机、IMU、红外深度相机、电源，所述多目相机包括两个 FLIR 工业网口相机和两个 USB3.0 相机，其组成多模态感知融合系统的步骤为：

S1：安装硬件：将激光雷达以以太网接口连接的方式连接到上位机，将两个 FLIR 工业网口相机以以太网接口方式连接到上位机，将两个 USB3.0 相机、IMU 以及红外深度相机分别连接到上位机的 usb3.0 接口，将各部分连接好后通过数据线与电源相连接；

S2：软件的安装和数据的获取：打开 Linux Ubuntu 系统，安装配置好各个模块的驱动和软件，使用 Robot Operating System 启动各个模态的节点，并且使用 RVIZ 将获取到的激光雷达的点云、多目相机的 RGB 图像、IMU 的加速度计以及陀螺仪信息以及红外深度相机的景深图的这些数据都显示出来；

S3：模型构建：接着使用 SLAM 理论体系将获取到的数据进行处理，该处理流程分为两步，分别是前端和后端，前端负责各个模块的特征提取和特征之间的相关性的表示，后端负责参数的优化和三维重建以及定位，使用建模辨识中的麦夸特算法对外参数进行迭代优化，得到最优估计，从而得出融合的最终模型和效果图。

优选的，所述多模态感知融合系统采用的操作系统为 Linux Ubuntu 系统，采用的中间件为 Robot Operating System，使用的编程语言为 c++和 python。

优选的，所述激光雷达采用镭神智能 c16-151B。

优选的，所述红外深度相机的数目为两个，所述红外深度相机和 IMU 采用 Intel Real Sense D435i。

优选的,所述激光雷达投影到地面的距离为 10m,投影后下方会有锥形盲区,所述红外深度相机工作距离为 0.2-10m,可以弥补激光雷达投不到的盲区。

优选的,所述激光雷达、多目相机、IMU、红外深度相机都分别具有独立的传感器。

与现有技术相比,本发明的有益效果是:使异类传感器自主联合,能够快速标定,并且进行采集信息的匹配和三维空间下的融合,由点云生成面片模型,再进行迭代优化,最终得到能够达到精度的三维重建模型,从而得到最精确地模型和效果图,使得融合更加精准,能够达到实时的感知环境,为以后的识别检测技术提供精准的技术数据,并且该多模态感知融合系统小巧,重量较轻,可用于无人车车载,无人机机载,医疗行业,军事无人环境的建模,也可用于室内室外等各种复杂环境,为规划导航奠定基础。

### 附图说明

图 1 是全场景的多模态感知融合系统的外观图。

图 2 是全场景的多模态感知融合系统的安装步骤图;

图中:1-激光雷达;2-第一 FLIR 工业网口相机;3-第一 USB3.0 相机;4-第一红外深度相机;5-第二红外深度相机;6-第二 FLIR 工业网口相机;7-第二 USB3.0 相机;8-多目相机;9-IMU。

### 具体实施方式

下面将结合本发明实施例的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

以下对本发明做进一步描述:

实施例:

如附图 1 所示,一种用于全场景的多模态感知融合系统,所述多模态感知融合系统采用的操作系统为 Linux Ubuntu 系统,采用的中间件为 Robot Operating System,使用的编程语言为 c++和 python;所述多模态感知融合系

统包括上位机、激光雷达 1、第一 FLIR 工业网口相机 2、第二 FLIR 工业网口相机 6、第一 USB3.0 相机 3、第二 USB3.0 相机 7、多目相机 8、IMU 9、第一红外深度相机 4、第二红外深度相机 5、电源；

具体的，如附图 3 所示，其组成多模态感知融合系统的步骤为：

S1：安装硬件：将激光雷达以以太网接口连接的方式连接到上位机，将第一 FLIR 工业网口相机 2、第二 FLIR 工业网口相机 3 以以太网接口方式连接到上位机，将第一 USB3.0 相机 3、第二 USB3.0 相机 7、IMU 9 以及第一红外深度相机 4 和第二红外深度相机 5 分别连接到上位机的 usb3.0 接口，将各部分连接好后通过数据线与电源相连接；

S2：软件的安装和数据的获取：打开 Linux Ubuntu 系统，安装配置好各个模块的驱动和软件，使用 Robot Operating System 启动各个模态的节点，并且使用 RVIZ 将获取到的激光雷达 1 的点云，多目相机 8、第一 FLIR 工业网口相机 2、第二 FLIR 工业网口相机 6、第一 USB3.0 相机 3 以及第二 USB3.0 相机 7 的 RGB 图像、IMU 9 的加速度计以及陀螺仪信息以及第一红外深度相机 4、第二红外深度相机 5 的景深图的这些数据都显示出来；

S3：模型构建：接着使用 slam 理论体系将获取到的数据进行处理，该处理流程分为两步，分别是前端和后端，前端负责各个模块的特征提取和特征之间的相关性的表示，后端负责参数的优化和三维重建以及定位，使用建模辨识中的麦夸特算法对外参数进行迭代优化，得到最优估计，最后得出精确地融合后最终模型和效果图。

具体的，所述激光雷达 1 采用镭神智能 c16-151B。

具体的，所述第一红外深度相机 4、第二红外深度相机 5 和 IMU 9 均采用 Intel Real Sense D435i。

具体的，所述激光雷达 1 投影到地面的距离为 10m，投影后下方会有锥形盲区，所述第一红外深度相机 4、第二红外深度相机 5 的工作距离为 0.2-10m，可以弥补激光雷达 1 投不到的盲区。

具体的，所述激光雷达 1、第一 FLIR 工业网口相机 2、第二 FLIR 工业网口



## 说明书

---

相机 6、第一 USB3.0 相机 3、第二 USB3.0 相机 7、多目相机 8、IMU 9、第一红外深度相机 4、第二红外深度相机 5 都分别具有独立的传感器。

如附图 2,所示为多模态系统的图结构表示。图结构中的顶点表示激光雷达,相机, IMU 等传感器。图结构中的边表示传感器之间的相对位姿变换推导式。

工作流程图如图 3 所示。

需要说明的是,在本文中,而且,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。

尽管已经示出和描述了本发明的实施例,对于本领域的普通技术人员而言,可以理解在不脱离本发明的原理和精神的情况下可以对这些实施例进行多种变化、修改、替换和变型,本发明的范围由所附权利要求及其等同物限定。



图 1

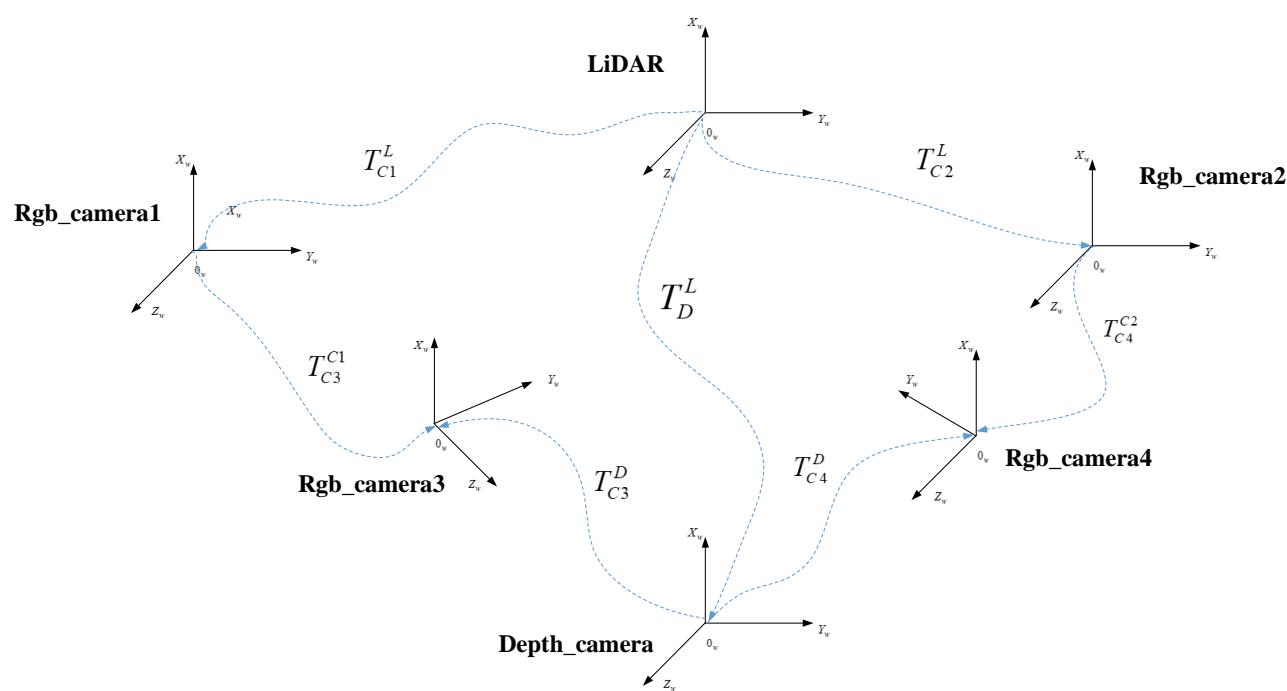


图 2



图 3 工作流程图