A Flexible Service Robot Named Markovito

L. Enrique Sucar and Eduardo F. Morales and Hector H. Avilés

INAOE, Luis Enrique Erro No. 1
Santa María Tonantzintla, Puebla, 72840, México

Extended Abstract

The development of service robots has recently received considerable attention. Their deployment, however, normally involves a substantial programming effort to develop a particular application. With the incorporation of service robots to daily activities, it is expected that they will require to perform different tasks. Fortunately, many of such applications share common modules such as navigation, localization and human interaction, among others. In this paper a general framework to easily develop different applications for service robots is presented. In particular, we have developed a set of general purpose modules for common tasks that can be easily integrated into a distributed, layered architecture, and coordinated by a decision–theoretic planner to perform different tasks. The coordinator is based on a Markov decision process (MDP) whose reward is set according to the task's goal, the states are represented by a set of variables affected by the general modules, and the actions correspond to the execution of the different modules. In order to create a new application the user only needs to define a new MDP whose solution provides an optimal policy that coordinates the different modules for performing the task. The rest of the paper provides a general description of the different modules that the robotics group at INAOE have developed.

For map building we construct occupancy grid-based maps, using information from laser scans, sonar readings and odometry. Our latest development uses Rao-Blackwellised particle filters, and is an extension of PMAP [5] that is able to work on line, includes information from laser as well as sonar readings, and have some parallel processing included to increase efficiency (see [6] for details).

For global and local localization we extract natural landmarks based on discontinuities [4]. A discontinuity is defined as an abrupt variation in the measured distance of two consecutive readings of the laser. Given a set of landmarks (discontinuities), a triangulation process is performed between all the visible landmarks to estimate the robot's position. The information from all the visible landmarks is combined considering the angle between landmarks, the distance between the robot and its farthest landmark, and if there are landmarks at both sides of the robot or only one, to give more accurate estimates. For the global localization problem, a ray tracing approach is used to simulate laser readings on the map and each cell is associated with all its visible landmarks and their values.

For navigation we have implemented a module that uses a dynamic programming algorithm, with exponential costs near obstacles, to find the least expensive path. In order to avoid new obstacles, the robot is sensing its environment while moving. In case a new obstacle is placed in front of the robot the module finds an alternative path.

We also developed an alternative navigation module based on TOPs (Teleo-OPerators) and probabilistic roadmaps. TOPs are a set of reactive rules that sense the environment continuously and apply actions whose continuous execution eventually satisfy a goal condition [3]. TOPs are automatically constructed from human traces using inductive logic programming techniques and a simple grammar learning algorithm. The learned TOPs are able to go to a goal position in dynamic and unknown environments. In order to navigate a probabilistic roadmap module is used to create collision free paths using a random generation of points in the configuration space [2, 7]. This points are joined and a smoothing algorithm is used to try to find shortcuts.

Markovito has a speech synthesis and recognition module to synthesize many pre-defined phrases in Spanish language. Speech recognition is based on the *Dimex* project [9], that aims to develop general acoustic-phonetic models and pronouncing dictionaries for Mexican Spanish language inside the *Sphinx2* framework for continuous and multi-speaker recognition. A directional microphone SHURE model SM81 was incorporated to the robot and allows a user to talk to the robot at a distance that ranges from 30 cm. to one m. We also apply a noise removal filter adapted from the Audacity's open source code.

On visual perception, we have worked on face detection, object detection and gesture recognition. Face detection is based on the AdaBoost algorithm proposed in [11]. Object recognition is carried out using the SIFT or Scale-Invariant Feature Transform algorithm [8]. This algorithm aims to detect and describe distinctive features of the objects. These features are stored for a future match with features extracted from incoming images to recognize an object or a scene. A gesture recognition module was developed using an alternative model for hidden Markov models (HMMs), that we call *dynamic naive Bayesian classifiers* (DNBCs) [1]. The training equations for DNBCs are derived in a similar manner as for HMMs [10]. To capture our gestures we implemented a monocular visual system using OpenCV libraries for Linux. Face detection is carried out by the Adaboost algorithm. The position of the right-hand and torso regions are estimated using anthropometrical measures, and the tracking system, trained under different lighting conditions of various users, is able to track the hand accurately in several environmental conditions. Our visual system is able to process up to 30 f.p.s. with image resolutions of 640×480 pixels.

All the modules and the coordinator have been instaled in Markovito. Markovito is based on a PeopleBot robot platform (ActivMedia). It has two rings of sonars, a Laser SICK LMS200, one video camera Canon VCC5, two infrared sensors, a directional microphone SHURE SM81, a gripper, the robot's internal computer, a Laptop, bumpers and a frame grabber WinTV USB 2. The effectiveness of our approach has been experimentally demonstrated in four different service robot tasks specified as MDPs and solved to find optimal policies: (i) following a human under user commands, (ii) navigating to several places in the environment designated semantically, (iii) finding one of a set of different objects in a *house*, and (iv) delivering messages and/or objects between different people.

References

1. H. Aviles and L. Sucar. Recognizing similar gestures. In *Proc. IEEE Inter. Conf. on Pattern Recognition (ICPR)*, China, August 2006.
2. J. Barraquand and J. C. Latombe. Robot motion planning: A distributed representation approach. *Journal of Robotics Research*, 1991.
3. S. Benson and N. Nilsson. Reacting, planning, and learning in an autonomous agent. *Machine Intelligence*, 14:29–62, 1995.
4. S. Hernández-Alamilla and E. Morales. Global localization of mobile robots for indoor environments using natural landmarks. In *Proceedings IEEE Conference on Robotics, Automation and Mechatronics (RAM-2006)*, 2006.
5. A. Howard. Simple mapping utilities (pmap), 2004.
6. V. Jáquez. Construcción de mapas y localización simultánea con robots móviles. Master's thesis, Instituto Tecnológico y de Estudios Superiores de Monterrey - Cuernavaca, 2005.
7. J. C. Latombe. *Robot motion planning.* Kluwer Academics Publishers, 1991.
8. D. Lowe. Distint image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
9. A. Pineda, L. V. nor, J. Cuétara, H. Castellanos, and I. López. Dimex100: A new phonetic and speech corpus for mexican spanish. In C. Lemaitre, C. Reyes, and J. González, editors, *Advances in Artificial Intelligence, Iberamia-2004, Lectures Notes in Artificial Intelligence 3315*, pages 974–983, 2004.
10. L. Rabiner. Tutorial on hidden markov models and selected applications in speech recognition. In *Readings in Speech Recognition*, pages 267–296. Morgan Kaufmann Publishers, 1990.
11. P. Voila and M. Jones. Robust real-time object detection. *International Journal of Computer Vision*, pages 137–154, 2001.