Intentional, accurate and natural object placement in virtual reality based neuropsychological assessment

Paul Chojecki[1], Mykyta Kovalenko[2], Detlef Runde[3], David Przewozny[4], Oliver Schreer[5]
Vision & Imaging Technologies
Fraunhofer Heinrich Hertz Institute HHI
Berlin, Germany
{paul.chojecki[1], mykyta.kovalenko[2], detlef.runde[3], david.przewozny[4], oliver.schreer[5]}@hhi.fraunhofer.de

***Abstract*—In this poster, we discuss the main challenges that need to be addressed to make immersive virtual reality (VR) assessment and training tools valid and usable for neuropsychological applications and clinical uses. We especially argue the importance of enabling natural, easy to use and precise interaction in order to tap into the full potential of VR advantages. We then address some of the main identified challenges and propose solutions and approaches for contact-free, hand-object manipulations for assessment and training VR applications. Our experiences are based on our research and development within the VReha project.**

Keywords—virtual reality, contact-free hand-interaction, natural user interfaces, neuropsychological training

I. Introduction

The increasing relevance of immersive virtual reality (VR) in neuropsychological clinical settings (e.g. for assessment and training), leads to the growing demand regarding natural interaction with the user's hands. Because current VR headsets allow realistic, high quality presentations of VR environments, users also expect object interactions in VR to be natural and realistic [1]. Furthermore, training and diagnostic tasks simulated in VR usually require a realistic representation of the real-world mechanics, providing high ecological validity and transferability of the results to the real world.

In our opinion, the use of game controllers is only an intermediate step on the way to natural, immersive and non-intrusive interactions. The aim is a contact-free detection of hand movements and gestural commands [2], without worn devices. This would enable precise, real-time representation of the users' hands and an easy to use interaction without additional cognitive load wasted on the recall of the controller's configurations.

II. current Situation

Commonly used devices for contact-free interaction in VR environments are currently the Microsoft Kinect® and the Leap Motion® (LM). The Kinect usually operates in an outside-in mode and is suitable for larger distances. The LM, which is typically mounted on head mounted devices, provides an inside-out mode for short distances and is currently the state-of-the-art solution for touch-less VR interaction [3]. However, there are still some issues concerning the use of the LM in professional settings, because it is designed for playful interactions [4]. In such game-based scenarios, a highly accurate, controllable and repeatable object interaction is not necessary and tracking errors are less important. Judging from our experience the LM programming interface has predefined interaction logics and physics, that mostly fail to meet the requirements in professional applications.

In the field of neuropsychological VR applications, errors caused by interaction technology frustrate and reduce user's acceptance of the overall system. In addition, interaction errors falsify the measurements in VR (e.g. processing times); and thus, falsify participant's performance outcomes.

III. Our Approaches

In the joint research project, VReha [5], we have developed various adaptations and extensions to overcome the shortcomings of the current standard LM implementations. Users should be able to grab, carry and release virtual objects of different sizes effectively and precisely, even if the objects are close to each other. Participants should also be able to perform the task (Fig. 1) quickly and effectively (i.e. handling errors, such as accidentally dropping objects, should not result in other previously positioned objects being moved or in time-consuming picking-up of objects from the ground). All these demands affect various hand interaction components, ranging from interaction design to gesture recognition algorithms. The challenges and solutions are listed below:

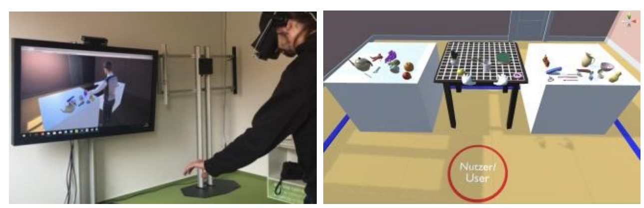

Fig. 1. Immersive visual memorization task in VR (left: first person view)

A. VR objects design and physics

Some common issues with the physics of VR objects occur, especially when releasing grabbed objects. In contrast to everyday real experiences, it is challenging in VR applications to release objects in a controlled way due to the lack of haptic feedback [6]. Furthermore state-of-the-art implementations of contact-free release commands, such as provided by LM, cause errors if the hand is not opened with all fingers fully extended at the same time. Belatedly extended fingers can accidentally give objects a little jolt during the release procedure leading to unintended object movement after the supposed release. This effect is unnatural, irritating and in neuropsychological applications, can lead to erroneous results. This is particularly relevant for spatial memory tasks, where participants are asked to place objects to memorized location.

This work is supported by the Federal Ministry of Education and Research (BMBF) of the Federal Republic of Germany (project VReha, grant no. 13GW0206A).

Note: This is not the official copyright released version of the IEEE proceedings paper. When citing this paper, use the following format: Kovalenko M, Runde D, Chojecki P, Przewozny D and Schreer O, "Intentional, accurate and natural object placement in VR based cognitive rehabilitation tasks", Proc. 13th Int'l Conf. on Virtual Rehab., WG Wright, S Subramanian, G Fluet, M Agmon, RM Proffitt, M Roberts (Eds), Tel Aviv, Israel, 21-24 July 2019.

Another issue arises with the object grabbing due to the computational cost of recognizing when a virtual hand is touching a virtual object [7]. An optimised bounding box around the object is needed to allow performant and precise object-hand interaction. A balance between size and level of detail of this bounding box is important. A large box simplifies the grab detection, but also increases the aforementioned release problem. Therefore, we optimised object bounding boxes for grabbing and object physics to stop any momentum after object release and ensuing unintended object movements.

B. Design of interaction logics and feedbacks

To improve the interaction logics and feedbacks of the common LM implementations, we:

- introduced a smart object colouring to indicate when an object is in the grabbing distance;
- set the hand representation to be semi-transparent, when grabbing or holding an object, thus the user is able to see even small grabbed objects. Although this implementation contradicts the everyday experience, it is expected to be intuitive and to lead to an easier execution;
- solved the error-prone grabbing of small objects by joining the object to the user's hand as soon as the hand is in grabbing distance and a grabbing intent is detected;
- introduced audio feedback when grabbing an object and when an object touches the table surface.

C. Error handling

In order to avoid time-consuming and unnecessary picking up of objects from the ground, if an object accidentally falls off from the table, the object jumps back to the table where it can be grabbed again easily. If an object falls on another object on the table surface, the already positioned object is not affected, thus its location does not have to be corrected again.

Another issue arises when the LM hand tracking fails and the virtual hand representation with the grabbed object is lost. Here, various solutions are possible: hand and object could disappear together and, as soon as the tracking works again, be faded in together again. Another solution approach is to use an additional sensor and take over the lost hand tracking and provide an uninterrupted interaction, as described in subsection *D.*

D. Hand tracking using artificial intelligence (AI)

Because LM has a limited working distance and field of view, it might lead to some tacking errors or incorrect hand representation when the hand leaves the field of view or extends away from the headset. Therefore, a visual tracking approach from the third-person outside-in perspective can also be employed. Specifically, the Kinect® sensor can be used for the hand tracking and interaction: both in tandem with LM to address its weaknesses, or just by itself.

However, Kinect does not have any built-in functionality for interaction, or, in our case, for the precise detection of grabbing and releasing gestures. That is why we also introduced a machine learning-based algorithm for video-analysis to accurately recognize various hand-postures associated with gripping and releasing objects of different sizes and shapes [8]. Specifically, we have developed a model that recognizes the static hand-gestures, as well as intermediary states between a release and a grab gesture (in a $[0; 1]$ range) for better and more precise control.

E. Self-representation by multimodal acquisition

LM implementations in VR applications allow users to see a representation of his hands, but no other body parts. This leads to irritations for inexperienced users, which are especially targeted in neuropsychological applications addressed here. To solve this problem and to implement a more realistic VR experience, we seamlessly fuse body tracking provided by Kinect® with LM hand tracking to generate a realistic, full body self-representation. Raw data of both sensors is filtered to avoid unnatural body and hand poses.

IV. Summary and Outlook

For the realization of neuropsychological VR procedures for assessment and training within the framework of the research project VReha, we developed a contact-free hand-interaction with self-representation for intuitive and user-friendly task executions, in which the user can concentrate on the actual task and is not distracted by irritating or unfamiliar experiences. To that end, we combined the Kinect and LM sensors, adapted the physics and interaction logics, introduced additional feedbacks and used artificial intelligence techniques to detect grab and release intentions. VR experiments using the abovementioned procedures in an immersive virtual memory task and a supermarket scenario are currently being developed.

References

[1] P. Chojecki and D. Przewozny, 'Herausfordserungen und Lösungsansätze der Interaktion und Selbstrepräsentation in VR-basierten, neuropsychologischen Diagnose- und Rehabilitationswerkzeugen', *znp*, vol. 29, no. 3, p. 182, 2018.

[2] A. Georgiadis and S. Yousefi, 'Analysis of the User Experience in a 3D Gesture-based Supported Mobile VR Game', in Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, New York, NY, USA, 2017, pp. 47:1–47:2.

[3] Z.-R. Wang, P. Wang, L. Xing, L.-P. Mei, J. Zhao, and T. Zhang, 'Leap Motion-based virtual reality training for improving motor functional recovery of upper limbs and neural reorganization in subacute stroke patients', Neural Regen Res, vol. 12, no. 11, pp. 1823–1831, Nov. 2017.

[4] J. Guna, G. Jakus, M. Pogačnik, S. Tomažič, and J. Sodnik, 'An Analysis of the Precision and Reliability of the Leap Motion Sensor and Its Suitability for Static and Dynamic Tracking', Sensors, vol. 14, no. 2, pp. 3702–3720, Feb. 2014.

[5] VReha project website, https://www.vreha-project.com/, (last visited: 2019-02-04)

[6] C.-M. Wu, C.-W. Hsu, T.-K. Lee, and S. Smith, 'A Virtual Reality Keyboard with Realistic Haptic Feedback in a Fully Immersive Virtual Environment', Virtual Real., vol. 21, no. 1, pp. 19–29, Mar. 2017.

[7] J. Jacobs and B. Froehlich, 'A soft hand model for physically-based manipulation of virtual objects', in 2011 IEEE Virtual Reality Conference, 2011, pp. 11–18.

[8] M. Kovalenko, S. Antoshchuk, V. Brovkov 'Becoming a Member of a Virtual Orchestra with Motion Recognition' in Proceedings of the XVI International Conference Culture and Computer Science, Berlin, 2018, pp. 131-145.