Robotics and Autonomous Systems 60 (2012) 1093–1107

Contents lists available at SciVerse ScienceDirect

Robotics and Autonomous Systems

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

Topological spatial relations for active visual search

Kristoffer Sjöö *, Alper Aydemir [1], Patric Jensfelt

Centre for Autonomous Systems, Royal Institute of Technology, SE-100 44 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 25 July 2011
Received in revised form
24 May 2012
Accepted 4 June 2012
Available online 9 June 2012

Keywords:
Visual search
Spatial relations
Active vision
Spatial reasoning
Autonomous robots

ABSTRACT

If robots are to assume their long anticipated place by humanity's side and be of help to us in our partially structured environments, we believe that adopting human-like cognitive patterns will be valuable. Such environments are the products of human preferences, activity and thought; they are imbued with semantic meaning. In this paper we investigate qualitative spatial relations with the aim of both perceiving those semantics, and of using semantics *to* perceive. More specifically, in this paper we introduce general perceptual measures for two common *topological spatial relations*, "on" and "in", that allow a robot to evaluate object configurations, possible or actual, in terms of those relations. We also show how these spatial relations can be used as a way of guiding visual object search. We do this by providing a principled approach for *indirect search* in which the robot can make use of known or assumed spatial relations between objects, significantly increasing the efficiency of search by first looking for an intermediate object that is easier to find. We explain our design, implementation and experimental setup and provide extensive experimental results to back up our thesis.

1. Introduction

While fiction has for a long time painted a picture of robots walking the Earth alongside us humans, in reality robotics has to date mostly been about industrial robots. The industrial robot revolutionized the manufacturing industry with its speed and precision and is still penetrating new markets, although at a somewhat slower pace. Now, we are seeing the beginning of a new era, an era where robots will actually walk, or at least move about, among people. What has been brewing in labs around the world for decades is now very slowly starting to see the light of day. One example of this new breed of robots is the service robot, intended to help us in the home or office, be it with cleaning, giving us a hand getting up or reminding us to take our medicine. These new robot systems require a completely new level of versatility and adaptability if they are to be able to operate side by side with humans.

There are many important issues that need to be addressed before we have service robots that go beyond vacuum cleaners and floor scrubbers. Some of these issues come from the fact that the robot will be mobile, with all attendant complications such as safety or power supply. There is, however, a whole set of very challenging problems that arise from an unstructured environment where not everything is known in advance, either regarding the properties of the environment or the kinds of tasks to be performed. Humans are superbly adapted to these kinds of conditions; not just physically (such as having legs to negotiate stairs and thresholds, and arms for opening doors and using appliances), but in terms of perceptual and mental abilities as well.

In this context, being able to perceive and act upon *semantic* information about the environment is key. This entails the ability to go beyond simple, hard-coded decision paths operating on low-level, metric, numerical data. The robot must be endowed with the capability to handle the real world in all its complexity, combining the actual sensory input with "commonsense" knowledge such as the typical location of objects, their function and their relation to other entities.

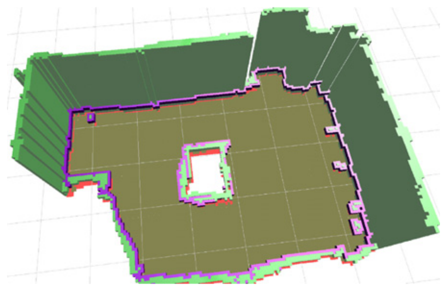
This paper shows how to use semantic knowledge to allow robots to efficiently locate objects in their environment, so that they can interact with or talk about those objects. We contribute quantitative measures for two crucial topological spatial relations that provide a natural hierarchical organization of space for both robots and humans; we then show how to make use of these relations for object search in a principled way, and experimentally demonstrate the gains in efficiency that come from including such semantic information in the search. These contributions are important steps towards the realization of semantic perception for robots, both in terms of "perceiving semantics" and of "perceiving, *using* semantics".

Fig. 1 shows an example of the kind of search behavior aimed for in this paper. The robot is looking for a book; knowing that the book is on the table it is more advantageous to first search for the larger table rather than the book directly; having located the table the robot can narrow its search down and locate the target more efficiently.

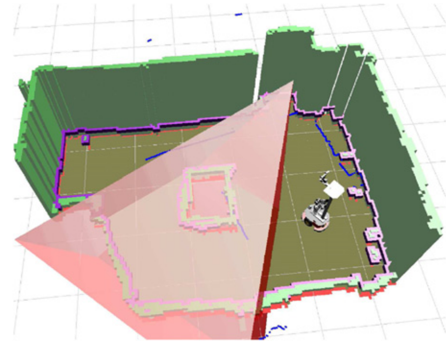
* Corresponding author. Tel.: +46 8 43747930.
E-mail address: krsj@nada.kth.se (K. Sjöö).

[1] The first two authors have contributed equally to this paper.

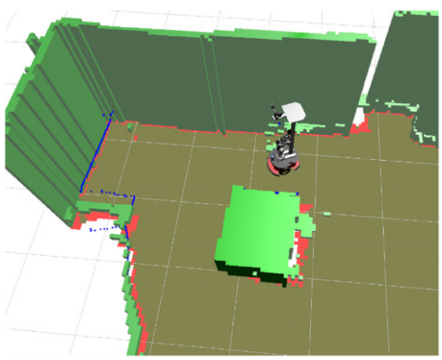
0921-8890/$ – see front matter
doi:10.1016/j.robot.2012.06.001



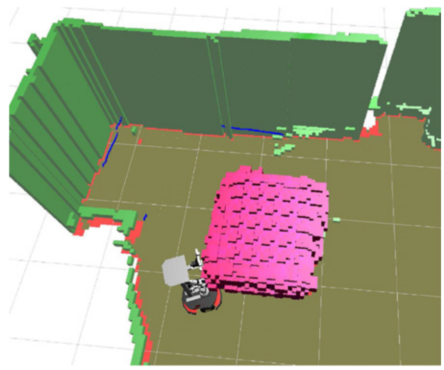
(a) Robot starts by projecting probable locations of the table.



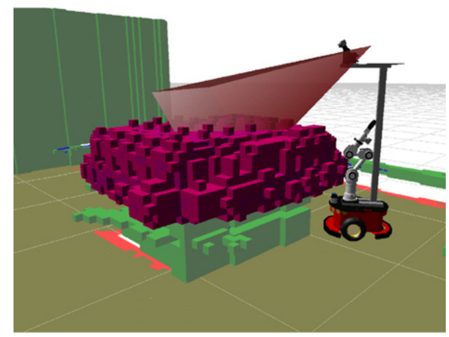
(b) A view is selected.



(c) The table is detected and inserted into the map.



(d) Searching for the book On the table.



(e) Selecting a view for the book.

(f) Book detected.

Fig. 1. Searching for a book by making use of spatial relations.

1.1. Motivation

For a discussion on how our work relates to prior work in the respective area, please see the dedicated sections. Here we will instead motivate why and how these two issues (spatial relations and search) are important.

1.1.1. Why spatial relations

The motivation to study topological spatial relations comes from the insight that adopting human-like cognitive patterns is likely to help robots approach human-like performance in the context of homes, offices or other environments that are the products of human preferences, activity and thought. Linguistic concepts may provide an enlightening insight into the nature of those cognitive patterns, as words and expressions must be supported by underlying mental representations.

In particular, spatial concepts are of great importance to robotic agents, especially mobile ones, as they:

- are a necessary part of linguistic interaction with human beings, both when interpreting utterances with a spatial content and when formulating such utterances.
- allow knowledge transfer between different robots, or from databases—for instance, the Open Mind Indoor Common Sense database [1], which contains "commonsense" information about indoor environments provided by humans, such as where objects may be found.
- provide qualitative abstractions that facilitate learning, planning and reasoning. By making use of spatial relations to model a scene in abstract terms the required amount of data will be drastically reduced. In planning and reasoning one can move away from a continuous metric space which is hard to deal with.

Language reveals that there are a great many spatial relations that humans use to organize space. Different relations serve different purposes and are relevant to different tasks; this paper deals with the task of searching for objects, and for that specific purpose we have elected to examine the most prominent topological