

Hello Robot, Can You Come Here?

Using ROS4iOS to Provide Remote Perceptual Capabilities for Visual Location, Speech and Speaker Recognition

François Ferland

Ronan Chauvin

Dominic Létourneau

François Michaud

IntRoLab - Intelligent, Interactive, Integrated and Interdisciplinary Robotics Lab

3IT - Université de Sherbrooke (QC) Canada J1K 0A5

[francois.ferland, ronan.chauvin, dominic.letourneau, francois.michaud]@usherbrooke.ca

1. INTRODUCTION

Mobile devices such as smartphones and tablets can provide additional sensing and interacting capabilities to a mobile robot, even extending its senses to remote locations. To do so, we developed ROS4iOS¹, a native port of ROS [6] allowing to seamlessly use data from mobile iOS devices to be processed on a robot. To demonstrate this capability, this video presentation illustrates how ROS4iOS has been used to implement an assistance scenario: a person in a remote location asks our IRL-1 [1] robot for assistance, and IRL-1 must recognize the person's voice, identify the remote location using images taken from the mobile device, navigate to the identified location and interact vocally with the person through the mobile device. When communication with the robot is established by launching a specific iOS application, audio from the mobile device is published on a single topic and directed toward two ROS nodes on IRL-1: PocketSphinx [4], a speech recognition toolkit, is used to obtain the person vocal commands. WISS [3], a speaker identification system, is used for speaker identification. From a vocal request made through the mobile device, the robot can identify the person and request to get images of the remote location using the rear-facing camera of the mobile device. RTABMap [5], a loop-closure detection system for visual location recognition, is used to locate the person. A map was previously built with the robot's laser range finder using the gmapping SLAM [2] algorithm, thus permitting reuse of the ROS navigation stack to plan a path and follow it safely from its current location to the requested one. Depending on its current state, the robot can also display the most suited view (i.e., phone, camera, navigation) on the iOS application. ROS4iOS can also be used to teleoperate IRL-1 through the mobile device or to discuss remotely with the robot, using the same ROS modules.

ROS4iOS opens up a rich set of possibilities for HRI, making ROS-compatible code accessible on mobile devices. In future work, we plan to integrate a dialog management system [7] for other assistance scenarios, such as image-based object fetching and delivery.

¹https://github.com/introlab/ros_for_ios

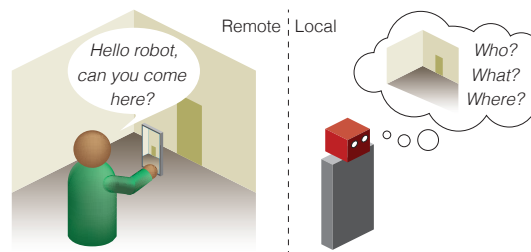


Figure 1: A person in a remote location interacting with the robot.

Acknowledgments

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada.

2. REFERENCES

- [1] F. Ferland, D. Létourneau, A. Aumont, J. Frémy, M.-A. Legault, M. Lauria, and F. Michaud. Natural interaction design of a humanoid robot. *Journal of Human-Robot Interaction*, 1(2):14–29, 2012.
- [2] G. Grisetti, C. Stachniss, and W. Burgard. Improved techniques for grid mapping with Rao-Blackwellized particle filters. *IEEE Transactions on Robotics*, 23(1):34–46, 2007.
- [3] F. Grondin and F. Michaud. WISS, a speaker identification system for mobile robots. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1817–1822, 2012.
- [4] D. Huggins-Daines, M. Kumar, A. Chan, A. W. Black, M. Ravishankar, and A. I. Rudnicky. PocketSphinx: A free, real-time continuous speech recognition system for hand-held devices. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, pages 185–188, 2006.
- [5] M. Labbé and F. Michaud. Appearance-based loop closure detection in real-time for large-scale and long-term operation. *IEEE Transactions on Robotics*, pages 734–745, 2013.
- [6] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Ng. ROS: an open-source robot operating system. In *ICRA Workshop on Open Source Software*, 2009.
- [7] C. Rich and C. L. Sidner. Using collaborative discourse theory to partially automate dialogue tree authoring. In Y. Nakano, M. Neff, A. Paiva, and M. Walker, editors, *Intelligent Virtual Agents*, volume 7502 of *Lecture Notes in Computer Science*, pages 327–340. Springer Berlin, 2012.