THE DESIGN AND DEVELOPMENT OF A VIRTUAL REALITY ENVIRONMENT FOR THE OPERATOR OF A TELEOPERATED ROBOT IN A TEST SETUP

J.K. (Joost) Buursink

BSC ASSIGNMENT

Committee:
dr. ir. D. Dresscher
C. van der Walt
dr. ing. G. Englebienne

July, 2024

018RaM2024
Robotics and Mechatronics
EEMCS
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Contents

Abstract

In the context of human work, there is a need for telerobotics. Telerobotics allow one person to rapidly help multiple people from one location. One of the issues encountered in telerobotics is time delay which reduces the sense of embodiment and therefore control over the robotic avatar. So, models need to be created to allow for mediation of the effects of time delay. In this research, a virtual environment is created to mediate this effect of time delay observed and experienced by the operator of the avatar. Using ROS we can find the states of the real setup that the environment approximates, such that the virtual environment behaves appropriately. In the design of the virtual environment, it is important to consider the sense of embodiment. This is done by taking into account the correct morphology of the human body. We find that these senses of embodiment are complex to quantify and as such, they are both measured implicitly and explicitly. During the evaluation, the environment was tested using a whack-a-mole game and we found that the average scores were 21.8 [%] higher when using the virtual environment, displaying its effectiveness. We also find an average embodiment score of 0.66 [-] on the 7-point Likert scale showing that embodiment is experienced. We find that overall the virtual environment provides better precision and accuracy over the movement of the robot arm. Further improvements for the virtual environment could be to improve the input control scheme, make the rest of the avatar move according to the rest of the operator's body and provide a more human-like arm and avatar.

1 Introduction

In this day and age, we are physically getting further and further apart. At the same time, we want to remain as close to what we desire as possible. The solution for this is the telemanipulation. The PACOF (Predictive Avatar Control and Feedback) project of the RaM research group at the University of Twente is creating a telemanipulation set-up. In the PACOF project, a construction is being designed for the teleoperation of a robotic avatar, with the idea of controlling that robotic avatar from a remote location. There is a (wireless) connection between the operator and the robotic avatar. A big challenge in this project is that this connection generates time delays. While classical methods succeed in providing a solution for small time delays, big time delays are a challenge. Model-mediated teleoperation addresses this by letting the operator interact with a model of the environment, which is updated by measurements of the robot. For the integration experiment of the PACOF project, a mock-up pipeline is needed for the visual information. Normally the visual pipeline should show a direct camera view of the surroundings of the robot arm. However, we want to visually take into account the time delays eventually using this visual pipeline. For now, this visual pipeline should only directly display the set-up and therefore is only a mock-up. This mock-up pipeline will be made in virtual reality (VR). The biggest benefit of using VR is that it provides the operator with both a sense of embodiment, the feeling that the robot is (an extension of) the operator, and telepresence, the feeling that they are in a remote location. This sense of embodiment and telepresence provides greater control over the robot, because, after all, it is easier to move your own body right where you are than it is to move someone else's body.

1.1 PACOF

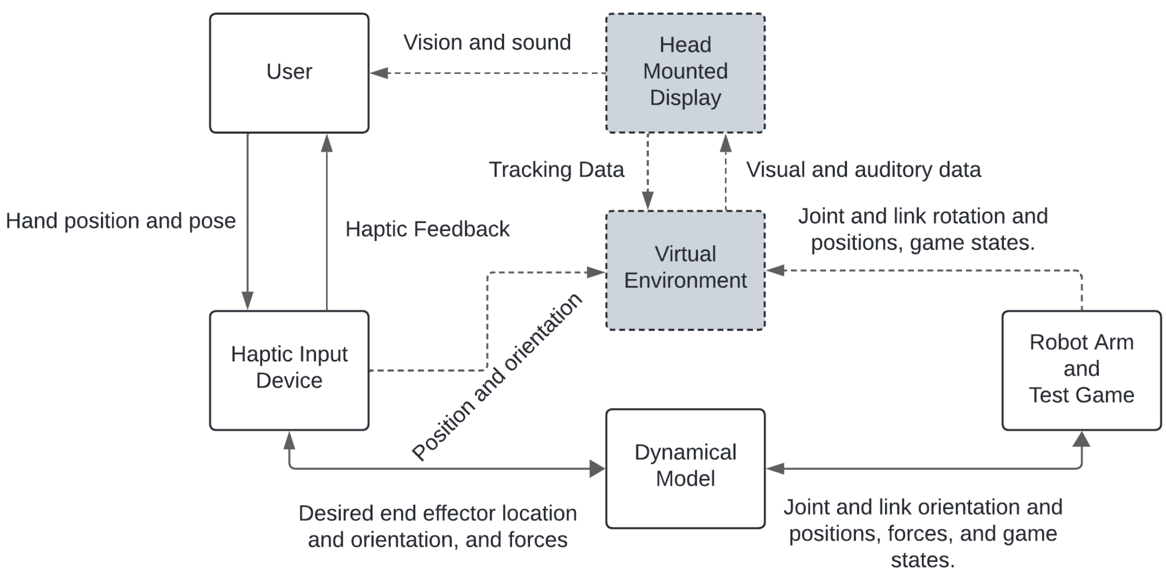

Figure 1: High-level overview of the PACOF project with data paths. Greyed out is the area that still needs to be created, of which it will be explained how that will be done in this report.

The current system of the PACOF project looks like the system seen in fig. 1, greyed out in the image is what is described in this report. The full system contains the following subsystems:

- The head-mounted display (HMD): The HMD is a display that sends visual information from the virtual environment to the user and sends information about the orientation and position of the user's head to the virtual environment.

- Haptic input device (HID): The HID is a grabbable object that the user can move to control the robot arm through the dynamical model.
- Dynamical model: The dynamical model converts HID inputs to movements in the robot arm and it converts feedback received by the robot arm to feedback in the HID.
- Robot arm: The robot arm is what needs to be controlled and what will be used to play the test game.
- Test game: The test game is a game of whack-a-mole that serves as a performance benchmark.
- The virtual environment (VE): The VE is the mock-up visual pipeline that creates the visual information based on the set-up that needs to be sent to the HMD.
- The user: The operator of the set-up.

This research will be concerned with the creation of the visual pipeline in the PACOF project. As such, it will be interfacing with the robot arm, the test game, the HID, and the user.

1.2 The set-up

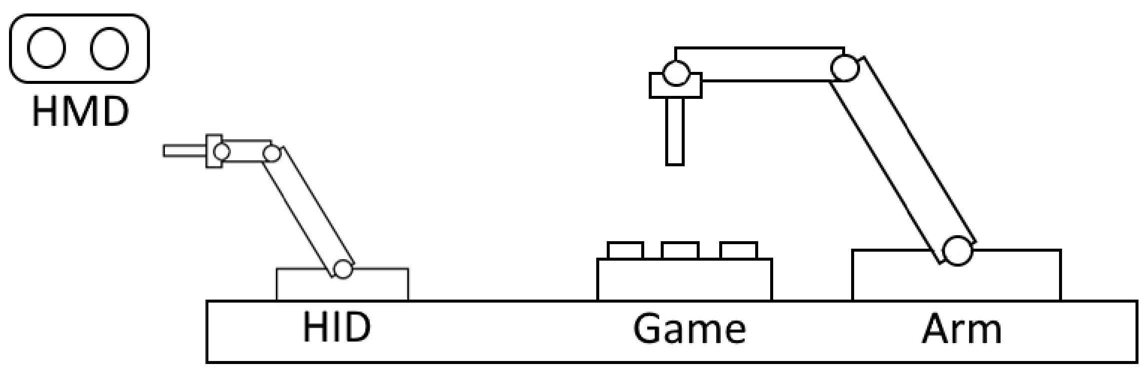

Figure 2: Diagram of the test set-up. Displayed are the head-mounted display (HMD), the haptic input device (HID) Force Dimension Omega 7, the whack-a-mole game, and the robot arm FRANKA EMIKA Panda.

Currently, the PACOF project is still in development and as such not all components are ready for deployment and usage yet. Therefore, a mock-up visual pipeline will be built for, and tested on a test set-up. This test set-up can be seen in fig. 2 The test set-up for the visual pipeline consists of 4 major components. These are:

- The robot arm. This is a FRANKA EMIKA Panda.
- A wooden box with 6 buttons programmed to play a whack-a-mole game.
- A head-mounted display. This is the HTC Vive.
- A haptic control device. For the haptic control device, we can either use the Force Dimension Omega 7 or the HTC Vive controllers.

The operator will move the HID which, through the previously mentioned dynamical model, will move the robot arm. This is moved in such a way that it can interact with and play the game of whack-a-mole. When the virtual environment is developed, it will be updated by the states of the game and the arm. The operator will be able to see a model of this test set-up through the HMD.

1.3 Problem statement

In the PACOF project, there is a need for a visual pipeline to display a live feed of the test set-up seen in section 1.2. This is such that the operator will be able to control the robot arm without necessarily having sight of the robot arm and the test game, something that is crucial in teleoperation. This is because currently there is no way for the operator to see what is happening at the robot arm without the use of cameras. Cameras can not see into the future, so they suffer a lot from this issue of time delay. For cameras, it is then really difficult to create a predictive model, which is what we eventually desire to be able to do. But when we use a virtual reality environment we can use the predicted states of the future instead of the current states. In a final version, these predicted states can be placed on a sensor-based scan of the test environment. However, due to time constraints and implementation complexity, a full implementation based on visual sensors is unfeasible right now. So, the virtual environment will be hand-crafted in Unity using 3D models and will be a mock-up.

Finally, this visual pipeline will need to be designed to maximize feelings of embodiment and telepresence. The sense of embodiment and telepresence, further discussed in section 2 is a feeling that improves the ability to control and perceive as the avatar of a virtual environment. This ability to generate sensations of embodiment and telepresence will remain as the visual pipeline is improved from a mock-up version to a full version.

1.4 Goal

The goal of this bachelor assignment is to create the visual pipeline of the model-mediated teleoperation set-up of the PACOF project in virtual reality for the operator of the robot arm. This visual pipeline will consist of a virtual environment which will be a digitalized copy of the test set-up. We know the exact test set-up, so it will be rebuilt from scratch in Unity. The virtual set-up will contain a box with a game of whack-a-mole. The virtual environment will be animated based on measurements from the test set-up, such as the configuration of the robot and the state of the game. This virtual environment will be designed with the sense of embodiment and telepresence of the operator in mind. Therefore, considerations will need to be made to optimize these senses, such as reducing overall latency and providing a virtual body with high correspondence in spatial configuration.

1.5 Research questions

In order to properly perform this assignment, the following research questions have been formulated:

1. How do we measure and quantify telepresence and the sense of embodiment, and how can we use that to evaluate the virtual environment?
2. What are relevant considerations in the design of a virtual environment?
3. How can the connection between Unity and the test set-up be established to modify the virtual environment based on real-world events?

2 Background

This section will discuss what sense of embodiment and telepresence are and their benefits.

In [1], the sense of embodiment is described as a feeling that occurs when a different body's sensations are felt as coming from a person's own biological body. This sense can be split into three different sub-senses: sense of self-location (SoS), sense of agency (SoA), and sense of body ownership (SoO). These three sub-senses will be briefly discussed.

The sense of self-location can be described as the feeling of being inside the avatar [1]. This does not necessarily refer to the feeling of being at the same location as the avatar, but rather being connected with the (robotic) body of the avatar. While similar, these ideas are not necessarily the same. This feeling of being there at the same location would be called the sense of presence. In the context of teleoperated robots, this sense of presence is specifically referred to as telepresence.

The sense of agency is the feeling of being in control of one's own body [1]. In the context of telepresence, this is the sense of being in control of the avatar. Sense of agency consists of the feeling of being able to have motor control over the avatar and feeling an active will to move the avatar.

The sense of body ownership is described as the feeling that the avatar creating stimuli feels like the biological body creating those stimuli [1]. It describes the feeling that the body you are experiencing is yours.

In [2], research was done where participants were given different sensory cues while in a virtual environment and their memory for objects in that environment was evaluated. The researchers found that with tactile, auditory, and olfactory cues, memory of objects in the virtual environment was significantly improved. These cues also improved the sense of presence, showing the important relation between the sense of presence and memory. As mentioned in section 2.2, higher agreeance between the avatar's movements and the operator's movements improves the sense of agency. This concept also works the other way around: a higher sense of agency allows a person to better control the avatar. And, with better control over the avatar, the operator has an easier time completing their tasks. In [3], it is found that a heightened sense of embodiment improves the ability to create positive social experiences with robot avatars. This is a crucial result, as the end result of the PACOF project is that the robots can be used in social situations. Therefore, it is really important to improve the sense of embodiment as much as possible.

3 Analysis

In this section of the report, answers will be given to all the research questions given in the introduction based on a literary analysis.

3.1 How Do We Measure and Quantify Telepresence and the Sense of Embodiment, and How Can We Use that to Evaluate the Virtual Environment

Telepresence and the sense of embodiment are complex feelings. It can often be difficult to disentangle feelings like SoE and telepresence, and it can be difficult to properly test them. There is not one physiological quantity that we can simply measure to get information on, for example, how embodied a person may be feeling. On the other hand, we can not simply ask a person how embodied they feel, because each person's answer to that question will be different. This is because each person has a different complex and diverse background, which makes their answer to that question subjective. To answer this question of how to measure and quantify telepresence and SoE, we will take a closer look at implicit and explicit assessment and try to find the best way to approach this problem for the current research.

3.1.1 Implicit Assessment

Research, such as seen in [4, 5, 6] has shown that certain physiological qualities of the human body can be related to an increase in telepresence or SoE. This is because, while feeling more connected to an avatar, it feels like events that happen to the avatar are instead happening to the user. When things happen to the user, they may feel more involved and thus feel their feelings more intensely. More intense feelings invoke more intense physiological responses. Examples of these physiological responses are responses such as a change in skin conductance due to nervous sweating or an increase in heart rate because of fear. These physiological responses can be used to evaluate the level of telepresence and SoE. An example of this would be the research done in [4], where skin conductance and heart rate are used to determine the importance of latency in SoE. In this research, a stressful situation was used such that skin conductance and heart rate were more relevant. Another example would be [5] where functional magnetic resonance imaging is used to measure brain activity to assess the sense of ownership over a prosthetic hand. Another example would be [6], where bodily temperatures are used to measure SoE. While physiological measurements provide very objective results, it can be difficult to differentiate the different feelings that comprise SoE from these physiological responses alone. This is especially the case for physiological responses such as body temperature, skin conductance, and heart rate as they can describe multiple different feelings besides what we are looking for.

3.1.2 Explicit Assessment

The alternative to this implicit assessment through physiological responses would be explicit assessment. The SoE and telepresence are constructs and by the fundamentals of psychological testing [7], constructs can always be measured. Specifically, this refers to measuring explicitly. By measuring explicitly, we can remove some of the unknowns regarding what we are exactly measuring. This, however, comes at the cost of introducing subjectivity. We can assess telepresence and SoE explicitly through the use of methods such as questionnaires and surveys. For SoE specifically, the use of questionnaires is very common, as is found in [8]. Performing a survey allows for freedom and adaptability during the creation of the questions. The questions can be tuned exactly to what is required. One thing that is important to note is that even questionnaires can be made more objective. This can be done by having multiple test subjects answer the same questions and answer in a way that allows a comparable outcome. Such as in [9], where the participants answer in a seven-step scale. Then, an average can be taken to minimise the subjectivity of the results.

3.1.3 A Toolbox to Measure SoE

In [8], a generalized toolbox is provided to assess the SoE in teleoperation and VR environments, in which these implicit and explicit assessment techniques are discussed. In this literary research it is found that even though there are definite correlations between SoE and physiological qualities of the human body, most physiological quantities would be too difficult to disentangle into telepresence and SoE. However, more explicit measures provide concrete results for the psychological feeling in question. In this research, it is more relevant to get these concrete results, so we will measure SoE and telepresence through use of a questionnaire. This questionnaire will contain questions answerable using a 7-point Likert scale, where -3 would correspond with no feeling of SoE or telepresence at all and +3 would correspond with the feeling that the user is the avatar.

We will both implicitly test and explicitly test the VE. For implicit testing, we can use the performance of the participant playing the game in the VE and compare it to the performance of the participant playing the game without the use of a VE. We can do this to determine how well the VE works in general, and thus how well SoE and telepresence are applied. However, it does not give insight into how well the different subsenses of SoE are experienced such as SoA, SoO, and SoS. Neither does it differentiate between SoE and telepresence. This will need to be done using the explicit test.

The explicit test is in the form of a questionnaire. In literature, there is already a plethora of questionnaires used to evaluate these senses. In [10] the writers provide a standardized questionnaire for embodiment. One of the benefits of this questionnaire is that it is at least partially applicable for most if not all cases of embodiment. It can be adapted to the exact case of embodiment that we have and will be done so. Answers to the questionnaire are given in a seven-step scale and using this seven-step scale, the feelings of embodiment can be quantized as well.

During the evaluation procedure, the virtual environment should provide at least an accurate copy of the end effector of the robot arm and the whack-a-mole game whose locations are updated live. The operator should be able to see both the game and the end effector at the same time such that they can play the game as well as possible. After this, the participants should be asked to fill in a questionnaire about their experiences in the virtual environment.

3.2 What Are Relevant Considerations in the Design of a Virtual Environment?

In this section, some relevant considerations will be given that will help in designing the virtual environment.

3.2.1 Reduced Control and Instruction Complexity

By moving the control interface to a VE, we allow a significant reduction in the complexity of movement. An example of this can be seen in a VE made for forestry machines in [11]. Normally, the operator of a forestry machine needs to operate each individual joint. However, due to the introduction of the VE, the operator can simply instruct the system where the claw of the crane should go and the VE will show a path of how to get there and move the claw to the instructed position. We see here that this significantly reduces instruction complexity. We can also take this into account in the design of our VE, as the use of a VE allows us to use different control methods that would not be possible without a VE. We can for example move the end effector of a robot arm directly by clutching it in the virtual environment, using VR controllers or VR hand tracking, and moving it directly instead of moving the haptic input device. The use of a VE allows us to use control methods that are more intuitive and less complex to use [11, 12]. In the design of the VE, this opportunity should be taken into account by introducing less complex control methods.

3.2.2 Modeling of the Operator

Proprioception is the concept of knowing the configuration and status of one's limbs without necessarily seeing, hearing, or feeling them. Proprioception is a very relevant topic when it comes to embodiment as it is heavily dependent on the SoO. With a reduction in SoO, a person's proprioception decreases. And with decreased proprioception the control over one's limbs also decreases. This is because, without proprioception, a task has to be done fully by other senses, without internal feedback from the body.

The research done in [13] has found that in the planning of movement, both visual information and proprioception are crucial. Proprioception is a feeling that comes with embodiment. So, to take this into account, we have to improve the feeling of embodiment in the VE. However, [13] also states that in certain cases, vision might supersede proprioception and vice versa. To avoid conflicting information between visual and proprioceptive senses, one can use a morphologically accurate body, meaning a body that structurally matches a human body, in the VE [1]. This way, what is seen will match what is sensed internally through proprioception. The research done in [14] also shows that having robot hands that match the movements of the operator improved the sense of body ownership.

There are multiple options for the visualisation of the operator in the VE. The hands can be modelled without any arms and the hands can be modelled with arms connected to them. Each of these solutions has its benefits. Using only a hand is the simplest solution, however, the SoO is reduced because of the mismatch in morphology. Using motion capture is a very accurate method to provide information about the configuration of the arms and because of that it greatly improves the SoO, and thus SoE. However, with this method, we need external trackers to track the arms of the operator. In the final option, where we use inverse kinematics, we do not need external trackers. Research done in [15] even shows that using inverse kinematics might be more beneficial to the SoE than using a motion capture suit. This research also provides an open-source solution which can be used in Unity. In the design of the VE, a choice needs to be made regarding how to approach the modelling of the operator.

3.2.3 Auditory and Vibratory Feedback

The next consideration is the feedback the operator receives in VR. This can be in the form of either auditory feedback, for example through the speakers of the head-mounted display, or in the form of tactile feedback with, for example, vibrations from the controllers. In the full setup that uses the Force Dimension Omega 7, a haptic device that allows control of the robot arm in 3D space, we can also use the controller to provide resistance to the movements, dependent on the implementation of the control scheme. If the control scheme uses the Force Dimension Omega 7 there is force feedback built into the device. However, if we use the virtual reality controllers to move the robot arm we can use the vibratory motors of the controllers. Either way, it is very important to use tactile feedback in the VE, because, as stated in [16], coherent haptics improve the SoA. An example of how this would practically be achieved is the use of vibrations in the VR controllers when an action is performed correctly. The human mind is then smart enough to connect the dots and realize that the action is performed correctly just by the vibration. We mentioned sound previously as a method to provide the operator feedback in the VE. The work done in [17] suggests that by providing auditory cues, we can improve the sense of presence. Providing audio cues can be achieved by using the built-in speakers of the head-mounted display. These audio cues can be used to provide the user with knowledge of the area surrounding the robot, this will improve telepresence as the observed environment of the operator will better match that of the test setup. Besides this, we can use audio cues to improve the SoA. This can be done by generating audio when a task is done successfully. The vibration and sound design of the VE needs to carefully be considered as a method to improve the SoA and feelings of telepresence.

3.2.4 Latency and Visual Fidelity

There are some additional considerations to be made in the design of the virtual environment, such as the effects of latency and visual fidelity. In [4], research is done on the effects of low latency in stressful environments. During the research, it was found that latency is a big factor in telepresence. Lower latency would increase telepresence, and higher latency would decrease telepresence. So, for a virtual environment that provides a high feeling of telepresence, latency should be reduced as much as possible. In [18], it was inspected what the impact of visual fidelity of the avatar to telepresence and the ability to perform certain tasks in a virtual environment is. This was done by comparing a low visual fidelity avatar in a virtual world, a generic self-avatar in the real world, and a visually faithful hybrid. It was found that while handling real objects made the task easier, the visual fidelity of the avatar's body played no significant role in the process. A similar research was done in [19] where the impact of pictorial realism on telepresence was researched. In this experiment, it was found that pictorial realism played no significant role in telepresence

3.2.5 Conclusion

There are a lot of relevant considerations in the design of a VE as it is a very broad concept that allows for a lot of freedom in the design process. At the same time, there are definitely some considerations that will strictly improve certain aspects and senses during the use of the VE. Such as the ability to use less complex control schemes to reduce complexity, or the use of artificial limbs in the design to improve SoO. We can also use auditory and vibrational cues to provide feedback and improve the SoA and telepresence during operation. There are some other considerations such as the visual fidelity of the environment which is not as impactful as, say, the latency of the environment. Still, these considerations are all relevant to some degree and will be taken into account for the design of the VE.

3.3 How Can the Connection Between Unity and the Test Setup Be Established to Modify the Virtual Environment Based on Real World Events?

The biggest practical question regarding this research is: "How can we best obtain and display information about the robot arm?" The robot arm is controlled using a dynamical controller that runs on a separate Linux computer. This computer makes all the calculations required to determine the rotations the robot arm joints should have for the end effector to reach a certain position and rotation. This is done through a process called inverse kinematics. When these rotations have been calculated, they are told to the robot and then the robot arm moves into position and allows the end effector to reach the requested destination. During this process, the robot arm sends data about current positions rotations and any sensor information such as forces back to the computer such that the computer gains feedback. This feedback is then used to properly and more accurately control the robot arm.

This communication between the robot arm and the computer is done through a Robotic Operating Software (ROS) network. ROS is a set of tools and software libraries that can be used for all sorts of applications in the field of robotics [20]. In this case, ROS is used to create a network that allows for communication through what is known as ROS topics. A ROS topic is a node that allows other computers to publish and read messages to and from. Messages then contain the necessary information relevant to that topic. In our case, we have two relevant topics that we would like to read information from. These topics are:

- `/franka_state_controller/joint_states`
- `ledstate`

The topic named `/franka_state_controller/joint_states` provides us with the states of the joints in local rotations and the topic named `ledstate` provides us with the states of the LEDs in the buttons of the whack-a-mole game. These ROS topics get updated with live data of the states of the test setup.

These topics need to be read in Unity in order to display such information. Unity is run on a separate computer so a network connection needs to be established between these computers. The ROS TCP Endpoint package in [21] provides an endpoint that can be run on the ROS network for the computer that runs Unity to connect with. Since Unity does not support connecting to ROS networks by default, [21] also provide the ROS TCP Connector package to connect to this endpoint from the Unity environment. The ROS TCP Connector package also provides a way to generate the necessary translation files to interact with the topics and messages published to the ROS network. These packages also allow us to publish to the ROS network from Unity if necessary.

When we have the live data of the state of the whack-a-mole game in Unity we need to apply it to the relevant objects to create an accurate visual representation of the test setup. For the robot arm we can use the URDF importer found in [21] and the URDF file found in [22] to generate a game object in Unity that has the same shape and configuration as the robot arm and can be controlled using the messages obtained through ROS. This way we can determine the position and rotation of the end effector with respect to the base of the robot arm, or we can display the full robot arm. The whack-a-mole game does not have a pre-existing model, so an approximation for this game will have to be made.

As for the virtual reality aspect of things. The Unity environment provides packages such as the XRI Development kit to allow for development using virtual reality. This package also allows for input using the controllers and for the camera to be moved using the virtual reality glasses.

3.4 Three Virtual Environments

We use the considerations discussed in **??** to create three different theoretical environments. A lot of these considerations come with specific design choices and can only be implemented while other considerations are disregarded. Therefore we designed each of these virtual environments with a focus on one of the three senses of embodiment.

3.4.1 Virtual Environment 1: Focus on Body Ownership

The first virtual environment design choice will contain a virtual body for the operator to embody with. This virtual body will be seated on a chair. The base of the robot arm will be connected to where the shoulder of the avatar would be. It will replace the avatar's right arm. The VE is designed like this to improve the sense of body ownership as there is a clear body to embody with that both matches the user and the test setup.

3.4.2 Virtual Environment 2: Focus on Agency

The second virtual environment design choice will contain a grabbable ball at the position of the end effector that the robot arm will try to follow as the operator moves it. This would change the control method to truly take advantage of virtual reality. This will likely significantly improve the sense of agency at the cost of the sense of body ownership and self-location. This environment could contain a more traditional VR avatar to improve the sense of body ownership.

3.4.3 Virtual Environment 3: Focus on Self Location

The final VE design choice is similar to the first VE design choice in that it contains a seated avatar. However, in this design, the end effector of the robot arm is visualised using a human hand. This hand is then connected to the avatar using arms whose states are calculated using inverse kinematics. The robot arm itself is then made fully invisible. The design idea here is that it will improve the sense of self-location by improving the connection with the body by having it be visually accurate to the operator's body.

4 Design

In this section of the report, an overview will be given of what we consider to be the most important parts of the final design and from this a minimum design will be created. Finally, an environment will be chosen from section 3.4 and that VE will be designed.

4.1 Key Components of the Design

In the design of the VE, there are multiple minimum requirements for the design to be considered operable. The most important parts of this design are the representation of the robot arm and the whack-a-mole game. The total requirements for the design are given in the following list.

The minimum design should contain:

- A connection with ROS to allow for reading of topics.
- A visual representation of at least the end effector of the robot arm.
- A visual representation of at least the buttons of the whack-a-mole game.
- A method to guarantee both the end effector and whack-a-mole game have their relative positions in the VE the same as in the real world.
- A camera that is moved by the operator using a pair of VR glasses.
- A method to move the position of the camera in case of a mismatch between the real head location of the user and the head location in the VE.

4.2 The Minimum Design

These minimum requirements were used to create a base VE which can be used as a starting point to develop specialized VEs as is considered in section 3.4.

4.2.1 Connection with ROS

Arguably the most important part of the entire virtual environment is the connection to the physical setup, as without it would be impossible for Unity to know what is going on. So, a ROSConnection object Unity is created with the use of the ROS TCP Connector package seen in [21]. In this ROSConnection game object the IP of the endpoint is set such that the Unity environment can obtain data from the ROS network situated at that IP address. This endpoint is run specifically on the ROS network to create connections to Unity services. The package used for this endpoint is the ROS TCP Endpoint package found in [21]. This connection will allow communication through topics with the Franka Emika Panda and the whack-a-mole game.

4.2.2 Implementation of the robot arm

As mentioned before, we can use the ROS Connection to obtain data from the robot arm, but we also need a place for that data to be used. For this, we need a model of the robot arm. To get this model we can take the URDF file from [22] and import that into Unity using the URDF importer from [21]. There were some inconsistencies, but after changing the URDF file these were all removed for proper importing. The changed URDF file that imports correctly to Unity can be found in [23]. The URDF importer automatically applies all sorts of control scripts and colliders to the model. The arm, however, is merely a visual representation. So, these control scripts and colliders are not necessary for the final VE and can therefore be removed from the game object leaving only the visual models.

To correctly parse the information from the `franka_state_controller/joint_states` topic on the ROS network into the VE, a C# class is generated using the ROS TCP Connector package from [21]. This C# class retrieves the serialized data generated by the ROS TCP endpoint located on the ROS network. Then it translates it to usable variables.

The `ArmSubscriber.cs` script was written which updates the local rotations of the joints of the virtual robot arm according to the data obtained using the previously mentioned C# class. The local angles are updated rather than the total angles as this is what the joint state message provides. This also allows us to use simple x, y, and z coordinate systems rather than using quaternions.

4.2.3 Implementation of the whack-a-mole game

For the implementation of the whack-a-mole game, there was not a preset model that can be used. So, the game object in the VE is made from scratch using cylinder and cube 3D models. This physical whack-a-mole game is measured and these measurements were taken and used for the virtual game.

A C# class is generated to retrieve the serialized data from the ROS network based on the `ledstate` topic for this whack-a-mole game model. The `ledstate` topic provides a binary state for each button in the whack-a-mole game representing whether that button's light is turned on. The `GameSubscriber.cs` script uses the contents of these messages to update the material of the button in the VE to show a lighter colour when a button is lit up on the physical test game.

Next, we need a way to place the whack-a-mole game correctly in the VE. The `GamePositionCalibratior.cs` script is created to allow for this placement. This calibration script works as follows. In the real setup, the tip of the end effector should be placed on the top left button of the game and then the calibrate button (C on the keyboard) should be pressed. This moves the game such that the tip of the end effector is also pressing the top left button in the VE. Then the tip of the end effector should be moved to the bottom right button of the game. Once again, the calibrate button should be pressed. This second step then rotates the game into the right position.

4.2.4 Implementation of VR

The VE also contains a camera that can be moved through the movement of a VR headset. This is done by using Unity's XRI toolbox. This toolbox provides camera movement using a VR headset. The most important part of camera control is the camera's position relative to the virtual environment. So, to put the camera in the right relative place, the `RecenterOrigin.cs` can be used. When the camera calibration button (space bar on the keyboard) is pressed the camera is moved towards a camera origin. This camera origin can be placed however the designer would like, such as at the head position of the avatar. This can be done by pressing the camera calibration button whenever the operator is fully seated in the physical world to make sure the operator looks through the eyes of the avatar.

4.2.5 Summary

The VE now contains a robot arm that moves according to the physical robot arm. The VE also contains a representation of the whack-a-mole game with changing buttons. This virtual game can be placed in the same position relative to the virtual arm as the physical game to the physical arm. Finally, the VE contains a camera that can be looked through using virtual reality glasses. This camera can be centred if necessary. The result of this virtual environment can be seen in fig. 3

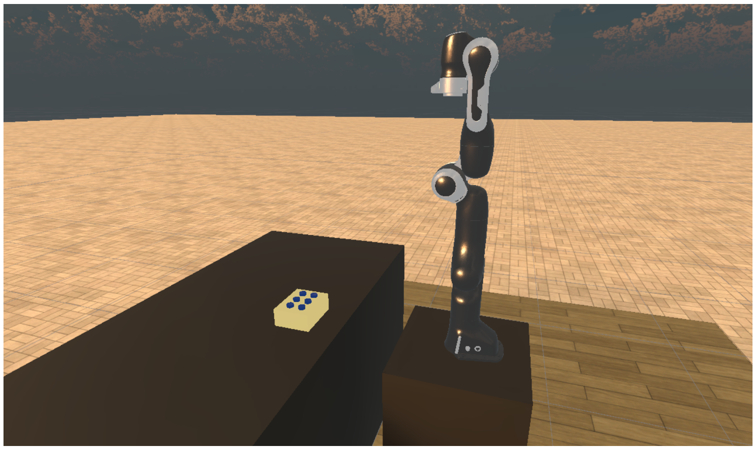

Figure 3: The intermediate VE.

4.3 The Chosen Virtual Environment and Its Extra Improvements

From the options discussed in section 3.4, the chosen VE is environment 1. Taking into account the limited remaining time, this would be the most feasible to implement while still developing a VE that improves the sense of embodiment by focusing on the sense of body ownership.

The environment from section 4.1 is further developed by adding a blue crash test dummy avatar from the Unity Asset Store to the scene. This avatar is posed such that it looks like it is sitting down. This avatar is seated on a chair that is also from the Unity Asset Store. I also made the robot arm in this environment blue to better match the colour of the avatar. The robot arm was placed by placing the base of the robot arm where the shoulder of the avatar would be. Finally, the robot arm is made semi-transparent to improve the vision over the whack-a-mole game.

The final result of the VE can be seen in figs. 4 to 6. The corresponding project can be found in the GitHub repository on [23]. The scripts that were written can be found in the Assets/Scripts folder.

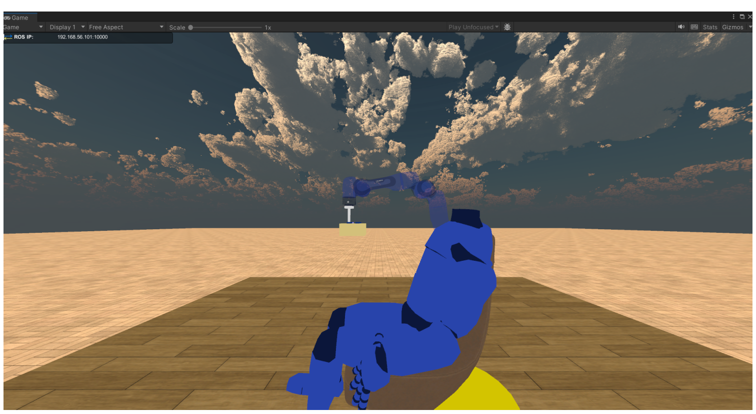

Figure 4: Side view of the final VE.

Figure 5: Top view of the final VE.

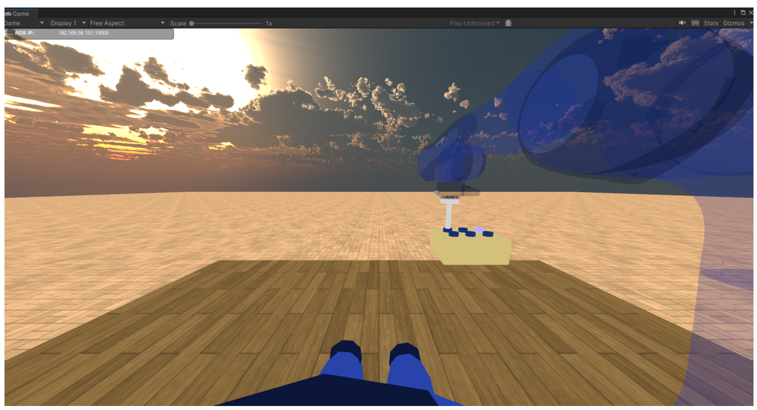

Figure 6: Player view of the final VE.

5 Experiments

This section describes the setup and procedure of the evaluation used to test the virtual environment. It will be discussed how ethical considerations are taken into account, and how the evaluation procedure has taken place.

5.1 Ethics

A test protocol, information letter, and consent form have been written for the evaluation procedure. The test protocol describes how the evaluation took place and serves as a guideline for the researcher during the evaluation. The information letter and consent form serve as a way to ask for informed consent from the participants. All of these documents can be seen in appendix B and have been approved by the CIS Ethics Committee of the University of Twente.

5.1.1 Experiment design

In the experiment, the participants will be asked to play a game of whack-a-mole in two different scenarios. In this game of whack-a-mole, there are six buttons in a 2x3 pattern which light up randomly. To score a point, the player should hit a button when it lights up, and before it goes out again. In both scenarios, the participant is asked to get the highest score they can within two minutes. In the first scenario, the participant controls the robot arm through the haptic input device while having vision of the physical whack-a-mole game and arm. The participant is first given five minutes to become familiar with controlling the robot arm. After these five minutes, the participant is ready to play the two-minute whack-a-mole game. In the second scenario, the participant is in the VE. Again, the participant is given five minutes to become comfortable with the controls of the robot arm and the haptic input device. Then the participant is asked to play the two-minute whack-a-mole game from within the virtual environment. The order of these scenarios is switched for each participant to avoid the bias of experience. In both scenarios, the robot arm and game are placed in front of the participant.

5.2 The Evaluation Procedure

The evaluation procedure serves as a way to analyse the VE both on the ability to generate the feeling of embodiment and improve performance during the whack-a-mole gameplay. The goal of using these two methods is to analyse the VE both implicitly and explicitly. This is done in an attempt to reduce the downsides of the explicit measurement, namely subjectivity, and to reduce the downsides of implicit measurement, namely lack of concrete information. The evaluation procedure therefore consists of two different sections: An implicit measurement and an explicit measurement

5.2.1 Questionnaire

After the experiment, the participants will be given a questionnaire about the experiences of the participant during the usage of the virtual environment. In this questionnaire, 16 questions are asked to classify the feelings of embodiment of the participant. These questions are taken directly or adapted from the standardized questionnaire provided in [10]. The answers to these questions will be combined into individual scores in the following categories: "Ownership, Agency, Location, Appearance". These four scores can be combined into a total score for embodiment. As appearance is less relevant for the purposes of this research, we also provide an embodiment score that does not take appearance into account. These scores are obtained according to the following equations where Q# refers to the question numbers from the questionnaire.

$$Ownership = \frac{Q1 - Q2 - Q3}{3} \tag{1}$$

$$Agency = \frac{Q4 + Q5 + Q6 - Q7}{4} \tag{2}$$

$$Location = \frac{Q8 - Q9 + Q10}{3} \tag{3}$$

$$Appearance = \frac{Q11 + Q12 + Q13 + Q14 + Q15 + Q16}{6} \tag{4}$$

$$Embodiment = \frac{2 \cdot Ownership + 2 \cdot Agency + 2 \cdot Location + Appearance}{7} \tag{5}$$

5.3 Expected Outcomes

With the described experiment and the virtual environment designed in section 4, I expect a certain performance. To start, I expect that the environment will make it easier to play the whack-a-mole game, so the scores when using the virtual environment will be higher. Second, due to the semi-transparent arm which allows for better vision, I expect an improved sense of agency. Finally, I expect that this virtual environment will score well on body ownership given the addition of the avatar.

6 Results

This section of the report describes the results of the evaluation procedure described in section 5.

6.1 Whack-a-mole results

The evaluation procedure described in section 5 has been followed and for the whack-a-mole section of the evaluation this results in the data seen in table 1. This table shows the scores the participants obtained playing whack-a-mole for two minutes per score. All odd-numbered participants started using the No-VR setup and used the VE afterwards. All even-numbered participants started by using the VE and used the No-VR setup afterwards.

	Person 1	Person 2	Person 3	Person 4	Person 5	Person 6	Mean Value
No-VR	7	29	50	48	40	19	32.1667
VR	24	43	52	49	49	18	39.1667

Table 1: Results of the whack-a-mole test with mean values.

Three interesting things can be noticed from these results. Before we look at those things, however, it should be stated that it is important to look at these results incredulously to avoid the fallacy of composition. These results might not be representative as there were only six participants. Nevertheless, they do provide further insight and reason to continue the research.

The most interesting result is that the VR set-up's mean value is higher by 21.8 [%] than when the VE was not used. We also observe this in the individual results where five out of six participants scored higher in the VE compared to not using the VE. These two observations in the results align with the hypothesis that using the VE will improve the ability to control the robot arm. However, it should also be mentioned that there is not a significant difference in the results for the two scenarios for three out of six participants.

6.2 Questionnaire

The second section of the evaluation procedure has also been performed with the participants providing the results seen in appendix C. The results of the equations discussed in section 5 and therefore the combined scores can be found in table 2.

	Person 1	Person 2	Person 3	Person 4	Person 5	Person 6	Mean Value
Ownership	-1	1.6667	1	0.66667	1.6667	2.3333	1.0556
Agency	2.25	0	1	1.25	1.5	1	1.1667
Location	1.3333	-0.66667	0	0.33333	1.3333	0.66667	0.5
Appearance	-0.83333	0.83333	-0.83333	-1.1667	-1.3333	-1.6667	-0.83333
Embodiment	0.61905	0.40476	0.45238	0.47619	1.0952	0.90476	0.65873
Appearance-free SoE	0.86111	0.33333	0.66667	0.75	1.5	1.3333	0.90741

Table 2: Results of the embodiment questionnaire with mean values.

These results score the participants on the different senses of embodiment. Where the senses of ownership, agency, and location are most relevant for this research. The scores are given on a 7-point Likert scale where -3 is the lowest score and 3 is the highest score. Again, these results should be taken with a grain of salt as they are based on a relatively small sample set of six participants.

From these results, we find that the experiences of embodiment differ significantly per person. However, even though the experiences of embodiment significantly differ per person we also find that every person experiences embodiment to some degree in the VE. The sense of embodiment that is on average experienced most strongly is the sense of agency. And the sense that is experienced weakest is that of appearance. Taking into account appearance in the overall embodiment score changes the result significantly.

7 Discussion

In this section of the report the findings in section 6 will be discussed and considerations for further research will be given.

7.1 On the whack-a-mole results

In section 6, we found that the average value for the whack-a-mole game was 21.8 [%] higher when using the VE compared to when the VE was not used. While this is a very significant result it says surprisingly little about the VE except for its overall effectiveness. With this result, given that we can extrapolate to more participants, we can state that the VE is effectively doing what it is designed to do: Allowing for easier performance of the whack-a-mole game. When we do not consider embodiment yet there are a couple of reasons for this. The whack-a-mole game is easier to play in the VE because the perspective from the VE is closer to the actual game compared to the distance between the haptic input device and the physical arm and whack-a-mole game. There is also the fact that the robot arm is semi-transparent except for the end effector in the virtual environment, making it easier to observe the end effector and its position in relation to the whack-a-mole game. As mentioned in section 3, implicit testing makes it difficult to say anything about the individual senses of embodiment. So with these results, we can not say much about the degree to which embodiment plays a role in the effectiveness of the VE.

7.2 On embodiment results

However, we also evaluated the experiences of embodiment during the use of the VE. These experiences of embodiment as mentioned in section 6 are evaluated using a questionnaire. During the questionnaire, the participants were asked to give their reasoning for their answers such as to provide insight into the experienced feelings of embodiment.

7.2.1 Ownership

Ownership was a feeling of embodiment that was experienced positively with a score of 1.1 [-] largely due to the direct link between the movement of the participant's arm and the movement of the robot arm in the VE. Not in all experiences this link is perfect, however.

Some participants describe that due to the manner of input, it feels as if the participant was controlling the arm of someone else or a mouse. This was experienced differently for different people because other participants mentioned that it still feels like your own arm. However, their arm is harder to move and control. One participant mentioned not being embodied with the robot arm at all. They felt like it was truly a robot arm that they were moving and not their own human arm. On the contrary, another participant mentioned feeling their muscles contract while in the virtual environment and therefore experienced their mind actively creating a connection between their arm and the robot arm. The same participant also mentioned experiencing the robot arm as an extension of their arm as opposed to experiencing the robot arm as an individual arm.

All participants mentioned that due to the manner of input, it is difficult to embody with the arm fully. The input is done completely by the connection between the haptic input device and the robot arm. The virtual environment is not designed with the ability to change this. This has caused several problems when trying to generate a sense of embodiment. First of all, it is relevant to mention that the input space of the haptic input device is both spherical and significantly smaller than the output space that the robot arm would desirably be able to reach. This limited space caused two main issues. There would be incorrect feedback given from the haptic input device. Feedback would be given as the edge of the input device is reached, while the edge of the output space has not been reached yet. And, the participant would have to let go of the input device and grab it again to cover more space in the output space if certain buttons were not reachable due to the limited input space. There is theoretically a very small window of space in the center of the output space, which if used as a starting position allows the participant to reach the entire output space without having to re-grab the robot arm. However, this was difficult to find during play.

Another reason for the difficulty in embodiment is the delay between the input device and the robot arm. While there is hardly any delay between the physical robot arm and its virtual model, it takes time for the dynamical model of the robot arm to catch up with the input given to it by the haptic input device. An interesting observation is how this slight delay affects how the participants interact with the robot arm in a way that displays a sense of ownership. The participants take into account the delay as if it's part of their movement. Over time the participants naturally started adjusting for that delay.

7.2.2 Agency

The sense of agency is the feeling of embodiment which is experienced the strongest out of all senses of embodiment with a score of 1.2 [-]. The participants all feel like they can control the robot arm to some degree.

One of the reasons for this is that the movements of the input device, when the button is pressed, directly match up with the robot arm. However, due to the need to re-grab the robot arm and therefore the occasional lack of response, as multiple participants described, the feeling is created that the robot has a slight mind of its own. This occasional loss of control makes it more difficult to fully control the robot arm. Those participants also mentioned it did not significantly impact their ability to experience a sense of agency over the robot arm and the overall feeling of agency is still experienced. However, it should be noted that it did affect the sense of agency and should be considered when looking at the results for the sense of agency.

Some participants mention that the robot arm can be moved more smoothly through the virtual environment as opposed to the stuttery movement experienced when the virtual environment is not used. Where this stuttery movement is mostly experienced due to a loosely fitted button that doesn't always work as intended and therefore "lets go" of the robot arm. We think this is experienced less in virtual reality because, as one of the participants mentioned, there is less thought necessary in the virtual environment. This is something we can only attribute to the sense of embodiment.

Another thing that supports the virtual environment in creating the sense of agency is that the connection in the virtual environment is a lot shorter. This is largely due to the change in perspective. In the virtual environment, the whack-a-mole game and the end effector of the robot are much closer to the face of the operator as the operator can phase through the robot arm. Another reason for this shorter connection would be the fact that in the VE, the robot arm is made to be transparent. This transparency allows the user to see through the robot arm and makes it so it doesn't obstruct the operator's view of the game and end effector.

7.2.3 Self-Location

From the important senses of embodiment, Self-location was experienced the weakest with an average score of 0.5 [-]. First, we will discuss the positive generation of the sense of location. Five out of six participants mentioned feeling an improved connection to the virtual environment due to the addition of the virtual body. This is because the avatar was located where the participant was. For most participants, it also overlapped with their posture. These participants mentioned that it improved their sense of self-location within the position of the avatar. One of the participants also mentioned that their "movements" overlapped as they felt they weren't moving and neither was the avatar. So we can state that the addition of an avatar in the same pose as the participant improves their sense of self-location. However, not every participant has the same length, not every participant has the same width and not every participant has the same posture while sitting. So, some mentioned these points negatively affecting their experiences of embodiment.

Another interesting observation is that two out of six participants mentioned being too focused on the game to notice the avatar. There are a couple of ways to interpret this and neither of those ways are conclusive. However, they are still relevant to mention. Often it is stated that good design is invisible. So being too focused on a game would in this case mean that the participants are embodied to such a degree that they do not realise how embodied they are. More likely, however, the participants were simply too distracted by the whack-a-mole game and did not pay attention to the virtual body they were attached to. As there is only a visual connection to this body it is easy to ignore.

With regard to embodiment with the avatar, multiple participants mentioned feeling less embodied with the virtual robot arm than the rest of the virtual avatar. This is because the virtual robot arm did not overlap with a human arm in a morphological sense. The virtual arm has only rotational joints whereas the human arm has much more complex joints such as the shoulder that can do more than just rotate. While these are mimicked by using more joints they are not morphologically accurate.

7.2.4 Appearance

While appearance is not the main concern of this research it is interesting to see how it relates to total embodiment. In the results of the questionnaire, we found that appearance was the lowest rated and only negative feeling of embodiment. This is because of multiple things.

- The participant is not a crash test dummy.
- The participant is not blue.
- The participant does not have a consistent size.

While they also mentioned this mattered less as the overall pose still matches theirs, it is relevant to mention that participants have mentioned feeling less embodied with the overall setup due to the reduced embodiment caused by the blue avatar.

8 Conclusion

In this report, a virtual environment was designed and developed to control a robotic arm from virtual reality. One of the reasons this is done is that it bridges the physical gap in distance between workstations and workplaces, especially in humanitarian work. So, a virtual environment is developed that serves as a workplace copy. In the case of this research, the virtual environment is a copy of a test setup containing a robot arm and a whack-a-mole game. This virtual environment is also the visual link between the (haptic) input device and the robot arm (or output device).

The virtual environment is designed in such a way as to attempt to generate positive senses of embodiment to improve control over the robot arm in the test setup. This is done by looking at what creates embodiment such as the addition of a human avatar in the virtual environment and the connection it makes with the virtual representation of the robot arm. Or, for example, the morphological connection between the operator and the virtual body. The hypothesis was that these considerations would positively influence the senses of embodiment and thus allow for better control over the robot arm in the physical world.

The virtual environment was then developed with these, and more, considerations in mind. In the final design, I created an environment where the user is represented by a blue crash test dummy sitting on a chair. The arm of the user is replaced by a semi-transparent version of the robot arm, allowing for better sight over the virtual whack-a-mole game and better sight of the game's relative position to the robot arm.

The designed virtual environment was eventually tested and evaluated by letting six different participants play a game of whack-a-mole using the robot arm in two different scenarios. In one scenario the participant is not in virtual reality and has direct sight of the robot arm, the game, and the haptic input device. In the second scenario, the participant was in virtual reality and could not see the haptic input device. Afterwards, the participants were asked to fill in a questionnaire regarding their senses of embodiment during the virtual experience.

Using this experiment I found that the average experience of embodiment was 0.91 [-] on a seven-point Likert scale. I also found that the scores of the whack-a-mole game were improved by 21.8 [%] during the use of the virtual environment.

The participants described multiple reasons explaining how embodied they felt inside the virtual environment. Such as needing to re-grab the robot arm during play reducing the overall feeling of embodiment or the reduced distance between the participant and the play area in the virtual environment improving the sense of agency by providing a clearer view and thus more control. From this, multiple future improvements were considered to improve the virtual environment for future use such as improving the input space such that it better aligns with the output space. Another improvement that is considered would be the addition of a scan of the participant as a replacement for the avatar in the virtual environment.

All in all, it can be said that, while the virtual environment still has a long way to go in terms of providing a proper sense of embodiment, the virtual environment does in fact provide some sense of embodiment. And that improved sense of embodiment improves the ability to play the whack-a-mole game and thereby improves control over the robot arm.

8.1 Recommendations

Given the findings from section 7, there are a lot of things that can be done to improve the virtual environment.

On the topic of body ownership, we can at least do the following things to improve the VE: We can scan and track the body of the operator to create an avatar that is virtually identical to that of the user. This will improve the sense of ownership as it is already your body. Any morphological inaccuracies will then be significantly reduced based on how good the scanning methods are. Having a human body that matches the operator's real body will also improve the appearance score which may contribute positively to the total embodiment experience. At the same time, there will not be an inaccurate pose to move towards. The pose of the avatar will already match that of the operator. Additionally, it may be relevant to replace the robot arm with a human arm connected to the end effector. This would likely improve the sense of self-location. However, it would need to be researched how much impact this would make exactly.

It would also be beneficial for the sense of ownership to change the input system for the test setup. How this would be done is to make the movement of the robot arm more sensitive to the movement of the haptic input device. By doing this we can mitigate the issues that make it difficult to embody with the robot arm such as the incongruent workspace and the need to re-grab the end effector to properly reach every point in the output space. This solution will also improve the sense of agency as it will make it easier to control the robot arm.

One more change that can be made to this once the robot arm can reach every position in the output space is that we can remove the need to press the loosely fitted button to move the robot arm. This will reduce the feeling that the user is holding someone else's arm. One participant mentioned that the rotation they applied to the haptic input device was not applied to the rotation of the robot arm. This participant specifically mentioned that they felt like there were fewer connections between their arm and the real arm because of that. While the task is still performable, the robot arm is less manoeuvrable than a normal arm. So this is also something that can be changed to improve the sense of agency.

References

[1] Konstantina Kilteni, Raphaela Groten, and Mel Slater. “The Sense of Embodiment in Virtual Reality”. In: *Presence Teleoperators &amp Virtual Environments* 21 (Nov. 2012). DOI: 10.1162/PRES_a_00124.

[2] H.Q. Dinh et al. “Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments”. In: *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*. 1999, pp. 222–228. DOI: 10.1109/VR.1999.756955.

[3] David Farizon, Peter Ford Dominey, and Jocelyne Ventre-Dominey. “Social Attitude Towards a Robot is Promoted by Motor-Induced Embodiment Independently of Spatial Perspective”. In: *IEEE Robotics and Automation Letters* 7.4 (2022), pp. 9036–9042. DOI: 10.1109/LRA.2022.3189150.

[4] M. Meehan et al. “Effect of latency on presence in stressful virtual environments”. In: *IEEE Virtual Reality, 2003. Proceedings.* 2003, pp. 141–148. DOI: 10.1109/VR.2003.1191132.

[5] H. Henrik Ehrsson, Charles Spence, and Richard E. Passingham. “That’s My Hand! Activity in Premotor Cortex Reflects Feeling of Ownership of a Limb”. In: *Science* 305.5685 (Aug. 2004), pp. 875–877. ISSN: 1095-9203. DOI: 10.1126/science.1097011. URL: http://dx.doi.org/10.1126/science.1097011.

[6] Jakob Hohwy and Bryan Paton. “Explaining Away the Body: Experiences of Supernaturally Caused Touch and Touch on Non-Hand Objects within the Rubber Hand Illusion”. In: *PLoS ONE* 5.2 (Feb. 2010). Ed. by Antoni Rodriguez-Fornells, e9416. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0009416. URL: http://dx.doi.org/10.1371/journal.pone.0009416.

[7] Sandra A. MacIntire and Leslie A. Miller. *Foundations of psychological testing*. McGraw-Hill, 2000.

[8] Jan Van Erp Sara Falcone Gwenn Englebienne and Dirk Heylen. “Toward Standard Guidelines to Design the Sense of Embodiment in Teleoperation Applications: A Review and Toolbox”. In: *Human–Computer Interaction* 38.5-6 (2023), pp. 322–351. DOI: 10.1080/07370024.2022.2039147. eprint: https://doi.org/10.1080/07370024.2022.2039147. URL: https://doi.org/10.1080/07370024.2022.2039147.

[9] Jan Van Erp Sara Falcone Gwenn Englebienne and Dirk Heylen. “Toward Standard Guidelines to Design the Sense of Embodiment in Teleoperation Applications: A Review and Toolbox”. In: *Human–Computer Interaction* 38.5-6 (2023), pp. 322–351. DOI: 10.1080/07370024.2022.2039147. eprint: https://doi.org/10.1080/07370024.2022.2039147. URL: https://doi.org/10.1080/07370024.2022.2039147.

[10] Mar Gonzalez-Franco and Tabitha C. Peck. “Avatar Embodiment. Towards a Standardized Questionnaire”. In: *Frontiers in Robotics and AI* 5 (June 2018). ISSN: 2296-9144. DOI: 10.3389/frobt.2018.00074. URL: http://dx.doi.org/10.3389/frobt.2018.00074.

[11] Simon Westerberg and Anton Shiriaev. “Virtual Environment-Based Teleoperation of Forestry Machines: Designing Future Interaction Methods”. In: *Journal of Human-Robot Interaction* 2 (Sept. 2013). DOI: 10.5898/JHRI.2.3.Westerberg.

[12] Theodoros Togias et al. “Virtual reality environment for industrial robot control and path design”. In: *Procedia CIRP* 100 (2021). 31st CIRP Design Conference 2021 (CIRP Design 2021), pp. 133–138. ISSN: 2212-8271. DOI: https://doi.org/10.1016/j.procir.2021.05.021. URL: https://www.sciencedirect.com/science/article/pii/S2212827121004807.

[13] Fabrice R. Sarlegna and Robert L. Sainburg. “The Roles of Vision and Proprioception in the Planning of Reaching Movements”. In: *Progress in Motor Control*. Springer US, 2009, pp. 317–335. ISBN: 9780387770642. DOI: 10.1007/978-0-387-77064-2_16. URL: http://dx.doi.org/10.1007/978-0-387-77064-2_16.

[14] Maryam Alimardani, S. Nishio, and Hiroshi Ishiguro. “Humanlike robot hands controlled by brain activity arouse illusion of ownership in operators”. In: *Scientific Reports* 3.1 (Aug. 2013). DOI: 10.1038/srep02396. URL: https://rdcu.be/dD1Wc.

[15] Mathias Parger et al. “Human upper-body inverse kinematics for increased embodiment in consumer-grade virtual reality”. In: Nov. 2018, pp. 1–10. DOI: 10.1145/3281505.3281529.

[16] Camille Jeunet et al. ""Do You Feel in Control?": Towards Novel Approaches to Characterise, Manipulate and Measure the Sense of Agency in Virtual Environments". In: *IEEE Transactions on Visualization and Computer Graphics* 24.4 (Apr. 2018), pp. 1486–1495. ISSN: 1077-2626. DOI: 10.1109/tvcg.2018.2794598. URL: http://dx.doi.org/10.1109/TVCG.2018.2794598.

[17] Huong Dinh et al. "Evaluating the Importance of Multi-sensory Input on Memory and the Sense of Presence in Virtual Environments." In: Apr. 1999, pp. 222–228. DOI: 10.1109/VR.1999.756955.

[18] B. Lok et al. "Effects of handling real objects and avatar fidelity on cognitive task performance in virtual environments". In: *IEEE Virtual Reality, 2003. Proceedings.* 2003, pp. 125–132. DOI: 10.1109/VR.2003.1191130.

[19] Robert Welch et al. "The Effects of Pictorial Realism, Delay of Visual Feedback, and Observer Interactivity on the Subjective Sense of Presence". In: *Presence* 5 (June 1996), pp. 263–273. DOI: 10.1162/pres.1996.5.3.263.

[20] *ROS: Home.* URL: https://www.ros.org/.

[21] Unity-Technologies. *GitHub - Unity-Technologies/Unity-Robotics-Hub: Central repository for tools, tutorials, resources, and documentation for robotics simulation in Unity.* URL: https://github.com/Unity-Technologies/Unity-Robotics-Hub.

[22] StanfordASL. *PandaRobot.jl/deps/Panda/panda.urdf at master · StanfordASL/PandaRobot.jl.* URL: https://github.com/StanfordASL/PandaRobot.jl/blob/master/deps/Panda/panda.urdf.

[23] TrueFocus. *GitHub - TrueFocus/VR-Environment: Contains the VR environment for the PACOF project.* URL: https://github.com/TrueFocus/VR-Environment.

A AI Statement

During the preparation of this work, I used LanguageTool and a free Grammarly trial to correct spelling mistakes and grammar. After using this tool/service, I thoroughly reviewed and edited the content as needed, taking full responsibility for the final outcome.

B Documents for Evaluation

The following pages describe the relevant ethical documents used before and during the evaluation process of the virtual environment. First, the test protocol will be shown, then the information letter and finally the consent form. We also show the questionnaire and the corresponding answer form used to examine the experiences of embodiment of the participants.

Test Protocol

Author: Joost Buursink (s2790882)
Bachelor Assignment: The Design and Development of a Virtual Reality Environment for the Operator of a Teleoperated Robot in a Test Setup.
Date: May 17, 2024
Contents: Contains the testing protocol for the VE. The testing protocol contains an introduction to the assignment, the objective of the test, and the design and method of how the testing will take place
Status: Draft

1 Background for the Assignment

In this day and age, we are physically getting further and further apart. At the same time, we want to remain as close to what we desire as possible. The solution for this is the PACOF project. In the PACOF project, a construction is being designed for teleoperation of a robotic avatar, the idea of controlling that robotic avatar from a remote location. There is a (wireless) connection between the operator and the robotic avatar. A big challenge in this project is that this connection generates time delays. While classical methods succeed in providing a solution for small time delays, big time delays are a challenge. Model mediated teleoperation addresses this by letting the operator interact with a model of the environment, which is updated by measurements of the robot. For the integration experiment of the PACOF project, a mock-up pipeline is needed for the visual information. This mock-up pipeline will be made in virtual reality (VR). The biggest benefit of using VR is that it provides the operator with both a sense of embodiment, the feeling that they are (an extension of) the robot, and telepresence, the feeling that they are in the remote location. This sense of embodiment and telepresence provides greater control over the robot, because, after all, it is easier to move your own body right where you are than it is to move someone else's body.

2 The Bachelor Assignment

The goal of this bachelor assignment is to create an environment in virtual reality for the controller of the teleoperated robot. This virtual environment (VE) will be a digitalized copy of the test setup. The test setup will contain movable objects whose location is sent to the headset to properly synchronize the VE with the real-world situation. This VE will be designed with the sense of embodiment and telepresence of the operator in mind. Therefore, considerations will need to be made to optimize these senses, such as reducing overall latency and providing a virtual body with high correspondence in spatial configuration.

3 The Test

At the time of testing there will be a VE designed in Unity that connects the test subject with the avatar in the test setup through a head-mounted display and a set of VR controllers. There will be tests done with 8 participants of approximately an hour each where the goal is to provide insight into how well the VE is designed in terms of how well it creates the seperate senses of embodiment. This will be done in the form of scenarios. Each participant will perform the same action three different times. They will play a game of whack-a-mole, but in different scenarios. Once without a virtual environment, once with a virtual environment where the robot arm can be moved by moving the end effector of the arm in VR, and one scenario where the robot arm is the participants "arm". At the start of each set of scenarios a number of things will be done, which should take approximately 10-20 minutes.

1. The participant will be asked to read the information letter and sign the consent form.
2. An inventory will be made about the participant's previous experiences with virtual reality.

3. A short discussion will be held about the different senses of embodiment and telepresence such that any unclear elements are removed.
4. The participant will be briefed on the set-up of the experiment and will be given a short overview of the different scenarios.
5. The participant is briefed on how to control the robot in -and outside of VR.

After this introduction, the scenarios will take place, each taking approximately 10-20 minutes. While the tasks to be done are different, each scenario will consist of the same three steps.

1. Before putting on the VR equipment the participant is briefed on the task that needs to be done.
2. After that, the participant is asked to describe freely what they are doing and experiencing while doing the following things in order:
   (a) The participant is asked to put on the VR glasses and grab the controllers (not for the scenario where no VR interface is used)
   (b) The participant is asked to play the game of whack-a-mole. The time in which the participant performs the task is noted down.
   (c) The participant is asked to remove the equipment.
3. After taking off the VR equipment, the participant is interviewed and asked to answer a couple of questions regarding the feelings of embodiment and telepresence during the scenario. Some of which can be ranked on a seven-step scale, and others will be more open questions.

4 Data Collection and Usage

During the test, I aim to collect three sets of data. I will make notes on what the participants are experiencing during testing. The time in which the participant performs the requested tasks is recorded. And, the responses to the questions asked during the interview sections of the scenarios are recorded.

- The responses given while performing the requested tasks will be noted down if they are insightful. Quotes may be taken for improved insight and/or use in further writings.
- The time in which the participant performs the tasks described in the scenario is saved and anonymized to: Person x Task y Time/Score z [s].
- Voice recordings will be made during the interview, which will be transcribed and then destroyed. Once again, quotes may be taken and used for improved insight and/or use in further writings.

All data such as quotes and times will be anonymized and may be used in the report and/or final presentation.

Information Letter

Author: Joost Buursink (s2790882)
Date: 15/05/2024

1 Purpose of the research

Thank you for your support in the research of the PACOF project. In this project, a construction is being designed for the teleoperation of a robotic avatar from a remote location. There is a (wireless) connection between the operator and the robotic avatar. A big challenge in this project is that this connection generates time delays. While classical methods succeed in providing a solution for small time delays, big time delays are a challenge. Model mediated teleoperation addresses this by letting the operator interact with a model of the environment, which is updated by measurements of the robot. For the integration experiment of the PACOF project, a mock-up pipeline is needed for the visual information. This mock-up pipeline is made in virtual reality (VR) and is called a virtual environment (VE). This VE allows the participant to control the robotic arm. This VE needs to be tested and evaluated on the sense of embodiment and telepresence. This is done using the upcoming user-test.

2 The evaluation procedure

The test will ask the participant to play a game of whack-a-mole. Once without a VE, and once with a VE. The performance of the participant during this game is tracked in order to obtain an idea of whether the VE improves the ability to control robotic arms. After this test, a questionnaire will be held asking the participant about their sense of embodiment and telepresence during the test.

3 Benefits and risks of participation

The benefit of participating is that it will provide insight into the senses of embodiment and telepresence. It will also allow further development in the PACOF project. There are some risks attached to participation in this evaluation process. These risks are as follows:

- If the participant has any form of injury to their hand, it may worsen that discomfort through use of one of the controllers.
- Chances are, the participant may feel nauseous being in virtual reality.
- While the participant is in virtual reality, there will be a moving robot arm also in the room. To prevent this, there will be an emergency stop button which would completely disable the robot that is always within arms reach of the researcher.

This research project has been reviewed and approved by the CIS Ethics Committee.

4 Considering withdrawal

At any point, the participant is allowed to withdraw from the study without having any reason for doing so.

5 Regarding data collection, processing and storage

During the test, three sets of data are collected. Notes will be taken on what the participants are experiencing during testing. The time in which the participant performs the requested task is recorded as an indicator of performance. And, the responses to the questions asked during the interview sections of the scenarios are recorded.

- The responses given while performing the requested tasks will be noted down if they are insightful. Quotes may be taken for improved insight and/or use in further writings.

- The time in which the participant performs the tasks described in the scenario is saved and anonymized to: Person x Task y Time z [s]. The performance of each participant is compared and used to evaluate the VE.
- Video recordings and photos may be made during the evaluation, which may be used in a further presentation. Once again, quotes may be taken and used for improved insight and/or use in further writings.

All data such as quotes and times will be anonymized and may be used in the report and/or final presentation.

Study contact details for further information:

name: J.K. Buursink email: j.k.buursink@student.utwente.nl

Contact Information for Questions about Your Rights as a Research Participant

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the CIS Ethics Committee University of Twente by ethicscommittee-cis@utwente.nl

Consent Form for The Evaluation of a Virtual Reality Environment for the Operator of a Teleoperated Robot in a Test Setup

YOU WILL BE GIVEN A COPY OF THIS INFORMED CONSENT FORM

Please tick the appropriate boxes	**Yes**	**No**
Taking part in the study		
I have read and understood the study information dated 15/05/2024, or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	○	○
I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.	○	○
I understand that taking part in the study involves data being collected about my thoughts, opinions, and experiences on/with the test setup and testing procedure. I understand that this data is collected through notes made by the researcher and an audio-recorded interview. I understand that all of this information will be anonymized. I understand that timing data regarding my performance during the testing procedure will be recorded, anonymised, and analysed.	○	○
Risks associated with participating in the study		
I understand that taking part in the study involves the following risks: I may experience physical discomforts, such as motion sickness during the usage of virtual reality devices. I will not be able to see while there is a moving robot arm in the room. I understand that the robot arm has a stop button which will be used by the researcher in case anything goes wrong.	○	○
Use of the information in the study		
I understand that information I provide will be used for the report and presentation of the research.	○	○
I understand that personal information collected about me that can identify me, such as [e.g. my name or where I live], will not be shared beyond the study team.	○	○
I agree that my information can be quoted in research outputs	○	○
Consent to be Audio/video Recorded		
I agree to be audio recorded.	○	○
I agree to be video recorded.	○	○
Future use and reuse of the information by others		
I understand that all data outside of the anonymised data in the report and presentation will be destroyed after the research is completed.	○	○

Signatures

______________________ ______________________ ________
Name of participant Signature Date

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

J.K. Buursink ___________________ ________
Researcher name Signature Date

UNIVERSITY OF TWENTE.

Questionnaire for the Assesment of Sense of Embodiment and Telepresence in the Virtual Environment

Author: Joost Buursink (s2790882)
Date: June 14, 2024
Contents: This document contains 16 questions scaled on a 7-point Likert scale regarding the experience of embodiment and telepresence during the testing procedure. These questions are directly copied, or adapted from [1].

1 Questions

"During the experiment there were moments in which...

1.1 Body ownership

1. I felt as if the virtual arm was my arm."
2. It felt as if the virtual arm I saw was someone else's arm."
3. It seemed as if I might have more than two arms."

1.2 Agency and motor control

4. It felt like I could control the virtual arm as if it was my own arm."
5. The movements of the virtual arm were caused by my movements."
6. I felt as if the movements of the virtual arm were influencing my own movements."
7. I felt as if the virtual arm was moving by itself."

1.3 Location of the body

8. I felt as if my body was located where I saw the virtual body"
9. I felt out of my body"
10. I felt as if my (real) body were drifting toward the virtual body or as if the virtual body were drifting toward my (real) body"

1.4 External appearance

11. It felt as if my (real) body were turning into an 'avatar' body"
12. It felt as if my (real) arm were turning into an 'avatar' arm"
13. At some point, it felt as if my real body was starting to take on the posture or shape of the virtual body that I saw"
14. At some point, it felt as if my real arm was starting to take on the posture or shape of the virtual arm that I saw"
15. At some point it felt that the virtual body resembled my own (real) body, in terms of shape, skin tone or other visual features."
16. I felt like I was wearing different clothes from when I came to the laboratory"

2 Scores for participant number ...

question	strongly disagree	disagree	somewhat disagree	neither agree nor disagree	somewhat agree	agree	strongly agree
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							

3 Total scores for participant number ...

1. Ownership = (Q1 - Q2 - Q3)/3
2. Agency = (Q4 + Q5 + Q6 - Q7)/4
3. Location = (Q8 - Q9 + Q10)/3
4. Appearance = (Q11 + Q12 + Q13 + Q14 + Q15 + Q16)/6

Ownership	Agency	Location	Appearance

4 Total scores

Total Score = (Ownership*2 + Agency*2 + Location*2 + Appearance)/7

Total embodiment score

References

[1] Mar Gonzalez-Franco and Tabitha C. Peck. "Avatar Embodiment. Towards a Standardized Questionnaire". In: *Frontiers in Robotics and AI* 5 (June 2018). ISSN: 2296-9144. DOI: `10.3389/frobt.2018.00074`. URL: `http://dx.doi.org/10.3389/frobt.2018.00074`.

C Questionnaire results

	Person 1	Person 2	Person 3	Person 4	Person 5	Person 6	Mean Value
Q1	1	-1	1	0	1	1	0.5
Q2	3	-3	0	1	-2	-3	-0.66667
Q3	1	-3	-2	-3	-2	-3	-2
Q4	1	1	1	-2	1	-2	0
Q5	3	3	3	2	2	3	2.6667
Q6	2	-2	1	2	2	1	1
Q7	-3	2	1	-3	-1	-2	-1
Q8	1	-2	2	2	2	2	1.1667
Q9	-2	-3	1	-2	-2	2	-1
Q10	1	-3	-1	-3	0	2	-0.66667
Q11	0	2	1	-3	-1	2	0.16667
Q12	1	2	1	0	1	-3	0.33333
Q13	1	2	0	1	-2	-3	-0.16667
Q14	-1	1	-2	1	-1	-2	-0.66667
Q15	-3	0	-2	-3	-3	-1	-2
Q16	-3	-2	-3	-3	-2	-3	-2.6667

Table 3: Answers to embodiment questionnaire with mean values.