Semantic Mapping and Reasoning Approach for Mobile Robotics

Caner GUNEY, Serdar Bora SAYIN, Murat KENDİR, Turkey

Key words: Semantic mapping, 3D mapping, probabilistic, robotic surveying, mine surveying

SUMMARY

The mobile robots need to have semantic information in their map representation about the entities in the environment in order to reason about their surroundings. Hence the mobile robots can act intelligently in the environment and solve autonomously a variety of robotic tasks. In this study, semantic mapping framework is established to give mobile robots the ability to perform high-level robotic tasks based on the semantic information. GeoRoSS is an autonomous mobile robot equipped with a reliable and precise 3D laser scanner that digitalizes environments. High quality geometric 3D maps with semantic information are automatically generated after the exploration by the robot.

Semantic Mapping and Reasoning Approach for Mobile Robotics

Caner GUNEY, Serdar Bora SAYIN, Murat KENDİR, Turkey

1. MOTIVATION

The mobile robots need to have semantic information in their map representation about the entities in the environment in order to reason about their surroundings. Hence the mobile robots can act intelligently in the environment and solve autonomously a variety of robotic tasks.

GeoRoSS is an autonomous mobile robot equipped with a reliable and precise 3D laser scanner that digitalizes environments. High quality geometric 3D maps with semantic information are automatically generated after the exploration by the robot.

In this study, semantic mapping framework is established to incorporate geometric representation, navigational graph, topological representation, place recognition (scene interpretation) and object classification through Multiple Representation Database (MRDB) and finally linked up all those representations with knowledge representation so that 3D semantic map will be built and semantic reasoning will be realized. The figure 1 displays the proposed semantic mapping framework approach. In this study, the mobile robot acquires semantic information from its sensors and uses this semantic information for mapping, navigation and mining tasks.

The ultimate goal in the proposed framework to enable the mobile robots to deal with high-level functionality, such as performing complex types of reasoning, inferring from objects in the vicinity, improving human-robot interaction, carrying out complex navigational tasks, path planning, change detection, etc. For instance, 3D semantic maps can be used by an inference mechanism to devise contingency plans that allow the robot to recover from exceptional situations.

2. THE SEMANTIC MAPPING FRAMEWORK

2.1 Spatial Component

3D models of the environment provide a volumetric representation of space which is important for a variety of robotic applications [Hornung et al., 2013]. 3D geometric representation of the environment allows the mobile robot to store spatial information of features from the robot environment so that the robot localizes itself in the environment and avoids collision.

Robotic metric maps or occupancy grid maps of the surroundings of the robot can be built from several sensor data. Most of the approaches use depth information, obtained by laser range finders, time of flight, structured light sensors or stereo data, or camera information obtained by one or more cameras or RGB-D sensors such as the Kinect. The probabilistic representation of the environment of the mobile robot is obtained from Simultaneous Localization And Mapping (SLAM) method based on such sensor data.

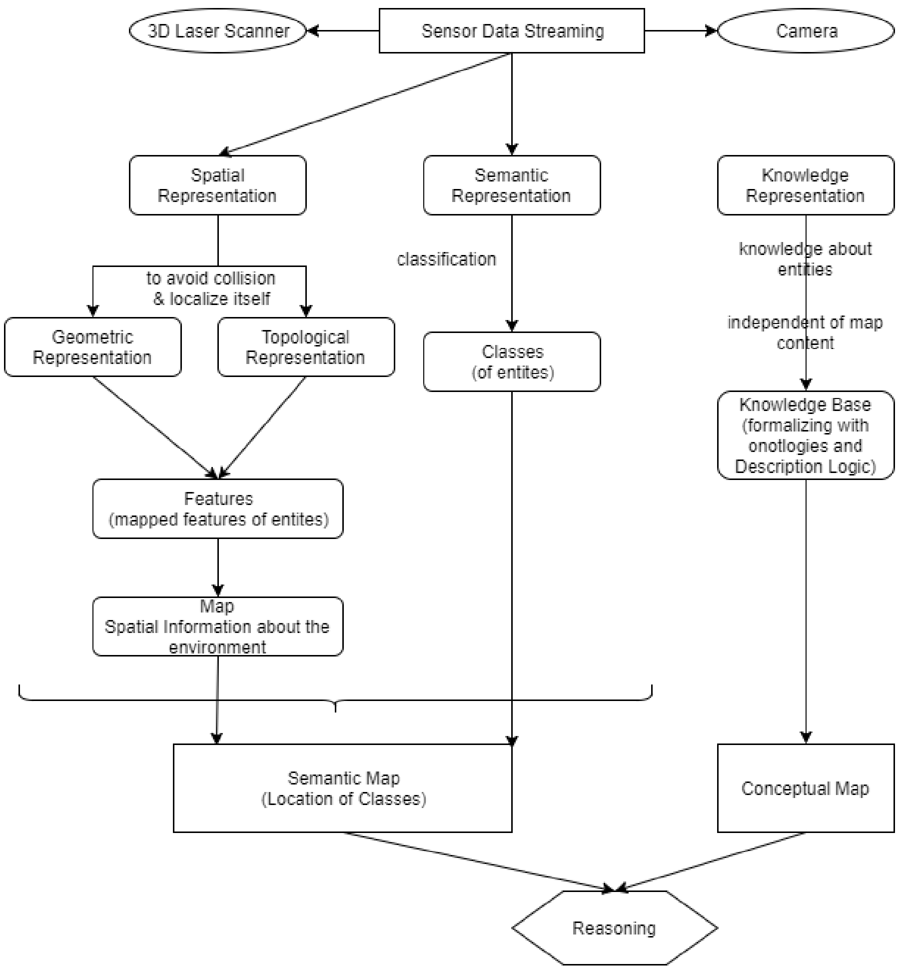

Figure 1: The main idea of the proposed semantic mapping framework

The prime sensor of the GeoRoSS is a 3D Terrestrial Laser Scanner (TLS) (Z+F Imager 5010). The semantic mapping process starts with acquiring the point clouds. As a wheeled robot GeoRoSS, with its 3D scanner, gathers successive point clouds.

Scan matching and mapping algorithms are applied for creating a large-scale globally consistent 3D map. In this step, individual scans (in a stop-and-go fashion) are registered to convert all clouds into one reference coordinate system. Simultaneously, corresponding point clouds are registered by a SLAM frontend based on the Iterative Closest Point (ICP) algorithm and a graph is constructed. This graph can be optimized continuously by a SLAM backend especially after loop closing. Thus, the environment representation is achieved with the initial 3D geometric map of the environment.

The metric maps created from laser range data are represented as a mixture of metrical and topological representations. The metric map is a line based map and represents the part of the space that can be described by lines. The navigation map connects nodes and edges representing the trajectory of the robot. Nodes represent open areas and arcs the possibility of navigation from one to another. The topological map divides a set of nodes in the navigation map into areas. [Zender et al., 2008]

Depending on the sensors, the point clouds contain texture information. In this study, semantic information is automatically extracted from 3D models built from a laser scanner since 3D TLS generates range data and intensity in a fused form for environment mapping. In addition to that, GeoRoSS has a thermal camera (T-Cam), which means point clouds contains texture information.

A lot of methods have been established for object classification or place recognition in camera images, point clouds or fused data for indoor and outdoor scenes. Most of the approaches use supervised learning, where the object classes are modeled. [Lang et al., 2014] Mainly, the results are represented as point clouds or images colored according to the classification result, in other words, objects are recognized by shape and color, and does not take into account the properties of objects. Therefore, the classification results are often not used for tasks based on the map. In order to achieve this, the relationship between the perceived environment in the sensor data and the common-sense knowledge has to be established. [Lang et al., 2015] [Lang et al., 2014]

2.2 Semantic Component

2.2.1 Contextual Classification of Sensor Data

Having the sensor data stream from 3D laser scanner, contextual (semantic) classification of 3D point clouds is performed by applying different classification methods, ranging from a simple linear model to a more complex one based on Probabilistic Graphical Models (PGM). PGM captures the contextual relationship among 3D points by training point cloud statistics [Lang et al., 2015]. Entity classification can be divided into different classes on real-world 3D point cloud data, such as the classification of spaces and/or the classification of surface elements.

In this step, scene features are determined by semantic labeling. The semantic labeling of point clouds is solved by classifying 3D point clouds. The object recognition is implemented with a trained classifier or previously learned objects.

2.2.2 Data Structure for Map Representation

There are different approaches to represent the environment based on 3D point clouds. Popular point cloud representations include point clouds, voxel grids, octrees, surfels, and cluxels.

For semantic mapping, a probabilistic representation is required to model free, occupied, and unmapped areas. Additionally, a data structure is needed to support such a probabilistic representation and the data structure should be memory-efficient, allow multi-resolutions and the integration of semantic labels.

In this study, the registered and classified point clouds can be converted into a multi-label and resolution octree map data structure based on OctoMap mapping framework. The goal in this step is not only using a data structure to hold the space-related information about this environment, but also subdividing the point cloud into octree voxels and classifying them contextually by means of semantic labels.

An octree is representing a hierarchical data structure dividing the 3D space into spatial subdivisions. The OctoMap mapping framework is based on octrees and creates a voxelized 3D map for registered 3D point clouds. Each node of the octree is a cubic volume named voxel. The whole volume is recursively subdivided into eight partial voxels with the same size until a minimum size for each voxel is reached. This minimal size defines the resolution of the octree. [Hornung et al., 2013] For each semantic label and object, a new OctoMap is created, such that multi-resolution object maps can be created and objects can be represented in a finer resolution than less interesting ones. The different OctoMaps are connected by a tree. [Wurm et al., 2011] However, the approach cannot handle 3D points with different labels in one voxel [Lang et al., 2013].

The multi-label problem in one voxel is solved using the classification methods, such as Markov network, Conditional Random Field (CRF) based on OctoMap data structure.

3. CONCEPTUAL COMPONENT

The main purpose in this study is to perform semantic reasoning based on individual entities in the map and/or their classes. To enable this reasoning, some background knowledge about entities is required rather than a semantic map including only features labeled with tags. The knowledge may come in any suitable knowledge representation format, as needed for the type or types of reasoning to be associated with the entities in the map. Given that such knowledge is typically independent of space, it is not strictly part of the map; however, the goal of the study requires that it exists for entities represented in the semantic map.

The conceptual component represents concepts, relations between those concepts and instances of spatial entities. The knowledge representation allows the mobile robot to model semantic knowledge about the robot environment and to use inference capabilities on reasoning about the functionalities of objects and environments. [Galindo et al., 2005]

Attributes are the concepts represented in the conceptual map, which can be derived by previously human-generated models (most often hand-coded into the system) or human-computer interaction (linguistic interaction with a human) and/or or can be inferred during the mapping process (using the robot's own sensors and real-valued measures) based on acquired classification results and knowledge. Conceptual ontology defines abstract categories for places and objects and how they are related. [Lang and Paulus 2014]

In this study, a well-known system for knowledge representation and reasoning, called Protégé, is used in order to provide the robot with inference capabilities. The conceptual knowledge is encoded as an OWL-DL ontology and a Description-Logic (DL) reasoner is used to classify spatial areas and objects. The voxels are combined by reasoning over a predefined knowledge (ontologies) into objects.

The conceptual knowledge connects sets of classified entities of the metric, navigation and topological map to attributes which allows for connecting the conceptual map and the other maps.

4. CONCLUSION AND OUTLOOK

In this study, the goal was to design an approach for real-time 3D semantic mapping system for mobile robotics in order to give mobile robots the ability to perform high-level robotic tasks based on the semantic information.

The use of knowledge representation techniques allows the mobile robot to represent and reason about situations of ambiguity, which may arise when the robot's information about the environment is incomplete. Semantic information can be exploited to autonomously devise strategies to resolve these ambiguities, for instance, by using an AI planner. In the future, to cope with this type of situations, GeoRoSS will be equipped with a state of the art AI planner, to reason about uncertainty.

REFERENCES

Armin Hornung Kai M. Wurm Maren Bennewitz , Cyrill Stachniss Wolfram Burgard, 2013, OctoMap: An Efficient Probabilistic 3D Mapping Framework Based on Octrees, Autonomous Robots (2013)

C. Galindo, A. Saffiotti, S. Coradeschi, P. Buschka, J.A. Fernandez-Madrigal, J. Gonzalez, (2005), "Multi-Hierarchical Semantic Maps for Mobile Robotics" Proceedings of the IEEE/RSJ Intl. Conf. Intelligent Robots and Systems, IROS '05, Edmonton, Canada, 2005, pp. 3492-3497.

Dagmar Lang, Susanne Friedmann, Dietrich Paulus, 2013, Semantic 3D Octree Maps based on Conditional Random Fields

Dagmar Lang, Susanne Friedmann, Marcel Haselich and Dietrich Paulus, 2014, Definition of Semantic Maps for Outdoor Robotic Tasks

Dagmar Lang and Dietrich Paulus, IROS 2014 Workshop on AI Robotics, Chicago Illinois, Semantic Maps for Robotics

Dagmar Lang, Susanne Friedmann, Jens Hedrich and Dietrich Paulus, 2015, Semantic Mapping for Mobile Outdoor Robots

Andreas Nüchter, Joachim Hertzberg, Robotics and Autonomous Systems 56 (2008) 915-926

K.M.Wurm, D. Hennes, D. Holz, R. B. Rusu, C. Stachniss, K. Konolige and W. Burgard, Hierarchies of Octrees for Efficient 3D Mapping, In Proc. of IROS, pages 4249–4255, 2011

CONTACTS

Caner GUNEY
Institution: Istanbul Technical University
City Istanbul
COUNTRY Turkey
Tel. +90 212 285 3825
Fax +90 212 285 6587
Email: bora.sayin@gmail.com
Web site: http://www.geo.itu.edu.tr/geoross/

Serdar Bora SAYIN
Institution: Istanbul Technical University
City Istanbul
COUNTRY Turkey
Tel. +90 212 285 3825
Fax +90 212 285 6587
Email: guneycan@itu.edu.tr
Web site: http://www.geo.itu.edu.tr/geoross/

Murat KENDİR
Institution: Istanbul Technical University
City Istanbul
COUNTRY Turkey
Tel. +90 212 285 3825
Fax +90 212 285 6587
Email: muratkendir@gmail.com