Digital Communications and Networks (2015) 1, 253–266

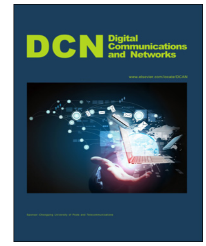
HOSTED BY

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

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



Building a grid-semantic map for the navigation of service robots through human-robot interaction

Cheng Zhao[a,*], Weixing Mei[b], Wei Pan[b]

[a]School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, United Kingdom
[b]Cognitive Science Department, Fujian Key Laboratory of the Brain-like Intelligent Systems, Xiamen University, Xiamen 361005, The People's Republic of China

Received 30 July 2015; accepted 6 September 2015
Available online 30 September 2015

KEYWORDS
Visual-voice interface;
Grid-semantic map;
Global localization;
ROS

Abstract
This paper presents an interactive approach to the construction of a grid-semantic map for the navigation of service robots in an indoor environment. It is based on the Robot Operating System (ROS) framework and contains four modules, namely Interactive Module, Control Module, Navigation Module and Mapping Module. Three challenging issues have been focused during its development: (i) how human voice and robot visual information could be effectively deployed in the mapping and navigation process; (ii) how semantic names could combine with coordinate data in an online Grid-Semantic map; and (iii) how a localization-evaluate-relocalization method could be used in global localization based on modified maximum particle weight of the particle swarm. A number of experiments are carried out in both simulated and real environments such as corridors and offices to verify its feasibility and performance.

1. Introduction

The navigation of service robots traditionally relies on geometrical maps that are either priory constructed or dynamically built from sensor data. This imposes a big challenge for general public to use service robots in their daily life since they have to be trained beforehand. On the other hand, humans use the semantic map to navigate around, which is intuitive and easy to learn. It would be beneficial if a sematic map could be deployed in the robot navigation system so that users can operate service robots using semantic information such as voice and gesture. Therefore, interactive mapping and navigation for service robots is currently an active research area in the robotics community. In order to add semantic information to the map, there are two mainly branches: (i) generate the

*Corresponding author.
E-mail address: IRobotCheng@gmail.com (C. Zhao).
Peer review under responsibility of Chongqing University of Posts and Telecommunications.

http://dx.doi.org/10.1016/j.dcan.2015.09.002
2352-8648/

semantic map automatically by extracting or classifying the data from sensors, and (ii) generate the semantic map via human–robot interaction using visual and voice signals.

Buschka and Saffiotti proposed an approach to segment and detect the room spaces for navigation using range data in an office environment [1]. Using the anchoring technique, Galindo et al. labelled the topological map with semantic information for navigation [2]. The work in [3] and [4] introduced an approach to learn topological maps from geometric maps by applying a semantic classification procedure based on associative Markov networks and AdaBoost. In addition, many works use visual features from camera sensors to extract semantic information via place categorization [5].

Since semantic information automatically extracted from sensors is limited and not robust enough, more and more researchers turn to the second branch: semantic mapping based on human-robot interaction, which extracts richer semantic information. A wearable gesture interface in an ear-mounted FreeDigiter device was proposed in [6] to label doors in the topological map. Several IMUs are used in [7] to detect movements of a person and door opening and closing events labelled as topological nodes in a graph-based SLAM framework. A contextual topological map was built by making the robot follow the user and verbal commentary [8]. A probabilistic model was proposed in [9] for recognition and classification of spaces into separate semantic regions and can use this information for generation of the topological map.

Kruijff et al. built the semantic map through natural language dialogues between human and robot [10]. The system in [11] integrates laser and vision sensors for place and object recognition as well as a linguistic framework, which creates a conceptual representation of the human-made indoor environment. Randelli et al. summarized many multi-modal interactions such as speech, gestures and vision for semantic labelling, which assists the robot in obtaining rich environment knowledge without many pre-requisites features [12]. Pronobis and Jensfelt presented a probabilistic framework combining heterogeneous, uncertain information such as object observations, shape, size, appearance of rooms and human input for semantic mapping [13]. It abstracts multi-modal sensory information and integrates it with conceptual common-sense knowledge, which makes the semantic map more descriptive, and the system is more scalable and better adapted for human interaction. The work in [14,15] proposed similar approaches for learning environmental knowledge about the grounding of expressions from task-based human-robot dialogs. In a survey [16], many semantic mappings for mobile robotics tasks are summarised in detail.

Fig. 1 The platform snapshot in this paper.

This paper presents the construction of a Grid-Semantic map for the navigation of service robots, which is based on human-robot interaction. A novel localization–evaluate–relocalization for global localization is applied to navigation. The rest of the paper is organized as follows. Section 2 introduces the platform, the software and the basic system configuration. Section 3 outlines the methodologies deployed in this research, including interaction, control, mapping and navigation. In Section 4, a number of experimental results are presented to verify the feasibility and performance of the proposed approach. Finally, a brief conclusion and future work are given in Section 5.

2. System overview

2.1. Platform and software configuration

The platform in this paper includes a Pioneer 3-AT robot (P3AT in short), a Sick LMS100 laser, a Kinect sensor, a notebook and a cell phone, as shown in Fig. 1. The whole system is based on the ROS [17] framework. The related libraries include: Fuerte ROS (Gmapping and navigation Package), PrimerSensor (5.1.2.1), NITE (1.5.2.21), OpenNI (1.5.4.0), JDK (1.6.0_20), Android SDK (2.3.3), IFLYTEC SDK (1013), Ekho (5.6), Aria (2.7.5.2).

2.2. Workflow

In our daily life, when guests come to our house for the first time, we show them around the house so that they could find a way around during their stay. Inspired by this, we propose an interactive navigation system for service robots, as shown in Fig. 2. It contains four modules and operates sequentially as follows:

1. The robot moves in an unknown indoor environment, following the user's gestures gathered by the Kinect-based skeleton tracking.
2. During the following process, a real time grid map is generated based on the RBPF algorithm using laser and odometer data.
3. At the same time, the user can use the vocal APP based on IFLYTEK on a cell phone to label places on the grid map.
4. Then the robot can combine the semantic names from voice recognition with the coordinate values from the Grid map together to build a Grid-Semantic map.
5. The robot obtains its current location through a novel localization–evaluate–relocalization method and acquires the destination from the vocal APP sent by user.
6. The robot transforms a Grid map to a Cost-Map, then makes a path planning using Dijkstra's algorithm, and finally reaches the destination.

All the experiment videos can be watched on Youtube [18].

3. Methodology

3.1. Interactive Module

3.1.1. Skeleton Tracking

In terms of visual interaction, the Kinect based Skeleton Tracking [19] is adopted because of the following four reasons.