Demonstration:Remote 3D Scanning with VR and Robotic Arm

Ryuto Usami
Nagoya University
Nagoya, Aichi, Japan
usami@ucl.nuee.nagoya-u.ac.jp

Kisho Watanabe
Nagoya University
Nagoya, Aichi, Japan
kisho@ucl.nuee.nagoya-u.ac.jp

Yuki Gushi
Nagoya University
Nagoya, Aichi, Japan
gushi@ucl.nuee.nagoya-u.ac.jp

Shuto Tsutsui
Nagoya University
Nagoya, Aichi, Japan
tsutsui@ucl.nuee.nagoya-u.ac.jp

Yoshiki Watanabe
Nagoya University
Nagoya, Aichi, Japan
yoshiki@ucl.nuee.nagoya-u.ac.jp

Kazuma Kano
Nagoya University
Nagoya, Aichi, Japan
kazuma@ucl.nuee.nagoya-u.ac.jp

Yuya Aikawa
Nagoya University
Nagoya, Aichi, Japan
aichi@ucl.nuee.nagoya-u.ac.jp

Kaiya Shimura
Nagoya University
Nagoya, Aichi, Japan
kaiya@ucl.nuee.nagoya-u.ac.jp

Nozomi Hayashida
Nagoya University
Nagoya, Aichi, Japan
linda@ucl.nuee.nagoya-u.ac.jp

Kenta Urano
Nagoya University
Nagoya, Aichi, Japan
vrano@ucl.nuee.nagoya-u.ac.jp

Takuro Yonezawa
Nagoya University
Nagoya, Aichi, Japan
takuro@nagoya-u.jp

Nobuo Kawaguchi
Nagoya University
Nagoya, Aichi, Japan
kawaguti@nagoya-u.jp

ABSTRACT

Although VR headsets have gained widespread acceptance, their potential for real-world actuation remains underexplored. We propose a system that allows to work in VR space using a robotic arm and VR headset, focusing on 3D scanning as the task. The robotic arm is equipped with an RGB-D camera to achieve remote 3D scanning. This system enables immersive remote work and enhances the link between real space and VR space by reproducing the real objects into the VR space.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; **Ubiquitous and mobile computing systems and tools**; **Virtual reality**; • **Computer systems organization** → *Robotics*; **Robotics**.

KEYWORDS

3D scanning, point cloud, robotic arms, Simultaneous Localization

ACM Reference Format:

Ryuto Usami, Kisho Watanabe, Yuki Gushi, Shuto Tsutsui, Yoshiki Watanabe, Kazuma Kano, Yuya Aikawa, Kaiya Shimura, Nozomi Hayashida, Kenta Urano, Takuro Yonezawa, and Nobuo Kawaguchi. 2024. Demonstration:Remote 3D Scanning with VR and Robotic Arm. In *Companion of the 2024 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Companion '24), October 5–9, 2024, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3675094.3677548

UbiComp Companion '24, October 5–9, 2024, Melbourne, VIC, Australia

ACM ISBN 979-8-4007-1058-2/24/10
https://doi.org/10.1145/3675094.3677548

1 INTRODUCTION

Robotics has become one of the effective applications of advancing VR(Virtual Reality) technology. In previous studies, immersive robot control systems using VR have been shown to provide users with a realistic experience of remote locations and improve the quality of operation [1], [2], [3], [4], [5]. VR robot technology, which involves remote control of robots via VR headsets, allows people to perform various tasks intuitively and remotely. With the increasing capacity of VR robots to manage more complex tasks, there arises the potential to create a new form of employment in VR-based labor from remote locations.

To accomplish complex tasks through VR robot technology, it is essential to align the necessary information between the real space and the VR space and to provide immersive visual experiences in remote locations. Furthermore, for VR robots to become widespread, using commoditized devices is also crucial [6].

In this study, we propose a system for performing tasks in remote locations using a VR headset, stereo cameras, and a robotic arm. This system provides an immersive visual experience through VR180 videos from stereo cameras, enabling more intuitive operations. Additionally, communication is established through WebXR and WebRTC, making it accessible with most existing VR headsets. It offers low latency and easy switching between operators.

In our demo, participants will remotely operate a robotic arm equipped with an RGB-D camera to experience 3D scanning of remote objects. A 3D model of the scanned object is reconstructed in the VR space based on the RGB data and depth data. This technology, which reproduces the object in the real space into the VR space is called digital twin. It has the potential to apply real-world data more deeply and bring additional value[7].

3D reconstruction is expected to be applied in various fields such as digitizing the objects in logistics warehouses and products in e-commerce. Realizing remote 3D scanning with VR and a robotic

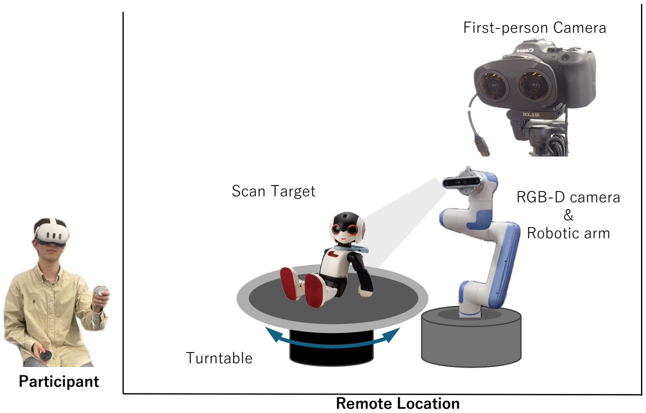

Figure 1: Demo Overview

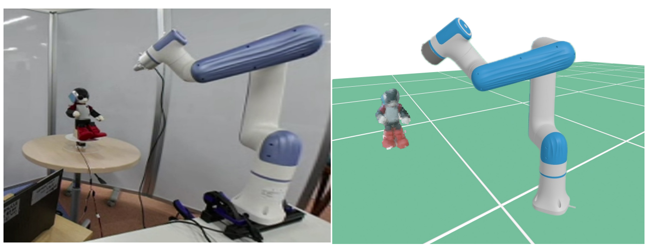

Figure 2: Two Perspectives Available in VR Goggles(Left: Image from First-person Camera, Right: Image of VR Space)

arm enables real-world objects and spaces to be instantly viewed in the VR space.

In studies of 3D reconstruction utilizing point clouds from RGB-D cameras, it is imperative to align point clouds from each frame to construct the model accurately [8]. Performing this task manually can lead to problems such as misalignment or loss of position [9]. Conversely, employing a robotic arm enables precise calculation of end-effector positions and angles from the angles of each axis. This further highlights the usefulness of utilizing a robotic arm for 3D reconstruction.

Our contributions are summarized as follows:

- Propose a system for intuitive robotic hand operation at remote locations using widely available VR headsets.
- Combined a robotic hand and RGB-D camera to enable remote 3D scanning and reconstructing the model to VR space.

2 SYSTEM CONFIGURATION

Figure 1 provides an overview of our demo. In this project, we utilize a robotic arm equipped with an RGB-D camera to scan objects on a turntable from various directions. Using the controllers included with the VR headset, we remotely operate the robotic arm and turntable for object scanning. The position and orientation of the controller are synchronized with the tip of the robotic arm. While observing footage from a first-person camera and the VR space, the operator controls the camera's position and orientation and the turntable's rotation angle.

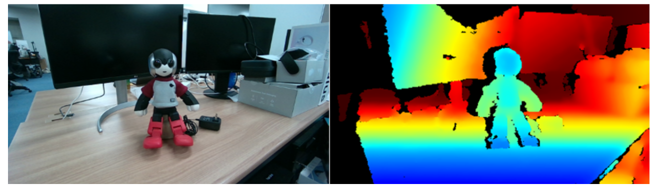

Figure 3: RGB-D Camera Image (Left: RGB Image, Right: Depth Image)

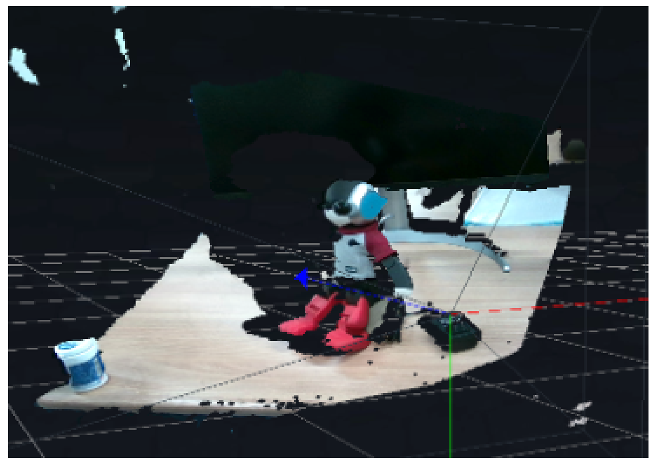

Figure 4: Reconstructed 3D Model

2.1 VR Space

In the VR space, there are two perspectives available, as shown in Figure 2. One shows the first-person camera view and the other displays a 3D model reconstructed from the same viewpoint. After advancing the scan from the first-person camera view, the operator can switch views to display the reconstructed 3D model. This allows the operator to continue scanning while checking insufficiently scanned or low-resolution areas.

In addition to the reconstructed 3D model, the latter perspective also shows a virtual robotic arm, separate from the actual robotic arm. The acquired position information from the controller is used to control the virtual robotic arm. Meanwhile, only the joint angles are transmitted to the actual robotic arm to synchronize the actual robotic arm with the virtual one. This ensures that controller operations are immediately reflected on the virtual robotic arm, providing instant feedback to the operator even from a remote location. Additionally, by first aligning the controller inputs with the virtual robotic arm, potential malfunctions of the actual robotic arm are prevented.

2.2 Robotic arm

We use the DOBOT Nova2 as a 6-axis robotic arm and employ Python API provided by DOBOT to transmit control information. The 6-axis robotic arm allows free movement and orientation of its end-effector in three-dimensional space, enabling the RGB-D camera attached to its end to scan the target object from various

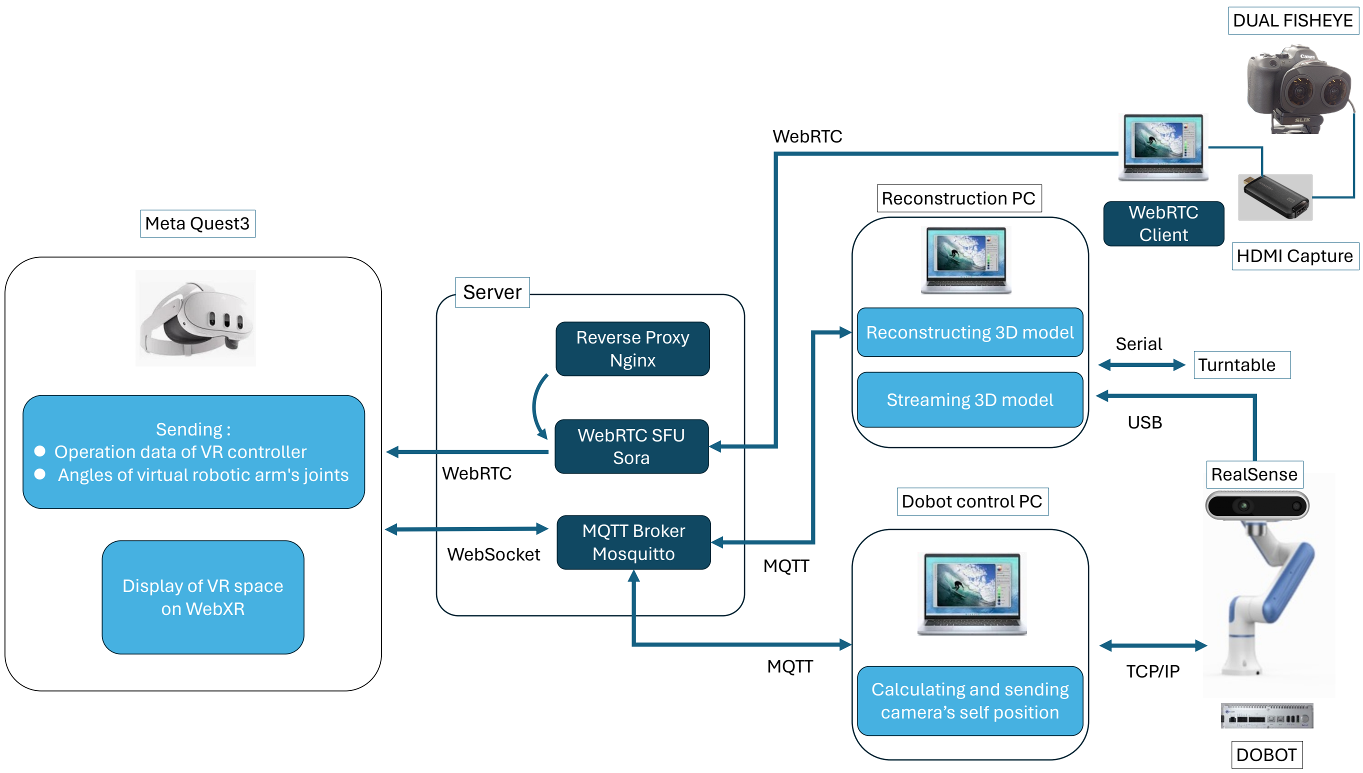

Figure 5: Detailed System Configuration

directions. The position and orientation of the robotic arm's end-effector are calculated using forward kinematics based on the angles of each joint, allowing for the determination of the self-position of the RGB-D camera attached to the end-effector. The combination of depth and color information from the RGB-D camera with its self-position facilitates the three-dimensional reconstruction of the scan results.

When performing scanning and three-dimensional reconstruction using only the RGB-D camera, it is necessary to estimate the camera's self-position. Typically, RGB-D SLAM (Simultaneous Localization and Mapping), a type of SLAM, is used for this purpose. RGB-D SLAM reconstructs the space from the acquired RGB images and depth information while simultaneously calculating the camera's self-position within that space. However, when capturing featureless, flat walls, it becomes challenging to calculate the self-position within the reconstructed space, and the camera may lose track of its self-position.

In this project, the self-position of the camera is calculated based on the angles of the robotic arm's joints, ensuring that the camera does not lose its self-position even when moved rapidly or when capturing featureless images.

2.3 RGB-D Camera

To scan objects, we use the Realsense Depth Camera, an RGB-D camera. The camera captures RGB images and depth images. Figure 3 show the RGB image and depth image, respectively, captured by this camera. Figure 4 illustrates the 3D model reconstructed based on the RGB and depth images.

2.4 Turntable

Since the range of motion of the robotic arm is limited, we prepared a turntable using a servo motor to enable scanning from all directions. The operator rotates the turntable and performs the scan using the VR headset controller.

2.5 First Person Camera

We use a camera with RF5.2mm F2.8 L DUAL FISHEYE lenses as a first person camera. This camera can capture VR180 images. VR180 images allow the operator to control the robotic arm and the turntable while keeping track of the depth of the remote location.

2.6 Communication Protocol

Figure 5 shows an overview of system configuration. In this system, we use WebRTC to transmit stereo camera images. These images are first sent through a WebRTC SFU server and then displayed on VR headsets' browsers using WebXR. The control information from the VR headset's controllers is transmitted to an MQTT broker via WebSocket, converted to MQTT format, and then sent to a PC connected to the robotic arm. With WebRTC, it's possible to have low-latency VR experiences with stereo camera images, allowing multiple users to view them simultaneously. Processing within the

VR headset is completed on the web browser, enabling access to this system with most existing headsets immediately.

3 FUTURE WORK

To achieve intuitive and accurate tasks using a remote robot arm, a smooth control system is essential. Developing and implementing mechanisms that can absorb human-induced tremors when moving the controller is one method to ensure smoother control. Additionally, by implementing feedback mechanisms, such as changing the color of the virtual robotic arm or vibrating the controller when the joint angles of the robotic arm approach their limits, task performance in remote MR environments can be improved.

Furthermore, to achieve higher degrees of flexibility in remote MR space operations, mounting the robotic arm and stereo camera on a mobile robot allows scanning from arbitrary angles without moving the objects. The mobile robot can accurately determine its position, eliminating the risk of losing the RGB-D camera's self-positioning. Incorporating self-positioning into the system removes the need for human intervention on-site. This development opens significant possibilities for tasks performed by robots in remote and hazardous locations where human access is restricted.

4 SUMMARY

This study proposes a system that enables working in a VR space using a VR headset, a stereo camera, and a robotic arm. Furthermore, as an example task using this system, we demonstrated a method for remote 3D scanning with an RGB-D camera on the robotic arm. The intuitive operation is facilitated by feedback from the virtual robotic arm in the VR space and the immersive visual experience provided by the stereo cameras. Additionally, since this system allows access to the remote robotic arm through a web browser, it is compatible with most existing VR headsets, making it applicable to a wide range of fields.

ACKNOWLEDGMENTS

This research was supported in part by JPNP23003, commissioned by the New Energy and Industrial Technology Development Organization (NEDO) in Japan, Council for Science, Technology and Innovation, "Cross-ministerial Strategic Innovation Promotion Program (SIP), Development of foundational technologies and rules for expansion of the virtual economy"(JPJ012495)(funding agency: NEDO), commissioned research (No.22609) by National Institute of Information and Communications Technology (NICT), and JST, CREST Grant Number JPMJCR22M4,

REFERENCES

[1] Patrick Stotko, Stefan Krumpen, Max Schwarz, Christian Lenz, Sven Behnke, Reinhard Klein, and Michael Weinmann. A vr system for immersive teleoperation and live exploration with a mobile robot. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, November 2019.

[2] Chen-Yu Kuo, Chun-Chi Huang, Chih-Hsuan Tsai, Yun-Shuo Shi, and Shana Smith. Development of an immersive slam-based vr system for teleoperation of a mobile manipulator in an unknown environment. *Computers in Industry*, 132:103502, 2021.

[3] Murphy Wonsick and Taskin Padir. A systematic review of virtual reality interfaces for controlling and interacting with robots. *Applied Sciences*, 10(24), 2020.

[4] Sven Behnke, Julie A. Adams, and David Locke. The $10 million ana avatar xprize competition advanced immersive telepresence systems, 2023.

[5] Max Schwarz, Christian Lenz, Raphael Memmesheimer, Bastian Pätzold, Andre Rochow, Michael Schreiber, and Sven Behnke. Robust immersive telepresence and mobile telemanipulation: Nimbro wins ana avatar xprize finals, 2023.

[6] Eric Rosen and Devesh K. Jha. A virtual reality teleoperation interface for industrial robot manipulators, 2023.

[7] Mengnan Liu, Shuiliang Fang, Huiyue Dong, and Cunzhi Xu. Review of digital twin about concepts, technologies, and industrial applications. *Journal of Manufacturing Systems*, 58:346–361, 2021. Digital Twin towards Smart Manufacturing and Industry 4.0.

[8] Michael Zollhöfer, Patrick Stotko, Andreas Görlitz, Christian Theobalt, Matthias Nießner, Reinhard Klein, and Andreas Kolb. State of the art on 3d reconstruction with rgb-d cameras. *Computer Graphics Forum*, 37(2):625–652, 2018.

[9] Yan-Pei Cao, Leif Kobbelt, and Shi-Min Hu. Real-time high-accuracy three-dimensional reconstruction with consumer rgb-d cameras. *ACM Trans. Graph.*, 37, sep 2018.