Interacting with a Conveyor Belt in Virtual Reality using Pointing Gestures

Jérôme Guzzi, Gabriele Abbate, Antonio Paolillo, Alessandro Giusti
Dalle Molle Institute for Artificial Intelligence, *USI-SUPSI*, Lugano, Switzerland
jerome@idsia.ch

***Abstract*—We present an interactive demonstration where users are immersed in a virtual reality simulation of a logistic automation system. Using pointing gestures sensed by wrist-worn inertial measurement unit, users select defective packages transported on conveyor belts. The demonstration allows users to experience a novel way to interact with automation systems, and shows an effective application of virtual reality for human-robot interaction studies.**

***Index Terms*—interaction, virtual reality, pointing, conveyor**

I. Introduction

Humans use pointing gestures to directly communicate spatial information to each other (e.g., “look at *that* bird“ or “park the car *there*”). In robotics, one can use pointing to communicate with and control mobile robots [1]: pointing-based control is very intuitive, needs little to no training, and requires minimal infrastructure, such as a bracelet equipped with an inertial measurement unit (IMU), to track the arm orientation. These advantages are also relevant in industrial automation settings; they allow hands-free interaction with the system from anywhere and without complex spatial reasoning: what you can see, you can point.

We recently introduced a pointing-based toolkit [2] to interact with co-located robotic and automation systems. Because testing in real facilities is not always feasible, we often relied on simulations during development, which is a common practice in robotics. In this context, how can we include humans in a simulation to test human-robot interaction algorithms? One very promising answer is to use *virtual reality* (VR) to immerse users in the simulation. In our case, interactions experienced in VR feel very similar to real-world interactions because: 1) pointing-based interfaces do not require any physical contact; 2) VR and real-world implementations rely on the same sensor (an IMU) to track the arm orientation. More generally, VR opens new possibilities to experiment with robotic systems with humans-in-the-loop: researchers can test systems that would be too dangerous, too expensive, or just too large to fit in their labs.

Our **contribution** consists in: 1) a compelling demonstration of pointing-based interfaces; 2) an example of a VR testbed for human-robot interaction studies. First, in Section II, we describe the tools used in the demonstration. Then, in Section III and in the accompanying video, we illustrate how the demonstration takes place.

This work is supported the European Commission through the Horizon 2020 project 1-SWARM, grant ID 871743.

II. Setup

We focus on the interaction with an automated logistic system that transports packages on conveyor belts. We target scenarios where users, possibly while performing other duties, notice a damaged (or otherwise relevant) package and want to communicate it to the system by pointing at it. Our demonstration lets users experience such a scenario in simulation by combining three technologies: a pointing-based interaction modality, a robotic simulator, and a virtual reality headset.

a) Pointing-based Interaction: Users wear a bracelet with an IMU on their wrist. From the IMU orientation, we reconstruct the hand position with respect to the user’s eye. Then, following the pointing ray originating from the eye and passing through the index finger [3], we infer where the user is pointing. In particular, we can compute whenever users are pointing at any object of known location, provided we also know where the user is. In fact, by inverting the procedure, we can estimate the user’s location: we ask the user to point at know locations and then compute the user’s pose that minimizes the error between inferred and actual pointed locations [4]. When users are pointing at packages, we provide instant visual feedback: a cursor drawn on LED strips placed alongside the conveyor belts.

b) Simulation: The simulation uses CoppeliaSim [5] to dynamically simulate packages being transported on, and diverted from, conveyor belts, as well as individual LEDs and LED strips used for feedback. The simulated automation system publishes the position of packages and receives a list of selected packages from the pointing-based interaction module using ROS2. Selected packages are automatically diverted to a specific bay. The simulation state is updated 100 times per second.

c) Virtual Reality: The simulation is synchronized (via TCP socket, by a custom tool we developed) and rendered in real-time using Unity, with minimal latency, at 120 fps on the wireless VR headset (Oculus Quest 2). We only send changes in the simulated world state to the headset (e.g., a change of color or pose), resulting in low bandwidth requirements (less than 1 Mbit/s) that don’t require any advanced wireless setup. We rely on the headset’s internal hand- and finger-tracking capabilities exclusively to render the user’s hands; this improves user experience as it makes pointing objects in VR a very similar experience as pointing them in the real world. Moreover, for a faithful interaction, also the visual feedback

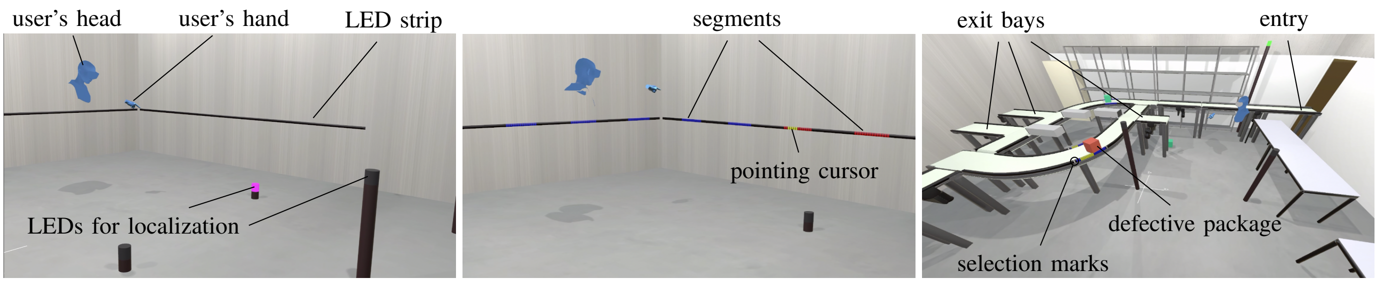

Fig. 1. Three virtual reality scenes: (left) pointing at LEDs while performing the localization procedure; (center) selecting all red segments on the LED strip; (right) selecting a defective package on the conveyor belt. These images, with the user's avatar (part of the torso and hands) in blue, are captured in VR from an external viewpoint while the user is performing the demonstration. The user is seeing the scene from their internal viewpoint (not displayed) through the headset.

from the system needs to be similar to the one in the real world. To this end, in VR, we render LED strips as masked textures, such that simulated pixels have the same size and similar perceptibility as real pixels. Both the VR and the real-world implementations only rely on measurements from the IMU bracelet and run the same ROS2-based software [2] for pointing-based interaction.

III. Demonstration

The goal of the demonstration is to let users actively experience pointing-based interaction with an automation system. First, users wear the VR headset and the IMU bracelet. Several users can participate at the same time. As there are no cables, users are free to move as they like. Then, the VR headset connects to one of several simulation scenes in which users are guided by voice to interact with the automation system. The scenes, illustrated in Figure 1, let the users progress from a simple setup with few elements, to the final realistic setup. A few minutes are enough to get accustomed to virtual reality (even for first-time users) and learn the basics of the interaction modality.

a) First scene, localization and pointing feedback: The user starts interacting by pressing a button on the bracelet. Then, the system lights up two LEDs in sequence, each for a couple of seconds, which the user has to point. The system uses this procedure to localize the user in the environment. After this short initial phase, the system provides visual feedback about the reconstructed pointing target by displaying a yellow cursor on a long L-shaped LED strip. Users are free to point anywhere to experience how the system tracks their pointing. They rapidly learn to take advantage of the feedback to increase pointing precision.

b) Second scene, selecting segments on a LED strip: The LED strip now displays moving red and blue segments, and the user has to select all red segments. A segment is selected when the user points at it for 2 seconds; then, the system plays a sound as feedback, and white dots are added to the sides of the segment. Users experience how the interface allows them to select (and deselect) even rapidly moving objects with little cognitive effort.

c) Third scene, selecting packages on a conveyor belt: The user is now immersed in an environment that accurately reproduces part of a real logistic automation system with packages transported on conveyor belts and automatically routed to exit bays. LED strips are installed along the belts to provide visual feedback to the user. Some packages are identifiable as defective (colored in red); the user has to signal them to the system, which will divert them to a specific bay. The interaction takes place as follows: 1) the user notices a defective package; 2) the user pushes the bracelet button; 3) the user performs the localization procedure by pointing at the two LEDs that light up in sequence; 4) the user points at the defective package until the system visually and acoustically signals that it selected the package. During the demonstration, the user can experience the interaction from different locations, some of which are more challenging than others, due to distance or occlusions.

IV. Conclusion

We presented a demonstration where users interact with a simulated automation system in virtual reality. The demonstration allows users to experience a pointing-based interaction modality to select packages transported on conveyor belts. The tool is very flexible and can be extended to visualize in VR, and possibly interact with, any system that can be simulated in CoppeliaSim.

References

[1] B. Gromov, G. Abbate, L. Gambardella, and A. Giusti, "Proximity Human-Robot Interaction Using Pointing Gestures and a Wrist-mounted IMU," in *Proceedings of the 2019 IEEE International Conference on Robotics and Automation (ICRA)*, 2019, pp. 8084–8091.

[2] G. Abbate, A. Giusti, A. Paolillo, L. Gambardella, A. E. Rizzoli, and J. Guzzi, "PointIt: A ROS Toolkit for Interacting with Co-located Robots using Pointing Gestures," in *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2022 (to appear).

[3] S. Mayer, V. Schwind, R. Schweigert, and N. Henze, "The Effect of Offset Correction and Cursor on Mid-Air Pointing in Real and Virtual Environments," *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2018.

[4] B. Gromov, L. Gambardella, and A. Giusti, "Robot identification and localization with pointing gestures," in *Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 3921–3928.

[5] E. Rohmer, S. P. N. Singh, and M. Freese, "CoppeliaSim (formerly V-REP): a Versatile and Scalable Robot Simulation Framework," in *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013, pp. 1321–1326.