Safety and Security Challenges for Collaborative Robotics in VR

Soroosh Mortezapoor Khrystyna Vasylevska
Institute of Visual Computing and Human-Centered Technology, TU Wien

Abstract

Virtual reality (VR) security and privacy are not limited to existing software solutions and applications. In this article, we present to the community the challenges of VR systems with robot integration. Integrating robots under ROS poses a massive risk in terms of data security. At the same time, using a robot for simulations in VR requires, first and foremost, the user's safety - hence redundant data collection and sharing. We want to draw the community's attention to these problems through our example in order to ensure that such systems are thoroughly developed in all directions and well prepared for further deployment to the consumer market.

1 Introduction

With the introduction of the consumer virtual reality goggles together with the mobile chips capable of high-quality real-time rendering, the development of the VR applications and the variety of their use-cases flourished. The physically correct rendering and interaction can be seen in almost every application. The developers' ultimate goal is to bring VR to the level of the famous Holodeck from the Star Trek series. This golden sci-fi standard defines the key properties of the ideal VR system:

- free user-driven navigation in the virtual environment (VE) relying on real locomotion,
- multi-user support: for locomotion and interaction with VE and other users in a realistic way,
- full-sensory illusion including but not limited to smell, taste, and touch.

USENIX Symposium on Usable Privacy and Security (SOUPS) 2021.
August 8–10, 2021, Vancouver, B.C., Canada.

Navigation and interaction are belong to the fundamental tasks in VR [1]. The most natural and realistic way to implement locomotion for the user is via actual walking [2]. Walking is also naturally limited with the boundaries of the VR workspace and a number of redirection techniques address this issue. Co-located mobile robots, with their superior precision, speed, and strength, provide a possibility to create hassle-free believable sensory illusions and use space efficiently. During the locomotion, users should be either aware of the possible collisions to be able to avoid those themselves or discreetly redirected to the safe areas using perceptional illusions. For an efficient redirection, the system should reliably predict the next step of the user and consider all possible obstacles ranging from static walls and semi-static objects like chairs to the highly dynamic ones such as other users and robots.

Users in VR are somewhat unpredictable factors with varying speed, movement direction, and limited field of view. In the typical home setup, only the user's head and sometimes hands are tracked. Therefore the VR system is often unsure about the exact body pose of the user.

At the same time, the position of each element of the robotic body is well known at each point in time. Unlike humans, robots are capable of extreme movement speeds and forces relative to one joint to the whole mechanical body. Therefore, the safety considerations for their use are of utmost importance.

Both robots and users can move around at different speeds and interact with other objects within the workspace, i.e. changing the "landscape," for instance, relocating a chair, knock off a small object, or displace a tracking marker. Therefore constant monitoring of the workspace and positions of its contained objects is essential to keep the experience safe.

In this paper, we want to present an outlook on the future of VR and the ongoing conflict between safety and security considerations in advanced VR systems with integrated collaborative robots. We provide a brief overview of existing robotic setups and how they impact safety and security if integrated into a VR system.

2 Robotics for VR

Collaborating robots (cobots) are still at the development stage for both real and virtual world use-cases. Thus it is vital to bring awareness to the VR community and formulate the needs of VR systems with integrated robotics to start a discussion to ensure the full development of the standards concerning all possible use-cases.

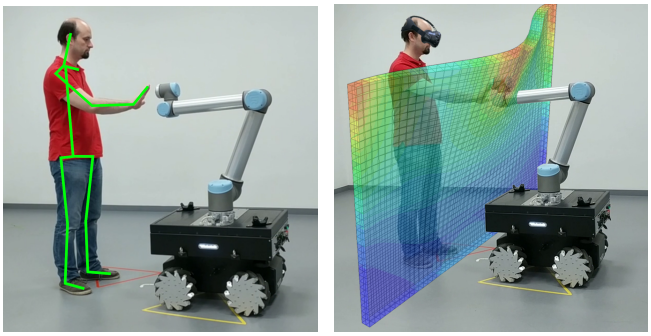
Overall, robots can be split into two groups - stationary mounted and mobile. Stationary robots are typically presented by robotic arm rigidly attached to its supporting base with a limited action space. The mobile robots do not have a fixed position in space and can freely move within the designated workspace. Mobile robots are often equipped with manipulators ranging from simple grippers to fully-fledged robotic arms. They, in turn, significantly increases the size of the action space and consequently the safety requirements for both users and robots.

Cobots might be used by being in direct contact with the user: either mimicking the behavior of a remote user supporting the telepresence [3] with a collocated user or by providing passive or active haptic feedback, like unmovable furniture or dynamic objects such as doors that might be opened. Furthermore, there are cases where the robot is directly attached to the human body like an exoskeleton for support during rehabilitation or providing extensive force feedback in simulation [4]. Alternatively, cobots might be used indirectly - to rearrange the workspace while the user is busy with a task at hand. Naturally, the direct interaction poses higher safety risks for the user.

In the most common scenario, the user always sees the robot with which he interacts, as shown in Figure 1a. However, in VR, a user does not necessarily see the surroundings unless a see-through functionality or 3D representations of real objects are available. This introduces a new possibility of human-robot interaction, such as covert interaction. In this case, the user interacts with the haptic environment without knowing about a robot in the workspace, as shown in Figure 1b. Indeed, not every VR scenario with haptics might allow for a robot visualization without causing the breaks in presence or ruining the experience completely. Relying on a covert approach, requires an even tighter control over space, objects and actors within as it raises the safety risks significantly.

3 VR Cobot Integration

A wide variety of robots, whether stationary or mobile, employ a middleware called ROS [5], which stands for Robot Operating System. Regardless of its name, ROS is not an operating system but a set of software libraries and tools that provide means for robot hardware control and abstraction, an easy yet efficient inter-process message passing system, and libraries ranging from low-level movement controls to high-level planning and execution algorithms. Extensive community contributions resulted in ready-to-use tools for both development and deployment. The critical feature of ROS is that it is not a real-time framework, even though real-time code can be integrated.



(a) (b)

Figure 1: Two use-cases of a cobot: a - interaction in real world, cobot and workspace are fully visible to the user; b - interaction in VR, cobot and workspace are not necessarily visible to the user.

ROS is designed to be lightweight, language-independent, and modular for scalability. One of the severe drawbacks of ROS is that the original design approach does not consider security. Most robotic systems are recommended to be isolated from the Internet. ROS takes a further step and recommends isolating the ROS devices' network, including the robots and their corresponding remote computers, from any other possible actors. The main reason is that once connected to the Internet, or a local network with other devices, not only the data exchange is exposed to these devices, but also malicious actuation commands become possible without any authentications, which might cause unexpected behavior and harm to robot, the objects in its surroundings and more importantly the collocated users.

Although cobots are made for collaboration, their mechanical power and energy still make them dangerous for users. Industrial safety standards such as ISO 15066 [6] address the risks by setting requirements for the robot to move at a very slow pace in the user's proximity. However, in VR, slow movements of the mechanical elements will make the user wait until a specific haptic interaction becomes available in a predefined point of space. That will lead to a break in presence and cause disproportionate prolongation of the VR exposure time, reducing the benefits of an integrated cobot. Interactive VR requires a fast, precise, and safe operation of a cobot in proximity to users, thus will require definition of new standards for both safety and security.

4 Safety vs Security

To ensure safety, the actual workspace should be closely monitored as a whole, not just on the global level of the system

but also locally on each client, to enable timely handling of a dangerous situation and independent disabling mechanisms. Moreover, to handle the possible occlusions by the parts of the environment, a fast and full exchange of information regarding the situation in the workspace is necessary.

Current equipment already offers a number of different tracking solutions that might help with immediate user's safety and state prediction for collocated interaction with a robot. Let us list a few that have been used so far:

- head tracking
- motion capture
  - partial with inverse kinematics (limited number of sensors)
  - full body (professional)
- eye tracking
- bio signals
  - brain activity - electroencephalogram (EEG)
  - muscle activity - electromyograph (EMG)
  - heart rate - for readiness evaluation
- hand tracking
- RGBD sensors
- mechanical tracking with force feedback (hands)

While employing all of these tracking solutions simultaneously is not possible, a combination of a few is quite common and reveal a lot of information. For instance, in our case of the direct interaction for haptics, we plan to utilize the standard sensors for the robot, such as laser scanners, odometry, torque/force sensors, and RGBD cameras for 3D scene reconstruction and perception in the direct proximity of the robot. The user's body will be tracked entirely with a motion capture suit or exclusively with laser-based tracking for hands and head. Consequently, the information should be sent out to ensure the synergy of the systems from both the robot and users. For the user's safety, a lot of information will be shared over the network with the robot to create redundancy for fast, precise, and safe planning of the robot's movements and predicting user's free movements.

In telepresence or teleoperation, the robot is driven by a remote user and might interact with the collocated users. In this case, the information will most likely be shared over the Internet rather than a specialized dedicated network. Unlike others, this use-case suggests a whole new level of security and safety risks even with the use of a VPN.

Apart from the raw sensor data, the custom VR software provided by the manufacturers or a third-party provider might collect the information that, while formally anonymized, might lead to user identification. For instance, Facebook recently published a work showing that user's direction might be predicted based on the head and arms movements [7]. Alternatively, the motion capture data can be identifying the user almost as reliably as a fingerprint [8]. Moreover, analogous to the findings of Facebook, the the data collected by the robot might give away the specifics of the of an object or material it is simulating and location where it is operating.

```
\tf (transformation tree)
    \robot
        \pose
            \vector3
            \quaternion
        \velocity
        \lidar (laser scaner data)...
        \odometry...
        \arm
            \joints (poses of each)...
            \gripper...
            \moveit (motion planning
            package data)
            \newPoses...
            \currentPoses...
            \velocities...
    \user
        \pose...
        \predicted future pose...
        \skeleton...
    \...
```

Listing 1: Topics Structure in ROS. The "..." signify that there is an further underlying data structure.

The hierarchical structure of the ROS data channels called ROS topics exposes all the data available in the system in plain text. By agreement the topic names and the data structure of the published messages in every topic are explicit and visible through out the ROS. Listing 1 shows a basic example how the topics structure looks like. ROS is a peer-to-peer network of processes called nodes that are loosely coupled and may even be distributed over several machines which are connected through the network. The primary communication mechanism is implemented using the Publish-Subscribe pattern. In rare cases, synchronous communication might be used with RPC-based ROS services. The main ROS Master node trusts any other ROS node that connects to it and will reveal all the information upon receiving a few standard requests. Although anonymous publishing of the data to topics is presented as a feature, it makes the safety control and troubleshooting process difficult and poses a huge security threat.

Part of the security issues of the current ROS-based robotic systems, such as plain text communications, unprotected TCP ports, and unencrypted data storage, will be addressed in ROS 2 that was published recently. ROS 2 promises multiple security features completely missing in the existing ROS by utilizing the DDS-Security specification: authentication, ac-

cess control, and encryption. ROS 2 can offer these security features thanks to its underlying DDS (Data Distribution Service). On the other hand, the very fundamental changes of ROS 2 from ROS made it backward-incompatible; the existing community-built libraries of ROS cannot directly be utilized with ROS 2. Although after release of a reasonably stable version of ROS 2, attempts are made toward migrating important ROS libraries to ROS 2, yet due to its rich development resources, the majority of ROS-based robot manufacturers have kept shipping their products with regular ROS on them.

Even when the ROS 2 will replace ROS, the integration packages should also support the secure data transfer. Currently, the communication of ROS with a game engine such as Unity 3D is supported via web sockets with no encryption in place. WebSockets over SSL/TLS are supported by Rosbridge suite, however this optional feature is turned off by default.

Another side of the risk, is the development and experimental processes that are often preceding the final product. ROS 2 relies on the number of plugins that require certificates that for the non-security oriented personnel might pose certain difficulties in setup and is likely to be skipped until the product is finished. This notion is supported by the findings of DeMarinis et al. [9]. In 2018, they scanned the IPv4 finding over a hundred of publicly-accessible hosts that were running a ROS master node with a default TCP port. With permissions of the owners, DeMarinis et al. demonstrated the simplicity of a takeover of the robots operated with ROS.

Proprietary software like Steam or Oculus often requires an internet connection, which goes wholly against the security recommendations for ROS. Thus even if the ROS PC is not accessing the Internet directly, there is a possibility to exploit its weaknesses via a connected VR client. By scanning for the default ROS master TCP port 11311, Rosbridge TCP port 9090, Unity integration TCP port 10000 and so on, attacker can gain access to the whole data tree accessing the camera data, laser scan point clouds, full robot description and even control the robot by publishing instructions to the moveit topics anonymously by design of ROS.

Ultimately, the VR systems employing robotics consider the safety requirements but often struggle to meet the security standards due to the reasons mentioned above.

5 Summary

The ongoing pandemic showed how remote work and automatic production lines might decrease the risks for the business. Robots and VR offer an ideal combination for remote work while keeping the human still involved in the process. Therefore, the use-cases described above might make it to the market in the nearest future. This paper aimed to extend the future outlook from the basic VR setup to a more extensive system with an integrated robot that can drastically impact a VR system's overall safety, security and privacy levels.

The approach of extensive data sharing adopted in robotics and implemented in ROS goes against the current security point of view, where it is crucial to minimize the data shared and transfer it as securely and reliably as possible. The current crisis stimulates the new surge of innovation and optimization, speeding up the adoption of virtual workplaces, remote task performance, and automation. Therefore the trade-offs between the system flexibility and safety on one hand and security and privacy on the other should be carefully examined by the community.

References

[1] Josef J. LaViola Jr. et al. *3D User Interfaces: Theory and Practice*. Addison-Wesley, 2017. ISBN: 978-0-13-403432-4.

[2] Frank Steinicke et al. *Human Walking in Virtual Environments*. Springer, 2013. ISBN: 978-1-4419-8431-9.

[3] Jonathan Steuer. "Defining Virtual Reality: Dimensions Determining Telepresence". In: *Journal of Communication* 42.4 (1992), pp. 73–93. ISSN: 00219916. DOI: 10.1111/j.1460-2466.1992.tb00812.x. URL: http://doi.wiley.com/10.1111/j.1460-2466.1992.tb00812.x.

[4] Christopher N. Schabowsky et al. "Development and pilot testing of HEXORR: Hand EXOskeleton Rehabilitation Robot". en. In: *Journal of NeuroEngineering and Rehabilitation* 7.1 (July 2010), p. 36. ISSN: 1743-0003. DOI: 10.1186/1743-0003-7-36. (Visited on 05/30/2021).

[5] *ROS/Introduction - ROS Wiki*. URL: http://wiki.ros.org/ROS/Introduction (visited on 05/30/2021).

[6] *ISO/TS 15066:2016(en), Robots and robotic devices — Collaborative robots*. URL: https://www.iso.org/obp/ui/#iso:std:iso:ts:15066:ed-1:v1:en (visited on 05/30/2021).

[7] Kara J. Emery et al. "Estimating Gaze From Head and Hand Pose and Scene Images for Open-Ended Exploration in VR Environments". In: *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. Mar. 2021, pp. 554–555. DOI: 10.1109/VRW52623.2021.00159.

[8] George Williams et al. "Body Motion Analysis for Multi-modal Identity Verification". In: *2010 20th International Conference on Pattern Recognition*. ISSN: 1051-4651. Aug. 2010, pp. 2198–2201. DOI: 10.1109/ICPR.2010.538.

[9] Nicholas DeMarinis et al. "Scanning the Internet for ROS: A View of Security in Robotics Research". In: *arXiv:1808.03322 [cs]* (July 2018). arXiv: 1808.03322. URL: http://arxiv.org/abs/1808.03322 (visited on 06/22/2021).