Completing Robot Maps by Predicting the Layout of Rooms Behind Closed Doors

Matteo Luperto[1], Federico Amadelli[2], and Francesco Amigoni[2]

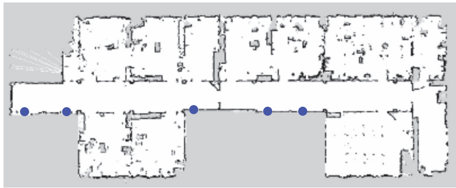
***Abstract*—The availability of maps of indoor environments is often fundamental for autonomous mobile robots to efficiently operate in industrial, office, and domestic applications. When robots build such maps, some areas of interest could be inaccessible, for instance, due to closed doors. As a consequence, these areas are not represented in the maps, possibly limiting the activities robots can perform. In this paper, we provide a method that completes 2D grid maps by adding the predicted layout of the rooms behind closed doors. The main idea of our approach is to exploit the underlying geometrical structure of indoor environments to estimate the shape of unobserved rooms. Results show that our method is accurate in completing maps also when large portions of environments cannot be accessed by the robot during map building.**

I. Introduction

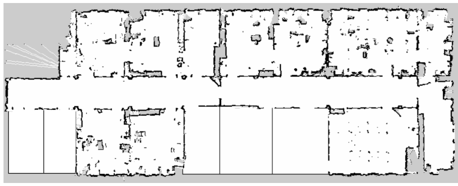
In recent years, ground mobile robots have been successfully employed in several indoor applications in industrial, office, and domestic environments [1]. When a robot is deployed in a new setting, it often autonomously builds a *map* representing the environment in which it operates. Then, the robot exploits the map to efficiently localize, navigate, and plan paths and tasks in the environment. Sometimes, the robot building the map and the robot using the map are different. 2D metric maps, like *grid maps*, are widely employed since they can be built from data coming from pervasive and relatively cheap sensors like 2D laser range scanners by using consolidated 2D SLAM methods [2]. Moreover, such maps are rather robust to events like day/night light changes, the presence of people, and objects moving around [1].

Ideally, a map should represent the entire operational environment of the robot. However, during the process of map building, it could happen that some areas of interest for the robot's activity are inaccessible, due to temporary conditions that are beyond the robot's control, like a blocked path or a closed door. As a consequence, these areas are not represented on the map, and this can limit the autonomy and operations of the robot exploiting the map. For example, if the robot is unaware of the presence of some rooms behind closed doors, it has no means to plan in advance the actions to be performed when the doors are opened (e.g., in order to map them).

In this paper, we provide an initial contribution towards solving the above problem, by presenting a method that completes robot 2D grid maps with the predicted layouts (i.e., the geometrical shapes) of unobserved rooms behind closed doors, which we call *closed rooms*. The main idea of our approach is to exploit the underlying geometrical structure of indoor environments that can be detected from the walls to provide knowledge about parts of the environment that are not directly observable at mapping time. This estimated knowledge, although approximated [3], could provide meaningful insights to the robot about the structure of the environment and could be exploited in tasks such as exploration [4], localization, task planning, and reasoning. Note that this paper focuses on how to predict and complete the map, while its uses will be addressed in future work.



(a) Map built when some doors are closed.



(b) Predicted layouts of closed rooms.

Fig. 1: An example run of our approach for predicting layouts of rooms that are behind doors closed at mapping time (blue dots), where we simulated 5 closed doors in a map from [5].

We assume a robot that is able to build a 2D grid map of an indoor environment in which some doors are closed and that is able to detect the positions of such closed doors. Since detecting doors (e.g., from vision) is not the purpose of this paper, we assume that the robot employs a method like [6], [7]. Given a grid map (Fig. 1a) and the positions in the map of the closed doors, our method identifies the main structural features of the environment by detecting the walls. The directions of the walls are associated with representative lines that are used to partition the map into a number of polygonal faces. Although we assume that most of the walls can be approximated by straight lines, which is the case for the vast majority of indoor environments, we do not enforce any Manhattan structure, but we use the walls' main directions, directly retrieved from the map, for making predictions. The predicted layout of a closed room is the set of faces that maximize an evaluation function that accounts for the consistency with the known portion of the environment. Finally, the predicted layouts of closed rooms are inpainted within the grid map (Fig. 1b). An interesting feature of our method is that it can jointly predict the layouts of multiple adjacent closed rooms (e.g., when all rooms along

[1]M. Luperto is with the Applied Intelligent Systems Laboratory, Università degli Studi di Milano, via Festa del Perdono 7, 20122 Milano, Italy matteo.luperto@unimi.it

[2]F. Amadelli and F. Amigoni are with the Artificial Intelligence and Robotics Laboratory, Politecnico di Milano, piazza Leonardo da Vinci 32, 20133 Milano, Italy francesco.amigoni@polimi.it

978-1-6654-1213-1/21/$31.00

the same side of a corridor are closed). Experimental evaluation is performed by considering both large-scale simulated environments and real-world grid maps from publicly available datasets. Results show that our method successfully predicts the layout of closed rooms and accurately completes grid maps even when large portions of the environments are not accessible at mapping time

II. Related Work

In this section, we survey some techniques developed across different fields to retrieve structural knowledge in indoor environments and to predict their unobserved parts.

Within the field of mobile robotics, a popular approach towards structure identification in indoor environments is room segmentation, where rooms are identified by dividing a metric map into semantically meaningful parts. A survey that compares 2D room segmentation methods is reported in [8]. Authors of [9] present a room segmentation method that uses structural line features similar to the representative lines we use in this paper. Another popular problem is the structure identification in indoor environments from 3D point clouds [10]–[14]. We take inspiration from the structure identification steps of [14], but we adapt them to noisy 2D grid maps. While the above methods identify structures within the map representing the observed part of the environment, in this paper we are interested in predicting the structure of unobserved rooms.

The method of [15], starting from a 2D metric map of an indoor environment, reconstructs a geometrical representation of (partially) observed rooms using Markov Logic Networks and data-driven Markov Chain Monte Carlo (MCMC) sampling. The shape of a room, approximated by a polygon, is obtained using a set of logic rules identifying the desired properties (e.g., perpendicular walls, box model). In [16], we propose a method to complete partially observed rooms. In this paper, we assume to have no knowledge about the closed rooms whose layout we predict.

The idea of obtaining knowledge on unobserved parts of environments has recently been addressed using heterogeneous approaches. Some of them predict unknown features of an environment by exploiting knowledge coming from other maps previously acquired in the same or in other environments. For instance, [17] predicts loop closures in a metric map. The parts of the metric map that are yet unknown are completed by superimposing matching maps from a database of previously observed environments. The method of [18] uses a library of map structures to predict the unknown parts of a map incrementally explored by a team of robots. Another recent method is that of [19], where a variational autoencoder (VAE) is employed to predict unobserved regions of an environment starting from a partial map. However, [19] considers buildings that are very similar to each other (see [20] for a discussion about a different, but related, setting), thus providing an approach that seems difficult to generalize to other different environments. Several methods predict the presence of specific elements in the unobserved parts of environments using neural networks trained on similar environments. For instance, [21] trains a convolutional neural network (CNN) on a set of images representing building floor plans and uses it to predict the locations of emergency exits. In [22], U-nets, a type of CNNs, are used to expand egocentric RGB-D observations to infer the occupancy state beyond the visible regions. In doing so, the robot can anticipate the next sensorial readings.

The method of [23] uses Conditional Neural Process for predicting the local map of the unobserved parts of an environment to improve online trajectories planning.

Some other approaches, like the one we present in this paper, do not rely on external knowledge but only on the content of the partial map. In this sense, a method that shares some similarities with our approach is that of [24], which reconstructs the neighborhood of a frontier between known and unknown portions of a map by identifying similar structures in the known map. If a local match is found, the matching portion of the map is superimposed to the frontier, thus providing an estimate of the structure of its neighborhood.

Some methods predict the existence and the semantic labels of unobserved rooms, but they do not predict their geometrical shapes nor update the metric map with the prediction, as we do in this paper. Examples are the systems proposed in [25], that uses semantic knowledge in the form of chain graphs to predict the existence of a room (and its label) in the unexplored space, and in [26], where the prediction of the existence of a new room is made by using sum-product networks. The approach of [27] predicts the topology and the labels of unobserved rooms by matching the observed part of the environment (represented as a labeled graph) to a database of environments. Also [28] predicts the presence of new rooms in partially observed environments by by reasoning on graphs using graph kernels.

To the best of our knowledge, no method addresses our specific problem, that of predicting the geometrical shape of rooms behind closed doors.

III. Our Method

Our method predicts the layout of rooms behind closed doors in a purely geometrical way, without learning models from other environments. It starts from a 2D *grid map* M of an indoor environment obtained by a robot through a SLAM mapping process using data acquired by a laser range scanner. This map is composed of identical square cells that are labeled as known or unknown according to the fact that they have been perceived or not by the robot. Known cells are further labeled as either free or obstacle, according to the occupancy of the corresponding area. The grid map is assumed partial in the sense that some rooms could not be accessed by the robot during the mapping process due to some *closed doors* and, as a consequence, are not included in the map. We propose a method that predicts the possible layouts of these *closed rooms*, which are then used to complete the grid map.

Our method assumes that the robot can detect the position of closed doors inside the environment (blue dots in Fig. 2a), for example by using existing computer vision methods like

(a) Partial map with closed doors. (b) Inpainted map with predicted closed rooms. (c) Full map.

Fig. 2: A partial map of a large-scale indoor environment with 14 closed doors (2a), the map completed by our method (2b), and the full map of the same environment (2c).

[6], [7]. Consequently, the initial input of our method is a grid map M and a set of closed door locations $d \in D$. Although doors are represented as line segments in 2D maps, our method considers their middle points. Hence, each $d = (x, y)$ represents the mid-point coordinates (coordinate system of M) of the line segment corresponding to a closed door in the map. We assume that each closed room has exactly one door.

Our method predicts the possible layouts of closed rooms by leveraging the fact that, due to the structured nature of indoor environments, their geometrical shapes have some common features with other rooms and walls as in the metric map. The method is based on a sequence of steps that are detailed in the next sections using the map of Fig. 2 as a reference.

A. Structural features

In the first step, we use the method of [16] to extract the structure of indoor environments by identifying the direction of walls in the metric map M and to partition M using those directions. The method starts by extracting a set of line segments from M by using Canny edge detection [29] and probabilistic Hough line transform [30]. Line segments are clustered together in two steps. First, the mean shift algorithm [31] clusters together line segments with similar angular coefficients. Then, for each angular cluster, all line segments that are also collinear (along the same line) are clustered together by performing spatial clustering. Full details are omitted for brevity, please refer to [16].

Each spatial cluster is then associated to a *representative line*, in red in Fig. 3, which indicates the direction of collinear, but possibly spatially separated, walls. The result is the detection of a (hopefully small) number of representative lines that describe the direction of all the walls within the environment. Four additional *boundary lines* are added at the extremity of M at a fixed distance from the bounding box of the map and with the same angular coefficients of the two largest angular clusters of line segments. We do not assume Manhattan environments, as the directions of the representative lines are directly recovered from the map. However, in many real-world indoor environments, most walls are perpendicular (e.g., see [32]) and, consequently, the representative lines used for map segmentation are often perpendicular. For example, the Intel Lab map from [5], shown in Fig. 4, contains curved walls, but representative lines corresponding to most walls are perpendicular. The resulting approximation is adequate for accurately predicting the geometrical shape of closed rooms, as shown in Section IV.

The intersections of all the representative lines partition the map into *faces*. A face $f \in \mathcal{F}$ is a polygon having as *edges* the line segments obtained by the intersections of the representative lines (Fig. 3). The faces with an edge belonging to a boundary line are called *border faces*.

Finally, we separate the faces that are inside the part of the environment observed by the robot from those that belong to the unobserved parts of the environment, as only the latter ones will be considered when predicting the layout of a closed room. Specifically, we keep only faces $f \in \mathcal{F}_u \subseteq \mathcal{F}$ where at least a 30% of their area is unknown.

B. Closed room locations

The prediction of the geometrical shape of closed rooms starts from the faces that are immediately behind closed doors. We associate each closed door $d = (x, y)$ to its closest edge e_d (on a representative line), thus determining the door orientation as collinear to e_d. As edges are the boundaries between faces, we consider the two faces that share an edge e_d: one of the two faces is inside the known part of M and the other one belongs to the unknown part of M. The latter face is the one that is behind the closed door d and is inserted in the set of initial faces $I \subseteq \mathcal{F}_u$, which are used as seeds to estimate the layouts of the rooms behind the closed doors. Initial faces of closed rooms for the map of Fig. 2 are shown in Fig. 3.

A particular case arises when there are two closed doors, d_1 and d_2, that are associated with the same edge e_d along a representative line ℓ (and, consequently, that have the same initial face). As we assume that there is only one door for each closed room, we artificially add a representative line perpendicular to ℓ and passing at equal distance from the two doors' positions d_1 and d_2. In this way, e_d is split into e_{d_1} and e_{d_2}. This allows us to address situations, as in the corridors of Fig. 6, with multiple closed rooms adjacent to each other.

C. Closed rooms expansion

The predicted layout of a closed room r in the environment is composed of one or more faces $f \in \mathcal{F}_u$ and is obtained by selecting the most likely set of faces from $\mathcal{F}_u$, adjacent to its initial face i_r, according to the surrounding environment.

We define as $d(f, f')$ the topological distance between faces f and f'. For instance, if two faces have one common edge, their distance is 1. The process of identifying the predicted layouts of closed rooms is performed greedily by jointly iteratively expanding them by considering an increasingly larger set of faces. More precisely:

(1) We initialize $k = 1$.

(2) For each closed room r, we select a set of candidate faces from $\mathcal{F}_u$ as:

$$F_r^k = \{f : f \in \mathcal{F}_u \mid \exists f' \in \hat{F}_r^{k-1} \text{ and } d(f, f') = 1\},$$

where $\hat{F}_r^0 = \{i_r\}$ (with $i_r \in I$) is the initial face behind door d of room r. Calling $\mathcal{P}(\cdot)$ the power set and $\Phi(\cdot)$

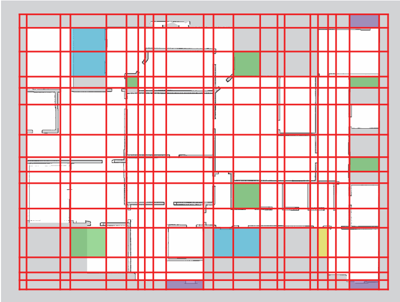


Fig. 3: Representative lines (red) and faces obtained from segmenting the map of Fig. 2. Initial faces of closed rooms are shown with different colors: green for independent rooms and light blue for dependent rooms, while initial faces of closed rooms on the border (border faces) are in purple. A particular case is the initial face in yellow, which represents a room that is initially independent but, expanding, becomes dependent by touching the predicted layout of a nearby closed room (in light blue at its left).

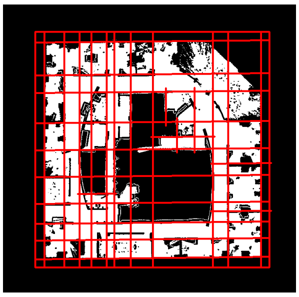


Fig. 4: Representative lines for the Intel Lab map from [5].

an evaluation function (described in Section III-D), we select the best layout for room r at step k as the set of faces that maximize $\Phi(\cdot)$, expanding the layout at step $k-1$:

$$\hat{F}_r^k = \arg\max_{F \in \mathcal{P}(F_r^k)} \Phi(\hat{F}_r^{k-1} \cup F).$$

We remove the faces in $\hat{F}_r^k$ from $\mathcal{F}_u$ (so that a face belongs to the predicted layout of at most one closed room) and we consider the next closed room r. For a given k, we consider closed rooms ordered from the smallest to the largest F_r^k (ties are broken randomly). However, we empirically observed that room ordering has a small impact on the final result.

(3) We increase $k \leftarrow k+1$ and we repeat from (2), until no faces are left in $\mathcal{F}_u$ or a threshold for k is reached. For each room, set of faces $\hat{F}_r^* = \hat{F}_r^k$ selected in the last step is considered as the predicted layout.

At a generic step k, the predicted layout $\hat{F}_r^k$ of a closed room r is thus updated from the predicted layout $\hat{F}_r^{k-1}$ at the previous step. The idea is that we jointly expand the predicted layouts of all closed rooms until a good estimate is found for each one of them. This is motivated by the fact that closed rooms can belong to two categories: *independent* closed rooms, whose predicted layout is not adjacent to the predicted layout of any other closed room; *dependent* closed rooms, which have at least a face of their predicted layout that is adjacent to a face of the predicted layout of another closed room. Examples of independent (dependent) rooms are shown in green (light blue) in Fig. 3. Note that, with the increase of k, some independent closed rooms may become dependent; this happens when the predicted layouts $\hat{F}_r^k$ and $\hat{F}_{r'}^k$ of two rooms r and r' are expanded in opposing direction, eventually sharing adjacent faces. An example of this is the room in yellow of Fig. 3. The reason for separating these two closed room types is that, while the predicted layouts of independent rooms should be consistent only with M, the predicted layouts of dependent rooms should be jointly estimated with that of the nearby closed rooms.

D. Predicted layout evaluation

A possible predicted layout of a room r, represented as a set of faces $\hat{F}_r^{k-1} \cup F$, is scored using an evaluation function $\Phi(\hat{F}_r^{k-1} \cup F)$. The function embeds competing objectives, like to maximize the area of the room and to maximize the coherence of the room structure wrt that of nearby rooms. Because of that, the objective function is a weighted sum of five different components, that are now described. In what follows, with a slight abuse of notation, we use F to denote the layout $\hat{F}_r^{k-1} \cup F$.

The first component, $\mathit{area}(F)$, is the room area.

The second component is the convex hull ratio *CHR*, that prefers regular room layouts:

$$\mathit{CHR}(F) = \mathit{CH}(F)/\mathit{area}(F),$$

where *CH* is the area of the room's predicted convex hull.

The third and fourth components are designed to maximize the similarity between the predicted layout of the room and the rest of the map. In particular, the third component minimizes the edges of a predicted layout that touch the unknown parts of the map. More precisely, the ratio of the free edges *FER* is defined as the ratio between the sum of the length of the edges of F that are also edges of a face $f \in \mathcal{F}_u$ (where $\mathcal{F}_u$ is the set of unobserved faces remaining after the application of the algorithm in the previous section) and the sum of the length of all edges along the external contour of F. The fourth component penalizes the predicted layouts that are not regular. More precisely, free faces penalty *FFP* is defined as the number of faces $f \in F$ that have at least *two* edges in common with faces $f \in \mathcal{F}_u$.

The fifth component is the room proportion P, the ratio between the two main dimensions of the room's bounding box, which is intended to favor regular predicted layouts.

We define two different evaluation functions, Φ_{ind} and Φ_{dep}, for independent and dependent rooms, respectively:

$$\Phi_{ind}(F) = \omega_1 \cdot \sqrt{\mathit{area}(F)} - \omega_2 \cdot \mathit{CHR}(F) - \omega_3 \cdot \mathit{FER}(F) + \omega_4 \cdot \mathit{FFP}(F),$$

$$\Phi_{dep} = \omega_1 \cdot \sqrt{area(F)} - \omega_2 \cdot CHR(F) - \omega_5 \cdot FER(F)$$
$$- \max(\omega_4 \cdot (FFP(F) - 1), 0) - \omega_6 \cdot (P(F) \cdot \min(FFP(F), 1)).$$

In the case of multiple adjacent closed rooms, the last term in Φ_{dep} tends to prevent that the expansion is stopped before all these rooms have a similar shape.

We do not enforce a square or rectangular shape for the predicted layouts (as shown, for example, Fig. 6), but our evaluation function aims at predicting accurate room shapes according to the observed map. However, since real-world indoor environments are inherently structured and most walls are perpendicular also in non-Manhattan environments (Section III-A and Fig. 4), good predictions are usually rectangular.

The next two steps address special cases.

E. Joint rooms layout prediction

At the end of the expansion (Section III-C) it could be the case that two adjacent dependent rooms have different shapes (e.g., this happens when one of the two rooms is by chance initially expanded in the direction of the other room's initial face, thus limiting the second room expansion). To adjust such situations, we allow adjacent dependent closed rooms the possibility to swap one or more faces between them. Given two sets of faces $\hat{F}^*_r$ and $\hat{F}^*_{r'}$, representing the predicted layouts of dependent rooms r and r', we compute E as the set of faces in $\hat{F}^*_r$ (or in $\hat{F}^*_{r'}$) that have an edge in common with a face in $\hat{F}^*_{r'}$ ($\hat{F}^*_r$) and that, consequently, could be exchanged between the two rooms. We jointly evaluate all the possible combinations of face assignments $\mathcal{P}(E)$ (in one assignment, some faces of E are assigned to r, the other ones to r') by evaluating the corresponding rooms' predicted layouts $\bar{F}^*_r$ $\bar{F}^*_{r'}$ using the following function:

$$\Phi_{joint}(\bar{F}^*_r, \bar{F}^*_{r'}) =$$
$$\omega_7 \cdot \sqrt{\min(area(\bar{F}^*_r), area(\bar{F}^*_{r'})) / \max(area(\bar{F}^*_r), area(\bar{F}^*_{r'}))}$$
$$- \omega_8 \cdot (FFP(\bar{F}^*_r) + FFP(\bar{F}^*_{r'}))$$

We eventually select the face assignment that maximizes Φ_{joint} and we swap the corresponding faces between r and r'. If the adjacent rooms are more than two, they are considered in pairs.

F. Closed rooms on the borders

The layout of a closed room r may extend outside the bounding box of the current map. In that situation, we cannot directly use faces and representative lines to predict the layout of r, as M does not provide any knowledge on one of the dimensions of the room. This happens when the initial face i_r of a room r is one of the border faces (e.g., that in purple in Fig. 3). To provide a layout also in this case with limited information, we roughly predict the shape of the room as a square (of the same size as the edge e_d of the initial face i_r, see Section III-B). If there are multiple dependent rooms in this condition, we adjust (by averaging) their outwards dimension to the same value.

# env	1	1	1	2	2	4	9
$\max \lvert D \rvert$	7	9	10	11	12	13	15

TABLE I: Number of simulated environments and corresponding $\max |D|$ number of closed doors.

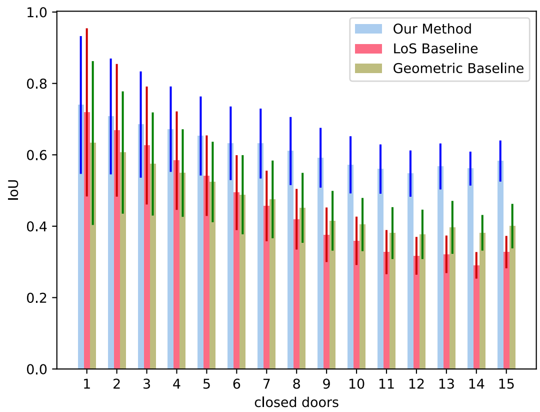


Fig. 5: Average and standard deviation of the IoU of the predicted layout of closed rooms, wrt number of closed doors in the environment.

G. Inpainting predicted layouts into the grid map

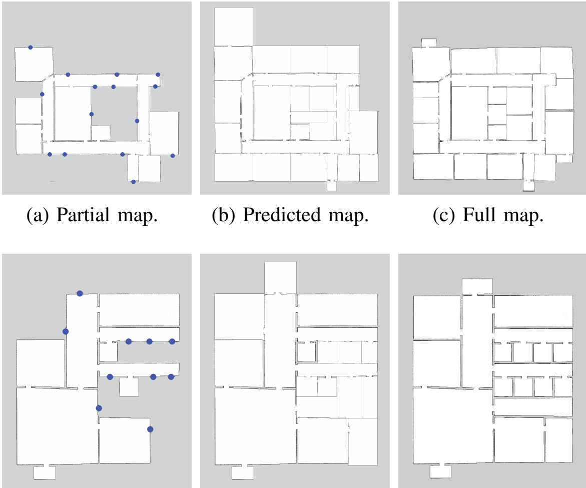
In this last step, we inpaint the predicted layouts of closed rooms into the map M. This is done by creating open passages corresponding to the positions of the doors D in the map (door width is a customizable parameter, which we set to 80 cm) and by changing the value of cells in M from unknown to free or obstacle, according to the fact that they correspond to the inner area of a predicted layout or to one of its external edges. As a result, a complete grid map $\hat{M}$ is eventually available to the robot. The predicted map for the partial map of Fig. 2a is shown in Fig. 2b. Note that the largest difference wrt the actual environment of Fig. 2c is in the rough predictions of closed rooms on the borders.

IV. Experimental Evaluation

In this section we present the experimental activities performed to evaluate the proposed method to predict the layouts of closed rooms in indoor environments. We present both quantitative results obtained in simulation and qualitative results obtained by applying our method to real-world maps from public datasets.

We start presenting results obtained in 20 simulated indoor environments (office and school environments) in which we consider up to 15 closed doors. Maps are obtained by running the ROS implementation[1] of the GMapping algorithm [33] on data collected by a robot equipped with a laser range scanner during the autonomous exploration of the buildings simulated in Stage[2]. The environments have different sizes and, accordingly, different maximum numbers of closed doors

[1]http://wiki.ros.org/gmapping
[2]http://wiki.ros.org/stage



(a) Partial map. (b) Predicted map. (c) Full map.

(d) Partial map. (e) Predicted map. (f) Full map.

Fig. 6: Two examples in which our method predicts the layouts of 15 and 10 closed rooms.

$\max |D|$ that a robot can find (see Table I). We limit the number of possible closed doors to 15 even for the larger environments in order to have a balanced evaluation of the method performance.

For each environment, we repeat 15 times the following operation: we build $N = \max |D|$ different maps by incrementally closing $1, 2, \ldots, N$ doors (if a door is closed in a map where i doors are closed, it is closed also for all maps in which $i+1, \ldots, N$ doors are closed). Closed doors are selected randomly. For each map obtained in this way, we run our method in order to predict the shape of the closed rooms. Overall, we evaluated $3,885$ maps (for a total of $24,045$ predicted room shapes). In each run, our method receives in input a grid map M and a set of closed doors D. We empirically set values of weights $[\omega_i]$ to $[0.06,\ 10,\ 7,\ 10,\ 2.5,\ 2,\ 1,\ 2]$ and the maximum number of expansion steps k of Section III-C to 9. Experiments are performed on a commercial laptop and each run requires less than 2 minutes for all maps.

Given the predicted layout of a closed room $\hat{F}_r^*$ and that of its ground truth counterpart F_r^* (obtained from the floorplan of the simulated environment), we compute their *Intersection over Union* (IoU) as:

$$\mathrm{IoU}(\hat{F}_r^*, F_r^*) = \frac{\hat{F}_r^* \cap F_r^*}{\hat{F}_r^* \cup F_r^*}.$$

An high IoU indicates that the geometric perdiction $\hat{F}_r^*$ accurately resembles F_r^* (IoU is commonly used for this type of evaluation, as in [12]).

Since, as discussed in Section II, we are not aware of any other method that predicts the layout of closed rooms, we compare our method against two baseline methods. The first one is called *line of sight* (*LoS*) *baseline* and predicts the layout of a closed room as the free area that could be observed in line of sight from the corresponding closed door d. This method is based on the assumption that all the unobserved area behind a closed door is part of the closed room. The predicted layout of the room is spatially limited by the bounding box of the map. The second method is called *geometric baseline* and adds to the predicted layout of a room all the faces $f \in \mathcal{F}_u$ that are in line-of-sight from the door d, until boundary faces are met.

Fig. 5 shows the performance of the proposed method against the two baselines. Our method obtains stable and accurate predictions of closed rooms' layouts even when a large number of doors are closed across the environments. On the other side, baseline approaches perform well when few rooms are closed (because they basically flood-fill gaps in the maps), but have a dramatic drop in performance as the number of closed doors increases.

Fig. 6 shows that our method can complete metric maps also when large parts of the buildings are not explored (15 closed doors). For instance, it provides a rather accurate prediction of all the closed rooms connected to the upper corridor in Fig. 6b. Fig. 6e shows a similar result where, despite the presence of multiple closed doors connected to the same corridor, our method provides a sound estimate of the environment map.

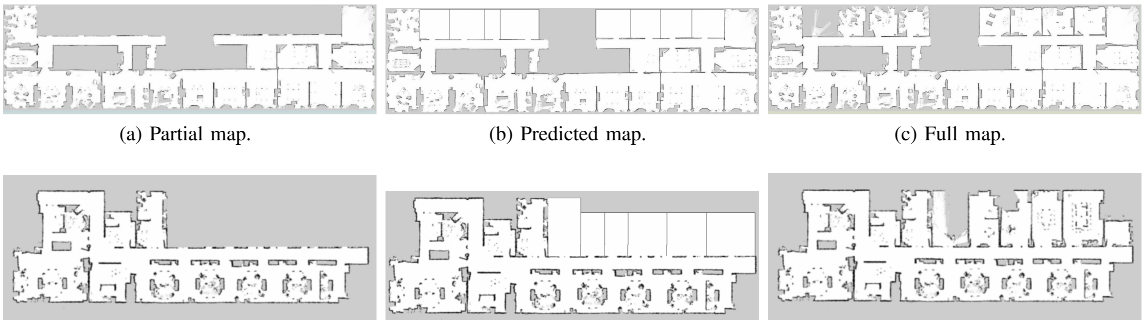
Finally, Figs. 1 and 7 show how our method can complete real-world partial maps (obtained from publicly available datasets [5], [34]) with multiple closed doors. For these results, we manually remove some rooms from the original map and we predict their possible layouts using our method. (Note that, although some maps in [5] have multiple closed doors, their locations are not provided.) Despite large missing portions of the map, our method provides a valid estimate of layouts of closed rooms even in the presence of clutter and inaccuracies. Further results on both simulated and real-world maps are available in a video[3].

V. Conclusions

In this paper, we presented a method for predicting the geometrical shape of closed rooms in indoor environments. The proposed method starts from a grid map in which the positions of closed doors that the robot could not enter are known and exploits the structural regularities of buildings to estimate the layouts of rooms behind such doors. The grid map is then completed by inpainting the layouts of closed rooms. Experiments show the effectiveness of our method, also compared against baseline methods, for large environments with up to 15 closed doors.

In future work we will lift the assumption that a closed room has only one door and we will integrate the proposed method in a deployed robot system, using a vision-based system for identifying closed doors and developing a way to update the predicted map as new knowledge is available. We are currently investigating the possible uses of the maps completed by our method for tasks as coverage, search, and exploration. Finally, we will apply the proposed method to domains, like collaborative and service robotics, where robots operate in environments in which doors could be closed to enhance the robots' understanding of their working environments.

[3]https://amigoni.faculty.polimi.it/research/ECMR2021-completing-maps-closed-rooms.html



(a) Partial map. (b) Predicted map. (c) Full map.

(d) Partial map. (e) Predicted map. (f) Full map.

Fig. 7: Application of our method to publicly available real-world maps from [5], [34].

REFERENCES

[1] L. Kunze, N. Hawes, T. Duckett, M. Hanheide, and T. Krajnik, "Artificial Intelligence for Long-Term Robot Autonomy: A Survey," *IEEE RA-L*, vol. 3, no. 4, pp. 4023–4030, 2018.

[2] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. The MIT Press, 2005.

[3] M. Luperto, M. Antonazzi, F. Amigoni, and N. A. Borghese, "Robot exploration of indoor environments using incomplete and inaccurate prior knowledge," *Robot Auton Syst*, vol. 133, p. 103622, 2020.

[4] M. Luperto, L. Fochetta, and F. Amigoni, "Exploration of indoor environments through predicting the layout of partially observed rooms," in *Proc. AAMAS*, 2021, pp. 836–843.

[5] A. Howard and N. Roy, "The robotics data set repository (radish)," 2003. [Online]. Available: http://radish.sourceforge.net/

[6] S. Prieto, A. Adan, A. Vazquez, and B. Quintana, "Passing through open/closed doors: A solution for 3D scanning robots," *Sensors*, vol. 19, no. 21, p. 4740, 2019.

[7] A. Llopart, O. Ravn, and N. A. Andersen, "Door and cabinet recognition using convolutional neural nets and real-time method fusion for handle detection and grasping," in *Proc. ICCAR*, 2017, pp. 144–149.

[8] R. Bormann, F. Jordan, W. Li, J. Hampp, and M. Hägele, "Room segmentation: Survey, implementation, and analysis," in *Proc. ICRA*, 2016, pp. 1019–1026.

[9] R. Capobianco, G. Gemignani, D. Bloisi, D. Nardi, and L. Iocchi, "Automatic extraction of structural representations of environments," in *Proc. IAS-13*, 2014, pp. 721–733.

[10] I. Armeni, O. Sener, A. Zamir, H. Jiang, I. Brilakis, M. Fischer, and S. Savarese, "3D semantic parsing of large-scale indoor spaces," in *Proc. CVPR*, 2016, pp. 1534–1543.

[11] S. Ochmann, R. Vock, R. Wessel, and R. Klein, "Automatic reconstruction of parametric building models from indoor point clouds," *Comput Graph*, vol. 54, pp. 94–103, 2016.

[12] R. Ambruş, S. Claici, and A. Wendt, "Automatic room segmentation from unstructured 3-D data of indoor environments," *IEEE RA-L*, vol. 2, no. 2, pp. 749–756, 2017.

[13] S. Oesau, F. Lafarge, and P. Alliez, "Indoor scene reconstruction using feature sensitive primitive extraction and graph-cut," *ISPRS J Photogramm*, vol. 90, pp. 68–82, 2014.

[14] C. Mura, O. Mattausch, A. Villanueva, E. Gobbetti, and R. Pajarola, "Automatic room detection and reconstruction in cluttered indoor environments with complex room layouts," *Comput Graph*, vol. 44, pp. 20–32, 2014.

[15] Z. Liu and G. von Wichert, "A generalizable knowledge framework for semantic indoor mapping based on Markov logic networks and data driven MCMC," *Future Gener Comp Sy*, vol. 36, pp. 42–56, 2014.

[16] M. Luperto and F. Amigoni, "Predicting the layout of partially observed rooms from grid maps," in *Proc. ICRA*, 2019, pp. 6898–6904.

[17] D. Perea Ström, I. Bogoslavskyi, and C. Stachniss, "Robust exploration and homing for autonomous robots," *Robot Auton Syst*, vol. 90, pp. 125 – 135, 2017.

[18] A. Smith and G. Hollinger, "Distributed inference-based multi-robot exploration," *Auton Robot*, vol. 42, no. 8, pp. 1651–1668, 2018.

[19] R. Shrestha, F. Tian, W. Feng, P. Tan, and R. Vaughan, "Learned map prediction for enhanced mobile robot exploration," in *Proc. ICRA*, 2019, pp. 1197–1204.

[20] M. Luperto and F. Amigoni, "Exploiting structural properties of buildings towards general semantic mapping systems," in *Proc. IAS-13*, 2014, pp. 375–387.

[21] J. Caley, N. Lawrance, and G. Hollinger, "Deep learning of structured environments for robot search," in *Proc. IROS*, 2016, pp. 3987–3992.

[22] S. Ramakrishnan, Z. Al-Halah, and K. Grauman, "Occupancy anticipation for efficient exploration and navigation," in *Proc. ECCV*, 2020, pp. 400–418.

[23] A. Elhafsi, B. Ivanovic, L. Janson, and M. Pavone, "Map-predictive motion planning in unknown environments," in *Proc. ICRA*, 2020, pp. 8552–8558.

[24] J. Chang, G. Lee, Y. Lu, and C. Hu, "P-SLAM: Simultaneous localization and mapping with environmental-structure prediction," *IEEE T Robot*, vol. 23, no. 2, pp. 281–293, 2007.

[25] A. Pronobis and P. Jensfelt, "Large-scale semantic mapping and reasoning with heterogeneous modalities," in *Proc. ICRA*, 2012, pp. 3515–3522.

[26] K. Zheng, A. Pronobis, and R. Rao, "Learning graph-structured sum-product networks for probabilistic semantic maps," in *Proc. AAAI*, 2018, pp. 4547–4555.

[27] A. Aydemir, P. Jensfelt, and J. Folkesson, "What can we learn from 38,000 rooms? Reasoning about unexplored space in indoor environments," in *Proc. IROS*, 2012, pp. 4675–4682.

[28] M. Luperto and F. Amigoni, "Predicting the global structure of indoor environments: A constructive machine learning approach," *Auton Robot*, vol. 43, no. 4, pp. 813–835, 2019.

[29] J. Canny, "A computational approach to edge detection," *IEEE T Pattern Anal*, no. 6, pp. 679–698, 1986.

[30] N. Kiryati, Y. Eldar, and A. M. Bruckstein, "A probabilistic Hough transform," *Pattern Recogn*, vol. 24, no. 4, pp. 303–316, 1991.

[31] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE T Pattern Anal*, vol. 24, no. 5, pp. 603–619, 2002.

[32] T. Kucner, M. Luperto, S. Lowry, M. Magnusson, and A. Lilienthal, "Robust frequency-based structure extraction," in *Proc. ICRA*, 2021.

[33] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with Rao-Blackwellized particle filters," *IEEE T Robot*, vol. 23, pp. 34–46, 2007.

[34] N. Hawes, C. Burbridge, F. Jovan, L. Kunze, B. Lacerda, L. Mudrova, J. Young, J. Wyatt, D. Hebesberger, T. Kortner *et al.*, "The strands project: Long-term autonomy in everyday environments," *IEEE RAM*, vol. 24, no. 3, pp. 146–156, 2017.