Robotics and Autonomous Systems 66 (2015) 86–103

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

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

Semantic mapping for mobile robotics tasks: A survey

Ioannis Kostavelis *, Antonios Gasteratos

Laboratory of Robotics and Automation, Production and Management Engineering Department, Democritus University of Thrace, Vas. Sophias 12, GR-671 00 Xanthi, Greece[1]

HIGHLIGHTS

- Two level navigation.
- Cognitive navigation.
- Spatial semantics.

ARTICLE INFO

Article history:
Received 6 February 2014
Received in revised form
7 December 2014
Accepted 12 December 2014
Available online 27 December 2014

Keywords:
Mobile robots
Semantic map
Topological map
Temporal coherence
Object recognition
Place recognition
Human–robot interaction
Knowledge representation
Planning

ABSTRACT

The evolution of contemporary mobile robotics has given thrust to a series of additional conjunct technologies. Of such is the semantic mapping, which provides an abstraction of space and a means for human–robot communication. The recent introduction and evolution of semantic mapping motivated this survey, in which an explicit analysis of the existing methods is sought. The several algorithms are categorized according to their primary characteristics, namely scalability, inference model, temporal coherence and topological map usage. The applications involving semantic maps are also outlined in the work at hand, emphasizing on human interaction, knowledge representation and planning. The existence of publicly available validation datasets and benchmarking, suitable for the evaluation of semantic mapping techniques is also discussed in detail. Last, an attempt to address open issues and questions is also made.

A local villager knows his way by wont and without reflection to the village church, to the town hall, to the shops and back home again from the personal point of view of one who lives there. But, asked to draw or to consult a map of his village, he is faced with learning a new and different sort of task: one that employs compass bearing and units of measurement. What was first understood in the personal terms of local snapshots now has to be considered in the completely general terms of the cartographer. The villager's knowledge by wont, enabling him to lead a stranger from place to place, is a different skill from one requiring him to tell the stranger, in perfectly general and neutral terms, how to get to any of the places, or indeed, how to understand these places in relation to those of other villages.

¬ Gilbert Ryle "Abstractions"

1. Introduction

The above quoted metaphor was used by the coiner of the phrase *logical geography* in his attempt to elucidate the term [1], however today's robotics specialists have realized that they face the same problem as the local villagers, yet the other way round. Nowadays one may argue that the problem of *simultaneously localization and mapping* (SLAM) has been solved, still the output of such a process is only perceivable by a man bearing compass and units of measurement. Accordingly, contemporary mobile robots behave like machine cartographers, unable to liaise with local villagers, that is the human inhabitants, who know by wont to navigate through the own environment. Thus, the majority of the existing mapping approaches aim to construct a globally consistent metric map of the robot's operating environment. The robots bear state of the art instrumentation that allows, on the one hand, the construction of the map and, on the other hand, the own localization with respect to this map and, thus, to determine their global pose with remarkable accuracy. Based on this capability, the robots

* Corresponding author. Tel.: +30 2541 079330.
E-mail addresses: gkostave@pme.duth.gr (I. Kostavelis), agaster@pme.duth.gr (A. Gasteratos).
1 http://robotics.pme.duth.gr

http://dx.doi.org/10.1016/j.robot.2014.12.006
0921-8890/

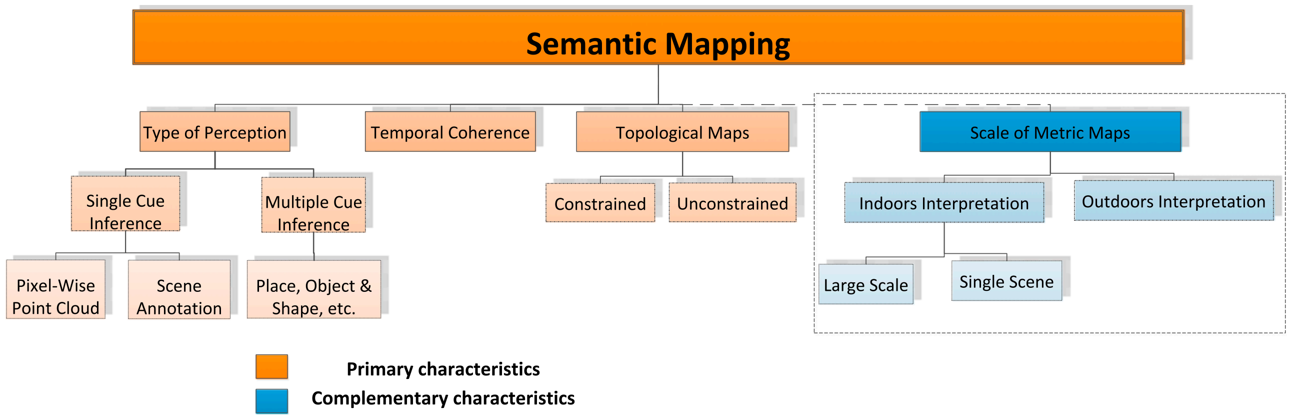

Fig. 1. Taxonomy of the semantic mapping methodologies. Note that the metric mapping is considered complementary attribute for the semantic mapping.

can plan a path and navigate towards a goal, which should be also a specified metric position in the global map reference frame. Howbeit, for a robot to apprehend the environment the way a human does and, consequently, to lead a stranger from place to place, a different skill than any geometrical map can provide is required. The robots to come should be endowed with capacities to understand their surroundings in a human-centric term, i.e. to be able to tell the difference between a room and a corridor or to discriminate the different functionality a kitchen and a living room have. Therefore, the formation of maps augmented by semasiological attributes involving human concepts, such as types of rooms, objects and their spatial arrangement, is considered a compulsory attribute for the future robots that should be designed to operate in environments inhabited by humans.

A solution to this problem is offered by semantic mapping, a qualitative description of the robot's surroundings, aiming to augment the navigation capabilities and the task-planning, as well as to bridge the gap in *human–robot interaction* (HRI), see e.g. [2–4]. Especially the work in [4] addresses semantic mapping with emphasis on HRI by using natural language, thus enabling the most direct way for robots to socialize with humans. Thence, semantic mapping is a flourished pioneering area encouraging the elaboration of several doctoral dissertations [5,6]. The term semantic derives from the Greek word $\sigma\eta\mu\alpha\nu\tau\iota\kappa\grave{o}\varsigma$ [*sēmantikos*], standing for significant, which in turn derives from the verb $\sigma\eta\mu\alpha\grave{\iota}\nu\epsilon\iota\nu$ [*sēmainein*], meaning to signify, that successively stems from the noun $\sigma\tilde{\eta}\mu\alpha$ [*sēma*], that is sign. Thus, semantics is related to the study between signs and the things to which they refer, that is their meaning. The latter is oriented to the identification of the way that two or more entities interact, behave towards, and deal with each other [7]. Thereby, the semantic mapping targets to the identification and the record of the signs and the symbols that contain meaningful concepts for humans, during the robot's wander in human-inhabited areas. Consequently, a semantic map is an enhanced representation of the environment, which entails both geometrical information and high level qualitative features. Speculating the ability of the artificial agents to semantically perceive the own environment and accurately recall the learned spatial memories, the fundamental communication link between human and robots can be established. Therefore, for a successful HRI the robots must retain cognitive interpretation capacities about space, i.e. they should involve semantic attributes about the objects and the places encountered, in association with the geometrical perception of the surroundings. Moreover, the semantic information existing in the abient need to be organized in a such a fashion that the artificial agent can appropriately perceive and represent its environment. The most suitable way to organize all these information is by means of a map, namely a semantic map. Due to the fact that contemporary robots use to navigate in their environments by computing their pose within metric maps, the vast amount of the semantic mapping methods reported in the literature use these metric maps to add semantic information on top of it [2,4]. Therefore, a semantic map comprises high level features that model the human concepts about places, objects, shapes and even the relationships of all these, whilst a metric map retains all those geometrical features the robot should be aware of in order to safely navigate within its surroundings. Yet, it should be further noted that works have been reported on semantic mapping, which do not use a metric map to determine the type of a place, specially the ones using vision [8,9].

The goal of the review paper in hand is to provide insights of the semantic mapping, to study the distinct components encompassing, to give a categorization of the related literature, to mention the possible applications in mobile robotics and, lastly, to refer to the methods and databases available for benchmarking. In order to support this goal, a quality-based taxonomy of the existing mapping strategies is attempted here, which should highlight the dominant attributes such methods retain. An illustrative representation of the described taxonomy of the most frequent components the semantic mapping approaches possess is depicted in Fig. 1. The primary characteristics constitute the condiciones sine quibus non a method producing a complete semantic map should satisfy. Of such are the modalities utilized to reason about the observed scene constituting an element apt to distinguish the abundance of different methods. In particular in many methods only single cues – e.g. objects – are utilized to infer about a place, while some other methodologies exploit multiple cues – such as objects, places and shapes – to produce semasiological clues about an area. Another frequented feature in many semantic mapping techniques is the temporal coherence such a map reveals, which renders it useful for high-level activities, viz. task planning or HRI. An additional important attribute a typical semantic mapping method possesses is the existence of a respective topological map, that is an abstraction of the explored environment in terms of a graph. The nodes of such a graph are organized in a geometrical manner, so as to simultaneously preserve conceptual knowledge about the explored places. These graphs could be either unconstrained ones retaining only geometrical characteristics or they could possess several constrains in accordance with the semantic attributes that they enclose. The existence of a 2D or a 3D metric map of the explored environment – either indoors or outdoors – is a complementary component, which frequently supplements the attributes implemented by the semasiological methods. According to the scale, to which each method is expanded, the metric map could be either a single scene or a progressively created map, that is the pose is referred to a local or a global coordinate system, respectively.