IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 10, 2020, accepted January 6, 2021, date of publication January 11, 2021, date of current version February 1, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3050881

The Effect of Semi-Transparent and Interpenetrable Hands on Object Manipulation in Virtual Reality

MICHIEL VAN VELDHUIZEN AND XUBO YANG, (Member, IEEE)
School of Software, Shanghai Jiao Tong University, Shanghai 200240, China

Corresponding author: Michiel van Veldhuizen (michielvanveldhuizen@gmail.com)

This work was partially supported by the National Key Research and Development Program of China (2018YFB1004902).

ABSTRACT Manipulating objects with the hand is a common task in Virtual Reality (VR). However, some issues can occur during these manipulations. Occlusion is one issue that happens when the virtual hand covers an object, and therefore the perception of that object is hindered. Semi-transparent hands could solve this problem. Another issue is the difficulty of delicate hand object manipulation. By using interpenetrable hands, turning off the physics of the hands, this difficulty might decrease. Still, there is a lack of research into how significant the impact is while using these methods. In this paper, with the designs of semi-transparent hands and interpenetrable hands, we present the results of our conducted user study focusing on the effect of semi-transparent and interpenetrable hands on hand object manipulation tasks in virtual reality. The user study includes a VR environment where participants are asked to perform tasks. These tasks are recorded in objective results of accuracy and speed. Afterwards, they fill out a questionnaire about their opinion on the difficulty while using the different methods in which the subjective findings are recorded. Additionally, we improved the semi-transparent hands by adding a feature that smoothly transitions the hand from opaque to semi-transparent. As an input device, we used the Leap Motion Controller (LMC) for this user study. However, any hand tracking sensor that tracks hands from the VR headset's point of view could be used. Semi-transparent and interpenetrable hands have shown significant improvement for precise manipulation, which was verified by user feedback from the questionnaire and the data from the tasks.

INDEX TERMS Hand Interaction, human–computer interaction, object manipulation, virtual reality.

I. INTRODUCTION

Over the last decade, we have seen virtual reality (VR) grow tremendously. At the start of the development, VR was mostly used for training simulations, which later transitioned to the gaming world. With innovations like the Oculus Rift and HTC Vive, large leaps were taken for VR development. These devices are currently still the main solutions on the market. Recently, we have seen more upcoming developments in VR research for educational purposes. This goes back to the origin of VR as a training and learning tool. A stepping stone is a virtual lab [1] where students can perform chemistry and physics experiments. There are many benefits, such as the cost of material, durability, and probably the most important, the possibilities. However, the experience in a virtual lab might be different from the real world. This is where VR can enhance this experience by creating the lab in VR [2]. In a VR environment, it is possible to experiment with and experience, for example, chemicals that would be considered dangerous or take a form that is not possible in the real world.

Hand object manipulation is a natural and intuitive interaction approach for virtual reality. There are still issues that occur during the manipulation of objects, such as occlusion of the hands and object or unwanted collision. Other issues can be that the size or shape of an object might make it challenging to grasp. Therefore, interaction with a VR environment can become problematic.

Occlusion happens during the manipulation of an object, for example, when placing the hand on an object to grasp it. In the real world, the issue of occlusion of a person's hand and an object is often solved by other means of perception, for example, touch. By the sense of touch, there is no direct need to see the object to perceive it correctly. When the user does not correctly perceive the object or the environment,

The associate editor coordinating the review of this manuscript and approving it for publication was Eunil Park.

the manipulations will likely become problematic. In VR, haptics such as touch is not always available. However, it is possible in VR to make the hand semi-transparent, which gives the user the ability to see the object through the hand. Hence, this will boost the capability of manipulation of the object in the way the user desires [3]. There is a way to enable haptics in the way of VR gloves. However, these are very expensive and therefore do not always fit the use case. VR gloves are often used just for research or by companies, not in an educational or personal environment.

The solution of semi-transparent hands has been introduced already [3], [4]. However, how big of an improvement will this have on the user to perform specific tasks? Is there a difference in improvement while interacting with small or big objects? How does the user feel about having an unnatural sight of their hands disappear? These are questions we want to answer with our user study. As an improvement for the semi-transparent hand, we introduced a method to make the hands smoothly transition from opaque to semi-transparent. This method is made to reduce the shock effect for the user seeing their hands disappear.

The other issue we looked into is "collision." Collisions happen while doing object manipulations. However, there are also unwanted collisions, for example, by touching previously placed objects while manipulating another object. Another possibility is pushing away an object when the user desired to grasp it. In most virtual environments, it is possible to manipulate objects both by virtually grasping or touching them. Grasping will, of course, stick the object to the hand, and touching will move the object in the direction that force is applied. However, when an object has a complex shape or is small, the force could be applied before the grasping mechanism is detected. Therefore, the interaction with these types of objects becomes complicated. To improve these interactions on the previously mentioned issues, we introduce the method of interpenetrable hands. This means that the hand can penetrate objects without applying force on them. However, it is still possible to grasp the object during this state. This method allows the user to get closer to the object to grasp it, rather than pushing it away. It also improves the grasping and releasing motion as the user is less likely to hit previously placed objects.

Except for the smooth transition into transparency, these methods have been introduced in a similar way already. However, there is no research into the impact and the effect of using these methods while performing different tasks. The tasks would have to be designed with diverse levels of difficulty. This way, it is possible to determine if the previously mentioned methods have a more significant or a small impact depending on the task's degree of difficulty. Besides just having objective results from performing a task, it is also essential to know what the user prefers to use. Therefore, in this work, we created an environment where participants had to perform different tasks with different levels of difficulty and recorded the speed and accuracy results. Their subjective findings of the tasks with the methods are recorded in the questionnaire. With this data, we can answer the following questions: Is there an improvement, and if so, how significant is it? Does the improvement differ on the difficulty of the task? Does the subjective feeling of the user match with the objective measurements? As the final part, with our user study, we were able to identify the finger occlusion problem while using a hand tracking sensor from the VR Headset to the hands. Additionally, we mention a future work solution that might solve this problem. However, this has yet to be tested.

We consider the following to be our key contributions:

- We conducted a user study and held a questionnaire to gather data on the user's opinion and performance while using the semi-transparent and interpenetrable hands methods. The performance data is analyzed and compared, which has shown an objective improvement of 8.80% in accuracy and saved 18.40% in time while performing the tasks. In the questionnaire's data, we see a subjective improvement in being in control by 51% while using the mentioned methods.
- We introduced the novel improvement to the semi-transparent hand to make the hand transition smooth into transparency, which was recommended by the vast majority of our participants.

II. RELATED WORK

A. SEMI-TRANSPARENCY

Making use of a semi-transparent selector is firstly proposed in the work of Zhai *et al.* [3]. Their work presents a significant improvement comparing a wireframe 3D selector to a semi-transparent selector. As they mentioned, this can be applied to any interaction within a 3D space. In VR, the virtual hand becomes the selector. Hence, translating the semi-transparent selector to VR means making the hands semi-transparent. Prior to making hands semi-transparent, but yet a similar example is the work of Miyasato [5]. They used a small screen on the hand's dorsal side and a camera on the palmer side. This way, a user was able to see through the hand without making the hand semi-transparent. An example of semi-transparency is seen in Buchmann [4], which shows the use of semi-transparent hands in augmented reality. This might differ from VR in the sense of realism. In their work, they make the real hands of the user appear semi-transparent. Their primary focus is to determine the user's preferred transparency of their hand. However, their works do not include performing tasks with semi-transparent hands and the possible change in accuracy or speed while performing these tasks. Furthermore, their method provides a fixed transparency level, while our work provides a smooth transition from opaque to transparency and therefore operates with a variable level of transparency.

The work of Knierim *et al.* [6] makes use of semi-transparent hands and compares the results for typing in VR with different degrees of transparency. In their work, they state that transparency did affect inexperienced typists. No hands gave the best results for the experienced typists as they do not need visual cues to type. However, semi-transparent hands did score better than realistic hands for the experienced typists. This does, however, differ from our work since a keyboard is a static object. It is not likely to move and will have the same layout every time. Therefore, an experienced typist does not need visual cues to type. While interacting with objects, as in our work, these visual cues will be required as the user is not familiar with the objects, and therefore we expect different results while using the semi-transparent hands.

B. VISUAL REPRESENTATION OF THE HANDS

The work of Schwind *et al.* [7] shows us the results of a user study to understand the integration of visual and haptic sensation of avatar hands while using a 2D Fitts's Law task. As they mentioned, they did not find any evidence that changing the depth cues increases or decreases the user's input performance for the task. However, the subjectively perceived ability to interact with the virtual world can potentially affect the performance, which is a baseline to expect different results from the data and the subjective feeling of the user.

The work of Lin *et al.* [8] presents six different distinct appearances of the virtual hand and with those conducts an experiment to see how the participants perceive them. Argelaguet *et al.* [9] show the results of the effects of the virtual hand representation while interacting with different objects. By the work of Jung *et al.* [10], we can see an improvement while using a personalized hand to estimate the object size. Alternatively, the works of Schwind *et al.* [11], [12] are studies in the visual representation of the hand based on different features or aspects. However, in all of these studies, a semi-transparent hand or a semi-transparent version of their appearance is lacking, nor is it measured if that will improve the task performance of a user, and by how much.

The visual representation of the hands also contributes to the embodiment as seen in the work of Pyasik *et al.* [13], which shows us that a realistic hand representation feels slightly more ownership over the hand than a fake hand representation. However, an object or no hand looking like a selector had the participants feel minor ownership. In both hand-looking representations, there was no disownership noticeable. The work of Martini *et al.* [14] used a 0, 25, 50, and 75% transparent representation of the hand and recorded the degree of ownership. From their results, we can see that the feeling of ownership over the hand starts to decrease slightly at 25% to 50% and decreases exponentially at 75%.

C. GRASPING METHODS

The works revolving around visual feedback for grasping, such as the work of Prachyabrued and Borst [16], explains the grasping types and the potentially misleading feedback information when applied. The previous work is based on their work [17] that focuses on the visual interpenetration trade-off in whole-hand virtual grasping. Another work focusing on virtual grasping feedback is the work of Canales *et al.* [18], which analyzed the participant's opinion on eight different visual feedback techniques. These works use the visual representation of the hands and therefore overlap with the previous sub-chapter as well. Their methods may seem similar to the interpenetrable hands. However, the difference is that their work is strictly a grasping method; the interpenetrable hands conversely *support* a grasping method rather than being a grasping method itself. Our work also involves interacting with multiple and different sized objects, which might improve the usage of the interpenetrable hands. Realistic grasping has been researched in the work of Borst and Indugula [19] in which they present a physically-based approach to realistically grasp and manipulate different kinds of virtual objects. This type of research was also conducted by Liu [20], which is more in-depth on this topic in computer graphics but does not include VR directly. Similar work in this field while using haptic feedback are works such as the work of Moehring and Froehlich [21] using a custom VR glove for effective manipulation of virtual objects. These works are similar to our work on improving the grasp of objects. The work of Höll *et al.* [22] presents the potential of efficient physics-based implementation for realistic hand-object interaction in VR. Their work has a similar base to Oprea *et al.* [23], which presents a visually plausible grasping system for object manipulation and interaction in VR, which is operated on different predefined objects. Most of these works use a glove based solution, however not all as in Kim and Park [24], which present a physics-based hand interaction while using the Leap Motion Controller. Besides focusing purely on grasping methods, a related focus is gesture interaction recognition. As has been researched in Li *et al.* [25], which summarizes the gesture interaction recognition methods and devices. Additionally, they also mention the problems with these. However, the problem we later mention in our paper is not listed in their work.

Additionally, comparing these works to ours, our work focuses on changing the different aspects of the hand, both visual and palpable, to help the user to improve their grasping capability rather than improving the grasping technique itself. Additionally, our study also focuses on the opinion of the user while doing so. Furthermore, our work does not focus on creating the most realistic grabbing method possible. This, however, could be potential future work to combine our current work with their realistic grasping method to see if there are changes in the results.

III. BACKGROUND

A. SEMI-TRANSPARENT HANDS

To give an overall understanding of why semi-transparent hands are functional for VR, imagine the following: in the real world, there is an apple in front of a user; the user can see it. Nevertheless, what if the user places their hand over the apple? He knows it is there because he can feel it, but he can no longer see it. However, he can make assumptions about the apple's position based on his touch. How would this work in VR? Say the user grasps an object that is smaller than their virtual hand. This action implies that his virtual hand will partially or fully occlude that object. He still knows it is there because he grasped it, but at what angle is the object? How is it grasped and connected to the hand? How can he place it? These are problems semi-transparent hands solve. However, it might create other problems, such as an unrealistic or unnatural feeling for the user. Plus, how significant is the actual impact?

By the work of Zhai *et al.* [3], we can see an improvement in accuracy and speed while using a semi-transparent selector. Translating this to VR, we should expect results to improve, but by how much? Furthermore, will this improvement differ depending on the degree of difficulty of the interaction? There are VR input devices that use this method for the hands or joysticks. However, before our user study, it has not yet been defined when the use of semi-transparent hands is most optimal or if the user even prefers it.

B. INTERPENETRABLE HANDS

Virtual object manipulation has been a widely discussed topic [26]–[29]. However, grasping and manipulating an object by hand can cause different issues. We noticed that accurately grasping an object can be a difficult task, especially for smaller objects or objects close to each other. The reason is that the virtual hand is a solid mass that might apply force to the object, and therefore move it during an attempt to grasp it. Additionally, the virtual hand's force could be accidentally applied to touch objects while manipulating another object. Grasping objects by intersecting an object after which a triggering technique, for example, a hand gesture, attaches the object to the hand has been mentioned used previously [30]. However, the improvement while using multiple objects plus solving a possible occlusion issue has not been recorded yet.

A solution for this is the interpenetrable hands, which removes the force the hand applies to objects, which means turning off the hand's physics. Hence, the hand can penetrate an object. However, grasping interaction is kept intact. This method is implemented by removing all collisions with any objects except the currently grasped object. Therefore, the grasping method has to overrule the no collision rule. This way, the hand can get as close to an object as needed to grasp it without moving it while making the grasping motion. This solution could also potentially be simplified, yet a more visually unrealistic, solution to the collision issues mentioned in Jacobs *et al.*'s paper on God-object [31].

Additionally, the method also supports interaction when the hand has to get close to previously placed objects without moving or misplacing those. A fair use case would be placing dominoes, where it is preferred to only interact with the current grasped domino but not accidentally push over the others. However, permanently removing the contact feature is not optimal. Contact with objects is often required in a virtual environment, for example, when pushing an object into a more accurate position or interacting with a button.

C. LEAP MOTION CONTROLLER

The Leap Motion Controller (LMC) is often used for research in hand gesture-based interaction [32]–[37]. The accuracy and robustness have been tested [38] as well. During our user study, we used the LMC as the hand tracking device and mounted it on the VR headset. It is known that this is often not the most optimal way to track hands as it has some issues, as mentioned later in the paper. However, an often more accurate hand tracking solution such as VR gloves does not always fit every use case as it dramatically increases the costs and the care with which to handle the hardware. The increase in costs makes the VR gloves often unsuitable for the standard consumer and is rather only used by companies and for research. A hand tracking device that is mounted on the VR headset, or for some VR headsets already internally installed, will significantly reduce the costs and eliminate the need for extra hardware. There is, of course, the option to use a physical controller over hand tracking. However, hand tracking shows a higher valence for tasks such as grasping [39] and enhance the sense of embodiment, which is beneficial in fields such as education, training, and rehabilitation [40], [41]. Studies have also shown positive results while teaching and training the VR interaction gestures to people with a disability such as Down Syndrome [42]. The LMC also has been used by Miranda *et al.* [43] to categorize the issues with mid-air InfoVis (Information Visualization) interaction. However, the issue we have noticed and explained later in this paper is not mentioned in their work due to the different use of the LMC.

IV. IMPLEMENTATION

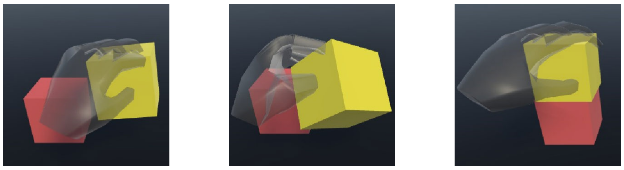
A. SEMI-TRANSPARENT HANDS

Our work uses this semi-transparent hands method, i.e., the user can see the object through his hands. See Fig. 1 and 2. However, this might create an uncomfortable and unnatural feeling as the user sees their hands disappear. As a solution, we present a method to make the hands smoothly transition from opaque to semi-transparent at the level the user desires. This transition happens at the moment the hand goes closer to an object. Therefore, it does not come as a shock to suddenly see the hands disappear. Therefore, the hand will operate on a

(a) Hand just in transparency range

(b) Hand closer to the object

FIGURE 1. The hand becoming transparent over the distance.



(a) Grasping interaction (b) Grasping interaction side angle on (c) Placing an object

FIGURE 2. **A set of three images showing an interaction sequence of grasping and releasing an object. Here you can see the benefits of semi-transparent hands. Figures (a) and (b) show a partial to almost full occluded red block visible due to semi-transparent hands. Figure (c) shows the final result of being able to stack the block because of the visual cues.**

variable level of transparency between opaque and the lowest transparency level corresponding with distance to the nearest graspable object.

The algorithm we used to make the hands smoothly semi-transparent is as follows: The hands have an *opacityLevel*, which defaults to 255. The range is the alpha of the RGB spectrum, where 0 is fully transparent and 255 is fully opaque. Let H be the hand and H_p the hand's position. From there, we can collide with all nearby objects while using the variable *opacityDis* to amplify the distance of the sphere collider and store all objects O by $O = spherecollider(H_p, opacityDis)$. For each element O, we calculate the distance between H_p and O_p and store it locally in D. While using D_n to store the object nearest to H_p, we store the nearest object as $O_n = \min(D, D_n)$. Then we check if $D_n < opacityDis$ to determine if the transparency should be active. If so, we set the shader of H from the standard shader to a shader that allows transparency. This shader makes fingers visible through the hand and, therefore, cannot be used during non-interaction. Vice versa, if $D_n > opacityDis$, the standard shader will be applied to H. To determine the value of *opacityLevel*, we map the values to a 0-255 range by:

$$opacityLevel = opacityMin + (255 - opacityMin) \times \frac{D_n}{opacityDis} \quad (1)$$

The value of *opacityMin*, opacity minimum, will be asked a question in the questionnaire (Q4) and represents the hand's most transparent state. The value of 255 could be called *opacityMax* and will remain 255 in this case. Now we can set the opacity of H to *opacityLevel*, and the hand becomes semi-transparent to a certain degree based on the distance to its nearest object.

B. INTERPENETRABLE HANDS

As mentioned before, the interpenetrable hands' implementation is to remove the physics/collision of the hands with other objects. However, the grasping possibility is still there. When a grasp motion is detected, the object will remain connected to the hand until a release motion is detected. How a grasp and release motion is detected will be explained in the next subchapter. The interpenetrable hands are implemented as an option that can be turned on or off. As mentioned before, interaction with a standard collision is sometimes wanted for situations as pushing something in a correct position or interacting with a button. By the gesture of opening the left-hand palm, a two interaction button window will be visible. One button to turn on the interpenetrable hands and one, the button with the cross through it, to turn it off. Which feature is currently active is visible by the color of the virtual hand. Turquoise means the interpenetrable hands are active; see Fig. 3.

(a) Option off (b) Option on (c) Result

FIGURE 3. **A set of three images. Image (a) shows the options in hand to turn on the interpenetrable hands; however, currently, it turned off. In the second image (b), the same as the first image is shown; however, the option is turned on, which is visible by the color of the hand. In the last image (c) the result of the interpenetrable hands is shown by the hand penetrating an object.**

Another possible implementation, rather than removing all collisions, could be that the hand can only collide with one or n objects, and all other objects will not collide. However, this might result in unexpected and inconsistent behavior for the user. Therefore, we did not choose to implement this.

C. GRASPING TECHNIQUE

As a grasping technique, the standard LMC grasping technique is used. The LMC can detect hands and fingers by using its two cameras and three infrared LEDs. The LMC provided SDK enables the developer to place the detected hands in a virtual environment. Once the hand and fingers are detected by the LMC, the state of flexion or extension for each individual finger is tracked with their joints angles. With this data, gestures can be programmed into the system. The LMC's standard grasping gesture means a flection of the thumb with the index and ring finger, if and only if they go towards each other. If an object collides with two or multiple points of different fingers from the opposite direction, then a grasp will be detected and executed on that object. This means that an object will stick to the hand. The index finger and the thumb are the primary contact points of the grasp. The object can be manipulated and placed elsewhere as long as the grasp motion is detected. Besides the grasp motion, it is also possible to manipulate or move an object by touching it. If and only if the interpenetrable hands are turned off. Correspondingly the object is released and no longer sticks to the hand once the release motion is detected. The release motion is a widening gap motion between the thumb and other fingers. This gesture is triggered by placing the thumb and contact fingers from flexion into extension.

D. VISUAL ASPECT OF THE HAND

To prevent a possible uncanny valley effect by having realistic bodies with inconsistent realism [44], [45], which might occur with transparency and changing the hand color, we choose to use a glove type of representation for the hands instead of a realistic version. A non-realistic hand but yet still hand looking like representation barely affects the sense of ownership over the hand [10], [13], [14] and therefore will not influence the test results.

V. THE USER STUDY

With our user study want to be able to answer the following research questions:

- RQ1: How big is the improvement of semi-transparent hands while performing tasks?
- RQ2: Is there an improvement in removing the collision (interpenetrable hands) while performing tasks?
- RQ3: Do the improvement differ based on the size of the objects?
- RQ4: How does the user feel about the unnatural sight of their hands disappearing during semi-transparency?

A. THE SETUP

For the setup of our user study, we used the HTC Vive as a VR headset. On the headset, we mounted the LMC. Therefore, the hands can be tracked in VR. To create an environment where the hands are tracked, and the tasks can be performed, we used the Unity 2020.1 real-time 3D development platform. Both HTC Vive (VIVEPORT with SteamVR) and Leap Motion (Ultraleap Hand Tracking V4 Orion) provide an SDK to use their hardware within Unity. The user study was conducted in a room where the participants had a 2 by 2 meter range to move around. There were two instructors to coordinate the safety of the participants and the hardware. These instructors also gave the user study instructions and operated the laptop containing the VR environment.

B. PARTICIPANTS

Fifteen participants participated voluntarily in the user study, which required about 30 minutes of their time to complete. The study started with a briefing to explain the research, the tasks, and the methods. The participants were aged 22 to 28 ($M = 24.7, SD = 1.49$) with the computer skills, on a Likert scale(1-7), ranging from 3 to 7 ($M = 5.6, SD = 1.51$) and experience in VR, on a Likert scale(1-7), ranging from 1 to 7 ($M = 3.2, SD = 2.15$).

Because of the experiment's duration, which might result in fatiguing of the participant, we let the participants sit in a swivel office chair. This means that when the participants feel the requirement to move, it is possible. However, a long duration of standing is not required and will not affect experiment results by exhaustion. The participants were also instructed to use their space if they felt necessary, and the environment was made that it was possible to do so without any danger of tripping over wires. Additionally, we have instructed the participants to let us know if they feel tired or needed a break. After every task, the participant was asked if they needed a break or anything was problematic. A task included: executing one task three times with different methods. Executing tasks with all three methods took on average 3 minutes and 43.59 seconds with a standard deviation of 35.1 seconds. All participants except one said they did not need a break and could straight away continue with the next set of tasks. The one participant that asked for a break was between her third and fourth set of tasks. The reason for the break was, as she stated: "The VR headset is becoming too heavy." After a break of about 5 minutes, she wished to continue the experiment. No abnormalities or significant differences were found in her experiment results neither before nor after the break.

To make the participants feel more comfortable with the environment, we allowed them to experiment during a test run. After the test run, we asked (Q4) the preferred state of transparency of the hands (*opacityMin*). Using increments of 10% transparency, we showed the result of that transparency level and asked which level was preferred. We used the individual user's preference of hands transparency setting throughout the user study. Also, the preference to use the smooth transition into transparency (Q5) was asked at this moment. The difference was shown by switching the mode on and off and seeing and feeling the results. All participants preferred this mode, and therefore was used for all participants.

C. THE TASKS

For our user study, we created sequences of tasks for the participants to perform. This environment is created in Unity. Therefore, the size scaling mentioned is those as used in Unity.

- Task one: stacking three blocks (Scale $xyz = 0.75$).
- Task two: stacking four smaller blocks (Scale $xyz = 0.5$).
- Task three: creating a hashtag shape with small sticks (Scale $xyz = 0.1, 0.1, 1$).
- Task four: placing lids on pots (Scale pot one $xyz = 1$, pot two $xyz = 0.5$).

Tasks one and four are more manageable due to the objects' increased size. Tasks two and three are more troublesome than the others because of the decrease in size and the precision required to place the objects correctly. See Fig. 4.

After providing the briefing and performing the test run, the actual user study began. The user performed each task three times: once with opaque hands and collisions (OC), once with semi-transparent hands (ST), and once with semi-transparent and interpenetrable hands (ST & IN). By performing the same task three times with different methods, we can compare the results.

Due to improvement over repeated trials, the results will most likely be best on the last trial. Therefore, a Latin-square design was applied, and we let one-third of the participants start with part OC, one-third with ST, and the final third with

(a) Task one begin (b) Task two begin (c) Task three begin (d) Task four begin

(e) Task one finish (f) Task two finish (g) Task three finish (h) Task four finish

FIGURE 4. **The four tasks while being performed in the user studies with their beginning and finished state.**

ST & IN. The order of the tasks was randomized to prevent any other statistical increase by the learning curve.

For the task, we measured the task completion time in seconds. This data is the input for the objective findings of RQ1, RQ2 and RQ3. The measurement starts from the moment the participant starts the current task until they say they are finished. Not making an automated finish when the blocks are stacked gives the participant the option to reposition the blocks to a more precise and accurate position if so desired. We measure the task's accuracy by subtracting the offset distance of the x and z axes in the space corresponding to the lowest object in the sequence, or the pre-made model as in task three. The accuracy algorithm for tasks one and two is as follows: Let O be all objects and O_0 the lowest block in the stack. We define two functions l, the location using the x and z axes, and s, the object's size. Additionally, we define M as the multiplier of the size. Therefore, we calculate the accuracy as:

$$100-\left(\sum_{i=1}^{O}\frac{|O_i l(X,Z)-O_0 l(X,Z)|}{O_i s(X,Z)\times M}\times 100\right) \qquad (2)$$

In tasks, one and two and four, we set M two. Task four has two pots and two lids. Therefore this algorithm is applied twice, and later both subtractions of the 100% are added together. In task three, we replace the O_0 of equation two with O_{in} to represent the nearest object of O_i, which is the example shape on the floor. Which then we can calculate as:

$$100-\left(\sum_{i=1}^{O}\frac{|O_i l(X,Z)-O_{in} l(X,Z)|}{O_i s(X,Z)\times M}\times 100\right) \qquad (3)$$

Also, M is set to 5 for task three. Otherwise, the task would be too difficult due to the size of the small objects.

D. QUESTIONNAIRE

The final part of the user study is the questionnaire. The first part of the questionnaire uses the Single Ease Question (SEQ) [15] method to gather the data after the participant performs each task. After each task was completed with a method, the following question was asked: "How difficult was this task with the current method?" This way, the participant can give a precise answer to the question. Asking all questions at the end of the tasks might have been difficult for the participant to recall answers. We notate these questions as $Qt1-1$ until $Qt4-3$ for four tasks and three methods. As standard in SEQ, we used a 7-point system, ranging from 1 = "very difficult" to 7 = "very easy." With these questions, we can conclude the subjective answer on RQ1, RQ2, and RQ3. Afterward, we asked the participants biographical and informational questions related to the task, again on a scale from 1 to 7. However, the end-points are "very bad" to "very good" for Q2, Q3, and Q7, and "strongly disagree" to "strongly agree" for Q5-Q6. Besides that, Q1 is collected as a number and Q4 as a percentage.

Q1 to Q3 is for gathering user information. Q4 is based on the work of Buchmann *et al.* [4]. Q5 is to determine the acceptability of the smooth transition into transparency. As suggested by the work of Lin *et al.* [8], Q6 compares the non-transparent hands (Q6) to the semi-transparent hands (Q6t). The answers of Q4, Q5, and Q6 provide the input for RQ4. Q7 separated by the three parts (Q7-1 to Q7-3) are based on findings on visual integration by Schwind *et al.* [11]. These are to determine a before and after use of methods state comparison. Therefore 7-1 is at the state of OC, 7-2 ST, and 7-3 ST & IN. Which provides an additional overall input for RQ1 and RQ2. The questionnaire consists of the following:

- (Q1) What is your age?
- (Q2) Experience with computers?
- (Q3) Experience with VR?
- (Q4) The preferred transparency level?
- (Q5) Do you prefer the smooth transition into transparency?
- (Q6) Does it seems like your own hands are located in the virtual world?
- (Q7) Were you able to interact with the environment the way you wanted to?

VI. RESULTS

A. TASK RESULTS

The results in Figures 5 and 6 show a decrease in performance time and an increase in accuracy while using the methods. To determine statistically significant differences in the

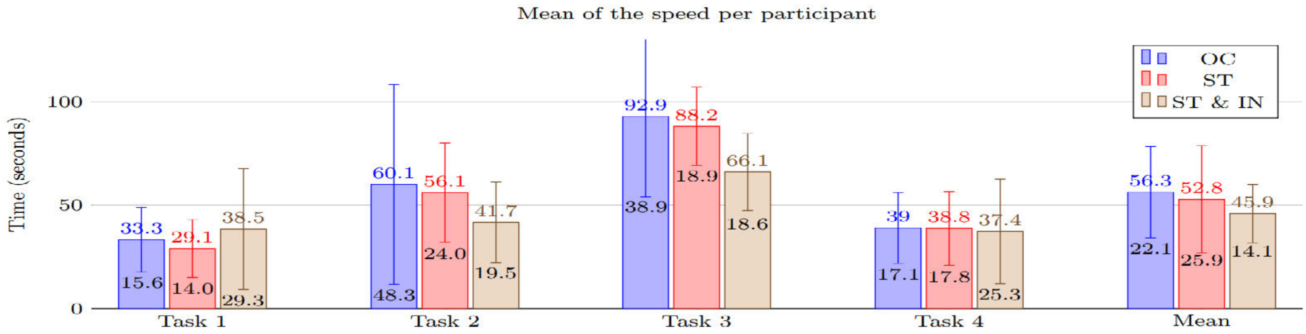


FIGURE 5. The combined per task speed results. Means are written in the color of the bar. SD's are written below the error bar in black.

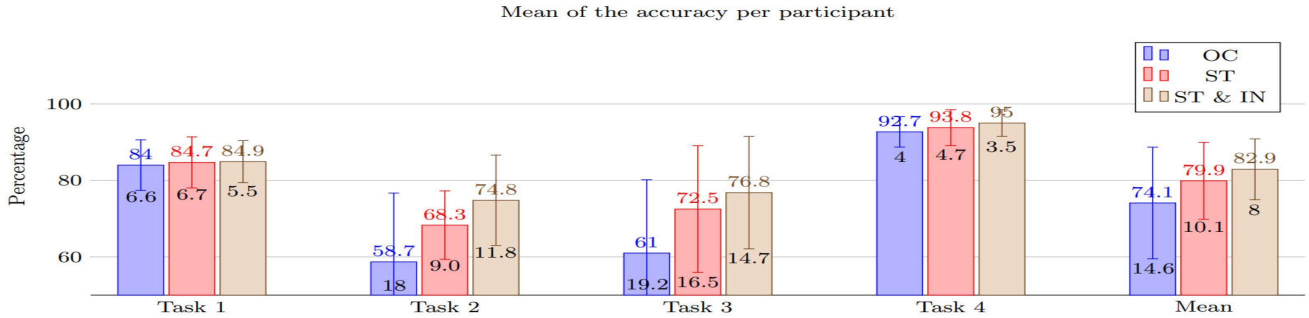


FIGURE 6. The combined per task accuracy results. Means are written in the color of the bar. SDs are written below the error bar in black.

results, a One-way ANOVA was applied. For each tasks in speed S no significance was found: $S-Task1$, $F(2,42) = 0.719, p < .493$. The same can be said for $S-Task2$, $F(2,42) = 1.203, p < .31$. We did find significance for $S-Task3$, $F(2,42) = 3.894, p < .05$. Finally, for $S-Task4$, $F(2,42) = 0.023, p < .977$. Furthermore, for each tasks in accuracy A, we found no significance in $A-Task1$, $F(2,42) = 0.084, p < .919$. We did find significance for $A-Task2$, $F(2,42) = 5.086, p < .01$ and $A-Task3$, $F(2,42) = 3.273, p < .05$. Finally, we found no significance for $A-Task4$, $F(2,42) = 1.332, p < .275$. Which is summarized in Table 1.

For the results where we found statistically significant differences, $S-Task3$ and $A-Task2\&3$, post hoc analyses were performed using the Holm-Bonferroni method as corrected pairwise t-tests to determine statistically significant differences the methods. First for time we found we significant differences for $S-Task3$ between ST and ST & IN ($p < .01$) and for $S-Task3$ between OC and ST & IN ($p < .025$). Secondly for accuracy we found significant differences for $A-Task2$ between OC and ST & IN ($p < .012$) and $A-Task3$ again between OC and ST & IN ($p < .011$). We found no significant differences when comparing the other method pairs (all with values $p > .025$ or $p > .05$ in their respective rank). Which is summarized in Table 2.

TABLE 1. The One-Way ANOVA applied to the task results which have shown a significant difference.

Accuracy/Speed	Task	F	P Value
Speed	3	F(2,42)	< 0.05
Accuracy	2	F(2,42)	< 0.01
Accuracy	3	F(2,42)	< 0.05

TABLE 2. The Holm-Bonferroni post hoc analysis applied to the task results which have shown a significant difference.

Accuracy/Speed	Task	Methods	P Value
Speed	3	ST-ST&IN	< 0.01
Speed	3	OC-ST&IN	< 0.025
Accuracy	2	OC-ST&IN	< 0.012
Accuracy	3	OC-ST&IN	< 0.011

As shown in the results, there are two more difficult tasks (two and three) and two more manageable tasks (one and four). To back up this claim we compared task one and four to task two and three for all methods both on time and accuracy with a One-way ANOVA resulting in: time $F(1,178) = 53.34, p < .000001$ and accuracy $F(1,178) = 113.53, p < .0000001$. The difference in difficulty is that participants interact with smaller objects in tasks two and three, which makes occlusion likely and interaction more difficult. This situation is also where the ST & IN method scores tremendously better than OC for accuracy at an increase of 16.12% on task two and 15.82% on task three. However, the increase in tasks one and four, interacting with larger objects, only receives a small improvement due to the methods. Hence we can conclude that these methods are most optimal for more precise tasks and interacting with smaller objects. The overall increase comparing ST & IN to OC saves 10.38 (18.4%) seconds, which excels up to 26.7 seconds (28.9%) in $Task3$, and increases the accuracy by 8.80%. Also, we can state that ST contributes towards 65% of the gained accuracy of ST & IN. However, ST adds less in time, only about 33% of the total increase. Therefore, we can conclude that the semi-transparent hands are a significant factor for accuracy, but not time. These results are as expected. Since transparent hands mainly focus on seeing the placement of the objects better and therefore increasing the accuracy. As noticeable in the increase between OC and ST in task three, the most challenging task has the most significant increase for accuracy.

On the other hand, the interpenetrable hands are a significant factor in speed and still a good factor for accuracy. These results are as expected. Since semi-transparent hands mainly focus on seeing the placement of the objects without occlusion and therefore increasing the accuracy. The interpenetrable hand's method mainly focuses on a more comfortable grasp and release sequence since there is less chance of unintentionally touching other objects. With these answers, we can conclude the objective measurement for RQ1: yes, there is an improvement while using semi-transparent hands. For RQ2: yes, there is an improvement while using interpenetrable hands. However, as answered for RQ3: These improvements differ by the objects' size and are strongest noticeable while using smaller sized objects.

As mentioned earlier, not all participants used the same level of transparency. The requested level of transparency by the question of Q4 was used. However, after analyzing this, we have stated no significant difference in the results based on the different levels of transparency. This might be occurring due to the chosen transparency level being relatively close to each other, as mentioned in the next chapter. Also, there was no significant difference in comparing the improvements between experienced and inexperienced VR users.

B. QUESTIONNAIRE RESULTS

As the final part of the user study, we had the participants fill in the questionnaire. See Fig. 7, which represents the SEQ questions. We analyzed the data for statistically significant differences by the non-parametric Friedman's tests. A significant difference was found in all tasks. For *Task*1 ($\chi^2(3) = 16.9, p < 0.0005$), *Task*2 ($\chi^2(3) = 16.2, p < 0.0005$), *Task*3 ($\chi^2(3) = 27.1, p < 0.00001$), *Task*4 ($\chi^2(3) = 10.2, p < 0.01$). Which is summarized in Table 3. Additionally a pairwise post-hoc analyses with a Holm-Bonferroni corrected Wilcoxon signed-rank test was conducted to determine statistically significant differences between the methods. For *Task*1 the difference OC and ST & IN was significant ($Z = -3.396, p < 0.001$). For *Task*2 there is a significant difference for OC and ST & IN ($Z = -2.951, p < 0.05$) and ST and ST & IN ($Z = -3.051, p < 0.01$). For *Task*3 there is a significant difference for OC and ST ($Z = -3.396, p < 0.001$) and ST and ST & IN ($Z = -3.235, p < 0.005$) and OC and ST & IN ($Z = -3.482, p < 0.001$). For *Task*4 there is significant difference for OC and ST & IN ($Z = -3.126, p < 0.01$). Which is summarized in Table 4. Other pairwise did not show statistical difference.

Again we can see the difference between the manageable and challenging tasks. Especially for challenging tasks, we can see the increase of score for ST & IN. According to the participant's feedback, IN gives a feeling that these tasks become more manageable than without this method. Which then answers RQ3. Yes, there is a difference in the improvement based on the size of the objects. As mentioned before, the interpenetrable hand's method has the most significant improvement in time (speed), but not accuracy. By this, we can conclude that the participants felt the task becomes more pleasant when performing it quicker rather than more accurately. This also answers the subjective feeling of RQ1 and RQ2. Yes, there is an improvement felt while using the semi-transparent hands and interpenetrable hands.

TABLE 3. Friedman's tests applied on the SEQ results which have shown a significant difference.

Task	$\chi^2(3)$	P Value
1	16.9	< 0.0005
2	16.2	< 0.0005
3	27.1	< 0.00001
4	10.2	< 0.01

TABLE 4. Pairwise post-hoc analyses with a Holm-Bonferroni corrected Wilcoxons test applied on the SEQ results which have shown a significant difference.

Task	Methods	Z	P Value
1	OC-ST&IN	-3.396	< 0.001
2	OC-ST&IN	-2.951	< 0.05
2	ST-ST&IN	-3.051	< 0.001
3	OC-ST	-3.396	< 0.001
3	ST-ST&IN	-3.235	< 0.005
3	OC-ST&IN	-3.482	< 0.001
4	OC-ST&IN	-3.126	< 0.01

For the questionnaire at the end of the user study, see Fig. 8. Q1 indicated the age of the participants. As shown, the participants' ages do not contain a large variation. Similar to Q1, in Q2, the participants' experience with computers does not vary much. However, Q3 does show a larger difference in experience with VR. There was no correlation found between age or experience with computers and their test results. However, the experienced VR users, with a score of 3+, performed better in both speed and time in all tasks with any method. Nevertheless, there was no significant difference found in the improvements between methods while comparing experienced and in-experienced VR users. Q4 shows the preference of semi-transparency numerically, with a mean of 44% semi-transparent, which is similar to the results of [4]. Most participants chose either 40% or 50% except for two who chose 30%. Q5 indicates that the vast majority agrees on preferring the smooth transition into semi-transparent hands. For realism, the non-transparent hands are slightly preferred over the semi-transparent hands, as visible in Q6 and Q6t, which answers the question of RQ4. Opaque hands are preferred. However, when semi-transparency is used, a smooth transition into transparency is recommended. Additionally, see Fig. 9, Q7-1 to Q7-3 show an increase in the confidence of interacting with the environment as the user wanted, with the most significant improvement in the ST & IN. This improvement from Q7-1 (3.27) to Q7-3 (4.93) is an increase of 51%. Applying the non-parametric Friedman's tests on Q7-1 to Q7-3: ($\chi^2(3) = 19.2, p < 0.0001$). Again applying Holm-Bonferroni corrected Wilcoxon signed-rank test as post hoc analysis do determine pairwise statistical difference. Where significant difference was only found between Q7-1

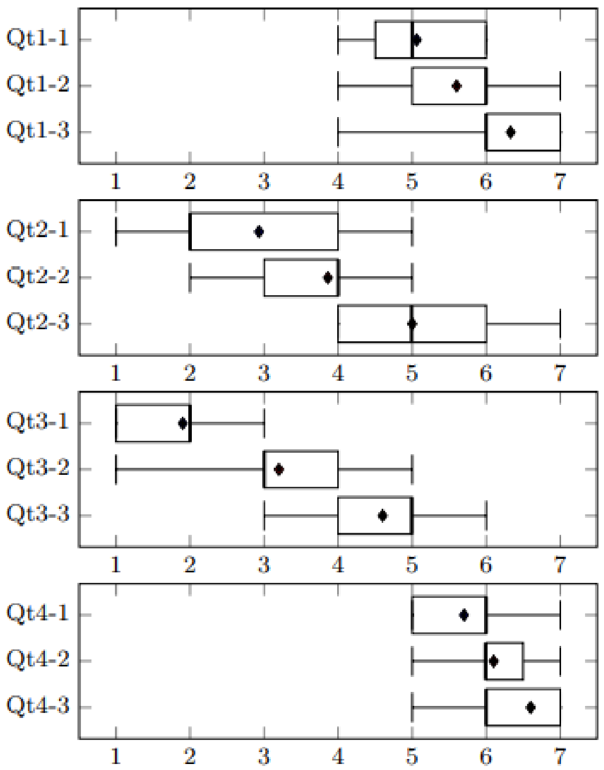


FIGURE 7. Each task rated per method, points ranging from 1 = Very difficult to 7 = Very easy.

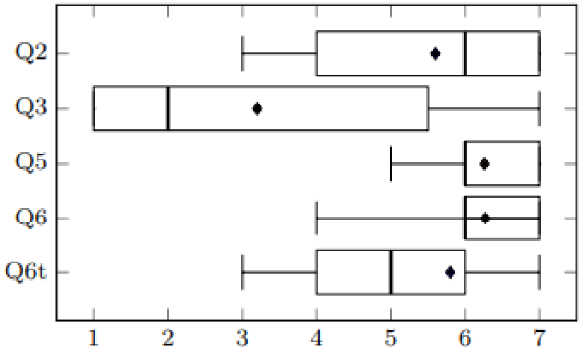


FIGURE 8. The remaining questionnaire questions on a scale of 1 to 7.

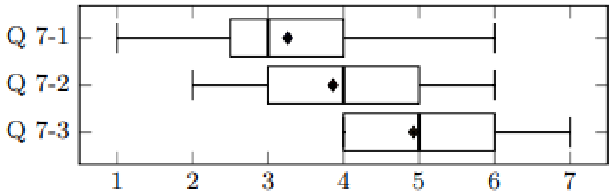
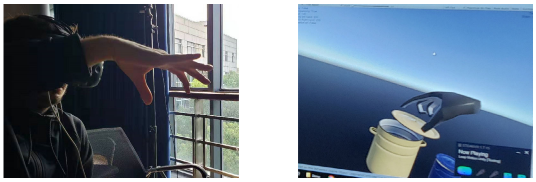


FIGURE 9. Overall subjective feeling of being in control while comparing the methods on a scale of 1 to 7.

and Q7-3 $= (Z = -3.399, p < 0.01)$ and between Q7-2 and Q7-3 $= (Z = -3.357, p < 0.001)$. This again concludes a positive answer towards RQ1 and RQ2.

C. RESULTS DISCUSSION

When reading or hearing about the interpenetrable hands, they may seem like a minor improvement or even none at all. We also noticed this during the briefing. The participants seemed to understand the impact of semi-transparent hands. Nevertheless, the interpenetrable hands' impact was often not recognized by the participants when being told about this method. However, this method scored significantly higher for both user preference and performance, as seen in the results. Moreover, more challenging tasks became significantly more comfortable to perform. Therefore, we can say that this method was perceived better after experiencing it rather than being explained.



(a) Real fingers in extension (b) Virtual fingers (middle finger to pinkie) in flexion

FIGURE 10. False finger detection during the release of a grasp because of its hand occluding the fingers. Note that figure a and b are recorded at the same time. However, because of occlusion from the sensor to pinky and ring finger a different hand gesture is shown in VR.

During the user study, we were able to state, in our opinion, the biggest issue with using a sensor, such as the LMC, from the VR headset point of view to track the hands, which is occlusion. This issue was not visual occlusion in VR, but the finger's occlusion by the hand for the hand tracking sensor in real life. This issue was also not mentioned in categorizing the mid-air interaction issues [43]. They use the LMC on the desk rather than mounted on the VR Headset. Therefore in their situation, the sensor's vision is unlikely to have the fingers or hand occluded. It is only possible when the sensor is mounted on the head, as in the scenario later mentioned. Neither was the issue mentioned in the summarization of gesture interaction recognition techniques and devices [25]. Xiao *et al.* [37] also mentioned "problems," one they stated as "inconsistency of the display device," nevertheless an exact issue, cause, or solution was not mentioned.

The previously mentioned occlusion issue is encountered in more specific situations than expected. Grasping an object is not the main issue. Releasing the grasp is the crux of the problem. This is because the angle of the sensor towards the hand will most likely have the ring finger and pinkie occluded by the distal part of the hand or other fingers. Therefore, the sensor often still detects those two fingers in flexion when they are actually in extension in the real world (see Fig. 10), which results in difficulty in releasing the objects. However, once these fingers are suddenly detected, a flicking motion will happen in VR because the fingers go from flexion to extension instantly, which is likely to hit the objects near the hand, including the one previously grasped. Whenever these situations occurred during the user study, we reset the current task and start again. For some participants, this issue never happened. However, for some others, this happened more often. We did not keep exact track of how many times this occurred to each participant. However, we assume that experience in VR does play a role in preventing this issue.

Another difficulty, again also not mentioned in [25], [43] due to different viewing point of the sensor, which we mainly encountered in task three, creating the hashtag, was the following. This task requires to place some objects at a 90deg angle. Therefore, the hand has to be in a pronated or supinated state. For example, picture a user that is holding their hand in front of their head. Then he makes a pronated (towards the right) grasping motion while keeping the wrist in the same place. As noticed, the user is unlikely to see his fingers because the distal side of his hand and his wrist are completely occluding the fingers. Correspondingly, the sensor mounted on the VR device will have issues detecting the fingers in this state. As we noticed with our more experienced participants, they partially overcame this issue by using the space to move around. By changing their body state, they overcome the pronation or supination of the wrist not being detected correctly. However, this will not be an optimal solution for a situation where a person is sitting stationary. Such movement might also be more fatiguing during prolonged use.

VII. LIMITATIONS

One of the limitations we noticed when implementing semi-transparent hands was that the center of the grasping motion for our sensor, the LMC, is the contact point of the thumb and index finger. Therefore, the semi-transparent hand method does not live up to its full potential since there is no hand palm over the object, and there is a smaller chance of occlusion. While using hardware that uses the center of the hand palm as the contact point of the grasping motion, we expect even better results for this method. This would also be a more realistic way of grasping an object rather than the contact point of the index finger and thumb in most cases.

Another limitation is that our current work is only limited to a sensor as a hand tracking device. Using other hardware such as VR gloves might give different results. However, we expect that the results do not vary tremendously.

VIII. CONCLUSION AND FUTURE WORK

In this paper, we have shown the results of the user study focusing on the effect of the users' performance while using semi-transparent hands, while using the newly introduced method to transition smoothly into transparency, and with interpenetrable hands. Additionally, with the questionnaire's data, we can conclude that using the semi-transparent and interpenetrable hands considerably improves the user's performance, especially with more challenging and precise tasks. Therefore, we can say that manipulating small objects and performing precise tasks benefit significantly from these methods, whereas interacting with larger objects only shows a small improvement. These findings can be used for developers to determine the interest of using these methods.

The work of Knierim *et al.* [6] using semi-transparent hands stated that "Inexperienced typists require hand visualizations to orient themselves in VR while transparency has no effect" and "Results show that experienced typists are less affected by different hand rendering." In our work, we did not see a distinct difference between the experienced and inexperienced computer or VR users (Q2, Q3). No correlation exists between users' past VR experience and their improvements in the tasks. Hence, we can conclude that these methods are an improvement for both experienced and inexperienced users.

A possibility for future work would be to use Deep Learning to create a dataset of labeled grasp and release motions. This dataset includes labeling falsely-detected release motions as an "occluded release motion." This way, it might be possible, even with the occlusion of a hands-on device, to improve an object's release. Consequently, it might be possible to detect or predict an "occluded release motion" and classify this behavior without actually detecting it. This classification can also be applied to the grasping motion. However, as we have noticed, these are often less likely to be occluded and, therefore, falsely detected. Afterward, the earlier presented user study can be performed again with the improvement of releasingi objects.

Another option could be to combine the realistic grasping methods of the related work section with this work. Additionally, instead of using a sensor (such as the LMC), use a glove based solution to see the change in the user's opinion and task performance.

REFERENCES

[1] D. Liu, P. Valdiviezo-Díaz, G. Riofrio, Y.-M. Sun, and R. Barba, "Integration of virtual labs into science E-learning," *Procedia Comput. Sci.*, vol. 75, pp. 95–102, 2015, doi: 10.1016/j.procs.2015.12.224.

[2] F. Wang, "Computer distance virtual experiment teaching application based on virtual reality technology," *Int. J. Emerg. Technol. Learn.*, vol. 13, no. 4, pp. 83–94, 2018, doi: 10.3991/ijet.v13i04.8472.

[3] S. Zhai, W. Buxton, and P. Milgram, "The 'silk cursor': Investigating transparency for 3D target acquisition," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. Celebrating Interdependence*, New York, NY, USA, 1994, pp. 459–464, doi: 10.1145/191666.191822.

[4] V. Buchmann, T. Nilsen, and M. Billinghurst, "Interaction with partially transparent hands and objects," in *Proc. 6th Australas. Conf. User Interface*, Newcastle, CA, Australia, vol. 40, 2005, pp. 17–20, doi: 10.5555/1082243.1082246.

[5] T. Miyasato, "See-through hand [virtual environment]," in *Proc. Australas. Comput. Hum. Interact. Conf.*, Adelaide, SA, Australia, 1998, pp. 142–143, doi: 10.1109/OZCHI.1998.732207.

[6] P. Knierim, V. Schwind, A. M. Feit, F. Nieuwenhuizen, and N. Henze, "Physical keyboards in virtual reality: Analysis of typing performance and effects of avatar hands," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Montreal QC, Canada, Apr. 2018, pp. 1–9, doi: 10.1145/3173574.3173919.

[7] V. Schwind, J. Leusmann, and N. Henze, "Understanding visual-haptic integration of avatar hands using a fitts' law task in virtual reality," in *Proc. Mensch und Comput.*, Hamburg, Germany, Sep. 2019, pp. 211–222, doi: 10.1145/3340764.3340769.

[8] L. Lin and S. Jörg, "Need a hand?: How appearance affects the virtual hand illusion," in *Proc. ACM Symp. Appl. Perception*, Anaheiem, CA, USA, Jul. 2016, pp. 69–76, doi: 10.1145/2931002.2931006.

[9] F. Argelaguet, L. Hoyet, M. Trico, and A. Lecuyer, "The role of interaction in virtual embodiment: Effects of the virtual hand representation," in *Proc. IEEE Virtual Reality (VR)*, Mar. 2016, pp. 3–10, doi: 10.1109/VR.2016.7504682.

[10] S. Jung, G. Bruder, P. J. Wisniewski, C. Sandor, and C. E. Hughes, "Over my hand: Using a personalized hand in VR to improve object size estimation, body ownership, and presence," in *Proc. Symp. Spatial User Interact.*, Berlin, Germany, Oct. 2018, pp. 60–68, doi: 10.1145/3267782.3267920.

[11] V. Schwind, L. Lin, M. Di Luca, S. Jörg, and J. Hillis, "Touch with foreign hands: The effect of virtual hand appearance on visual-haptic integration," in *Proc. 15th ACM Symp. Appl. Perception*, Vancouver, BC, Canada, Aug. 2018, pp. 1–8, doi: 10.1145/3225153.3225158.

[12] V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, and N. Henze, "SThese are not my hands!: Effect of gender on the perception of avatar hands in virtual reality," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Denver, CO, USA, May 2017, pp. 1577–1582, doi: 10.1145/3025453.3025602.

[13] M. Pyasik, G. Tieri, and L. Pia, "Visual appearance of the virtual hand affects embodiment in the virtual hand illusion," *Sci. Rep.*, vol. 10, no. 1, Dec. 2020, Art. no. 5414, doi: 10.1038/s41598-020-62394-0.

[14] M. Martini, K. Kilteni, A. Maselli, and M. V. Sanchez-Vives, "The body fades away: Investigating the effects of transparency of an embodied virtual body on pain threshold and body ownership," *Sci. Rep.*, vol. 5, no. 1, Nov. 2015, Art. no. 13948, doi: 10.1038/srep13948.

[15] J. Sauro and J. S. Dumas, "Comparison of three one-question, post-task usability questionnaires," in *Proc. 27th Int. Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, 2009, pp. 1599–1608, doi: 10.1145/1518701.1518946.

[16] M. Prachyabrued and C. W. Borst, "Visual feedback for virtual grasping," in *Proc. IEEE Symp. 3D User Interface (3DUI)*, Minneapolis, MN, USA, Mar. 2014, pp. 19–26, doi: 10.1109/3DUI.2014.6798835.

[17] M. Prachyabrued and C. W. Borst, "Visual interpenetration tradeoffs in whole-hand virtual grasping," in *Proc. IEEE Symp. 3D User Interface (3DUI)*, Costa Mesa, CA, USA, Mar. 2012, pp. 39–42, doi: 10.1109/3DUI.2012.6184182.

[18] R. Canales, A. Normoyle, Y. Sun, Y. Ye, M. D. Luca, and S. Jörg, "Virtual grasping feedback and virtual hand ownership," in *Proc. ACM Symp. Appl. Perception*, Barcelona, Spain, Sep. 2019, pp. 1–9, doi: 10.1145/3343036.3343132.

[19] C. W. Borst and A. P. Indugula, "Realistic virtual grasping," in *Proc. IEEE Virtual Reality*, 2005, pp. 91–98, doi: 10.1109/VR.2005.1492758.

[20] C. K. Liu, "Dextrous manipulation from a grasping pose," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 1–6, Jul. 2009, doi: 10.1145/1531326.1531365.

[21] M. Moehring and B. Froehlich, "Effective manipulation of virtual objects within arm's reach," in *Proc. IEEE Virtual Reality Conf.*, Mar. 2011, pp. 131–138, doi: 10.1109/VR.2011.5759451.

[22] M. Holl, M. Oberweger, C. Arth, and V. Lepetit, "Efficient physics-based implementation for realistic hand-object interaction in virtual reality," in *Proc. IEEE Conf. Virtual Reality 3D User Interface (VR)*, Reutlingen, Germany, Mar. 2018, pp. 175–182, doi: 10.1109/VR.2018.8448284.

[23] S. Oprea, P. Martinez-Gonzalez, A. Garcia-Garcia, J. A. Castro-Vargas, S. Orts-Escolano, and J. Garcia-Rodriguez, "A visually realistic grasping system for object manipulation and interaction in virtual reality environments," *Comput. Graph.*, vol. 83, pp. 77–86, Oct. 2019, doi: 10.1016/j.cag.2019.07.003.

[24] J.-S. Kim and J.-M. Park, "Physics-based hand interaction with virtual objects," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Seattle, WA, USA, May 2015, pp. 3814–3819, doi: 10.1109/ICRA.2015.7139730.

[25] Y. Li, J. Huang, F. Tian, H.-A. Wang, and G.-Z. Dai, "Gesture interaction in virtual reality," *Virtual Reality Intell. Hardw.*, vol. 1, no. 1, pp. 84–112, Feb. 2019, doi: 10.3724/SP.J.2096-5796.2018.0006.

[26] D. Mendes, F. M. Caputo, A. Giachetti, A. Ferreira, and J. Jorge, "A survey on 3D virtual object manipulation: From the desktop to immersive virtual environments," *Comput. Graph. Forum*, vol. 38, no. 1, pp. 21–45, Feb. 2019, doi: 10.1111/cgf.13390.

[27] T. T. H. Nguyen, T. Duval, and C. Pontonnier, "A new direct manipulation technique for immersive 3D virtual environments," in *Proc. 24th Int. Conf. Artif. Reality Telexistence*, 2014, p. 8. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01083215.

[28] L. Aguerreche, T. Duval, and A. Lãcuyer, "Thierry and Lécuyer, anatole," in *Proc. Conf. EGVE*, 2009, pp. 153–156, doi: 10.2312/EGVE/JVRC09/153-156.

[29] D. Mendes, M. Sousa, R. Lorena, A. Ferreira, and J. Jorge, "Using custom transformation axes for mid-air manipulation of 3D virtual objects," in *Proc. 23rd ACM Symp. Virtual Reality Softw. Technol.*, Nov. 2017, pp. 1–8, doi: 10.1145/3139131.3139157.

[30] D. A. Bowman, E. L. J. J. Kruijff, and P. Ivan, *3D User Interfaces: Theory and Practice*. New York, NY, USA: Addison-Wesley, 2004. [Online]. Available: https://dl.acm.org/doi/10.5555/993837

[31] J. Jacobs and M. Stengel and B. Froehlich, "A generalized God-object method for plausible finger-based interactions in virtual environments," in *Proc. IEEE Symp. 3D User Interface*, Mar. 2012, pp. 43–51, doi: 10.1109/3DUI.2012.6184183.

[32] Potter, Leigh Ellen and Araullo, Jake and Carter, Lewis, "The leap motion controller: A view on sign language," *Proc. 25th Austral. Comput.-Hum. Interact. Conf., Augmentation, Appl., Innov., Collaboration*, Adelaide, CA, Australia, 2013, pp. 175–178. [Online]. Available: https://dl.acm.org/doi/10.1145/2541016.2541072

[33] Y. Du, S. Liu, L. Feng, M. Chen, and J. Wu, "Hand gesture recognition with leap motion," 2017, *arXiv:1711.04293*. [Online]. Available: http://arxiv.org/abs/1711.04293

[34] Y. Kim, S.-G. An, J. H. Lee, and S.-H. Bae, "Agile 3D sketching with air scaffolding," in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, Montreal, QC, Canada, Apr. 2018, pp. 1–12, doi: 10.1145/3170427.3186522.

[35] B. Lee, D. Lee, and S. Chin, "Structural motion grammar for universal use of leap motion: Amusement and functional contents focused," *J. Sensors*, vol. 2018, pp. 1–15, Oct. 2018, doi: 10.1155/2018/6073786.

[36] Stoyan, Kerefeyn and Stoyan, Maleshkov, "Manipulation of virtual objects through a Leap Motion optical sensor," *IJCSI Int. J. Comput. Sci.*, vol. 12, no. 5, p. 6, Sep. 2015.

[37] W.-Y. Xiao, J.-Z. Duan, and C.-Q. Xue, "Research on visual programming system of gesture interaction interface in VR," in *Proc. the 4th Int. Conf. Big Data Internet Things*, New York, NY, USA, Aug. 2020, pp. 38–42, doi: 10.1145/3421537.3421554.

[38] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the leap motion controller," *Sensors*, vol. 13, no. 5, pp. 6380–6393, May 2013, doi: 10.3390/s130506380.

[39] J.-N. Voigt-Antons, T. Kojiá, D. Ali, and S. Möller, "Influence of hand tracking as a way of interaction in virtual reality on user experience," 2020, *arXiv:2004.12642*. [Online]. Available: http://arxiv.org/abs/2004.12642

[40] J. Nordin. (2017). *Realistic Virtual Hands: Exploring How Appearance Affects the Sense of Embodiment*. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-137672

[41] M. Khademi, H. Mousavi Hondori, A. McKenzie, L. Dodakian, C. V. Lopes, and S. C. Cramer, "Free-hand interaction with leap motion controller for stroke rehabilitation," in *Proc. CHI Extended Abstr. Hum. Factors Comput. Syst.*, Toronto, ON, Canada, Apr. 2014, pp. 1663–1668 doi: 10.1145/2559206.2581203.

[42] Guerra, Martin-Gutierrez, Acevedo, and Salinas, "Hand gestures in virtual and augmented 3D environments for down syndrome users," *Appl. Sci.*, vol. 9, no. 13, p. 2641, Jun. 2019.

[43] B. P. Miranda, "Categorizing issues in mid-air Infovis interaction," *Proc. 20th Int. Conf. Inf. Visualisation (IV)*, Lisbon, Portugal, 2016, pp. 242–246, doi: 10.1109/IV.2016.70.

[44] J.-L. Lugrin, J. Latt, and M. E. Latoschik, "Avatar anthropomorphism and illusion of body ownership in VR," in *Proc. IEEE Virtual Reality (VR)*, Arles, France, Mar. 2015, pp. 229–230, doi: 10.1109/VR.2015.7223379.

[45] M. Mori, K. MacDorman, and N. Kageki, "The uncanny valley [From the Field]," *IEEE Robot. Autom. Mag.*, vol. 19, no. 2, pp. 98–100, Jun. 2012, doi: 10.1109/MRA.2012.2192811.

MICHIEL VAN VELDHUIZEN received the bachelor's degree in software engineering from the Windesheim University of Applied Science. He is currently pursuing the master's degree in software engineering with Shanghai Jiao Tong University. His current research interests include virtual reality and computer–human interaction.

XUBO YANG (Member, IEEE) received the Ph.D. degree in computer science from the State Key Laboratory, CAD & CG, Zhejiang University, in 1998. He is currently a Professor with the School of Software, Shanghai Jiao Tong University. His research interests include computer graphics, virtual reality, and human–computer interaction.

* * *