文章编号：2095-2163(2020)08-0087-06　　中图分类号：TP391.9　　文献标志码：A

基于激光雷达和深度相机融合的视觉 SLAM 研究

李陆君，张　智，韩　蕊，李晨阳，丁亚男
（上海工程技术大学 航空运输学院，上海 201620）

摘　要：传统的障碍物探测主要是采用一些非视觉性器件，这些传感器在抗干扰能力方面突出，并且可以长时间工作，但是也存在一些缺点。本文提出一种基于 Bayes 推理的激光和视觉地图层面融合的建图与定位方法。将 Kinect 相机和二维激光雷达同时安装在移动机器人上，采用激光扫描和深度图像结合的方法来检测障碍物，从而达到探测范围广、精度高、实时性好的特点，同时在三维探测方面也存在一定优势。

关键词：Bayes 推理；激光雷达；kniect；建图；定位

Research on visual SLAMbased on fusion of lidar and depth camera

LI Lujun, ZHANG Zhi, HAN Rui, LI Chenyang, DING Yanan
(School of Air Transportation, Shanghai University of Engineering and Technology, Shanghai 201620, China)

【Abstract】Traditional obstacle detection mainly uses some non-visual devices, including laser sensor, ultrasonic sensor and infrared sensor. These sensors are outstanding in anti-interference ability, and can work for a long time, but there are also some shortcomings, such as single data acquisition, and it is difficult for the robot to accurately detect the three-dimensional information of the surrounding environment. The combination of vehicle mounted camera and two-dimensional lidar has good performance in detecting the position and shape of obstacles; the fusion of two-dimensional lidar and vision can improve the performance in solving the problem of position loss in wheel skidding and ground negative obstacle detection. It not only reduces the sensitivity to the environment, but also improves the detection accuracy. Compared with monocular and binocular cameras, Kinect camera is convenient and low-cost, and can provide environmental depth data.

In this paper, a mapping and positioning method based on Bayes reasoning is proposed. Kinect camera and two-dimensional lidar are installed on the mobile robot at the same time, and the method of laser scanning and depth image combination is used to detect obstacles, so as to achieve the characteristics of wide detection range, high precision and good real-time performance. At the same time, there are also advantages in three-dimensional detection potential.

【Key words】Bayes reasoning; lidar; Kinect; mapping; positioning

0　引　言

障碍物检测是众多应用场景的基础，同时也是移动机器人运动控制准确实现的重要依据。传统的障碍物探测主要是采用一些非视觉性器件，其中包括激光传感器、超声传感器以及红外传感器等。这些传感器在抗干扰能力方面突出，并且可以长时间工作，但是也存在一些缺点。如数据采集单一，不够丰富，机器人也难以准确的探测出周围环境的三维信息等等。虽然三维激光雷达在三维信息探测方面性能突出，但成本较高，难以广泛应用。目前，研究学者广泛研究的是基于视觉的环境感知技术，该技术成本低，信息丰富。但基于视觉的传感器对环境敏感度高、检测精度低。因此基于单目视觉的环境感知方法不够理想，而基于双目视觉的环境感知方法又容易引起多义性。基于以上问题，有研究者提出将视觉和激光检测方法结合起来，这样做既降低了对环境的敏感度，同时也提升了检测的精确度。例如：将车载相机和二维激光雷达[1]组合起来，在探测障碍物位置和形状上有较好的性能表现；将二维激光雷达和视觉融合起来，在解决车轮打滑时位置丢失问题，以及地面负障碍物检测上有较好的性能提升。对比单目和双目相机，Kinect 相机[2]方式便捷、成本低，可以提供环境的深度数据[3]。本文利用这一优点，将 Kinect 相机和二维激光雷达[4]同时安装在移动机器人上，采用激光扫描和深度图像结合的方法来检测障碍物，从而达到探测范围广、精度高、实时性好的特点，同时三维探测方面也存在一定优势。

1　基于激光雷达传感器的 SLAM

在计算激光雷达与障碍物之间的距离时，是通过其发出的激光束遇到障碍物反射回来的时间差而反应出来的。激光雷达因其在抗干扰能力、检测精

作者简介：李陆君(1993-)，男，硕士研究生，主要研究方向：交通运输工程、视觉定位和导航、多传感器组合导航。

收稿日期：2020-01-10

度上性能突出,并且对环境敏感度较低,计算速度快。故在移动机器人的定位、地图构建、导航等方面应用广泛。激光雷达分为一维、二维、三维等不同测量维度。近年来,面阵激光测距雷达和 2 自由度扫描激光雷达也相继发展起来。二维激光雷达在目前移动机器人[5]应用上较为广泛,数据获取速度快、数据量小,且便于处理,不足之处在于扫描平面数据单一。

本文采用的是 RPLIDAR A2 激光雷达,其测量半径可达 18 米,测量范围全角 360°,高精度、宽视角、高分辨率的优点有助于移动机器人的环境信息识别。除此之外,其设计小巧、易于安装,光照影响小,黑暗环境下也不影响使用。RPLIDAR A2 激光雷达实物如图 1 所示。

图 1 RPLIDAR A2 激光雷达

Fig. 1 RPLIDAR A2 lidar

激光雷达的具体参数见表 1。激光雷达不容易被其它因素干扰,只有在雨天、雾天等恶劣天气激光会发生折射,精度会下降。

表 1 RPLIDAR 激光雷达参数

Tab. 1 RPLIDAR lidar parameters

参数	最小值	典型值	最大值	说明
测距范围/M	0.15	0.15-12	12	反光物体
扫描角度/(°)		0-360°		
测距分辨率/mm		<0.5		被测物目标 1.5 m 内
角度分辨率/(°)	0.45	0.9°	1.35°	10 HZ 扫描时
单次测量时间/ms		0.25		
测量频率/Hz	2 000	4 000-8 000	8 000	
扫描频率/Hz	5	10	15	扫描一周获取 400 个点

RPLIDAR A2 激光雷达扫描效果如图 2 所示。扫描数据是以极坐标的方式展现出来的,对于环境的扫描会返回得到一组有序的二维数据。假设激光扫描的范围是 360°,经过扫描会得到 360 个极坐标点,表示为 $\alpha_i = (\rho_i, \theta_i)$, i = 1,2,3…360。其中,$\rho$ 表示距离信息;θ 表示扫描角度信息。若在测量范围之内,激光雷达没有接收到返回数据,则认定该值为最大值。将获得的扫描数据从极坐标转换到直角坐标系,如式(1)所示,由此得到数据集合 $b = \{(x_i, y_i), i = 1,2,3\cdots360\}$。

$$\begin{cases} x_i = \rho_i \cos\theta_i, \\ y_i = \rho_i \sin\theta_i. \end{cases} \tag{1}$$

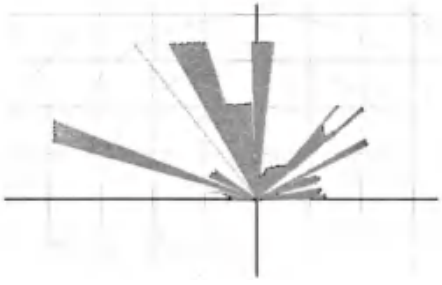


图 2 激光雷达扫描图

Fig. 2 Lidar scanning

图 3 是激光雷达采集数据过程。其通过发射激光束获取外部环境信息,并从环境中返回实时数据到激光雷达;利用 RS232 串口与上位机通讯,再把返回的数据传给上位机处理。这个过程视为完成一次完整的数据采集。

图 3 激光雷达采集数据过程

Fig. 3 Data acquisition process of lidar

2 基于深度相机的 SLAM

外部传感器在移动机器人的地图构建和导航控制中起着关键作用。本研究的深度摄像头采用的是 RGB-D 相机。RGB-D 相机有两大功能:采集 RGB 彩色图像和采集深度图像。RGB-D 相机代表型的为 Kinect 相机。Kinect 相机的实物图如图 4 所示。图中从左至右依次为:红外结构发射器、RGB 摄像头、红外线 CMOS 相机(红外结构光接收器)。中间的 RGB 彩色相机可以显示二维图像彩色信息,左右两侧的红外发射器和红外线 CMOS 相机相互配合,记录红外信息和深度信息。其中,图像的深度信息

指的是被拍摄物体到摄像头的实际距离。Kinect 底座设置有一个仰角马达，可以上下移动，并完成对焦物体能够实现较好的追踪效果。

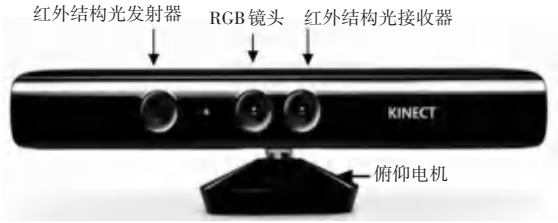


图 4　Kinect 相机

Fig. 4　Kinect camera

Kinect 相机的详细参数见表 2。

表 2　Kinect 详细参数

Tab. 2　Kinect detailed parameters

类型	参数	类型	参数
测量范围	0.8~3.5 m	物理倾斜范围	±27°
水平视角	57°	支持系统	Linux、windows
垂直视角	43°	俯仰角调整	电动马达
扫描频率	30 帧/s	接口类型	USB 2.0
RGB 像素大小	640 * 480	尺寸	28 * 6.5 * 6.5
红外相机像素大小	640 * 480	支持系统	Linux、windows

彩色图像信息和深度图像信息可以同时被 RGB-D 相机记录和获取，在深度图像中的像素值，其内容涵盖了一种远近的关系信息量。即为空间坐标点与深度相机所处平面距离的远近关系，以光编码技术呈现。主要表现为随机性发射出三维图像散斑，进行三维深度编码。这种技术突破了传统的结构光技术，与智能发射二维图像编码方式也有区别。

深度信息的获取主要是借助深度相机中红外发射器发射出随机性的、强的散斑，距离的变化也会带动散斑颜色的显著变化。进一步处理表现出强随机性的散斑，其中包括统计、测量、标定，并且将处理后的散斑作为参考对象，便于测量拍摄物体与深度相机之间的距离。依据物体上和已经处理后的两种散斑信息，计算出物体与深度相机所在平面的距离。

RGB-D 相机中的红外相机通过环境中的深度信息的探测，依次获取空间中某一个点的坐标。通常情况下，建立以 RGB-D[6] 红外摄像头为原点的坐标系，红外发射器的方向定义为 X 轴的正方向，Z 轴定义为垂直于图像平面的方向，Y 轴则由右手定则定义，正方向定义为图像平面指向场景中物体的方向。

三维坐标测量原理图如图 5 所示。图中标定了参考平面和物理平面。红外发射器方向为 X 轴方向，红外相机垂直于物体所在平面方向为 Z 轴。当环境中的物体处于参考平面时，可获得物体的散斑图像，当其处于物理平面时，在 X 轴上散斑图像会产生视差。根据三角形相似原理，得到式(2)、(3)：

$$\frac{D}{b} = \frac{Z_0 - Z_K}{Z_0}, \tag{2}$$

$$\frac{d}{f} = \frac{D}{Z_K}. \tag{3}$$

其中，D 为物体在物理平面的位移；b 为基线长度；f 为红外相机的焦距；Z_K 为物体在物理平面的深度值；d 是散斑图案在图像平面的视差。

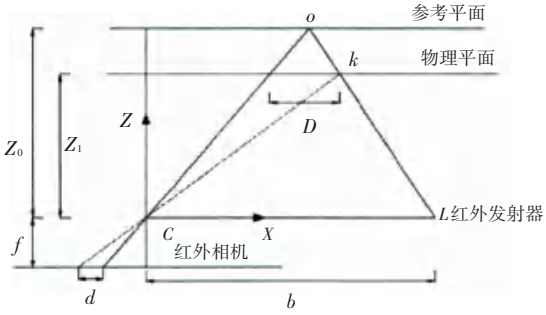


图 5　三维坐标测量原理图

Fig. 5　Schematic diagram of 3D coordinate measurement

$$Z_K = \frac{Z_0}{1 + \frac{Z_0}{f_b}} d. \tag{4}$$

其中，常量 Z_0，f 和 b 的值是通过校准获取的。其中，物体的某点成像尺度是由该点的深度值(即 Z 坐标值)和相机的焦距 f 共同决定的。进一步通过该点的图像坐标和成像尺度计算出 X、Y 轴的坐标，如式(5)、(6)所示：

$$X_K = -\frac{Z_K}{f}(x_k - x_0 + \delta_x), \tag{5}$$

$$Y_K = -\frac{Z_k}{f}(y_k - y_0 + \delta_y). \tag{6}$$

其中，X_K 和 Y_K 是该点在图像坐标系上的坐标；x_0 和 y_0 是原点坐标系；δ_x 和 δ_y 为透镜畸变的校正项。所以空间中某点的三维坐标可表示为(X_K，Y_K，Z_K)。

3　融合型的 SLAM

本文提出一种基于 Bayes 推理[7]，激光和视觉地图层面融合的建图与定位[8]方法，其融合流程如图 6 所示。单元格 $C_{i,j}$ 为空的概率，传感器模型为 P_s^o 通过 Bayes 更新单一传感器模型与它最近的概率估计值，更新后的数据引入到栅格 $C_{i,j}$ 中。当前传感器数据 $r_t = (t_1, t_{t-1}, \cdots, t_o)$ 与 Bayes 递推公式相结

合,可以计算出新的 P^o。地图在初始化状态时,$C_{i,j}$ 概率相等,也就是说地图[9]在初始化状态时,栅格单元先验概率 P_m^o 可表示为:

$$P_m^o = 1 - P_m^o - 0.5 \forall C_{ij}. \tag{7}$$

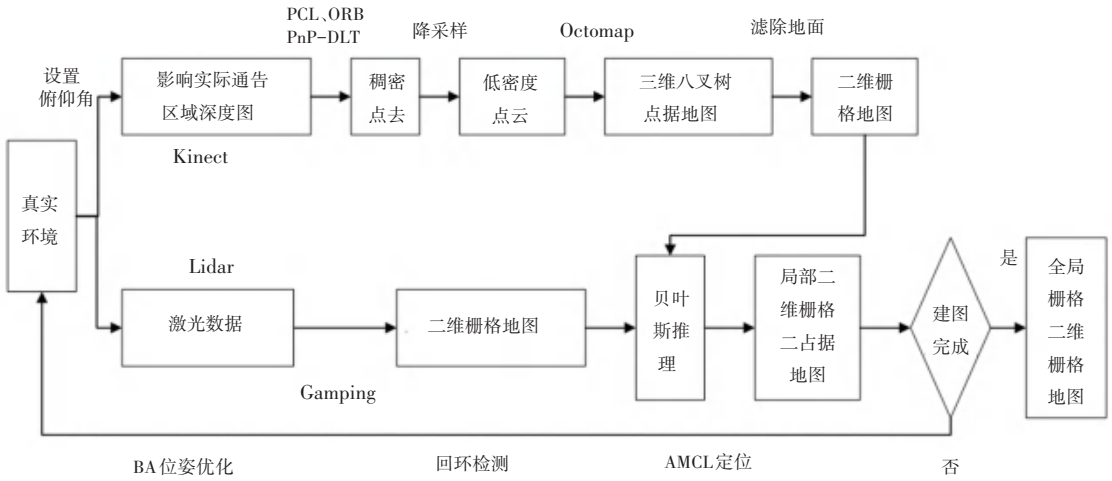


图 6 基于贝叶斯推理的地图层面融合

Fig. 6 Map level fusion based on Bayesian reasoning

通过融合方式转化为栅格占据模型[9],由式(7)变化为新的栅格[10]融合公式为:

$$P_o = \frac{P_s^o P_m^o}{P_s^o P_m^o + (1 - P_s^o)(1 - P_m^o)}. \tag{8}$$

其中,P_m^o 指先验栅格被占用的几率;$1 - P_m^o$ 是先验栅格未被占用的几率;P_s^o 为能够转换为占据栅格的条件概率;P_s^o 指传感器所测距离 R 更新 6 后的估计值大小或者逆传感器模型占用网格单元的条件概率大小。此外,每一个栅格单元均包括更新之前,以及基于 Bayes[11]融合传感器模型后的占用。更新之前,栅格占用不是确定的、精确的评估;更新后,局部地图都有确定的概率估计值。通过 Bayes 更新的局部地图,对于全局地图需要按顺序扩展,不断迭代更新全局地图,则旧的栅格被新的栅格值取代。更新过程如图 7 所示。

图 7 栅格地图更新过程

Fig. 7 Grid map update process

基于获取到的 Kinect 栅格占据和激光占据栅格,通过融合模型得到新的栅格占据,从而建立新的栅格地图。融合规则见表 3。

为了提高本方案的可靠性,考虑到两个传感器都会有失效的情况发生,通过设置阈值,保证在只有一个传感器正常工作时也能建立栅格地图[12]。通过决策,确定栅格是否被占据,得到新的栅格占据概率,生成栅格地图。

表 3 激光雷达和 kinect 视觉传感器融合规则

Tab. 3 Fusion rules of lidar and Kinect vision sensor

	Occupy(kinect)	empty(kinect)	Uncertain(kinect)
Occupy (lidar)	Occupy	Occupy	Occupy
empty (lidar)	Occupy	empty	empty
Uncertain (lidar)	empty	Occupy	Uncertain

设阈值为 T^0,如果占据概率高于阈值,则栅格值被设成 1。决策判断如公式(9):

$$P_{n=1,2}^o = \begin{cases} 1, & p^o > T^o; \\ 0, & \text{othe.} \end{cases} \tag{9}$$

其中:$P_{n=1}^o = P_{n=2}^o$ 分别对应视觉栅格占据概率值和激光雷达栅格占据概率。此方法也可以推广到其它传感器的融合。如 T^0 选择 0.5,更新后的栅格单元 $C_{i,j}$ 中,从占据栅格概率为:

$$p_f^o = \frac{p_1^o p_2^o}{p_1^o p_2^o + (1 - p_1^o)(1 - p_2^o)}. \tag{10}$$

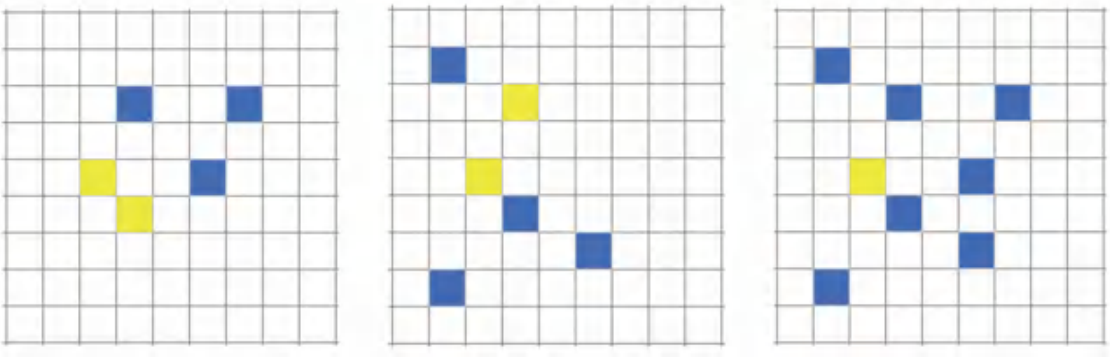
基于 Bayes 推理激光和视觉栅格融合过程如图 8 所示。

基于激光雷达和 Kinect 的贝叶斯推理模型完成地图层面[13]的融合,获得了包含更多环境信息的栅格地图。

4 实验仿真

本文提出基于 Bayes 推理的激光和视觉地图层面融合的建图与定位方法,将深度图中的三维数据通过降采样方法,构建而为栅格地图,基于获取到的 Kinect 栅格占据和激光占据栅格,通过融合模型得到新的栅格占据,从而建立新的栅格地图。通过决

策将确定栅格是否被占据，得到新的栅格占据概率，生成栅格地图。通过实验构建深度相机和激光雷达融合图像，验证方法的有效性。



（a）激光栅格　　（b）视觉栅格　　（c）融合栅格

图 8　激光和视觉地图层面融合过程

Fig. 8　The fusion process of laser and visual map level

实验所采用的主控计算机处理器为酷睿六代 i7、主频 2.5 GHz；操作系统运用 Ubuntu 的同时，启用了 ROS 系统；二维激光雷达为 RPLIDAR A2、最大测距值为 36 m、采样频率为 8 000 次/s、测距精度能达到 0.5 mm、最大扫描角度是 360°。实验环境障碍物实物设置如图 9 所示，实验仿真结果如图 10-图 13 所示。

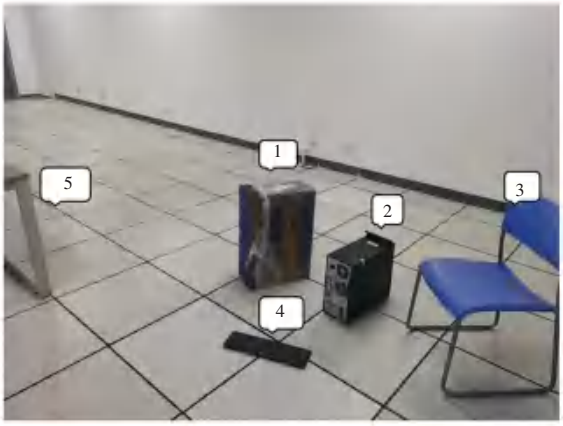


图 9　实验环境障碍物实物图

Fig. 9　Physical map of obstacles in the experimental environment

由图 10 可见，激光雷达获取的栅格地图中仅显示了障碍物 1、3、5 的分布。由于激光雷达安装位置限制，低于激光雷达扫描平面的障碍物 2、4 未被发现。

由图 12 可见，Kinect 相机获取的栅格地图中仅发现了障碍物 1、2、3。由于碍物 5 位于 Kinect 视野范围外，因此没有获取到该障碍物的分布；而障碍物 4 高度低于 1 cm（在八叉树生成二维栅格占据时设置了阈值），对于高度不足 1cm 的障碍物予以滤除，所以栅格地图上看不到障碍物 4 的分布。

图 10　激光雷达实验图

Fig. 10　Lidar experiment diagram

图 11　RGB 彩色图像和深度图像

Fig. 11　RGB color image and depth image

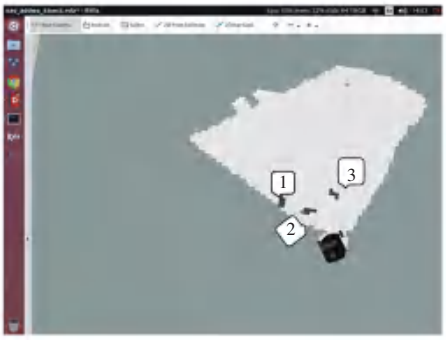


图 12　基于 Kinect 获取的栅格占据

Fig. 12　Kinect′s grid occupation map

图 13 是激光和视觉信息基于贝叶斯推理融合

后获得新的栅格地图,在栅格地图上可看到障碍物 1、2、3、5 的分布。由于障碍物 4 不会阻碍机器人的通行而被滤除。融合后的栅格地图中的障碍物会对机器人的通行产生真实影响。相比于单一传感器融合的栅格地图包含了更为丰富的环境信息。

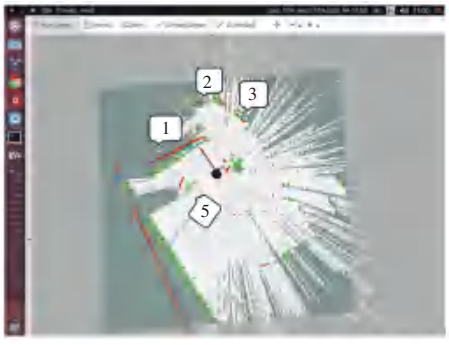


图 13 深度相机和激光雷达融合

Fig. 13 Fusion of depth camera and lidar

通过图 11 与图 9、图 10 的对比可以看出,单独使用激光雷达获得图像和深度相机获取的图像,都无法满足三维空间环境的表达要求。而融合后的图像,能够显示出障碍物椅子的三维图像。有效的避免单一传感器的不足,实现三维的地图构建,可以用于机器人在三维的环境下的的定位和导航。

5 结束语

本文提出一种基于 Bayes 推理的激光和视觉地图层面融合[14]的建图与定位方法。通过实验验证其有效性,相比于单一传感器多传感器方式,具有探测范围广、精度高、实时性好的特点,同时三维探测方面也存在优势。RGB-D SLAM[15]信息获取相对丰富,但容易受干扰、噪声大、无法投射透明材质、大规模建图容易失败。二维激光可以投射透明材质、抗干扰能力强,但仅仅只能扫描一个平面,由于安装位置和视野原因,又不可能装的特别低。基于此,基于贝叶斯推理完成两种传感器地图层面的融合建图,以及回环检测,具有较好的精度和可靠性。

参考文献

[1] 张力元,张法星,蔡爱玲,等. 三维激光雷达扫描技术在河流模拟中的应用[J]. 水电能源科学,2020,38(8):121-124.
[2] 司徒仕忠,邱广萍,王锦春. 基于深度相机的障碍物识别[J]. 科技创新与应用,2019(27):37-40.
[3] 李若鹏. 基于车载激光雷达的点云匹配定位[D]. 北京建筑大学,2020.
[4] 张恒,徐万红,张禹. 激光雷达与深度相机融合的 SLAM 技术研究[J]. 机械工程师,2020(7):22-24+29.
[5] 丁仁杰. 基于激光视觉信息的 AGV 即时定位与地图构建方法研究[D]. 哈尔滨工业大学,2019.
[6] 杨记周,胡广宇,许玉云,等. 室内移动机器人平台设计与实现[J]. 自动化仪表,2019,40(5):75-78.
[7] 缪文良. 基于改进 RGB-D SLAM 的移动机器人建图与导航研究[D]. 哈尔滨工业大学,2019.
[8] 刘笑飞,吴立平. 基于动态贝叶斯网络的反无人机处置决策研究[J]. 电子制作,2020(12):30-32.
[9] 陈超,徐军,张伟伟. 多传感器融合的移动机器人同步定位与建图[J]. 现代电子技术,2020,43(14):164-169.
[10] 徐美清,刘洞波. 神经网络信息融合创建机器人栅格地图研究[J]. 科技资讯,2019,17(21):11-12.
[11] 林鹏. 融合 IMU 与 RGB-D 相机的室内移动机器人地图创建算法研究[D]. 湖南大学,2019.
[12] 白云裳. 室内移动机器人定位技术的研究[D]. 沈阳理工大学,2020.
[13] 瑚琦,蔡文龙,卢定凡,等. 单目视觉惯性的同时定位与地图构建综述[J]. 软件导刊,2020,19(7):275-280.
[14] 张伟伟,陈超,徐军. 融合激光与视觉点云信息的定位与建图方法[J]. 计算机应用与软件,2020,37(7):114-119.
[15] 缪文良. 基于改进 RGB-D SLAM 的移动机器人建图与导航研究[D]. 哈尔滨工业大学,2019.

(上接第 86 页)

2 结束语

由图 3 和图 10 可以看出,中值滤波对高斯噪声的处理效果并不好。这是由于高斯噪声使用的是被污染的随机大小值,因此无论如何选择数据,获得的结果都是被污染的值。由图 4 和图 11 可以看出,均值滤波在处理图象时对图像细节的保护能力不好,虽然去除了噪声,同时也损失了很多图像细节,使获得的图像变得模糊,降低了去除噪效果。由图 7、图 14 可以看出,改进后的方法在提升去噪效果的同时,也减弱了了图像去噪后带来的模糊问题。与文献[6]方法相比提高了阈值门限选择的精度与稳定性。小波重构后的图像经过约束最小二乘方滤波减弱了图像边缘模糊的问题,提升了图像的信噪比。

参考文献

[1] 张莲, 秦华峰, 余成波. 基于小波阈值去噪算法的研究[J]. 计算机工程与应用, 2008,44(9): 172-173.
[2] 王蓓, 张根耀, 李智,等. 基于新阈值函数的小波阈值去噪算法[J]. 计算机应用, 2014, 34(5): 1499-1502.
[3] 秦冬冬, 陈志军, 闫学勤. 多层阈值函数下的小波图像去噪[J]. 计算机工程, 2017, 43(6): 202-206.
[4] 王瑞, 张友纯. 新阈值函数下的小波阈值去噪[J]. 计算机工程与应用, 2013, 49(15): 215-218.
[5] 王凤,赵佰亭,等. 图像的小波可变阈值去噪方法[J]. 光电子,2019,30(8):854-866.
[6] 刘永平,郭小波,等. 基于新阈值函数和小波分析的数字图像去噪方法[J]. 电脑与信息技术,2020,28(2):5-8.