Please Tell Me Where I Am: A Fundament for a Semantic Labeling Approach

Frank Bahrmann, Sven Hellbach, and Hans-Joachim Böhme

HTW Dresden, Fakultät Informatik/Mathematik, Dresden, Germany
{bahrmann,hellbach,boehme}@htw-dresden.de
http://www.htw-dresden.de/

Abstract. This paper presents a fast but yet simple solution to create areas on metrical occupancy grid maps which can easily be converted to topological maps. Those maps with their areas provide an understandable "non-expert" view on a robots environment and allow semantic labeling. This serves as a foundation for supporting navigation tasks like path planning, localization and human-machine-interaction which are not in the scope of this paper.

Keywords: Occypancy Grid Map, Metric Topological Map, Semantic Place Labeling, Human-Machine Interaction

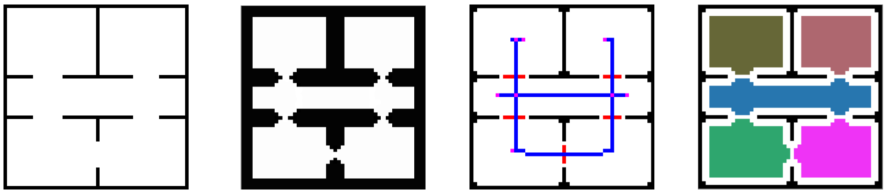
1 Introduction

The proposed method is part of a scenario where a mobile service robot platform is introduced to a new office or home environment. In this state the robot needs to learn how to interact with its new surroundings. There exists a lot of other research addressing that problem by detecting and classifying significant features through object detecting or similar [1,5,7,4], regardless of the ambiguity of the labels with respect to the features and vice versa. However, in our opinion the best way to accomplish this familiarization is by interacting with the robot in a very human-like manner. This idea is inspired by the human habit to show around new co-workers. To handle this scenario it is necessary for the robot to detect and follow humans, to interact with them through a dialogue and to recognize rooms by splitting a built map in semantic parts - a topological map [6]. The last will be the main subject in this paper. The proposed method could become beneficial to localization tasks, reactive motion control and "semantic" path planning.

2 Approach

Our method is divided into three subsequent steps (see figure 1b-d). To explain the single stages, a rather simple simulated occupancy grid map is used as raw material (see Fig. 1a), more complex maps are shown in the results section. The first step is filtering and dilating the original map (Fig. 1b), followed by thinning (blue line in Fig. 1c) and the last step is separating areas (Fig. 1c and d).

S. Wölfl (Ed.): Poster and Demo Track of the 35th German Conference on
Artificial Intelligence (KI-2012), pp. 120-124, 2012.



(a) (b) (c) (d)

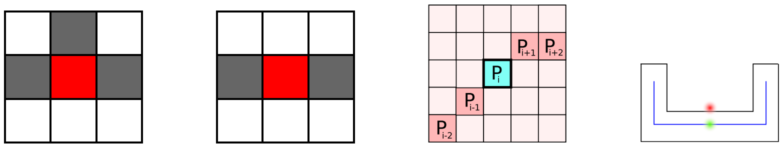
Fig. 1: (a) The raw map (where free space is represented by white pixels and occupied space by black pixels) and subsequent processing steps: (b) The dilated and gap closed binary map. (c) The map skeleton (blue) with starting and intersection points (purple) and critical lines (red). (d) The constructed areas.

Filtering and Dilating for gap closing and distortion removal: First of all the occupancy grid map is converted in a binary grid map in which the cell state can either be free or occupied by applying a threshold slightly below uncertainty. The next step is meant to close smaller gaps which can occur during the mapping process. A common way to achieve this is by dilating all occupied cells and eroding them back to their original state. To filter out smaller distortions like chair or table legs, a 8-directional flood fill algorithm is used on all occupied cells to find those which are not part of a larger structure like walls. If a filled area is smaller than a specified threshold, then this area will be deleted. Finally, to achieve a higher stability later in the processing chain, all occupied cells are dilated with the robot's radius (see Fig. 1b). After this kind of dilatation the robot can be assumed as having the size of a single grid cell for robot navigation.

Thinning to build a skeleton map: The established Zhang & Suen thinning algorithm is used like proposed in [8] to create map skeleton, which is one pixel in width. Fig. 1c shows the thinned example map with blue lines.

Classifying intersection- and start-cells to ensure a graph like structure: Ko et al. [2] described a way to use a map skeleton to build topological maps. Unfortunately, their algorithm defines only starting cells leaving out connecting edges, which is insufficient in our case. For the reconstruction of a topological map it is additionally necessary to determine the position of intersections of the skeleton edges. The fastest and most reliable method was to use predefined 3x3 pixel templates. Two cases are shown in Fig. 2, where the middle red pixel represents the tested one.

Detecting door hypotheses by finding Critical Lines on the skeleton: All detected starting cells are now used to store a predecessor and a successor per skeleton cell by following along the neighboring cells until each cell on the



(a) (b) (c) (d)

Fig. 2: Example of predefined templates (a) with and (b) without intersection. (c) Ordered neighborhood notation for grid cell P_i. (d) Area center - red dot marks the center of gravity, green the logically correct area center on the skeleton (blue line).

skeleton is visited. The result are connected grid-lines, which can be used to interpolate a map graph later.

It is now possible to calculate a normal on each cell P_i. That can be done by a local gradient approximation by averaging the point pairs (P_{i+1}, P_{i+2}) and (P_{i-1}, P_{i-2}) (see Fig. 2c). The averaging of these point pairs was chosen to achieve a higher local stability. The resulting normal is identified as a critical line (expected to be doors which are shown as red lines in Fig. 1c) by the total side clearance (average length of all normals) and scaling their length to serve as an adaptive threshold or by using a constant threshold.

Approximating room hypotheses: Determining the proportions of areas (room hypotheses) is done by using a 4-way flood-fill algorithm to fill free cells with IDs, in which the IDs serve as an identifier of the area. The area construction is completed when there is no remaining free cell anymore. With this method it is possible to cover areas regardless of their geometrical shape. To determine the center coordinate of an area, we calculate the center of gravity and look for the area related skeleton cell with the smallest euclidean distance. Hence we ensure to get a coordinate which lies on the skeleton within free space and additionally not outside of the area (shown in Fig. 2d).

Integration in path planning: The majority of path planning in mobile robotics is done by using a grid-cell map as a graph in which each cell is represented as a graph node. Where adjacent (orthogonal & diagonal) grid cells are connected with edges. To apply existing grid map based implementations of path planners, we simply transferred all of the resulting skeleton cells in a grid map.

Human guided labeling As indicated in the introduction, the map is labeled as follows: The mobile robot system will be delivered into a new environment where a human supervisor will show it around. Our system is already able to navigate collision free, detect and follow people [3] and robustly map indoor environments (SLAM). Even though the used speech dialog system is in an

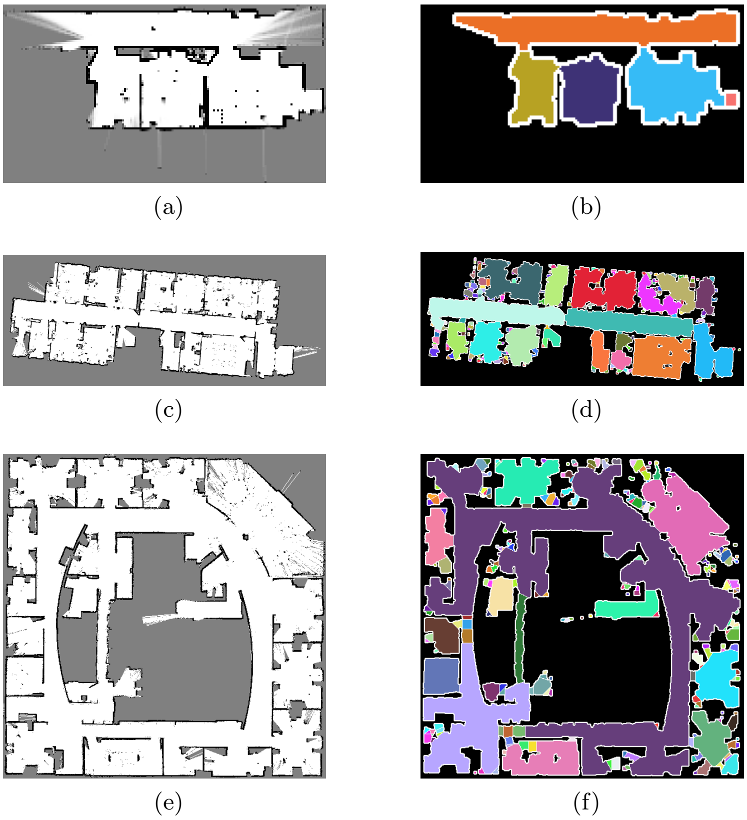


Fig. 3: Achieved results: (a)/(b) our research lab ($163x90\,px \,\widehat{=}\, 146.7\,m^2$ in $0.702\,s$), (c)/(d) the Intel research lab data-set ($579x581\,px \,\widehat{=}\, 3363.99\,m^2$ in $4.566\,s$), (e)/(f) the fr079 data-set ($911x368\,px \,\widehat{=}\, 3352.48\,m^2$ in $4.831\,s$)

early stage, a small context sensitive grammar has been built to achieve the proposed behavior. With all these abilities, the robot is able to follow its guide, who will verbally label the robot's current position as a coordinate within the grid map. After the tour is completed, the SLAM algorithm produced a grid map, which serves as starting point (raw map) for the proposed method. The verbally defined labels are then propagated throughout the constructed areas.

3 Results

This section presents our preliminary results with some qualitative evaluations (shown in Fig 3). Those first results indicate that the method is quite robust. Cluttered maps (see Fig. 3e) seem to demand further research. Furthermore, real-world experiments were successfully performed on an open door day at our university. A K-Team Koala robot was placed in a model apartment similar to the raw map shown in Fig. 1a. The visitors task was to verbally show the robot around by using a speech recognition framework to define the robot's goal and label the rooms. For such an event it is expected to provide a stable and reliable system.

4 Conclusion & Future Work

This paper covered first steps towards an approach for semantic map labeling in cooperation with a human teacher. A main disadvantage, which has to be addressed, is detecting false positive critical lines. Further steps are to be taken to drastically reduce this problem by applying a pattern recognition system on the detected critical lines and its surroundings. This system does not need to be scale or rotation invariant because of the prior knowledge of the critical lines width and orientation.

References

1. Fabrizi, E., Saffiotti, A.: Extracting topology-based maps from gridmaps. In: Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on. vol. 3, pp. 2972–2978. IEEE (2000)
2. Ko, B., Song, J., Lee, S.: Real-time building of a thinning-based topological map with metric features. In: Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on. vol. 2, pp. 1524–1529. IEEE (2004)
3. Poschmann, P., Hellbach, S., Böhme, H.J.: Multi-modal people tracking for an awareness behavior of an interactive tour-guide robot. In: Proceedings of the International Conference on Intelligent Robotics and Applications (ICIRA) (2012), in press
4. Sousa, P., Araiijo, R., Nunes, U.: Real-time labeling of places using support vector machines. In: Industrial Electronics, 2007. ISIE 2007. IEEE International Symposium on. pp. 2022–2027. IEEE (2007)
5. Tarutoko, Y., Kobayashi, K., Watanabe, K.: Topological map generation based on delaunay triangulation for mobile robot. In: SICE-ICASE, 2006. International Joint Conference. pp. 492–496. IEEE (2006)
6. Thrun, S.: Learning metric-topological maps for indoor mobile robot navigation. Artificial Intelligence 99(1), 21–71 (1998)
7. Topp, E., Christensen, H.: Topological modelling for human augmented mapping. In: Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on. pp. 2257–2263. Ieee (2006)
8. Zhang, T., Suen, C.: A fast parallel algorithm for thinning digital patterns. Communications of the ACM 27(3), 236–239 (1984)