Robotics and Autonomous Systems 62 (2014) 663–674

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Extracting semantic indoor maps from occupancy grids

Ziyuan Liu [a,b,*], Georg von Wichert [c,b]

[a] Institute of Automatic Control Engineering, Technische Universität München, Munich, Germany

[b] Institute for Advanced Study, Technische Universität München, Lichtenbergstrasse 2a, D-85748 Garching, Germany

[c] Siemens AG, Corporate Research & Technologies, Munich, Germany

ARTICLE INFO

Article history:
Available online 4 January 2013

Keywords:
Semantic models
Cognitive robotics
Indoor mapping
Monte Carlo methods

ABSTRACT

The primary challenge for any autonomous system operating in realistic, rather unconstrained scenarios is to manage the complexity and uncertainty of the real world. While it is unclear how exactly humans and other higher animals master these problems, it seems evident, that abstraction plays an important role. The use of abstract concepts allows us to define the system behavior on higher levels. In this paper we focus on the semantic mapping of indoor environments. We propose a method to extract an abstracted floor plan from typical grid maps using Bayesian reasoning. The result of this procedure is a probabilistic generative model of the environment defined over abstract concepts. It is well suited for higher-level reasoning and communication purposes. We demonstrate the effectiveness of the approach using real-world data.

1. Introduction

The primary challenge for any autonomous system operating in realistic, rather unconstrained scenarios is to manage the complexity and uncertainty of the real world. In robotics this holds, as soon as the robots leave the carefully engineered production environments in which they have been so successful in the past decades.

The typically high degree of uncertainty in real-world environments, that makes a robot's life so hard, comes from the following sources: the limited measurement accuracy and other limitations of the system's sensors, modeling errors and purposefully made simplifications in the system's internal representations, unobserved environment dynamics and random effects in action execution. While it is unclear how exactly humans and other higher animals master these problems, it seems evident, that abstraction plays an important role. The use of abstract concepts allows us to define the system behavior on higher levels and independently of the exact setting of the environment and the exact sensor readings.

In this study we address the first two of the problems mentioned above, in that we provide the system with a limited capability of abstraction allowing for a higher-level understanding of its environment. In addition, we directly address the uncertainty related issues by strictly following a probabilistic approach that explicitly models and keeps track of the uncertainty associated with any variables of the problem.

As a by-product, the system's capability to use predefined concepts will ease cooperation in mixed human–robot tasks, since a common language used by both the human and the robot is a precondition for efficient exchange of information between both parties. This is however not addressed in this paper.

To illustrate the general idea, we use an example from an indoor navigation scenario, namely the semantic analysis of the commonly used occupancy grid maps. The objective of the presented method is to provide an abstracted, semantically annotated but still probabilistic map of the indoor environment. For this purpose, we first use a robot – equipped with a 2D laser scanner – to build an occupancy grid of the environment using a standard SLAM method [1] and then employ the procedure described in the remainder of this document to extract the semantic information. To do this, we use a Markov chain Monte Carlo (MCMC) based sampling technique [2] to generate samples from the probability density function capturing the distribution of probable worlds the robot could encounter. The maximum posterior solution could then be used as an estimate of what the world semantically looks like.

2. Problem formulation

Most of todays' mapping approaches aim to construct a globally consistent, metric map of the robot's operating environments. See Fig. 1 for a typical result. Such maps enable the robot to localize itself with respect to the environment and thus determine

* Correspondence to: Karlstr. 45, room 3015, 80333, Munich, Germany. Tel.: +49 89 289 26900; fax: +49 89 289 26913.

E-mail addresses: ziyuan.liu@tum.de (Z. Liu), georg.wichert@siemens.com (G. von Wichert).

0921-8890/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.robot.2012.10.004

Fig. 1. A typical occupancy grid map of an indoor environment, obtained from the Robotics Data Set Repository (Radish) [3].

its global pose in an assumed flat world with an accuracy of typically a few centimeters in translation and below one degree in rotation. Based on this capability, the robot can also plan a path and navigate towards a goal, that will also be specified by its metric position in the global map reference frame. However, the robot does not understand its environment in terms of typical semantic concepts like rooms, corridors or functionally enriched concepts like a kitchen or living room. Furthermore, the robot does not understand relations like adjacency, connectivity via doors, or properties like rectangularity that – if known to be relevant to the given environment – could help to build the maps in the first place.

Our work aims at extracting such semantic models of the environment from the more or less raw sensor data. In the context of this paper, we assume, that a map, like the one depicted in Fig. 1, was already constructed using one of the proven methods available for this purpose [1].

Assigning semantics to spatial maps in robotics has not been looked at as intensely as the metric or topological mapping. Still, several important contributions to the field have already been made. They can be clustered into two major groups. The first group consists of methods based on place labeling, some notable examples are [4–11]. These methods assign semantic labels to places or regions of the accessible work space of the robot. They are very much in the tradition of [12] or [13].

A second group is formed by approaches assigning semantic labels to parts or objects of the perceived structure of the environment, like traversable terrain, trees or similar structures in outdoor environments or walls, ceilings, and doors in indoor settings [14–22].

In addition to the two groups mentioned above, there are also other approaches. The approach of [23] semantically models places via objects. In [24], a method is proposed, which explores the environment in a room-by-room style and fits the explored map part into polygons. Tapus and Siegwart [25] build a map of the environment based on so called fingerprints of explored places. Lim et al. [26] introduce an ontology-based method that integrates low-level data with high-level constraints to represent the knowledge as a semantic network.

Different from those methods mentioned above, we aim to construct a probabilistic generative model of the world around the robot, that is essentially based on abstract semantic concepts but at the same time allows us to predict the continuous percepts that the robot obtains via its noisy sensors. This abstract model has a form similar to a scene graph, a structure which is widely used in computer graphics. The scene graph (see Fig. 2(c)) in our case consists of rooms and doorways connecting the rooms and can be visualized as a classical floor plan (see Fig. 2(b)).

The scene graph and thus also the semantically annotated world state is denoted by a vector of hidden parameters W specifying the world state, that generated the occupancy map M we are currently looking at. In the Bayesian framework we can use a maximum

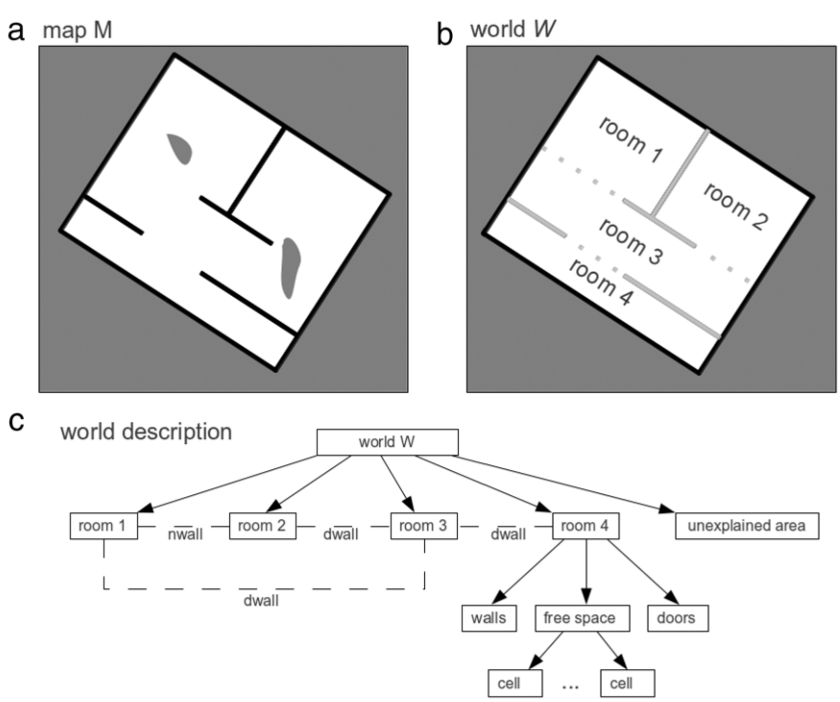

Fig. 2. (a) A simplified occupancy grid map: unexplained area is drawn in gray, free space is drawn in white. Occupied area is drawn in black. (b) A possible floor plan represented as a scene graph (W): the world is divided into four rooms and the corresponding unexplained area. The connectivity is given by the wall types (dwall: a wall that has one or more doors on it; nwall: a wall that separates two rooms but does not contain a door on it; bwall: a wall that just serves as boundary). A partially dotted line in light gray indicates a dwall, where the dotted part is the door, and the solid line part is the rest of the wall. A light-gray line (without dots) shows one nwall, and black stands for a bwall. (c) The semantic description of the world in the form of the scene graph: directed links connect nodes. The dashed lines represent connectivity. Like room 4, each room has three child nodes: walls, free space, and doors. Note that the lowest level of node in the tree structure is the grid cell that belongs to walls, free space and doors.