# Demo: Towards Universal User Interfaces for Mobile Robots

Dávid Rozenberszki
david.rozenberszki@nokia.com
Nokia Bell Labs
Budapest, Hungary

Gábor Sörös
gabor.soros@nokia-bell-labs.com
Nokia Bell Labs
Budapest, Hungary


**Figure 1: Co-localized robots and head-mounted displays enable auto-generated and context-aware user interfaces. The virtual UI elements (here a task list) extend the limited physical interface of the mobile robot above it and move together with it.**

## ABSTRACT

We demonstrate the concept and a prototype of automatically generated virtual user interfaces for mobile robots. Human(s) and robot(s) are co-localized in the space via their own on-board navigation and shared spatial anchors. The robot sends the description of context-aware user interface elements to a head-mounted display, which renders the virtual widgets around and seemingly attached to the physical robot.

## CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**; **Ubiquitous and mobile computing systems and tools**.

## KEYWORDS

mobile robot, user interface, mixed reality, spatial anchor

**ACM Reference Format:**

Dávid Rozenberszki and Gábor Sörös. 2021. Demo: Towards Universal User Interfaces for Mobile Robots. In *Augmented Humans International Conference 2021 (AHs '21), February 22–24, 2021, Rovaniemi, Finland*. ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3458709.3458996

## 1 INTRODUCTION

Smart home robots are gaining popularity in various tasks. Typical robots today offer a very limited physical user interface for cost and space reasons. For example, vacuum cleaner robots typically have only a start-stop button, and all other features are shifted to a companion smartphone app. While some models offer spoken language output, the communication between the robot and the human is severely limited. Defining a simple task like cleaning a specific area involves opening the app, navigating to a specific tab, and selecting an area on the floor map. The smartphone interaction is spatially disconnected from the controlled object and represents a barrier, in contrast to simply pointing to the floor. We address these issues by augmenting the robot with context-aware virtual interaction widgets extending the UI around the controlled object, enabled by co-localized robots and head-mounted displays.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
*AHs '21, February 22–24, 2021, Rovaniemi, Finland*
© 2021 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-8428-5/21/02.
https://doi.org/10.1145/3458709.3458996

## 2 VIRTUAL UI FOR MOBILE ROBOTS

In order to reduce the barrier between the controller and the controlled object, we need to spatially register the user interface elements with the controlled object. Assuming that augmented reality glasses will become mainstream in the future, we can extend physical interfaces with interactive virtual elements. Previous research on universal user interfaces [1, 2] has shown that with the help of head-mounted displays (HMDs), we can render arbitrary widgets around objects from a machine-readable UI description. Because the widgets are virtual, the user can even customize and personalize their appearance from the general functional description.

The virtual widgets can be spatially registered with the controlled object if we track the objects and the HMDs in the space. Autonomous mobile robots are able to build a map of their environment and localize themselves within that map, just like mixed reality glasses do. However, the maps of the devices are not necessarily aligned in any way, because typically the coordinate system is defined by the position and orientation at startup. Recognizing a particular area between multiple cameras and/or at different times is possible via so called AR anchors, which are a compact representation of a local map with a unique ID. Mixed reality devices and mobile AR platforms can anchor digital content to static physical locations, however, defining anchors on dynamic objects is not possible with today's technology. Alternatively, one could let the devices recognize and track each other, but doing so flexibly with a wide variety of appliances is computationally too heavy for an HMD and likewise for appliances.

To be able to attach virtual widgets to our mobile robots, we instead place anchors in the static environment, and let both the robots and the HMDs discover those. Given robust mapping and localization for all participants of the environment and the common anchors, the devices can align their maps and coordinate systems and calculate each other's pose. The common anchors in the space can be anything like Aruco markers or other common AR anchors available in many SDKs. The user and the robot are free to move in the space, the virtual widgets appear to be attached to the robot. Once co-localized, the devices share their widget descriptions with each other, an the HMDs render the appropriate primitives around the reported pose of the robot, and the user can interact with the virtual widgets with hand gestures recognized by the HMD. Figure 2 illustrates this interaction flow. In our first prototype, we do not cover important questions about secure communication and control permission levels, and problems with localization imperfections.

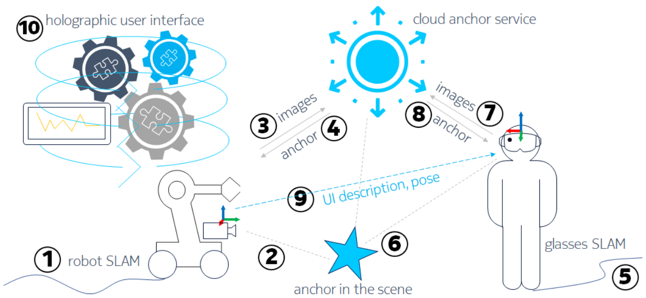


**Figure 2: Interaction flow: a robot and a human arrive to the same location each running their own SLAM, co-localize based on an anchor, and holographic user interface elements around the robot are rendered on the HMD.**

## 3 IMPLEMENTATION

In our prototype implementation, we used a Roomba i7 [3] vacuum cleaner and an OpenBot [5] robot car, and two HoloLens HMDs. We attached a small Intel NUC computer on the Roomba, a Jetson Nano on the OpenBot, and Intel RealSense D435 RGB-D and T265 (not using built-in pose tracking) stereo cameras on the robots (see Figure 3). For navigation, the robots were running UcoSLAM [6], an extended version of the popular ORB-SLAM2 which is able to recognize Aruco markers as landmarks. While the robots are able to co-localize based on markers in the space alone, this is not supported on the HMDs.

Instead of bringing a custom SLAM to HoloLens, we bring the relocalizer of HoloLens to the robots. Therefore, we apply Azure Spatial Anchors[1], a cloud service for vision-based localization, which recently got released also for the Robot Operating Systemon Linux. This makes the development of collaborative robotics applications easier than with other AR anchor technologies. Because the Jetson's arm64 CPU is currently not supported by Azure and it is not powerful enough for additional tasks, we chose to stream the camera images coupled with estimated poses to a host PC to perform the relocalization.

We extended both ends with ROS communication wrappers and implemented a three-level architecture, where (1) the SLAM algorithms are running on board and can be different on each device, (2) live pose streams are shared over the local network in a unified format, and (3) map alignment happens with cloud anchors.

On the HMD, simple widgets are rendered in a Unity app (similar to the work of Delmerico et al.[2]) using the Microsoft Mixed Reality Toolkit and they can be manipulated with standard HoloLens gestures. A simple state machine decides on which widgets to render as illustrated in Figures 1 and 4.

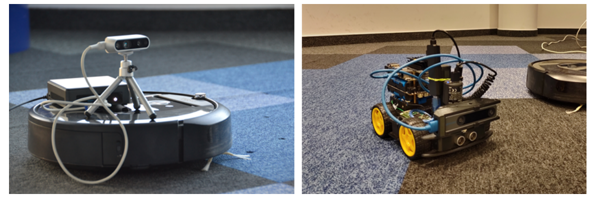

**Figure 3: Roomba and OpenBot with on-board navigation**

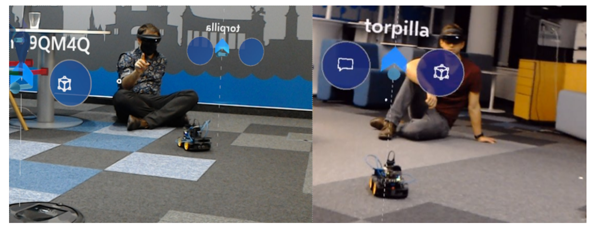


**Figure 4: Virtualized user interfaces for multiple robots in the same environment**

## 4 APPLICATIONS

We show two robots and two HMDs in the same physical space, all running their own SLAM, all searching for a common anchor, and all sharing their pose streams with respect to the common coordinate frame. An HMD can connect to the selected robot and retrieve its current user interface description, and render appropriate holographic elements above the reported position. In the future, we could also extend our system with other collaborative applications, for example the robots can avoid collision with each other at corners because they can expect each other coming. As the HoloLens can already recognize user clicks on the room geometry, we could define the robots target position by simply clicking on the floor. By having multiple tracked HMDs, we could also imagine extending the humans with virtual widgets in the air [4]. We hope that our work inspires further research on collaboration between augmented humans and augmented mobile robots.

## ACKNOWLEDGMENTS

We thank Lóránt Farkas for his help with the robot hardware.

## REFERENCES

[1] Vincent Becker, Felix Rauchenstein, and Gábor Sörös. 2019. Investigating Universal Appliance Control through Wearable Augmented Reality. In *Proceedings of the 10th Augmented Human International Conference 2019.*

[1]https://azure.microsoft.com/en-us/services/spatial-anchors/

[2]https://github.com/microsoft/mixed-reality-robot-interaction-demo

[2] Valentin Heun, Shunichi Kasahara, and Pattie Maes. 2013. Smarter Objects: Using AR Technology to Program Physical Objects and Their Interactions. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*.

[3] iRobot. 2020. Roomba i7. https://www.irobot.com/en/roomba/i7-series. [Online; accessed 08.01.2021].

[4] Mikko Kytö and David McGookin. 2017. Augmenting Multi-Party Face-to-Face Interactions Amongst Strangers with User Generated Content. *Computer Supported Cooperative Work (CSCW)* 26 (12 2017), 527–562. Issue 4. https://doi.org/10.1007/s10606-017-9281-1

[5] Matthias Müller and Vladlen Koltun. 2020. OpenBot: Turning Smartphones into Robots. *arXiv:2008.10631* (2020).

[6] Rafael Muñoz-Salinas and R. Medina-Carnicer. 2020. UcoSLAM: Simultaneous localization and mapping by fusion of keypoints and squared planar markers. *Pattern Recognition* 101 (2020).