Pumas [OPL] Team Description Paper

Hiroyuki Okada, Jesus Savage, Luis Contreras, and Marco Negrete

National Autonomous University of Mexico, Tokyo Information Design Professional University

Abstract. A service robot is a robot that can assist humans to perform common daily tasks in shared environment, such as houses, offices or hospitals. With this in mind, the final goal of a service robot must be making human life easier and more comfortable. Also, a robot can be an excellent companion, for example, for elderly or lonely people, making their life better and happier. To achieve this, a service robot must be capable of understanding spoken and visual commands in a natural way from humans, navigate in known and unknown environments while avoiding static and dynamic obstacles, recognize and manipulating objects, detect and identify people, among several other tasks that a person might request. This paper describes our current research topics and main findings as well as the efforts to implement all the developed software into our open platform Justina.. We have improved the abilities of robots with various techniques that have been applied to other robot and social IT systems. We briefly introduce them and our latest related research in this description paper.

1 Team Summary

Following our current collaboration, this year we are participating with the service robot Justina, developed at the Biorobotics Laboratory, where we are integrating the joint efforts of team eR@sers (Japan) and Pumas (Mexico). Both research groups have a long history participating in RoboCup@home and a recent history in the @home DSPL league.

Team eR@sers was formed around 2000 to participate in RoboCup 4 legged league. Thereafter, the team joined the @home leage where eR@sers achieved a first place at RoboCup 2008, 2010, second place in RoboCup 2009, 2012, 2017, and third place in RoboCup 2018, and its social robot HSR obtained the @Home Innovation Award in 2016. Furthermore, Team eR@sers was finalist in World Robot Summit(WRS) 2018. On the other hand, Team Pumas DSPL has participated in national and international competitions since 2006. In the Robocup 2018, the team obtained the second place in the categoty DSPL@Home with the robot "Takeshi" and was finalist in WRS 2018, while in the RoboCup 2019 the team got the fourth place in DSPL and second in OPL.

As a joint research group, we have participated in RoboCup 2021 and RoboCup 2022 – in the latter, we ranked 4th place and got the "Smoothest, Safest Navigation" Award.

We mainly focus on the adaptability to the environmental changes and on the integration between the sensory-motor data and symbolic representation, utilizing only the neuro-dynamical model. All developed functions can be packed in ROS modules and almost all training data comes from real sources; the system has been tested in real environments.

2 Innovative technology and scientific contribution

2.1 The VIrtual and Real roBOt sysTem (VIRBOT) [1] [2]

To deal with the challenges that a service robot has to perform, we propose VIRBOT, a robot architecture that combines traditional, reactive, and probabilistic techniques. In VIRBOT, the operation of a service robot is divided into four general layers: Input, Planning, Knowledge Management, and Execution, where each of them has several subsystems.

VIRBOT has a combination of basic artificial intelligence (AI) techniques, specifically the ones used in Natural Language Understanding (NLU), with devices and technology recently developed. By combining symbolic AI with digital signal processing techniques, a good performance in a service robot has been obtained. NLU is used in a service robot to interpret spoken language and then execute a task, where one of the main problems using NLU is determining the meaning representation. Once the application is defined, we have a framework that establishes the robot semantics, defined as a series of instructions that allow a robot to perform relevant operations.

In this section, we will describe most relevant VIRBOT modules categorized by layer.

Input Layer This layer encloses the robot's internal and external sensors, real or simulated, in a series of modules, as follows.

Human-Robot Interface: This module is responsible of recognizing and processing voice and gesture commands. Speech is processed here in the NLU module.

Symbolic Representation and Interpretation: Here, digital signal processing techniques are applied to the data provided by the internal and external sensors to obtain a symbolic representation of the environment.

Perception (Hypothesis Generation): This module generates a set of beliefs about the possible states of the environment. Beliefs are based on the symbolic representation of the sensorial information coming from internal and external sensors, as well as the processed user input from the Human-Robot Interface module. Such beliefs are validated later on to either trigger actions or update the robot's world model.

Planning Layer This layer is responsible of generating plans at a high level of abstraction and performing global reasoning. Beliefs generated by the perception module are validated in this module with information of the Knowledge Management layer. Once validated or recognized, a belief is considered knowledge and either stored or used to trigger the Action Planner, which will generate a plan of action or sequence of physical operations to achieve the desired goals. However, if something unexpected happens while executing a plan, the Goal Activator will be notified, interrupting the Action Planner and triggering the generation of a new plan.

Knowledge Management Layer This layer involves all modules that store and provide access to the robot's knowledge. Such knowledge, which may not be symbolic, ranges from raw and probabilistic maps, to semantic knowledge of the language. For high-level reasoning, a rule-based system is used. The facts and rules are written in CLIPS [3], a language developed by NASA, and represent the robot's knowledge while encoding knowledge of an expert.

Execution Layer This layer is responsible of executing generated plans and making local decisions. At its core, the Bank of Procedures encapsulates a set of hardwired functions the Action Planner combines to assemble more complex plans. These functions implement state machines to partially solve very specific problems, including robot motion and object manipulation. Such functions rely on low level Behavior Methods, a set of reactive algorithms to solve local, unforeseen situations like obstacle avoidance.

2.2 Multimodal feedback for active perception

We use a multimodal system for active robot-object interaction using laser-based SLAM, RGBD images, and contact sensors. In the object manipulation task, the robot adjusts its initial pose with respect to obstacles and target objects through RGBD data so it can perform object grasping in different configuration spaces while avoiding collisions, and updates the information related to the last steps of the manipulation process using the contact sensors in its hand.

We propose in [5] [6] an active object manipulation systems using a 3-DOF RGBD camera (height, pan and tilt movements) on top of a service robot and a 6-axis force sensor in the hand. Through this sensors, the robot is able to detect the obstacle's position and orientation in robot coordinates while the different states of the manipulation process take place.

In particular, the robot arrives near the target within an uncertainty given by the localisation system based on 2D laser scans, but with a localisation error big enough to affect the performance in the grasping step using only the arm's inverse kinematics. Therefore, we propose the use of the upper RGBD camera to update the robot's relative position to the furniture and to locate the target object, and then we use the contact sensor in the robot manipulator to detect when the robot reaches it.

On the other hand, in the recognition task, in [7] we propose a series of strategies for object recognition in human-made environments. We have proven the feasibility of the proposed methods by evaluating the performance in the object recognition task as part of the Storing Groceries and Clean Up tasks in the RoboCup at Home international competition and in the Tidy Up task in en World Robot Summit.

2.3 Multiview Object Recognition using HMM [8]

We present a very economical framework for multi-view object recognition based on Hidden Markov Models that can be trained using a really low number of images, using a flat computer, and in short times; the resulting model is able to recognise the objects and infer the camera trajectory from a new sequence of query images, as shown in Figure 1.

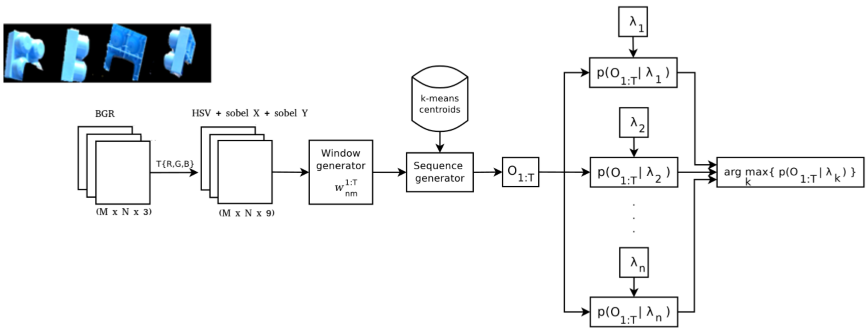

Fig. 1. Multiview object recognition system based on Hidden Markov Models.

We evaluate our system for object recognition with single shot and multi-view shots in the Columbia Object Image Library – a database of colour images of 100 objects taken from different points of view. The experiments include variations on the test images for validating the robustness of the method in the presence of Gaussian white noise, and object rotation in the yaw angle. In Table 1 and Table 2, you can see the results for different vector quantizer sizes using forward and backward sequences, respectively. The first column refers to a single shot inference while the second one shows an inference using 4 shots without noise. Similarly, the third and fourth columns include Gaussian noise in the query images. It can be noticed the improvement in classification using multiple shots.

2.4 Sparse-Map: Automatic Topological Map Creation [9]

We present a task-based map compression technique useful for path-planning and navigation in indoor environments for service robots where, from a point

	clean images		noisy images	
Cents	Acc	Acc fixed	Acc	Acc fixed
128	0.6456	0.7800	0.6368	0.74
256	0.7866	0.8708	0.7213	0.8175
512	0.8737	0.9600	0.8793	0.9575

Table 1. Accuracy results for different VQ sizes. Forward sequences

	clean images		noisy images	
Cents	Acc	Acc fixed	Acc	Acc fixed
128	0.6275	0.7000	0.6081	0.7175
256	0.7300	0.8150	0.6981	0.775
512	0.87	0.9475	0.8593	0.9375

Table 2. Accuracy results for different VQ sizes. Backward sequences

cloud of 3D map features, we calculate a number of clusters based on their spatial position and generate a sparse 3D representation of the environment (Figure 2).

Moreover, we proposed several metrics to asses the quality and performance of a map representation and we tested our proposal using a series of point-cloud benchmarks and clustering techniques, where our method has a comparable performance using a fraction of the memory footprint than the baselines. In the experiments, we created an Octomap-based occupancy grid and a Sparse-Map from the same point clouds. We sample 1000 start and goal points from these maps and request a plan between those points.

Our first metric, the path length L. We had a similar lengths on average, but our method computed them faster. This is associated to the smaller graph size we generate; Sparse-Maps have less than a couple hundred nodes while occupancy grids can have over a thousand nodes to plan. Then, we have the angular tortuosity T. Here, we observe that our method generated paths with smaller tortuosity. Next, the translational dispersion D. While Sparse-Map generates paths with small tortuosity, they presented a higher dispersion; in other words, our method generates paths with little rotational movement but they are more jagged than the paths from the dense planner.

Finally, we have released our system as a ROS-based open source library at `https://github.com/JesusCoyotzi/SparseMapper`

2.5 Movement planning for grasping

Justina's manipulators have 7-DOF each one which allow to perform more human-like movements for grasping. Currently, our manipulation system uses the point cloud to segment objects above a plane. The resulting sub-cloud associated to each object is processed with Principal Component Analysis to get its orientation and general shape. Object shapes are classified as boxes (all sizes

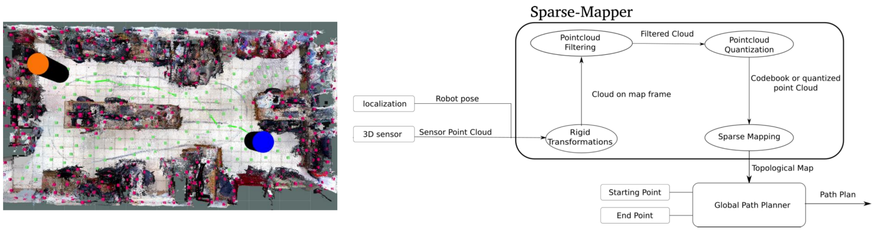

Fig. 2. Sparse-Map, a full three-dimensional task-specific sparse representation of a scene via clustering algorithms.

are similar), cylinders (one axis is much lager than the others) or planes (one axis is much smaller than the others) and an optimal gripper orientation is calculated. We test several orientation candidates and solve the inverse kinematics for each one using the Newton-Raphson method. Among all feasible orientation candidates, we choose the one with the smallest articular varlues. Figure 3 show an example of the grasping system.

Fig. 3. Example of movement planning for grasping

3 Contribution for RoboCup@Home

3.1 RoboCup@Home 2021 WorldWide

Following the challenges in the new-normality, we designed and hosted the RoboCup@Home 2021 WorldWide, as in [10] [11], where the platform is still online and available for interested people to test their service robots' solutions to the proposed challenges.

3.2 RoboCup@Home Education

In addition, starting from 2006, RoboCup@Home has been the largest international annual competition for autonomous service robots as part of the RoboCup initiative. However, it is observed that the development curve of the RoboCup@Home teams have a very steep start. The amount of technical knowledge and resources (both manpower and cost) required to start a new team has made the event exclusive to only established research organizations. For instance, in domestic RoboCup Japan Open challenge, the participating teams in RoboCup@Home were merely around 10 teams, which are about the same teams for the past few years. There were actually several new team requests however the development gap was huge for them to even complete the construction of the robots.

For this reason, RoboCup@Home Education initiative had been started at RoboCup Japan in 2015. RoboCup@Home Education is an educational initiative in RoboCup@Home that promotes educational efforts to boost RoboCup@Home participation and service robot development. Under this initiative, currently there are 3 projects started in Japan:

1. RoboCup@Home Education Challenge at RoboCup AsiaPacific2017 Bangkok.
2. RoboCup@Home Education Challenge at RoboCup Japan Open since 2014.
3. Development of an educational Open Robot Platform for RoboCup@Home
4. We hosted RoboCup@Home Education Workshop Rome, Italy,March 15–16,2017 `https://sites.google.com/dis.uniroma1.it/athomeedu-rome2017/home`
5. Outreach programs (domestic workshops, international academic exchanges, etc.)

(For more information, visit `http://www.robocupathomeedu.org/`)

4 Link to Team Video, Team Website

Team Pumas
Official website: `http://biorobotics.fi-p.unam.mx`

Team Video: `https://youtu.be/ozKxjrrxchI`

References

1. J. Savage, D. Rosenblueth, M. Matamoros, M. Negrete, L. Contreras, J. Cruz, R. Martell, H. Estrada and H. Okada: Semantic reasoning in service robots using expert systems, Robotics and Autonomous Systems, vol. 114, 2019.

2. Pressaco, A., Savage, J., Cruz, J., and Ortuño, S.: A Framework to Process Natural Speech in Service Robots Using a Combination of a Speech Recognition System, Universal Dependencies Extracted by Means of the Stanford Parser and an Expert System, Interactive Collaborative Robotics, Lecture Notes in Computer Science, vol 12998. Springer, Cham. 2021
3. G. Riley: CLIPS Reference Manual Version 6.0. Technical Report Number JSC-25012, Software Technology Branch, Lyndon B. Johnson Space Center, Houston, TX, USA, 1994.
4. T. Inamura and Y. Mizuchi: Competition design to evaluate cognitive functions in human-robot interaction based on immersive VR, The RoboCup Symposium 2017.
5. Contreras, L., Yokoyama, H., and Okada, H.: Visual feedback for active robot-object interaction, The 36th Annual Conference of the Robotics Society of Japan, 2018
6. Contreras, L., Yokoyama, H., and Okada, H.: Multimodal feedback for active perception in service robots, IEEE/RSJ IROS Workshops, 2018
7. Contreras, L., Muto, Y., Okada, H., and Sakamaki, A.: Robot-object interaction strategies for object recognition, The 37th Annual Conference of the Robotics Society of Japan, 2019
8. Nuñez, L., Negrete, M., Savage, J., Contreras, L., and Moctezuma, M.: Multi-view Object and View Sequence Recognition using Hidden Markov Models, IEEE International Conference on Automation Science and Engineering (CASE), 2022
9. Hernandez, J., Savage, J., Negrete, M., Contreras, L., Sarmiento, C., Fuentes, O., and Okada, H.: Sparse-Map: automatic topological map creation via unsupervised learning techniques, Advanced Robotics, 36:17-18, 825-835, 2022
10. Contreras, L., Matsusaka, Y., Yamamoto, T., and Okada, H.: sDSPL - Towards a benchmark for general-purpose task evaluation in domestic service robots The 39th Annual Conference of the Robotics Society of Japan, 2021
11. Matsusaka, Y., Contreras, L., Okada, H., Iwanaga, Y., Yamamoto, T.: A Continuous Integration Based Simulation Environment for Home Support Robot and its Application to RoboCup Competition, IEEE/SICE International Symposium on System Integrations (SII), 2023