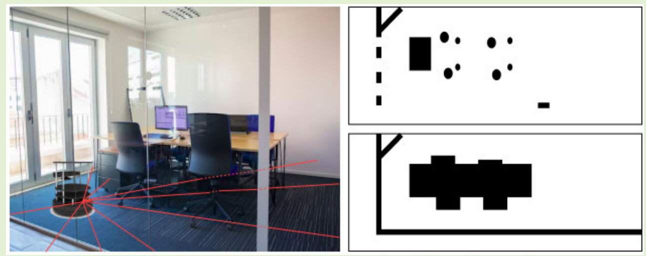
Leveraging 3-D Data for Whole Object Shape and Reflection Aware 2-D Map Building

Alicia Mora, Ramon Barber, *Senior Member, IEEE*, and Luis Moreno, *Member, IEEE*



***Abstract*—Two-dimensional laser scan sensors stand out as the preferred choice for robot mapping applications. However, these sensors have a significant drawback. Encountering objects with varying shapes at different heights, such as tables, poses challenges for these sensors due to their limited detection capability resulting from their dimensionality. This limitation increases the risk of potential collisions. Additionally, there are multiple polished materials that generate noise due to reflection. In order to have a robust occupancy grid map representation, these problems must be addressed. This article proposes the usage of a 3-D laser scan sensor to generate a 2-D occupancy grid map that incorporates the complete geometry of objects and effectively filters out noise from reflective materials like glass. The main novelty of the method is that it takes advantage of all the available 3-D data to avoid any information loss about objects' shapes. Additionally, a new approach for filtering reflection noise based on the analysis of indoor structural elements is proposed. Both approaches are merged for the creation of a robust indoor representation that allows to safely navigate the environment. Finally, a recursive Bayesian filter is applied for merging data, so noise due to dynamic elements that appeared during data collection is also filtered. Experimental evaluations in indoor environments with diverse objects and reflective surfaces, including dynamic elements like people, demonstrate the effectiveness of the proposed approach.**

***Index Terms*—3-D laser scan, occupancy grid map, reflection, robust map building.**

I. Introduction

TWO-DIMENSIONAL maps, more specifically occupancy grid maps, are the most popular environment representations for indoor robotic applications. This is because they are easy to build, maintain, and update in large-scale scenarios, in addition to allowing easy computations such as calculating optimized paths [1]. They are typically built using 2-D laser scan sensors, usually positioned on the robot mobile base a few centimeters above ground level, as can be observed in commercial mobile platforms such as TIAGo [2] and Pepper [3] and in some other noncommercial platforms designed for research [4]. However, this positioning is very limiting for fully detecting objects that greatly vary in shape depending on the height at which they are observed. Hence, planned paths could go through objects and the robot would collide when executing the task. An example of the problem is shown in Fig. 1, where information from a 2-D and a 3-D sensor is captured from the same scenario. Zones 1 and 2 correspond to office areas, with desks and chairs. While the 2-D sensor directly detects the background wall, since it is placed almost at ground height, the 3-D sensor is able to capture all object information. Using 3-D maps instead could solve the collision problem, as shown in works such as [5], [6], [7]. However, their construction and use is computationally expensive and memory requirements are very high for their storage. Another problem with these sensors is due to their operation with light. Polished surfaces reflect light in its entirety, and may not fall back on the sensor, so they are not measured. Even in the case of glass, elements behind it can be detected, so distance information is not accurate.

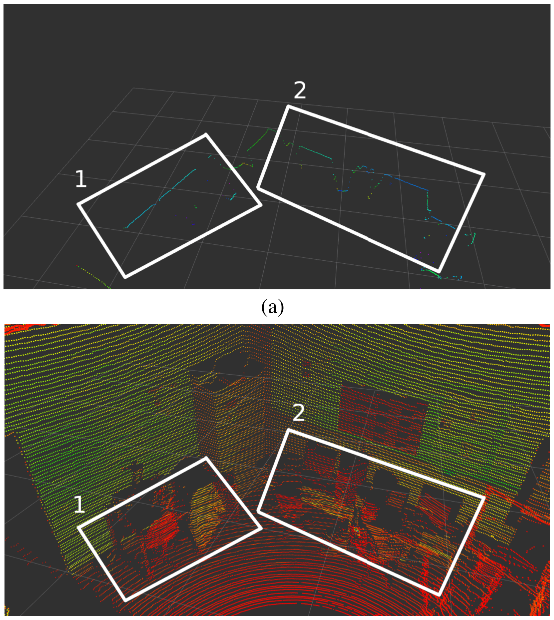
Based on the above information, this work proposes to use information from a 3-D laser scan sensor to generate a traditional 2-D occupancy grid map that takes into account both the complete geometry of objects and the filtering of noisy data coming from reflective materials like glass. In this way, all the advantages of 2-D maps are maintained with the addition of providing geometric and reflective robustness. Robustness is understood in this article as the ability to clearly differentiate between free, occupied, and unknown zones in an occupancy grid map.

Using 3-D information to map in 2-D has been considered before, typically obtaining information from an RGBD camera. However, these works only focus on simulating a 2-D laser scan sensor, so that for one viewing angle only the

Manuscript received 12 September 2023; accepted 26 September 2023. Date of publication 17 October 2023; date of current version 16 July 2024. This work was supported by the RoboCity2030 Madrid Robotics Digital Innovation Hub DIH-CM Project under Grant S2018/NMT-4331. The associate editor coordinating the review of this article and approving it for publication was Prof. Lan. *(Corresponding author: Alicia Mora.)*

The authors are with the Robotics Lab, Universidad Carlos III de Madrid, 28911 Leganés, Spain (e-mail: almorav@ing.uc3m.es; rbarber@ing.uc3m.es; moreno@ing.uc3m.es).

Digital Object Identifier 10.1109/JSEN.2023.3321936



(a)

(b)

Fig. 1. Data captured from the same scenario using (a) 2-D laser scan sensor and (b) 3-D laser scan sensor. The first data are unable to correctly locate objects, only the background wall and some noise corresponding to furniture legs can be seen. Three-dimensional data perfectly defines objects shapes.

closest obstacle to the sensor is maintained, resulting in a loss of valuable information. In addition, no fusion of this representation with reflective denoising has been performed so far. Reflective surfaces are normally detected by measuring reflectivity of the sensor light beams as they strike materials perpendicularly, producing a peak value. However, for a 3-D sensor, this is only applicable to the horizontal plane in which it is located, assuming that glass is placed in the vertical plane of the walls. Therefore, another approach is needed to solve the problem. In this article, we propose the detection of structural elements on buildings to filter out this noise.

This is a continuation of previous work performed in our research group. Three-dimensional information has been previously used to generate representations at various heights [8] and glass has been detected by 2-D sensors as a function of reflectivity [9]. With this new proposal, both objectives are merged and a robust occupancy grid map for robot navigation is achieved. As a summary, the following items outline the main contributions of the proposed work.

1) Three-dimensional data are used to model 2-D occupancy states of an environment considering elements at multiple heights and distances.
2) Indoor structures are analyzed to filter out reflection noise.
3) Both sources of information are merged to create a robust 2-D occupancy grid map for safe robot navigation applications.

Experiments have been performed in indoor environments with objects of multiple shapes and various reflective surfaces, as well as dynamic objects, e.g., people passing by. Results show that the proposal generates robust maps that filter these noises and consider the full geometry of objects.

II. Related Work

The following is an analysis of the most representative works on 2-D map generation and reflective surface detection. It is worth mentioning that the state of the art on these subjects is scarce and none of the presented works propose the fusion of both methodologies for the generation of a robust representation, as is done in this article.

A. Two-Dimensional Mapping Based on 3-D Data

One of the preferred 3-D sensors for indoor mapping are RGBD cameras. Kamarudin et al. [10] propose a method for generating a 2-D virtual laser scan out of depth data. First, depth values are transformed into XYZ coordinates. Then, for each Z column, the element that is closer to the camera is selected as the 2-D distance value in that direction for the virtual sensor. In [11] and [12] similar approaches are described. The work in [13] develops the same technique but using a 3-D laser scan sensor applied to self-driving cars. Wulf et al. [14] define a similar approach for mobile robots in outdoor environments. Although authors claim that this method minimizes computational time, simplifying 3-D data with only the nearest value to the sensor results in a loss of relevant information and an underutilization of the potential of 3-D sensors.

Other works have demonstrated the capabilities of using 3-D information in its full extent. In [15], the complete geometry of objects is considered for building a 2.5-D map, where each occupancy cell saves height information. This outcome is used to generate a 2-D grid map where collision threats are clearly marked. In [16], a 2-D map is augmented with 3-D data for obstacle avoidance tasks in real time. Missura et al. [17] propose to project a 3-D point cloud to the floor plane and delimit each region by a boundary. This representation is tested online and mounted on a robot for motion planning tasks. Similar to these works, our purpose is to improve the creation of 2-D maps based on 3-D information. Our work proposes the projection of all relevant 3-D information and its further probabilistic treatment for error elimination.

B. Detection and Filtering of Reflective Surfaces

Regarding the detection of reflective materials, some works propose to fuse information from multiple sensors, mainly laser scan and ultrasonic sensors [18], [19]. Even though these methods prove a better performance with respect to solely mapping with laser data, they have issues like a limited field of view or only detecting range differences. Other works focus on 2-D data where reflectivity peaks are analyzed [9], [20]. For this to occur, sensor light beams must strike perpendicularly on the surface. Therefore, as explained in [21], the extension to 3-D sensors of the use of reflectivity is only applicable to its horizontal plane, so using a 3-D sensor only contributes to unnecessary costs. In the same work, they use the dual return of a Velodyne device to detect discrepancies corresponding to areas with glass, but this is only possible in certain 3-D laser models. Another relevant aspect of these works is that many of them focus on detecting reflective surfaces but not on filtering their errors, so the resulting map is noisy.

Other works propose the detection of holes or window frames in walls [22], [23]. These approaches imply the necessity of processing 3-D point clouds to detect elements such as planes, which is computationally expensive. Our proposal is to analyze the structural elements of the environment similar to the work in [24], where 3-D clouds are projected to the horizontal plane to reduce computational complexity. Given that a scan cannot measure information through walls, in cases where this occurs, it is mostly due to noise created by reflective surfaces, so by comparing both information, we will be able to filter it out. Combining this idea with the projection of complete object point clouds and the probabilistic treatment of data will result in a robust 2-D occupancy grid map for robot navigation.

III. Methodology

The methodology is closely related to the traditional way of generating occupancy grid maps, where 2-D captured data is accumulated considering the robot pose. However, modifications are required to adapt the representation models to data captured from a 3-D laser scan sensor. The main idea is to make use of the spatial information obtained from a 3-D laser sensor to map the complete geometry of objects and eliminate reflective areas. According to [25], the following components for data interpretation are required for recovering spatial information from sensor data.

1) *Coordinate transformations*, which allow to concatenate laser scan measurements taken from different positions into a unified world model. This is represented as a set of robot poses, both position and orientation, from which scans are taken. The reference frame corresponds to the first pose from which scans are measured.
2) A *spatial interpretation model*, adapted to each different kind of sensor, which translates the sensor data into a statement about which areas are occupied or empty. By giving it a probabilistic approach, this turns into an inverse sensor model (ISM). One of the main contributions of this work is adapting the traditional 2-D scan ISM to the information collected from a 3-D scan without losing any relevant information. Additionally, it considers structural elements, mainly walls, to avoid representing noisy data.
3) A *map updating model*, which composes views provided by the sensor into a single grid representation.

In this section, the design of these three elements for the proposed application is further developed. The main objective of this work is creating an occupancy grid map that defines the state of the environment regions. Occupancy grid maps define the mapped scenario by a set of equally sized cells forming a grid. Each cell has an assigned probability that defines its occupancy probability. This value ranges from zero to one, with zero corresponding to empty regions, one corresponding to occupied regions, and 0.5 defining unknown zones. For an occupancy grid map cell, the state variable associated with a cell C, $s(C)$, has to fulfill that $P(s(C) = \text{OCC}) + P(s(C) = \text{EMP}) = 1$, where OCC refers to occupied and EMP to empty regions.

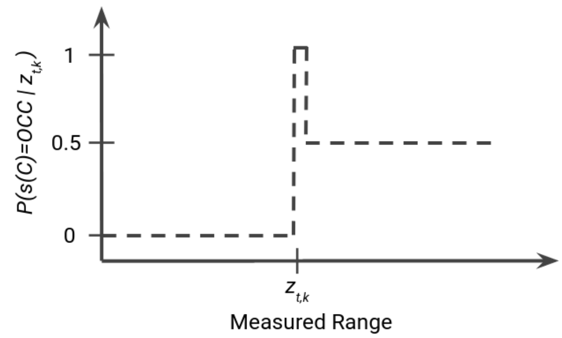


Fig. 2. ISM for a 2-D laser scan beam.

A. Coordinate Transformations

One of the most challenging problems in robot navigation is generating a map while locating the robot in it, which is known as simultaneous localization and mapping (SLAM). In order to know where to place laser information on a map, it is necessary to know where the robot is located. At the same time, sensor information must be compared to a map to know this information. SLAM solves this consistency by matching consecutive scans and recalculating the current robot position inside the map during scan registration. Although this information is essential, it is out of the scope of this article. We depart from the previous work presented in [8], in which Harmony search performs the optimization of robot poses to define a new approach for SLAM. The result is a set of optimized poses that will be used to iteratively update the map.

B. Inverse Sensor Model

One of the most important elements in the proposed methodology is the mathematical definition of data obtained from scan sensors. In this work, the probabilistic distribution of a cell C being occupied (OCC) given a sensor measurement z at a time t is represented with $P(s(C) = \text{OCC}|z_t)$, whereas for empty (EMP) regions it is expressed as $P(s(C) = \text{EMP}|z_t)$, where s refers to the state of the environment. In the case of 2-D laser scan sensors, each of their beams k measure a distance $z_{t,k}$ that indicates how far the nearest object is in the direction to which it is pointing. Hence, for each beam, the ISM is described as shown in Fig. 2. This figure represents that, for zones prior to the measurement taken, the probability of occupancy is zero. For the zone corresponding to $z_{t,k}$ the probability is one and beyond that region it is 0.5, unknown, since the zone has not been observed yet.

In the case of 3-D sensors, there is more than one horizontal plane on which distance information is measured. This implies that for each horizontal sector defined by resolution, there is a corresponding distance measurement on each height. The most commonly used approach in literature is to select the closest point at each horizontal resolution angle to generate a profile like the one shown in Fig. 2, where the highest probability would correspond to the closest obstacle. However, the other

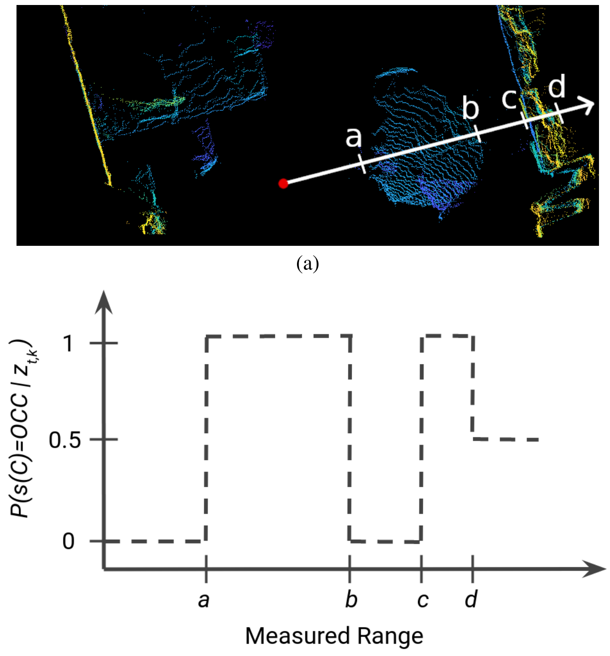


Fig. 3. Representative example of a 3-D ISM. (a) Bird's-eye view of a point cloud from which the ISM is derived. The red dot is the sensor origin and the white arrow represents the direction toward which the ISM is going to be defined. (b) ISM defined by the previous information. In comparison to the 2-D ISM, the profile has more and wider peaks.

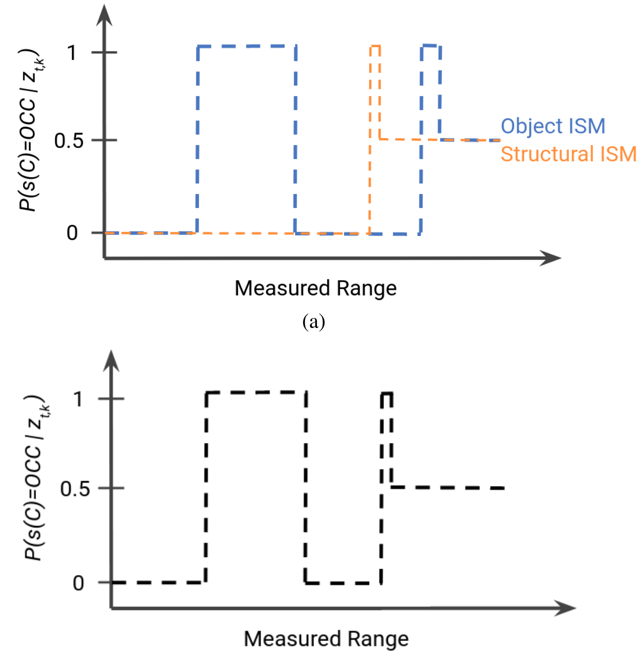


Fig. 4. Incorporation of structural elements into the sensor profile. (a) Object ISM (blue) and the structural ISM (orange) provide conflicting information, since they indicate opposing information. (b) Information is merged to create a robust object and reflection aware ISM.

measurements are also valuable information that should not be filtered out, as they provide a significant amount of spatial knowledge of the area.

In this work, the generation of a sensor profile that considers such information is proposed. The following procedure is performed on each measured point cloud. First, 3-D information is cropped from the robot's base to its maximum height, since it is in this range where potential collision objects exist. In this way, floor and ceiling points are removed as well. In this step, it is especially important to know the sensor's position with respect to the robot so that the trimming can be performed correctly. These cropped points are then projected to a horizontal plane. For each horizontal resolution angle, a virtual line is generated from the sensor origin to the maximum viewing range of the 3-D laser, and the information on the projected plane under it is analyzed. The profile starts as for a 2-D sensor, with the probability equal to zero. Each time an occupied region is encountered, the probability goes up to one, and goes back down to zero when it ends. This is repeated until the last occupied region is found, after which the profile remains at 0.5, unknown. This is visually represented in Fig. 3. Fig. 3(a) shows a bird's-eye view of a point cloud captured at a certain time t, where blue dots are points closer to the floor, yellow dots are closer to the ceiling and the point where the robot is placed is represented with a red dot. As an example, the white arrow indicates one of the horizontal directions at which the 3-D ISM has to be defined. This is shown in Fig. 3(b), where the probabilities from the origin to a and from b to c are zero, which means that there is free space, whereas for the regions between a and b and c and d the probability is one, which means that they are occupied zones. No information is found after d, so the probability is set to 0.5 after it.

C. Profile Modification for Filtering Reflection

The previous approach is valuable for considering 3-D information. However, it does not consider reflective surfaces. For this reason, we propose the study of the arrangement of structural elements to eliminate these effects. Similar to Section III-B, every point cloud is initially cropped, but this time at a greater height, below the ceiling but very close to it. In this way, everyday indoor environment elements such as chairs or tables do not appear. Dynamic elements such as people are not found either. Therefore, only the structural elements of the environment remain in the cutout. This approach only deals with reflective surfaces that are not directly touching the ceiling, so by cropping the cloud in the mentioned way, they are also avoided.

The method additionally assumes that a scan cannot pass through a structural element such as a wall. To detect these situations, a new profile is generated from the structural cutout. In this case, the relevant value is the first occupied cell, which indicates the limit of the known structural information. Beyond this point, distance information will not be reliable. Hence, this point is compared against the previously calculated profile. Any zone that satisfies $P(s(C) = \mathrm{OCC}|z_t) \neq 0.5$ after the mentioned point is removed, since there should not be any information after a structural element like a wall. Fig. 4 shows how these two profiles are merged rejecting unreliable data. Initially, two profiles are defined, one for object data (blue profile) and another one for structural data (orange line), as shown in Fig. 4(a). The structural ISM has only one peak, which corresponds to the structural element that is found in the defined direction. In the selected example, there is a discrepancy between this data and the object ISM, since the latter indicates free space beyond the wall. Both sources of information are merged by cropping object data from the structural ISM peak, leading to the profile shown in Fig. 4(b).

D. Map Updating Model

The problem of updating the occupancy grid map with sensor measurements located at a specified pose in space is performed using a recursive Bayesian filter as described in [25]

$$p(\mathrm{OCC}|z_{1:t}) = \frac{p(z_t|\mathrm{OCC}, z_{1:t-1})\,p(\mathrm{OCC}|z_{1:t-1})}{p(z_t|z_{1:t-1})} \tag{1}$$

where OCC is the simplification for $s(C) = \mathrm{OCC}$ and $z_{1:t}$ is the set of laser observations over time. This equation considers that the state of the world is defined by all the cell states in the map. Taking into consideration that each cell can either be empty or occupied, a total number of $2^{M\times N}$ conditional probabilities would be required to define an $M \times N$ grid map. In order to reduce complexity, the majority of state of the art works have modeled the problem as a Markov random field (MRF) of order 0, which means that each cell in the map is estimated as an independent variable [26]. In this way, the number of required parameters is reduced to $2 \times M \times N$. This assumption is not generally true, since it cannot be stated that an observation is independent of all previous observations. This fact can lead to conflicts for sensor beam models with a wide range of view, such as sonar. However, since laser scan sensors have very narrow beams, this natural error is not relevant. Hence, $p(z_t|\mathrm{OCC}, z_{1:t-1})$ can be assumed to be equal to $p(z_t|\mathrm{OCC})$. Applying Bayes' rule again, (1) becomes

$$p(\mathrm{OCC}|z_{1:t}) = \frac{p(\mathrm{OCC}|z_t)\,p(z_t)\,p(\mathrm{OCC}|z_{1:t-1})}{p(\mathrm{OCC})(p(z_t|z_{1:t-1}))} \tag{2}$$

where $p(\mathrm{OCC}|z_t)$ is the ISM.

The same procedure is applied for the empty probability, leading to the following equation:

$$p(\mathrm{EMP}|z_{1:t}) = \frac{p(\mathrm{EMP}|z_t)\,p(z_t)\,p(\mathrm{EMP}|z_{1:t-1})}{p(\mathrm{EMP})(p(z_t|z_{1:t-1}))}. \tag{3}$$

In order to obtain the belief of a cell being occupied over the belief of being empty, (2) and (3) are divided. Additionally, log odds are used because of numerical stability

$$\begin{aligned}\log\left(\frac{p(\mathrm{OCC}|z_{1:t})}{p(\mathrm{EMP}|z_{1:t})}\right) &= l_t(\mathrm{OCC})\\ &= \log\left(\frac{p(\mathrm{OCC}|z_t)}{p(\mathrm{EMP}|z_t)}\right) + \log\left(\frac{p(\mathrm{OCC})}{p(\mathrm{EMP})}\right)\\ &\quad + l_{t-1}(\mathrm{OCC}).\end{aligned} \tag{4}$$

The first term corresponds to the sensor model, the second one is the prior (which for $p(\mathrm{OCC}) = 0.5$ becomes zero), and the third one is the previous state of the map. By applying this formula, the map is updated recursively and probabilistically. The main advantage of this probabilistic formulation is that dynamic elements are filtered out, in addition to possible pose estimation and measurement errors.

Fig. 5. Example scenario in which ADAM was teleoperated to capture data. The 3-D sensor is located in its uppermost part.

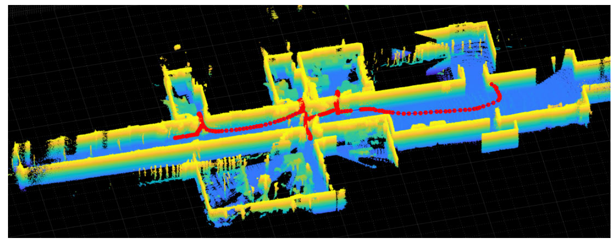


Fig. 6. Global point cloud corresponding to the test scenario. Robot poses, marked with red dots, are estimated using by Harmony search SLAM.

IV. Experimental Results

The objective of the presented experiments is to test the performance of the proposed method. A 2-D occupancy grid map is generated from data taken by a 3-D sensor, more specifically an Ouster 3-D LiDAR of 128 channels and 50 m resolution. The result is compared with other mapping techniques both qualitatively and quantitatively.

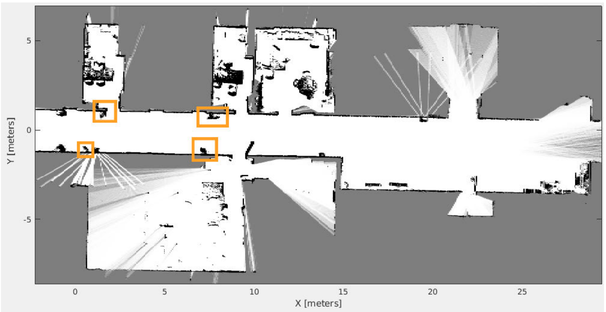
A. Experiment Setup

The 3-D laser scan sensor is mounted on Autonomous Domestic Ambidextrous Manipulator (ADAM), a bimanipulator mobile robot of 1.5 m height and 0.5 m diameter. A better description of the robot can be found in [28]. The robot was teleoperated through a hall with opened and closed doors, entered some rooms, and moved to a near open area. This real-world scenario has its own complexities, including glass windows, reflective materials like radiators and fire extinguishers and cluttered offices with several furniture pieces. Fig. 5 shows a scene of the scenario in which ADAM was deployed.

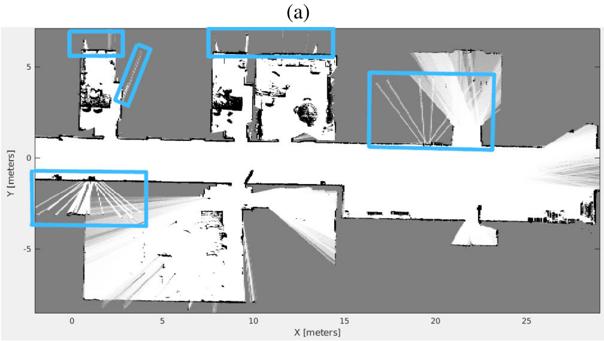
Data were captured using ROS melodic, saving 3-D point clouds and timestamps. Algorithms were developed and tested in MATLAB 2019b. Finally, before map building, Harmony search SLAM was executed to calculate estimated robot poses. The 3-D environment model is shown in Fig. 6, where red dots indicate points at which point clouds were recorded. Points corresponding to the ceiling have been removed to appreciate indoor details. It can be noted that the path does not enter rooms completely, and yet using 3-D information allows to reconstruct the rooms almost completely.

B. Map Building

The main objective of this section is to generate a 2-D occupancy grid map using the presented method. Starting from robot poses, 3-D geometric information is concatenated.



(a)



(b)

Fig. 7. Probabilistic mapping. (a) Mapping results when accumulating data. (b) Probabilistic mapping, where dynamic objects that appeared in the orange boxes are removed. However, noise due to reflection (blue boxes) is still present.

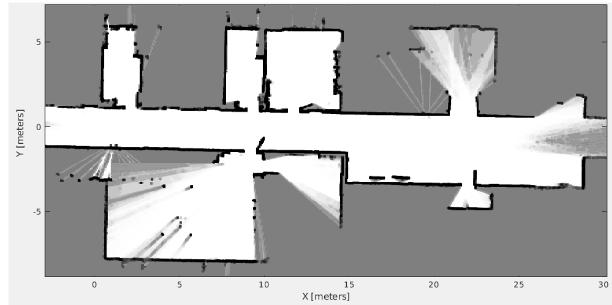


Fig. 8. Structural map built using data at a high distance above the robot.

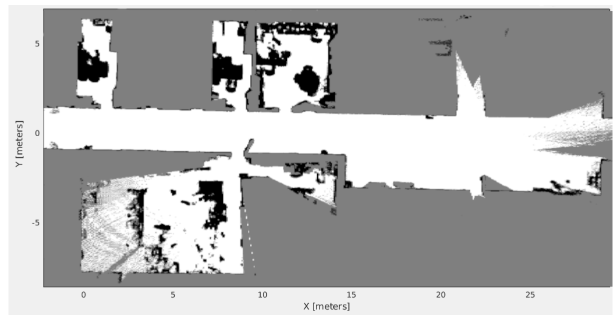


Fig. 9. Final occupancy grid map, where objects are included with their whole shape and reflection is removed. Occupied regions have been inflated for a better appreciation.

As explained above, each of the captured point clouds is cut out and projected in order to be included in the map scan by scan. For every scan, it was checked in which map cells the 3-D points coincided, and one unit was added every time this occurred. Finally, data were normalized. The main problem encountered with this method is that dynamic elements could not be filtered. This results in the classification of areas that are actually free as occupied, which limits the robot ability to move. After applying the probabilistic approach explained in Section III-D and including the sensor ISM, these elements are successfully filtered. A preliminary result before applying the proposed probabilistic approach and the result of modifying occupancy based on Bayes theorem is shown in Fig. 7. Dynamic elements corresponding to people (orange boxes) are removed in the final version of the map.

However, it can be noticed that noise caused by reflective surfaces still exists. This can be clearly noticed in Fig. 7(b), where blue rectangles indicate locations where this happens. The bottommost and rightmost areas correspond to fire extinguisher boxes. The rest are caused by glass windows and diverse metallic objects. To filter out this noise, structural data are analyzed. In order to visualize this information in a representative way, a structural map has been generated in which the 2-D scans generated from the 3-D information have been used, yielding the result shown in Fig. 8. Occupied regions have been inflated for a better visualization. It can be seen that the internal structure of the building is very clearly defined. The ISM based on the structural map is created as explained in Section III-C and it is compared against the ISM based on collisionable objects, so that the noise is eliminated. The final map is shown in Fig. 9, where occupied regions have been inflated for a better visualization. Areas mapped as thin lines, corresponding to reflections, are removed with respect to the previous version of the map shown in Fig. 7(b).

C. Comparison Against Other Map Building Techniques

The result obtained above has been compared against two other mapping techniques. The first one is one of the most widely used mapping techniques: gmapping [27]. Starting from data taken by a 2-D sensor, the method applies SLAM to aggregate scans and extract the occupancy grid map. Meanwhile, the second method is the most widely used to simplify data from a 3-D sensor to generate a 2-D occupancy map. For each range angle, the closest occupancy value is selected and projected to the ground plane to generate a new simulated 2-D laser. It should be noted that the execution time of both methods is faster than in our method since they work with 2-D information. Therefore, this metric is not decisive. The proposed comparison focuses on map performance.

The result of applying gmapping is shown in Fig. 10. Values from the robot's 2-D sensor have been captured at the same time as the 3-D sensor data, so they correspond to information from the same locations that were previously mapped. As can be seen, room geometries are approximately maintained. However, the number of mapped objects is significantly lower. In areas where chairs and tables can be found in the real scenario, empty regions are found with small noise zones due to the legs of these pieces of furniture. With the proposed method, this fact is corrected, showing the complete geometry of elements. Thus, the advantage of using the proposed method over the more popular one is illustrated. By considering the

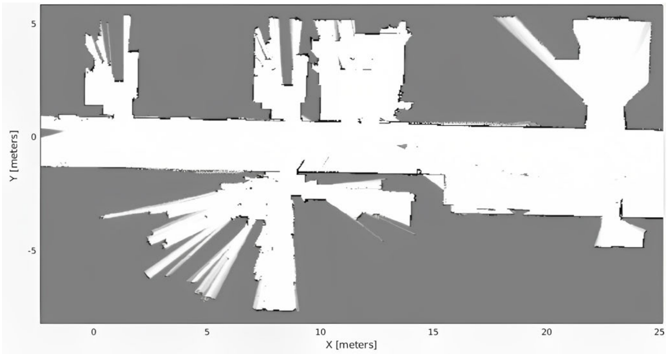


Fig. 10. Occupancy grid map built using a 2-D laser scan sensor and ROS gmapping.

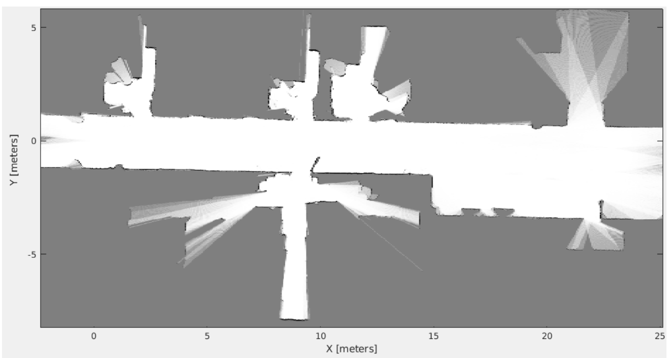


Fig. 11. Occupancy grid map built using the state-of-the-art approach in which only the closest element to the sensor is kept for each viewing angle.

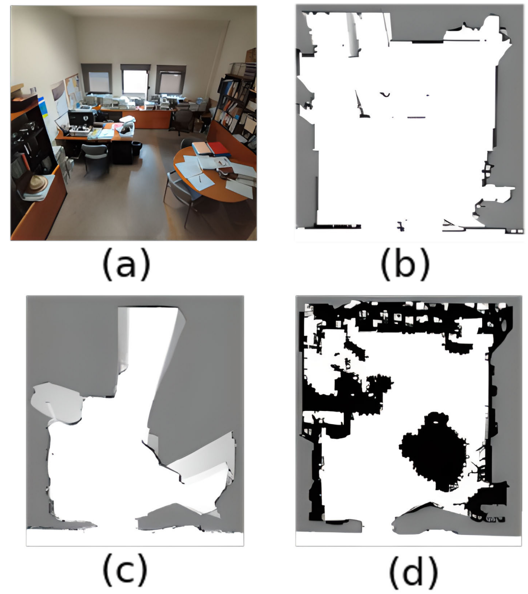


Fig. 12. Close-up view of room maps. (a) Real-world scenario. (b) Gmapping result. (c) Result using the closest occupied point. (d) Result with our proposed method.

TABLE I
PERFORMANCE METRICS

	Precission	Recall	F-Score
Gmapping [27]	0.528	0.810	0.631
Closest Point [10]–[14]	0.913	0.610	0.731
Our method	**0.957**	**0.917**	**0.937**

complete geometry of elements, a much safer navigation for robots is ensured, since the representation closer to reality.

The result of applying the simplification of keeping the closest point is shown in Fig. 11. In this case, as we need to start from data in three dimensions, the same scans that were recorded for the map proposed in this work have been used. In this case, the complete geometry of objects is indeed considered, since the information that is used to build the map takes into account that the region from the robot to the nearest occupied point is free. However, it can be observed that there is a large information loss. Even if a point is detected as occupied by the 3-D sensor, the information farther away from it provides us with more spatial information that does not have to be filtered out. As shown in the map built with this technique, the rooms are not completely reconstructed with the path traveled by the robot. While with our method, the rooms are mapped with a small glimpse of the robot near the door, for this other method, the robot would have to do much more traversing to complete the map.

For a better appreciation of these characteristics, a close-up view of one of the rooms is shown in Fig. 12. In this room, there is a large round table with chairs around it in the lower right corner, a desk in the upper left corner, and bookshelves attached to the walls. As it can be seen, none of the two tables are mapped on the gmapping case. There are only some small mapped regions for the desk placed at the top left corner, but they do not fully constrain navigation in that zone, so it is still a collision risk for the robot. In the second map, these pieces of furniture are delimited but information behind them are filtered out, eliminating relevant information. With our method, shown in the third map, free and occupied regions are well defined, being closer to reality.

Finally, quantitative results are shown in Table I. Precision indicates the ratio of correctly labeled cells out of the total number of cells marked as free or occupied. Recall indicates the ratio of cells that have been correctly labeled out of the total number of cells that should be labeled. F-score indicates the overall performance. As it can be seen, gmapping is not capable of clearly differentiating free and occupied regions (0.528 prec.), whereas closest point is missing relevant information (0.610 prec.). On the contrary, our method outperforms in both cases, with the highest value in all three metrics.

V. CONCLUSION AND FUTURE WORK

In this article, a method for generating robust 2-D grid maps using 3-D information has been proposed. As it has been observed, the generation of a new 3-D ISM has allowed the creation of maps without losing spatial information. In addition, by analyzing the environment structure, a new technique for filtering reflection noise has been proposed. Finally, thanks to the probabilistic data treatment, dynamic elements encountered during data collection are filtered out, resulting in a map that is more faithful to reality.

As future work, since we start from a robust geometric basis, we intend to carry out navigation applications for human assistance with domestic robots such as object search in which semantic elements are incorporated.

References

[1] S. Thrun, “Learning metric-topological maps for indoor mobile robot navigation,” *Artif. Intell.*, vol. 99, no. 1, pp. 21–71, Feb. 1998.

[2] PAL Robotics. *Tiago—Mobile Manipulator Robot*. Accessed: May 31, 2023. [Online]. Available: https://pal-robotics.com/robots/tiago/

[3] Aldebaran. *Pepper the Humanoid and Programmable Robot*. Accessed: May 31, 2023. [Online]. Available: https://www.aldebaran.com/en/pepper

[4] A. Eirale, M. Martini, L. Tagliavini, D. Gandini, M. Chiaberge, and G. Quaglia, “Marvin: An innovative omni-directional robotic assistant for domestic environments,” *Sensors*, vol. 22, no. 14, p. 5261, Jul. 2022.

[5] M. Grinvald et al., “Volumetric instance-aware semantic mapping and 3D object discovery,” *IEEE Robot. Autom. Lett.*, vol. 4, no. 3, pp. 3037–3044, Jul. 2019.

[6] A. Rosinol et al., “Kimera: From SLAM to spatial perception with 3D dynamic scene graphs,” *Int. J. Robot. Res.*, vol. 40, nos. 12–14, pp. 1510–1546, Dec. 2021.

[7] Y. Pan, Y. Kompis, L. Bartolomei, R. Mascaro, C. Stachniss, and M. Chli, “Voxfield: Non-projective signed distance fields for online planning and 3D reconstruction,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2022, pp. 5331–5338.

[8] P. Gonzalez, A. Mora, S. Garrido, R. Barber, and L. Moreno, “Multi-LiDAR mapping for scene segmentation in indoor environments for mobile robots,” *Sensors*, vol. 22, no. 10, p. 3690, May 2022.

[9] A. Mora, A. Prados, P. González, L. Moreno, and R. Barber, “Intensity-based identification of reflective surfaces for occupancy grid map modification,” *IEEE Access*, vol. 11, pp. 23517–23530, 2023.

[10] K. Kamarudin et al., “Method to convert Kinect’s 3D depth data to a 2D map for indoor SLAM,” in *Proc. IEEE 9th Int. Colloq. Signal Process. Appl.*, Mar. 2013, pp. 247–251.

[11] M. F. A. Ghani, K. S. M. Sahari, and L. C. Kiong, “Improvement of the 2D SLAM system using Kinect sensor for indoor mapping,” in *Proc. Joint 7th Int. Conf. Soft Comput. Intell. Syst. (SCIS), 15th Int. Symp. Adv. Intell. Syst. (ISIS)*, Dec. 2014, pp. 776–781.

[12] A. Oliver, S. Kang, B. C. Wünsche, and B. MacDonald, “Using the Kinect as a navigation sensor for mobile robotics,” in *Proc. 27th Conf. Image Vis. Comput. New Zealand*, Nov. 2012, pp. 509–514.

[13] X. Wang, P. Ma, L. Jiang, L. Li, and K. Xu, “A new method of 3D point cloud data processing in low-speed self-driving car,” in *Proc. IEEE 3rd Adv. Inf. Manage., Communicates, Electron. Autom. Control Conf. (IMCEC)*, Oct. 2019, pp. 69–73.

[14] O. Wulf, C. Brenneke, and B. Wagner, “Colored 2D maps for robot navigation with 3D sensor data,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep./Oct. 2004, pp. 2991–2996.

[15] L. Garrote, J. Rosa, J. Paulo, C. Premebida, P. Peixoto, and U. J. Nunes, “3D point cloud downsampling for 2D indoor scene modelling in mobile robotics,” in *Proc. IEEE Int. Conf. Auto. Robot Syst. Competitions (ICARSC)*, Apr. 2017, pp. 228–233.

[16] H. Gim, M. Jeong, and S. Han, “Autonomous navigation system with obstacle avoidance using 2.5D map generated by point cloud,” in *Proc. 21st Int. Conf. Control, Autom. Syst. (ICCAS)*, Oct. 2021, pp. 749–752.

[17] M. Missura, A. Roychoudhury, and M. Bennewitz, “Polygonal perception for mobile robots,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2020, pp. 10476–10482.

[18] S. Wu and S. Wang, “Method for detecting glass wall with LiDAR and ultrasonic sensor,” in *Proc. IEEE 3rd Eurasia Conf. IoT, Commun. Eng. (ECICE)*, Oct. 2021, pp. 163–168.

[19] A. J. Barreto-Cubero, A. Gómez-Espinosa, J. A. E. Cabello, E. Cuan-Urquizo, and S. R. Cruz-Ramírez, “Sensor data fusion for a mobile robot using neural networks,” *Sensors*, vol. 22, no. 1, p. 305, Dec. 2021.

[20] X. Wang and J. Wang, “Detecting glass in simultaneous localisation and mapping,” *Robot. Auto. Syst.*, vol. 88, pp. 97–103, Feb. 2017.

[21] X. Zhao, Z. Yang, and S. Schwertfeger, “Mapping with reflection-detection and utilization of reflection in 3D LiDAR scans,” in *Proc. IEEE Int. Symp. Saf., Secur., Rescue Robot. (SSRR)*, Nov. 2020, pp. 27–33.

[22] R. Wang, J. Bach, and F. P. Ferrie, “Window detection from mobile LiDAR data,” in *Proc. IEEE Workshop Appl. Comput. Vis. (WACV)*, Jan. 2011, pp. 58–65.

[23] G. Cui, M. Chu, W. Wangjun, and S. Li, “Recognition of indoor glass by 3D LiDAR,” in *Proc. 5th CAA Int. Conf. Veh. Control Intell. (CVCI)*, Oct. 2021, pp. 1–4.

[24] Z. He, J. Hou, and S. Schwertfeger, “Furniture free mapping using 3D LiDARs,” in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2019, pp. 583–589.

[25] L. Matthies and A. Elfes, “Integration of sonar and stereo range data using a grid-based representation,” in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 1988, pp. 727–733.

[26] A. Elfes, “Using occupancy grids for mobile robot perception and navigation,” *Computer*, vol. 22, no. 6, pp. 46–57, Jun. 1989.

[27] *Gmapping—ROS Wiki*. Accessed: Jun. 13, 2023. [Online]. Available: http://wiki.ros.org/gmapping

[28] R. Barber et al., “A multirobot system in an assisted home environment to support the elderly in their daily lives,” *Sensors*, vol. 22, no. 20, p. 7983, Oct. 2022.

Alicia Mora received the B.S. degree in industrial electronics and automation engineering and the M.S. degree in robotics and automation from the Carlos III University of Madrid, Leganés, Spain, in 2020 and 2022, respectively, where she is currently pursuing the Ph.D. degree with the Robotics Lab.

She is collaborating as a Researcher with the Robotics Lab, Carlos III University of Madrid. Her main research interests include mobile robots, including mapping considering geometric, topological and semantic representations, as well as navigation for assistive robots.

Ramon Barber (Senior Member, IEEE) received the B.Sc. degree in industrial engineering from the Polytechnic University of Madrid, Madrid, Spain, in 1995, and the Ph.D. degree in industrial technologies from the Carlos III University of Madrid, Leganés, Spain, in 2000.

He is currently an Associate Professor with the Department of System Engineering and Automation, Carlos III University of Madrid. His research interests include mobile robotics, including environment perception, modeling, planning, localization and navigation tasks considering geometric, topological, and semantic representations.

Dr. Barber is a member of the International Federation of Automatic Control (IFAC).

Luis Moreno (Member, IEEE) received the B.Sc. degree in industrial engineering and the Ph.D. degree in industrial engineering from the Universidad Politécnica de Madrid, Madrid, Spain, in 1984 and1988, respectively.

From 1986 to 1994, he was a Lecturer and an Associate Professor with the Universidad Politécnica de Madrid. Since 1994, he has been a Professor with the Polytechnic School of Universidad Carlos III de Madrid, Leganés, Spain, and since 2009, he has been a Full Professor. He has participate in more than 40 research projects with public funding (National and European) and 26 projects with private funding. He has authored four patents, five books, 18 books chapters, 80 Journal papers, and 160 conference papers. He has been a supervisor of 28 Ph.D. thesis. He has collaborated in creation of five technological startups companies originated in its research group. His main research interests include robotics, in particular mobile manipulators, lightweight robots, advanced actuators, path planning, perception, navigation, environment modeling, and exoskeletons.