Labeling for a learning mobile robot

Nicolas Bredèche* and Jean-Daniel Zucker***
*Université Pierre et Marie Curie – LIP6/CNRS – 8, rue du capitaine Scott – 75015 Paris - France
**LIMSI/CNRS – Bt. 508 – Université Paris Sud – 91403 Orsay cedex - France
nicolas.bredeche@lip6.fr - jean-daniel.zucker@lip6.fr

Abstract

This paper describes an approach to ground names of objects seen by a mobile robot in a real world environment. Names are given by a human teacher and are stored within a lexicon, along with images where the corresponding objects are present.

Here, the building of a high-level perceptual system is required to enable an efficient learning of the physical definition of objects. The paper briefly describes how an abstraction-based reformulation of real world data can highlight relevant information for a given concept.

Introduction

In this paper, we briefly describe our approach to ground a human lexicon into the environment of a mobile robot. Anchoring human symbols requires learning to build efficient perceptual capabilities out of raw perceptions of the environment (i.e. an image from the robot's video camera). Our approach addresses several field of AI such as artificial vision and machine learning.

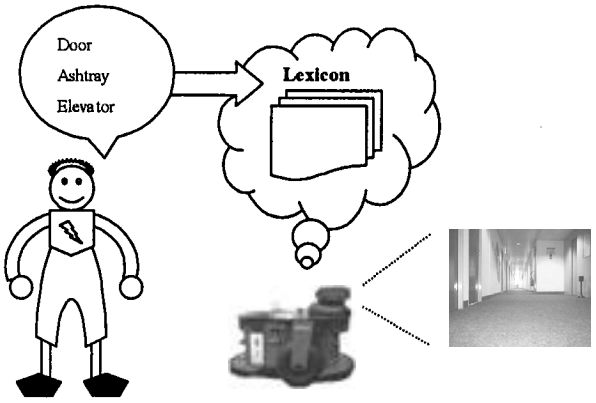

Figure 1 : The learning task – the teacher tells the robot what are the objects he should see now.

The lexicon contains names of objects given by a human teacher to the robot when the aforementioned objects are seen by the robot in the environment (*see fig.1*). The basic interest of using Machine Learning techniques to build the perceptual system of the robot is that there already exists efficient tools which can handle this supervised learning task and output understandable rules to define a concept from a set of descriptions. However in robotics, we face the problem of having few examples for a given concept, and the description of each example is rich in quantity and poor in quality due to few relevant information, noise, and numerous representations of an object (e.g. *the perception of an object can differ according to the angle we look at it,...*).

Therefore, we study the possibility to build an abstraction-based perception from the raw data given by the camera in order to provide a simpler and more relevant description to the learning process.

Towards a physically grounded lexicon

From the robot's point of view, learning to recognize an object named by a human teacher implies building a high-level perceptual system from raw perceptions. However the goal is to identify objects with a good efficiency, the structure of the teacher's and the robot's perceptual systems strongly differ. This is due to the fact that the video camera used by the robot cannot be compared to the human eye and that we do not intend to model a cognitive function of the brain, which should not prove to be possible for a real-time robot application.

Hence, we face many problems related to real-world situated robotics. Firstly, the image from the camera is noisy and complex and costly image manipulations cannot be performed due to the real time constraint. And secondly, since acquiring images by the robot depends on a time consuming exploring task in the environment, we are limited to few hundreds images to learn a given concept (which is consider to be a small training set in machine learning).

In order to cope with these problems, the robot has to provide a high-level description of what it sees. As a matter of fact, a learning task will suffer from too complex data and relevant information must be highlighted. Moreover, we are interested in recognizing objects within an image, which is different to an image indexing task (Stricker and Orengo 1995), where one has to classify a whole image. As a consequence, we cannot use global information on the image but rather local information which will describe small parts of it.

Thus, the task we are addressing is to build such a high-level perception for an autonomous mobile robot.

An abstraction-based approach to anchoring

Our approach is based on abstraction (Saitta and Zucker 2000), which relies on the introduction of a bias in representing data thanks to reformulation in order to get a more relevant representation of the same data for the given learning task. Abstraction has been widely studied in the fields of problem solving (Amarel 1968) and planning (Sacerdoti 1974) and is becoming a subject of great interest to the machine learning community since it offers a general framework to handle complex data (Giordana and Saitta 1990).

Our approach is based on a description of the images in terms of shapes connected to each others. This description is done in an appropriate representation language making possible explicit relations between connected shapes (Chevaleyre and Zucker 2001). These shapes are given by any region growing algorithm based on colour similarity, which are known to be a lot faster than model-based segmentation even though the resulting shapes are not meant to describe an entire object but merely small homogeneously coloured part of objects in the world. This description is used as a beginning to build the high-level perception.

The building of the high-level perceptual system depends on the concept to learn and the learning task will not be able to get the best results immediately due to limitation of the representation language (only 1-depth relations between shapes is handled). That's why we'll proceed iteratively and use the previous results from the learning algorithm to reformulate the description until we are satisfied with the prediction accuracy. We use an abstraction operator that is based on the building of new shapes from basic shapes according to specific rules to be defined during learning.

We mean to converge slowly to recurrent shapes of the objects to be identified (e.g. *a shape based on an aggregate of shapes each of a specific colour that can identify a given object 2 times out of 5*).

Important issues and conclusion

The work presented here addresses the problem of how a robot can create a high-level perception in order to perform an efficient anchoring for symbols which are known to be physically grounded in another perceptual system, that of a human teacher. The more the robot operates in its environment the more it encounters examples of a given concepts (i.e. *images with a specific object in it*) so that the anchoring must be continuously revised.

In this case, the role of knowledge representation is a key aspect to highlight relevant data. The anchoring depends on the capability of the abstraction operator to increase or maintain through time the prediction accuracy of a learning algorithm in an autonomous mobile robot environment.

Our work finds its application in the multi-robots Microbes projects (Picault and Drogoul 2000) and we use a Pioneer mobile robot with LCD camera.

References

Amarel S. On Representations of Problems of Reasoning about Actions. in Machine Intelligence 3, D. Michie (ed.), American Elsevier, New York (1968).

Chevaleyre, Y. and J-D Zucker. Solving multiple-instance and multiple-part learning problems with decision trees and rules sets. Application to the mutagenesis problem. Canadian AI 2001.

Giordana A., L. Saitta. *Abstraction: a general framework for learning*. In Working Notes of the AGAA-90 Workshop, pages 245-256, Boston, MA (1990).

Picault S., A. Drogoul. The MICRobES Project, an Experimental Approach towards "Open Collective Robotics", Proceedings of the 5th International Symposium on Distributed Autonomous Robotic Systems (DARS' 2000), SpringerVerlag Tokyo Inc.

Sacerdoti, E. D. Planning in a hierarchy of abstraction spaces. Artificial Intelligence, 5:115-135 (1974).

Saitta L., J-D. Zucker. Abstraction and Phase Transitions in Relational Learning. in Symposium on Abstraction, Reformulation and Approximation (SARA 2000).

Stricker M. A., M. Orengo. *Similarity of Color Images*. SPIE Proceedings Vol. 2420, 1995.