

Using Speech to Guide a Mobile Robot Manipulator

L. Bort and A. P. del Pobil
Robotic Intelligence Lab.
Universitat Jaume I, Castellón (SPAIN)
{lbort,pobil}@inf.uji.es

Abstract

Mobile manipulators are mobile robot bases with at least one mounted robot arm which function in an integrated manner. In this paper a single arm mobile manipulator guided by speech commands is presented. The purpose of the mobile manipulator is to reach concrete locations in its environment and to pick up objects. The user can interact verbally with the system to help it localize and identify new objects, as well as incorporate them to its database of learned objects. The overall system is capable of recognizing speech commands, interpret them and execute the corresponding action. Human-robot interaction via speech has also been tested using different examples. Our approach supports the fact that the key to successful service robots is an adequate human-robot interface.

1 Introduction

A mobile manipulator consist of a mobile base and one or two manipulators, usually called robot arms, or simply arms. Initially, arms were used to do tasks such as picking objects, move them and finally place them at another location. This simple task could be extended to a particular domain, for example, to do assembly tasks in a production line. However, this approach is limited by the arm workspace. One solution to this problem is to move the objects that the arm has to manipulate, this is the case of a conveyor belt. But this approach is quite rigid and restrictive to particular domains. Another better solution can be found in a mobile manipulator which takes advantage of the platform mobility to increase the arm workspace.

Currently, a growing number of mobile manipulators exists. There are behavior-based mobile manipulators [12], hybrid architecture mobile manipulators [14], Humanoid Robots with two arms [2], mobile manipulators applied to service robots and rehabilitation robotics [6, 7], and others [11, 5].

Our mobile manipulator, *Jaume*, consists of a Nomad XR4000 mobile robot and a Mitsubishi PA-10 manipulator. The mobile platform has two computers on-board, one controls the base and the other the arm. Both computers are networked in order to facilitate the collaboration between the base and the arm. Furthermore, the base is equipped with several sensors such as tactile, infrared and ultrasonic sensors that are used for navigation. The arm is a seven degree of freedom manipulator with a parallel jaw gripper. It is attached at the top of the base with an inclination of 45 degrees respect to the horizontal plane as shown in the picture below. Additionally, a color stereo pair and a force torque sensor are used for arm operation.

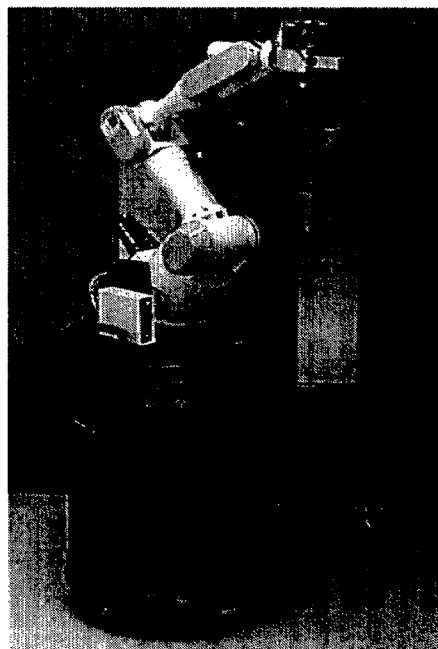


Figure 1: Jaume, our mobile manipulator

Actually, the architecture of this system is organized into four main modules: the navigation module, two

speech processing modules (one for arm commands, and the other for mobile robot commands) and the communication module. The arm module is still under development.

The overall system is capable of recognizing speech commands, interpret them and execute the corresponding action. This system has been tested successfully in an unstructured office environment turning out to be a reliable, efficient and robust approach. Human-robot interaction has also been tested using different examples.

In this paper, a mobile manipulator guided by speech commands is presented. The purpose of this robot is to reach concrete locations in its environment and to pick up objects that it has previously recognized. Both location and object are indicated by speech commands. The user can interact verbally with the system to help it localize and identify new objects, as well as incorporate them to its database of learned objects.

The goal of the work presented in this paper is to integrate different capabilities into a unique mobile manipulator architecture for a service robot. Interaction with natural human speech lends the system behave in a friendly and natural manner. A dialogue-based approach for guiding the robot is also possible, so that the user can help the robot solve the unexpected ambiguities, as well as increment its task-oriented knowledge.

This paper is organized as follows. Firstly, a detailed description of the system architecture is given. In this section, the main modules of the system are introduced as well as their performance. Secondly, in the results section, some capabilities of the mobile manipulator are shown. Finally, some conclusions and future work are given.

2 System Architecture

The architecture of the system is divided into the following modules:

- The Navigation Module
- Two Speech Processing Modules
- The Communication Module

The navigation module consists of three modules: the initiation module, the door entering module and the landmark-searching module. The first module is executed when the robot is turned on, and its purpose is to reach an initial location in the robot environment.

The door-entering module is responsible for crossing door-like openings and uses infrared sensors. Finally, the landmark-searching module allows the robot to move between landmarks avoiding obstacles at the same time. In this approach we have chosen doors as landmarks.

The two speech processing modules translate speech commands into instructions that the arm module or the navigation module can understand. New objects can be added to the robot database through the interaction with the user, so that they can be reached in the future.

Finally, the communication module let us interconnect the rest of the modules by using OMG CORBA (Common Object Request Broker Architecture). The resulting system can be seen as a set of conceptual objects that interact regardless of the programming language, operating system and the host of each object.

2.1 The Navigation Module

The navigation module is responsible for the movements of the mobile base. It consists of the following three modules:

- The Initiation Module
- The Door Entering Module
- The Landmark-Searching Module

2.1.1 The Initiation Module

The robot is able to know its actual position using its odometer. This device informs of the relative position with respect to an initial one. By default, this initial position is set up when the robot is turned on, but it can be changed during the robot operation in order to establish another initial position.

The purpose of our robot is to reach specific locations in its environment so it is necessary to know where it is in each moment. Besides, the locations the robot has to reach are expressed respect to an initial position so it has to start always at the same location. The purpose of the Initiation Module is to reach this initial position.

So as to reach this initial position, an algorithm based on infrared readings is used. This algorithm takes advantage of the infrared sensors precision in order to position the robot in a more accurate position. After a slow series of steps, the robot finally reaches its

initial position. At that moment, it can reach other locations.

2.1.2 The Door Entering Module

The robot environment consists of several rooms connected by doors. These particular doors are wide enough to allow the robot to go through them, but it is necessary a high level of accuracy that cannot be achieved with simple ultrasonic sensors so infrared sensors must be used. For this reason, doors are considered as special landmarks and the robot has to reach them when its final location is in another room different from the present.

Several strategies were tested in order to overcome this problem. They all consider that there are two obstacles in front of the robot representing the door frame, and then use the front readings of the infrared sensors to build an histogram.

The first strategy is the median-based method. In this method, a cumulative histogram is built to get the median value. Then the symmetric of this value is calculated representing the central point of the door frame. Finally the robot is driven to the direction of the sensor with this value.

Another strategy is the two local maxima method in which two local maxima representing the door frame are searched. The new direction of movement can be established as the central point between these two maxima. The problem of this method is that we cannot guarantee the two maxima.

The most reliable strategy is the double media filter. This method is based on applying two times a media filter over the initial raw histogram. This will extend the influence of the door frame to its neighborhood since the frame represents the highest values of the histogram. The result of this operation is that the minimum value of the histogram represents the central point within the door frame.

2.1.3 The Landmark-Searching Module

The Landmark-Searching Module is based on the VFH+ algorithm [16, 4] used for real time obstacle avoidance in several mobile robots [3]. This algorithm uses the concepts of Certainty Grids [13] and Occupancy Grids [8] to build a map of the robot environment. Each cell of this map represents a small region of 5x5 cm. of the environment. The value of this cell is the confidence of the algorithm in the presence of any obstacle.

The key idea is sensor fusion over time. Therefore, if there is an error in one of the sensor readings, the value of the corresponding cell will not be affected too much, so this algorithm improves the safety of methods based on single sensor readings.

As said in the door entering module, doors are used as landmarks that the robot has to reach. To ensure that the robot reaches a door very close to it, a rectangle with the door inside is defined so the robot will effectively reach the door only when it is in this rectangle.

2.1.4 Integration of the Three Modules

The purpose of the Navigation Module is to reach a concrete destination starting from an initial position. The problem of the initial position is solved by the Initiation Module, while the problem of reaching a destination is solved by a combination of several executions of the other modules.

For example, if the destination is in the room next to the initial one, the robot has to execute the Landmark-Searching Module to reach the door, then the Door Entering Module to cross it and finally the Landmark-Searching Module to reach the final destination.

2.2 The Speech Processing Modules

The user can control the robot using his or her voice since the communication between the user and the robot is made through the use of natural speech. A microphone is used to give orders to the robot while a voice synthesizer is used to play the decisions taken by the system. Furthermore, a commercial speech recognition software is used in order to translate a voice order into a string that can be directly processed by our system.

Despite natural speech is less reliable and robust than usual programming languages, voice commands are used to control the robot. This allows a better human-robot interaction through the use of speech rather than typing in commands on a keyboard.

The speech processing modules are organized into two modules, one for arm commands (arm speech processing module) and another for platform commands (base speech processing module).

2.2.1 The Base Speech Processing Module

The base speech processing module takes a string representing the voice command given by the user as input. This command has a part for the platform and a

part for the arm. Both parts have to be processed by the corresponding module. The arm part will be processed by the arm speech processing module and the platform part by the base speech processing module.

Once the part of the string corresponding to the platform is determined, a natural language processing has to be made. So that, the system can decide if the command is correct or not. In that case, a platform command will be sent to be executed in the base.

The location and the status of the robot is also taken into account by this module. Therefore, if the robot has picked up an object and the user sends a new command in order to pick up another object, the system will notice that and will give a message to the user informing of the problem. A similar situation may occur when the robot is unable to reach a concrete destination.

2.2.2 The Arm Speech Processing Module

The input string to the arm speech processing module is the part of the voice command corresponding to the arm. This string is given by the base speech processing module.

Three elements are searched to process this string: the object to manipulate, the verb specifying the action and the final destination of the object. However, if the object is not known, a learning process starts to introduce it in the database of learned objects. This process consists of a human-robot dialogue to help the robot localize and identify new objects. During this process, qualitative spatial reasoning techniques are used [9, 10] to guide the robot.

As in the previous speech processing module, a status of the arm is taken into account. So that, if a syntactically correct command is sent to the arm but it cannot be executed due to the current status, the user is warned about the problem.

2.3 The Communication Module

One of the problems when one builds a system made of several modules developed in different systems is the communication amongst them. In our system, there are modules programmed in C++ and running in Linux while others are developed in Java running in Windows NT. Moreover, the computers where these programs are executed are different, but interconnected by a LAN.

An implementation of CORBA (Common Object Request Broker Architecture) [1, 15] is used to overcome

the above problem. CORBA is an architecture that let us interconnect applications regardless of the operating system and the programming language. Besides, an application can communicate with another in a different host transparently. CORBA will manage all the communications for us.

In this context, the system developed consists of several objects that can communicate each other. However, the implementation of them can differ. In fact, the objects of the navigation module are implemented in C++ running in Linux and the objects of the speech processing modules in Java running in Windows NT. Finally, all the objects of the system are distributed in a heterogeneous LAN and can be moved from one host to another without any change in the system. All the communications work is done by CORBA.

3 Results

The mobile manipulator presented is capable of recognizing speech commands, interpret them and execute the corresponding action over the physical robot. No speech programming language is used, instead a natural human speech interface has been taken to lend the system behave in a friendly and natural manner.

In this way, the robot can interpret commands like: *Jaume, go to Pedro's Office and fetch the book.* After interpreting the human voice, the commercial speech recognizing program will pass the string to the base speech recognizing module. This module will split the string into two parts, one for its own and the other for the arm speech recognizing module. Both speech recognizing modules will analyze its string and the status of the robot. Finally, each module will respond with a command directly executable by the robot or a warning message to be played by the voice synthesizer in case something was wrong.

A dialogue-based approach for guiding the robot is also possible in case an object cannot be found in the environment. With this capability, the user can help the robot solve unexpected ambiguities as well as increment its task-oriented knowledge.

The system presented consists of four main modules: two speech recognizing modules, the navigation module and the communication module. However, the manipulation module is still under development, so that the system is only capable of doing navigation tasks. The manipulation module will let the robot perform manipulation tasks using stereo vision and a force sensor. Therefore, an integration of manipulation and navigation will be possible in the near future.

The scheme presented has been tested successfully in an unstructured office environment turning out to be a reliable, efficient and robust approach. Particularly, the robot can reach several offices in its environment after being ordered to do so. Then, a manipulation task is virtually executed, and finally the robot returns with the object grasped.

4 Conclusions

This paper has presented a mobile robot manipulator for service applications. The final system integrates a set of modules that control navigation, arm motion, speech processing and communication in a user-friendly manner, since the robot can be instructed verbally to pick up a distant object and to bring it back. Our approach supports the fact that the key to successful service robots is an adequate human-robot interface.

Currently, we are developing new grasping algorithms based for general 3-D objects. Furthermore, we are planning to introduce force control to them. We are also working on color stereo vision processing applied to reaching and grasping algorithms. Human safety in the mobile manipulator workspace is also being studied.

Acknowledgements

The authors would like to thank José Bellmunt for his contribution in the communication module and Neus Vallés and Santiago Albaladejo for their contribution in the speech processing modules.

The research has been carried out in the Robotic Intelligence Laboratory at Jaume I University and has been supported by Fundació Caixa Castelló under project P1B97-06, Generalitat Valenciana under project GV97-TI-05-8 and Comisión Interministerial de Ciencia y Tecnología under project TAP1998-0450. The first author was also supported by a grant of the Spanish Ministry of Education and Culture.

References

- [1] Baker, S., "CORBA Distributed Objects Using Orbix", *Addison-Wesley, ACM Press*, 1997.
- [2] Bishoff, R., "HERMES - A Humanoid Mobile Manipulator for Service Tasks", *International Conference on Field and Service Robots*, Canberra, 1997.
- [3] Borenstein, J. and Ulrich, I., "The GuideCane - A Computerized Travel Aid for the Active Guidance of Blind Pedestrians", *Proc. 1997 IEEE International Conference on Robotics and Automation*, pp. 1283-1288.
- [4] Borenstein, J. and Koren, Y., "The Vector Field Histogram - Fast Obstacle Avoidance for Fast Mobile Robots", *IEEE Journal of Robotics and Automation*, June 1991, Vol. 7, No. 3, pp. 278-288.
- [5] Connell, J., "A Colony Architecture for an Artificial Creature", *MIT Artificial Intelligence Lab. Technical Report 1151*, 1989.
- [6] Dario, P., Guglielmelli, E., Laschi, C., Guadagnini, C. Pasquarelli, G. and Morana, G., "MOVAID: A New European Joint Project in the Field of Rehabilitation Robotics", *Proc. of the 17th International Conference on Advanced Robotics (ICAR'95)*, Sant Feliu de Guixols, Spain, September 1995, vol. 1 pp. 51-56.
- [7] Daxwanger, W., Ettl, E., Fischer, C., Freyberger, F., Hanebeck, U. and Schmidt, G., "ROMAN: Ein Mobiler Serviceroboter als Persönlicher Assistent in Belebten Innenräumen", *In: Schmidt, G. and Freyberger, F.:12. Fachgespräch AMS 1996, Springer Verlag, Berlin*, pp. 314-333.
- [8] Elfes, A. "Using Occupancy Grids for Mobile Robot Perception and Navigation", *IEEE Computer*, 1989, Vol. 22, No. 6, pp 46-57.
- [9] Escrig, M. T., Toledo, F., "Qualitative Spatial Reasoning: Theory and Practice", *IOS Press*, 1998.
- [10] Escrig, M.T., Toledo, F., and del Pobil, A.P., "An Overview to Qualitative Spatial Reasoning", *Current Trends in Qualitative Reasoning and Applications*, 1995, pp. 43-60.
- [11] Khatib, O., "Mobile Manipulation. The Robotic Assistant", *Journal on Robotics and Autonomous Systems*, 1999, Vol. 26, pp 175-183.
- [12] MacKenzie, D. and Arkin, R. C., "Behavior-Based Mobile Manipulation for Drum Sampling", *Proc. 1996 IEEE International Conference on Robotics and Automation*, pp. 2389-2395.
- [13] Moravec, H. P., "Sensor Fusion in Certainty Grids for Mobile Robots", *AI Magazine*, 1998, pp. 61-74.

- [14] Petersson, L., Egerstedt, M. and Christensen, H. I., "A Hybrid Control Architecture for Mobile Manipulation", *Proc. 1999 IEEE International Conference on Intelligent Robots and Systems*.
- [15] Sai-Lai, L. and Riddoch, D., "The omniORB2 version 2.7.1 User's Guide", *AT&T Laboratories Cambridge*, 1999.
- [16] Ulrich, I. and Borenstein, J., "VFH+: Reliable Obstacle Avoidance for Fast Mobile Robots", *Proc. 1998 IEEE International Conference on Robotics and Automation*, pp. 1572-1577.