Seamless Hand-Based Remote and Close Range Interaction in Immersive Virtual Environments

Daniel Zielasko[1], Uta Skorzinski[1], Torsten W. Kuhlen[1], Benjamin Weyers[1]

Visual Computing Institute, RWTH Aachen University, Germany
JARA-HPC, Aachen, Germany[1]

Abstract

In this work we describe a hybrid, hand-based interaction metaphor that makes remote and close objects in an HMD-based immersive virtual environment (IVE) seamlessly accessible. To accomplish this, different existing techniques, such as go-go and HOMER, were combined in a way that aims for generality, intuitiveness, uniformity and speed. A technique like this is one prerequisite for a successful integration of IVEs to professional everyday applications, such as data analysis workflows.

1 Introduction

With the rise of consumer Head Mounted Displays (HMDs), the integration of Virtual Reality (VR) into everyday life is faster than ever before, for instance in gaming, training (Law et al., 2015), education (Borst et al., 2016), rehabilitation (Schultheis and Rizzo, 2001) and even data analysis (Zielasko, Bellgardt, et al., 2017). However, there is a large number of potential applications of VR where still more research is necessary to investigate whether the use of VR is beneficial in comparison with classic, e.g., desktop applications. In any case, more than an HMD and a 3D perception of the scene is necessary to evaluate this for a specific use case, because VR strongly benefits from multi-modal sensory input and interaction. Hence, the success of VR, e.g., in training may be explained by the fact that the motion sequence, which depends on the task, can be mirrored often to a very high degree in an immersive virtual environment (IVE). With this, the training transfer is improved, even when the task itself may be much faster solved when projected to a 2D problem and solved with a mouse. In cases where the mirroring of *natural* interaction does not obviously increase the overall performance, because the performance criteria is less anchored to reality or the *real* counterpart does not exist, such as in the analysis of abstract data (e.g., a network, represented as a node-link diagram) the benefits of a *natural* interaction, like grabbing, is more difficult to measure. However, there are study results suggesting that, e.g., a gesture system or grabbing can improve the user experiences in small controlled experiments (Huang et al., 2017). Whether this accumulates to a higher overall task performance, e.g., *better* answering a scientific question, is still hard to answer, as grabbing and thus *natural* hand-based close range interaction has disadvantages as well. An example is the limited arm reach and thus the necessity to travel to the point of interest (POI) you want to interact with (Bowman, Kruijff, et al., 2004) in a large map or data set. And when arrived, it is still not possible to stay there and do ad hoc comparisons to another POI. To tackle this, the go-go technique (Poupyrev et al., 1996) was developed, which extends the classic virtual hand metaphor (Bowman, Kruijff, et al., 2004) by stretching out the virtual hand and thus making remote objects reachable. Another possibility is to manipulate remote objects via a virtual pointer. Both approaches have advantages and disadvantages (Bowman, Kruijff, et al., 2004; Nedel et al., 2003) and, thus, hybrid solutions were created (see Section 2). In this work, we present a new hybrid method that enables close range and remote interaction. Our method merges the classical virtual hand metaphor with the go-go metaphor and a raycasted virtual pointer (see Section 3). In addition, the design has a strong focus on the use of uniform gestures and seamless mode transitions to maximize the learn ability, ease of use and error tolerance. The technique has been developed with the exploration of a large 3D graph in mind.

2 Related Work

While several egocentric interaction techniques have been proposed (Argelaguet and Andujar, 2013), and some have already been mentioned in the introduction, there are actually only a few virtual hand and virtual pointing hybrids. Maybe the most famous one is the *hand-centered object manipulation extending ray-casting* (HOMER) technique (Bowman and Hodges, 1997). This technique consists of two modes or phases. First, the user selects an object via raycasting. For this phase, the technique has to provide a trigger method, which is a challenge specifically in a hands-free interaction setting (Zielasko, Freitag, et al., 2015; Zielasko, Neha, et al., 2017). When the object is selected, a virtual hand is attached to the object that mimics the actual hand of the user. This technique brings more *natural* and direct manipulation to remote objects. Nevertheless, it might be uncommon and sometimes even difficult to point to very close objects. Additionally, with only a virtual pointer available, it can be difficult to select occluded objects. Furthermore, the manipulation gets more inaccurate with increasing distance. Therefore, the Scaled-World Grab technique (Mine et al., 1997) uses a different approach. Whenever the user grabs an object, the world around her is scaled down and scaled up again when the object is released. This is an excellent method for object placement. However, the scaling has to be done with caution due to cybersickness, and it does not work for infinite or circular maps, or in buildings. Furthermore, it might negatively influence the user's presence. Therefore, in this work we address the strengths and weaknesses of HOMER (Bowman and Hodges, 1997) by offering the virtual hand and virtual pointer metaphors simultaneously. This comes with the challenge of not confusing or overextending the user and avoiding the risk of unknown system states. In the following section we present our approach.

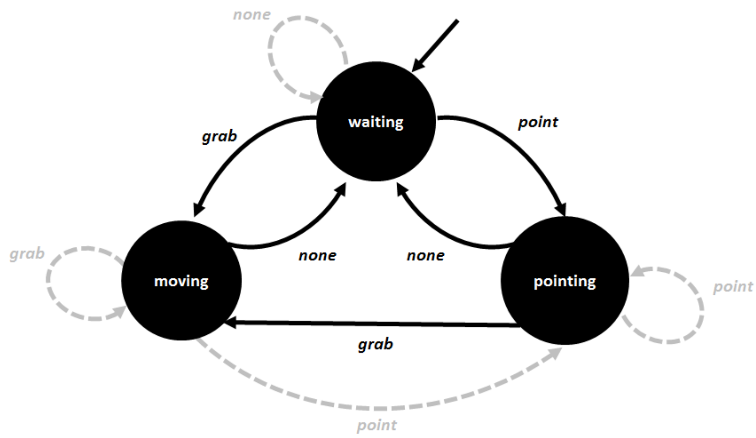

Figure 1: Interaction state machine. The gray transitions are not relevant in practice or for an implementation, but are depicted here to highlight that every transition condition always leads to the same state.

3 Seamless Remote and Close Range Interaction

HOMER (Bowman and Hodges, 1997) brings *natural* and direct interaction to remote objects maybe as close as it can get (see Section 2). But to reach this, it sacrifices some grade of naturalness and directness in the close range. This is the reason why we took a step back and decided to use the virtual hand and virtual pointer metaphor not sequentially, but simultaneously, such that the user can decide which technique to use in closer ranges, since it might feel better to actually grab an object when it is in reach, rather then first pointing at it. Sometimes it might even be hard to point to a close object as you need a minimal distance (length of the extended hand $+$ object radius $+ \epsilon$) to be able to point at it, where ϵ is the range you need to properly notice a pointer feedback (such as a ray). Thus, in our design, there are three instead of two modes or states, 1) the user grabbed an object and is actually moving/dragging it, 2) the user is pointing, or 3) the system waits for a transition to any of the other two (see Figure 1). This is not obviously better compared to HOMER, as an increase of states and thus, transitions usually increases the mental work load that is needed to use a system and the user may lose track of the current system state and how to leave it more easily. Therefore, the design introduces three types of transition conditions that behave uniformly in each state, i.e., each condition always leads to the same state independent of the current state (see Figure 1). For the user, this actually feels like a state free interaction, while internally those states may exist. The transition conditions are triggered by two gestures, *grabbing* and *pointing*. Formally, there is a third one, namely *none*, which is triggered in the absence of both of the others. The system description might read more complicated as it is. For the user it simply means, when I point at an object and close my hand it gets draggable and can be moved (or selected, see below) as long as my hand is closed, or when my hand is close to an object and I close my hand it gets draggable as well. Both interactions are possible at all times. Important to note is that the system does not need to detect hand gestures but instead only hand postures. Nevertheless, inaccurate tracking

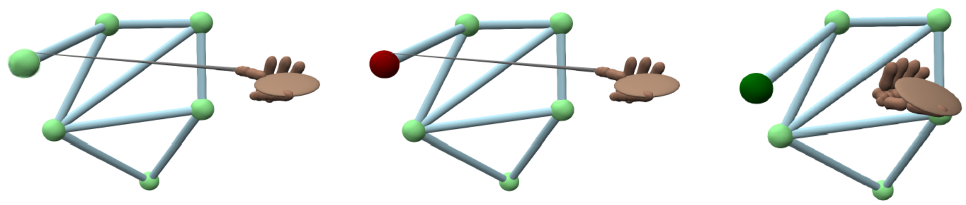

Figure 2: The user drags a distant object and selects it using a virtual pointer. From left to right, the user tries to point at a vertex in a graph. The pointer is active because the user's index finger is extended. The target is hit, here shown by the red focus color. Finally, the user moves in the index finger and thus grabs the object, highlighted with the dark green color.

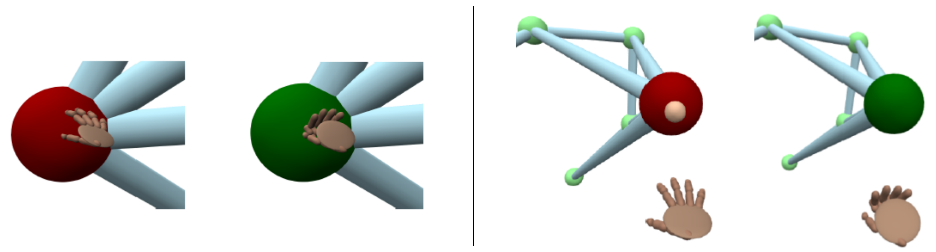

Figure 3: In the shown scenarios the user's index finger is not extended, thus nothing happens as long the user does not grab an object or extends an index finger (see Figure 2). ***Left****, the user's hand approaches a vertex and the vertex gets focused, highlighting it in red. When the hand changes into a grabbing posture and a vertex is focused, this vertex becomes draggable and green.* ***Right****, the user's hand has passed an extension threshold and from this point on, a sphere extends the hand's reach, w.r.t. the amount of threshold exceeding, while the virtual hand stays in the position of the real hand.*

and inter-user variability requires at least a minimum amount of trajectory evaluation and/or trajectory smoothing. This additionally makes the system more reliable, as it is only searching for two easily distinguishable postures. These postures are, either the hand is closed and all fingers of a hand are not extended (*grabbing*), or only the index finger is extended (*pointing*). All other hand posture are interpreted as no posture (*none*). Therefore, to grab a pointed at (focused) object, the user only has to move in the index finger to grab it and thus has not to open the hand and close it again to perform a complete grabbing gesture. Furthermore, an object gets focused and thus can be highlighted and grabbed, when the user points at it or a hand is close to it. An object that is additionally grabbed should be highlighted differently (see Figure 2 and 3). With the current design, objects can be focused and dragged. But sometimes the application requires that objects can be selected as well. In these cases, one could add a tapping gesture (Huang et al., 2017; Jang et al., 2015), or when a dragging operation is not required, one can interpret its incoming posture transitions as a trigger event. Additionally, it is interesting to note that the design allows a uniform and simultaneous interaction of both hands, i.e., ambidextrous users can use it without reconfiguration and the user does not have to think about which hand to take and can even use both hands at once.

3.1 Direct Go-Go & Redirected Virtual Hand

The described method already leads to a seamlessly reachable remote and close range interaction space, where an interaction space is the space the user can reach and interact with. However, the motor space (Argelaguet and Andujar, 2013), i.e., the space reachable with the hands and, thus the space in which we can benefit from direct interaction, stays small. To extend this a bit more, a variant of the go-go technique (Poupyrev et al., 1996) was additionally integrated. When the user's hand is close to her arm's maximum reach, a small sphere is spawned in the virtual hand and takes its role as a virtual pointer and starts leaving the hand to its pointing direction (see Figure 3, right). The distance to the hand thereby is linearly interpolated between 0 and a *maximum distance*, given the ratio from the threshold exceedance to its maximum, which is given by the physical arm reach. We did not translate the virtual hand and instead introduced the sphere to keep the user's virtual hand aligned with her real hand, to avoid interfering with the user's embodiment (Kilteni et al., 2012). Additionally, we did not choose the extension to be very large such that the user can use it precisely and fast. This comes with drawbacks, namely that this does not provide interaction to infinite distances, but that is OK our pointing does, and the user has to learn which objects are still reachable with this modified go-go technique and which are not and thus have to be pointed at instead. However, we think that by *cutting off* the original go-go technique in this way we can profit from its advantages, but simultaneously do not suffer from its disadvantages, such as the speed, when compared to pointing (Bowman, Kruijff, et al., 2004). Because even when distant objects with the go-go technique are reachable it takes a *long* time to steer the detached virtual hand, or in our case sphere, to its destination, to only then be able to interact with the object. In contrast, the direct mapping of the sphere's position to the position of the user's hand, which we are using, instead of a continuous control (Poupyrev et al., 1996) of the sphere, gives the user the opportunity to learn this behavior and directly grab for a distant object. This behavior raises another question. Why do not make this interaction uniform as well and introduce a gain g, which then always, i.e. from the beginning without first exceeding a threshold, translates the virtual hand g-times further than the real one, but without extending the user's threshold of noticing it? This is already done in other contexts and is called redirected reach (Suhail et al., 2017). The latter just came up while using a prototypical implementation of the described design and is really interesting to consider in the future.

4 Summary & Conclusion

In this work we described a hybrid, hand-based interaction metaphor that makes remote and close objects seamlessly accessible. To accomplish this different existing techniques were combined in a way that aims for intuitiveness, uniformity and speed. Thereby, the proposed design creates three different interaction spaces. The first is a little bit smaller than the user's motor space and the user can freely choose between pointing and grabbing in this space when interacting with an object. It is smaller than the motor space as a small part of it is used to reach the second space. The latter is larger than the motor space and due to a go-go behavior still reachable with grabbing. We describe those as two spaces as the behavior is not completely uniform between the two spaces. As soon as redirected reach with a linear behavior is used here, those two spaces are implicitly fused to one. The only disadvantage this could raise is that the precision is negatively affected, because the user has to control a larger space with the same motor resolution. The last and third space contains everything reachable with pointing. This metaphor together with, e.g., a free-hand travel technique (Zielasko, Horn, et al., 2016), can easily be included in a everyday desk-based analysis scenario (Zielasko, Weyers, et al., 2017) for immersive analytics, as a typical workflow includes different scales and both close and far away from the user. Although the first tests in this scenario worked well, the method obviously has to be formally evaluated to support our expectations.

Acknowledgments

The authors would like to acknowledge the support by the Excellence Initiative of the German federal and state governments, the Jülich Aachen Research Alliance – High-Performance Computing and the Helmholtz portfolio theme "Supercomputing and Modeling for the Human Brain". This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 720270 (HBP SGA1).

References

Argelaguet, F., & Andujar, C. (2013). A Survey of 3D Object Selection Techniques for Virtual Environments. *Computers & Graphics*, *37*(3), 121–136.

Borst, C. W., Ritter, K. A., & Chambers, T. L. (2016). Virtual Energy Center for Teaching Alternative Energy Technologies. *In Proc. of IEEE VR*, 157–158.

Bowman, D. A., & Hodges, L. F. (1997). An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. *In Proc. ACM Symposium on Interactive 3D Graphics*, 35–ff.

Bowman, D. A., Kruijff, E., LaViola Jr, J. J., & Poupyrev, I. (2004). *3D User Interfaces: Theory and Practice*. Addison-Wesley.

Huang, Y.-J., Fujiwara, T., Lin, Y.-X., Lin, W.-C., & Ma, K.-L. (2017). A Gesture System for Graph Visualization in Virtual Reality Environments. *In Proc. of IEEE Pacific Visualization Symposium*, 41–45.

Jang, Y., Noh, S.-T., Chang, H. J., Kim, T.-K., & Woo, W. (2015). 3D Finger CAPE: Clicking Action and Position Estimation under Self-Occlusions in Egocentric Viewpoint. *IEEE Transactions on Visualization and Computer Graphics*, *21*(4), 501–510.

Kilteni, K., Groten, R., & Slater, M. (2012). The Sense of Embodiment in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, *21*(4), 373–387.

Law, Y. C., Knott, T., Pick, S., Weyers, B., & Kuhlen, T. W. (2015). Simulation-Based Ultrasound Training Supported by Annotations, Haptics and Linked Multimodal Views. *In Proc. of Eurographics Workshop on Visual Computing for Biology and Medicine*, 167–176.

Mine, M. R., Brooks, F. P., Jr., & Sequin, C. H. (1997). Moving Objects in Space: Exploiting Proprioception in Virtual-environment Interaction. *In Proc. of ACM Computer Graphics and Interactive Techniques*, 19–26.

Nedel, L. P., Freitas, C. M. D. S., Jacob, L. J., & Pimenta, M. S. (2003). Testing the Use of Egocentric Interactive Techniques in Immersive Virtual Environments. *INTERACT*.

Poupyrev, I., Billinghurst, M., Weghorst, S., & Ichikawa, T. (1996). The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR. *In Proc. of ACM Symposium on User Interface Software and Technology*, 79–80.

Schultheis, M. T., & Rizzo, A. A. (2001). The Application of Virtual Reality Technology in Rehabilitation. *Rehabilitation Psychology*, *46*(3), 296.

Suhail, M., Sargunam, S. P., Han, D. T., & Ragan, E. D. (2017). Redirected Reach in Virtual Reality: Enabling Natural Hand Interaction at Multiple Virtual Locations with Passive Haptics. *In Pro. of IEEE 3DUI*, 245–246.

Zielasko, D., Bellgardt, M., Meißner, A., Haghgoo, M., Hentschel, B., Weyers, B., & Kuhlen, T. W. (2017). buenoSDIAs: Supporting Desktop Immersive Analytics While Actively Preventing Cybersickness. *Proc. of IEEE VIS Workshop on Immersive Analytics*.

Zielasko, D., Freitag, S., Rausch, D., Law, Y. C., Weyers, B., & Kuhlen, T. W. (2015). BlowClick: A Non-Verbal Vocal Input Metaphor for Clicking. *Proc. of ACM Symposium on Spatial User Interaction*, 20–23.

Zielasko, D., Horn, S., Freitag, S., Weyers, B., & Kuhlen, T. W. (2016). Evaluation of Hands-Free HMD-Based Navigation Techniques for Immersive Data Analysis. *Proc. of IEEE Symposium on 3D User Interfaces*, 113–119.

Zielasko, D., Neha, N., Weyers, B., & Kuhlen, T. W. (2017). A Reliable Non-Verbal Vocal Input Metaphor for Clicking. *In Proc. of IEEE 3DUI*, 40–49.

Zielasko, D., Weyers, B., Bellgardt, M., Pick, S., Meißner, A., Vierjahn, T., & Kuhlen, T. W. (2017). Remain Seated: Towards Fully-Immersive Desktop VR. *Proc. of IEEE Virtual Reality Workshop on Everyday Virtual Reality*, 1–6.