VE-SpaND: A Tangible User Interface using spatial understanding to navigate a 3D virtual environment

Sigurd Sæther Sørensen
School of Information Technology and Electrical Engineering
The University of Queensland, Qld., 4072, Australia

Abstract

This paper explores a tangible user interface (TUI) using peoples' spatial knowledge of the physical world to navigate a 3D virtual environment (VE) and whether it makes a useful alternative to conventional keyboard and mouse. Literature from TUI and navigation of VEs are presented, as well as studies conducted in the intersection of these; where few revolve around aligning peoples' spatial knowledge of the physical world with a TUIs functionality. The prototype VE-SpaND is presented, which was built based on data collected from a previous interview revolving peoples' spatial knowledge, their ability to navigate a 3D VE using conventional keyboard and mouse and their desired method of interacting with 3D VEs. Findings suggest that VE-SpaND aligns well with participants' spatial knowledge of the physical world and that users with limited previous experience of 3D VEs found it easier to navigate using the prototype.

1 Introduction

Over the last decade, our world has seen rapid development in the domain of virtual three-dimensional (3D) technology. In our pockets, we have devices capable of Augmented Reality and high definition 3D games. We have Virtual Reality headsets that we can connect to our computers, consoles and even our phones, which let us immerse ourselves in many a virtual world. Moreover, a wide array of industries ranging from education, design, construction and manufacturing to medicine and psychotherapy have adopted 3D VEs into their workflow [1]. With an ever-increasing number of applications for 3D technology across multiple platforms, a need for suitable domain-specific interfaces follows [1].

In the pursuit of such interfaces, researchers have looked to the field of TUI and related fields for solutions [2]–[5]. The hallmark of TUIs being that they provide a physical counterpart to digital information [6]. Despite researchers exploring various solutions, we still, for the most part, use conventional mouse and keyboards and handheld controllers to navigate 3D VEs. Burigat & Chittaro use Ruddle, Payne and Jones to argue that *"To navigate [a 3D VE] successfully, people must plan their movements using spatial knowledge they have gained about the environment and which they store as a mental map.";* and that such spatial knowledge is built slowly over time [1], [7]. As such, I believe there lie opportunities in bridging the gap between people's spatial understanding of the physical world and the navigation of 3D VEs. Therefore, throughout this study, I explore the following research question: *"Can user-centric tangible user interfaces provide a useful alternative for navigating 3D virtual environments?"*.

2 Background

This study draws upon literature in the spaces of TUI and navigation of 3D VEs, as well as studies on specialised input devices for navigating 3D VEs.

The term TUI was first coined by Ishii and Ullmer in 1997 and was revisited a decade later by Ishii in his paper Tangible Bits: Beyond Pixels [6], [8]. Tangible Bits introduce the fundamental principles and properties of TUIs and highlight their design challenges and advantages compared to conventional Graphical User Interfaces. One significant TUI advantage is the Double Interaction Loop, where the first loop provides immediate haptic feedback in the physical realm, followed by the second loop consisting of digital feedback. Meaning that users won't have to wait for digital computations to determine the outcome of their actions.

Where Ishii explore the benefits of a TUI when used, Bakker and Niemantsverdriet look at what levels of attention people can interact with a device [9]. They introduce the Interaction-Attention Continuum with three levels of attention, focused- peripheral- and implicit interaction; spanning from fully focused attention to entirely outside the field of attention. Additionally, the authors discuss examples of how a device can offer various levels of control spanning the continuum.

In the pursuit of designing user-friendly solutions, Benford *et al.* present a framework for designing expected, sensed and desired interactions. This framework considers what movements a computer can sense and to what degree it overlaps with what we have come to expect as users and the movements we desire.

As for navigation in 3D VEs, Burigat and Chittaro study the use of navigational aids and highlight how they benefit inexperienced and experienced users differently [1]. Also, they discuss spatial navigation and sensory feedback of our physical world and how it differs when experiencing a 3D VE through a screen.

Where Burigat and Chittaro explore navigational aids, Ousland and Turcato look at how people interact with- and navigate 3D VEs. They emphasise the adverse effects of input delays on user experience, user issues on remembering button combinations and that special 3D input devices are not required.

Despite this, researchers keep exploring alternative solutions of navigating VEs. One such interface is ZeroN, a levitating TUI capable of six degrees of freedom [2]. Lee *et al*. explore the challenges of- and opportunities made available by a levitating TUI that is not constrained to two-dimensional flat surfaces. Whereas it possesses unique interactions such as leaving a physical artefact in mid-air, they also found, due to a lack of physical relationships between objects, a decreased legibility of interaction.

Leflar and Girouard went in a different direction, exploring navigation of 3D VEs using a flexible handheld device. They found that participants preferred using the device similar to how they would use a console controller and that users became frustrated when the physical bend direction did not align with the direction of movement on the screen.

These works make up the pillars of this study, providing a solid foundation of TUIs, navigation in 3D VEs, and studies conducted in the intersection. Whereas Ousland and Turcato suggest that there is no need for specialised 3D input devices, they also give notice to issues revolving input delay, which the double feedback loop of a TUI, as explained by Ishii, could circumvent with immediate haptic feedback. Moreover, a TUI can help bridge the gap between what is expected and desired through using peoples' understanding of the physical world when navigating 3D VEs. By carefully mapping a controller's functionality, aligning it with our knowledge of the physical world, and implementing navigational aids, a TUI could offer interaction at various levels of attention, benefiting both experienced and inexperienced users. In constructing a TUI, it bears to remember the challenges and opportunities previously faced by similar projects. Among these is a lack of physical relationship between objects when elevated, as presented by Lee *et al.* and screen movement-controller mismatch as found by Leflar and Girouard.

3 Discussion

To explore the research question, I conducted a study using the research through design approach. A total of eight participants were recruited based on the following criteria; people who have no previous experience using a 3D software; and people with limited experience using 3D software who either actively use or have previously used 3D software. The former group would provide value by being unbiased and rate a solution's usability to navigate 3D VEs based on their conceptualisation formed on their existing knowledge of our physical world. The latter group would provide a general familiarity of 3D VEs and a frame of reference to critically review an interface per its efficiency performing desired tasks. All participants resides in Australia and comes from different cultural backgrounds in Europe, North America, Asia and Oceania.

In an initial interview, I set out to explore the participants' spatial knowledge of the physical world and 3D VEs and how these aligned. I also observed how participants naturally navigated using a conventional keyboard and mouse in the software Unity, and handed them a set of shapes of which they used to explain their understanding and to illustrate how they ideally would like to navigate. Based on the interview results, I constructed the prototype VE-SpaND (**v**irtual **e**nvironment **spa**tial **n**avigation **d**evice). VE-SpaND acts as a tangible token of the virtual camera in Unity, which decides what part of the 3D VE the user can see on their computer screen.

VE-SpaND consists of four parts, a sphere, base, projector and camera. The sphere sits loosely on top of the base, which together forms the physical controller that users can interact with, which represents the virtual camera. The projector is mounted to look down from above and projects a top-down view of the virtual camera's surrounding area to give context to the virtual camera. The physical camera is attached on top of the projector to have the same perspective as the projector. The camera's job is to track the position of the physical controller as it moves on the surface of the table.

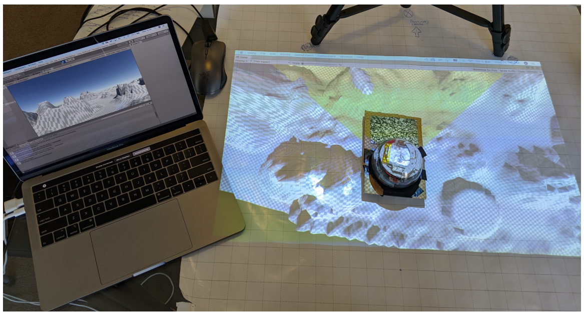

Fig 1. VE-SpaND

The prototype utilises the Unity library Vuforia and a Logitech C9-series web camera to track the planar movement of the physical controller. The camera is attached to an Asus projector placed opposite to the user, pointing down towards the tabletop. In the centre of the sphere of the physical controller, an MPU-6050 3-axis gyroscope + accelerometer using the MPU6050_tockn library collects and sends rotation data to an Adafruit Huzzah32 microcontroller. A 3.7v 1500mAh LiPo battery powers the Huzzah32 which forwards the data over Bluetooth to a 13' 2019 edition MacBook Pro that catches incoming data in a Unity script.

VE-SpaND is capable of horizontal movement in all directions, where the further away from the centre a user moves it, the more velocity it gains moving in said direction. When the physical controller is moved to the centre area, all camera movement stops. The sphere attached to the physical controller allows users to rotate the virtual camera on the x-, y- and z-axis (roll, pitch and yaw). VE-SpaND is, due to technical issues, currently not capable of vertical movement, where the intention is for users to lift or press the physical controller to move vertically.

Due to restrictions put in place by the Australian government in response to COVID-19, I chose to use two separate methods for data collection to avoid dangerous situations and ethical issues. For those participants where meeting- and conducting in-person studies were in line with restrictions, I opted to run an in-person prototype test and interview. Where in-person meetings were not in line with regulations or deemed unethical, I instead opted for a second approach. Here I sent participants a digital folder containing videos, images, open-ended questions and instructions for them to look at and then respond. I interviewed three participants using the first method, consisting of two non-users and one regular user. I sent the folder to Four participants, two regular users and two non-users, and the last participant, a regular user, was unable to participate at this stage of development.

Whereas an obvious limitation with this study is that fewer than half of the participants were allowed physical access to the prototype, a set of intriguing findings still emerged. When asked how it felt to navigate using VE-SpaND, two regular users said it felt fine to use, and a third thought it would take some getting used to. As for non-user, three participants felt that VE-SpaND made more sense to them, where one person said: *"This new way of navigation feels natural and maps to the way I would be thinking of navigating around a 3D space."*. The last non-user said it seemed fluid but that the projected screen lagging behind the physical movement would distract a lot. Similar data came through when I asked participants which alternative they would prefer, conventional keyboard and mouse or VE-SpaND, and why. Three non-users highlighted how the prototype makes more sense having no previous experience and aligns well with the way they think, whereas a fourth would pick keyboard and mouse due to familiarity. As for the regular users, two said they would prefer using VE-SpaND, albeit one of them suggested it would not do well in environments that would expand or require a lot of vertical movement. The third regular user would pick a keyboard and mouse.

In response to whether VE-SpaND would prove most beneficial by itself or in combination with existing input devices, all participants said it would do best in combination with a keyboard. The reason of which was how a keyboard would provide additional functionality such as keybinds, where the prototype could handle all navigation. Moreover, when pointing out the advantages and disadvantages of both alternatives, a majority of participants highlighted how VE-SpaND brought everything to one hand and that it is intuitive to use. As for the keyboard and mouse, participants commented on its precision, familiarity and its space efficiency, as it does not require a projector setup.

In terms of what alternative would best help them understand their 3D position, two non-users stated how the prototype made more sense, where one said they didn't suddenly find themselves under the environment's surface or inside mountains using the prototype. Two regular users mentioned how *"...the prototype gives a better overview with the minimap."* and *"The prototype seems like it is very connected*.". Moreover, three participants, one regular user and two non-users said both alternatives were fine or about the same, where one non-user mentioned that understanding one's 3D position is not solely reliant on the controller, but also environmental cues. When asked which alternative would make for a better option for teaching beginners, all participants decided on the prototype. One user stated that *"To learn 3D software, I would definitely go with the prototype. One reason I haven't properly attempted 3D design or software is because of the learning curve."*. Other participants mentioned how the prototype would not overwhelm users with options; it would be easier to learn and move around and could act as a stepping stone.

In addition to the findings above, participants made several suggestions to make for an overall better experience- and others made a note of issues using VE-SpaND. One user mentioned that the device would be confusing to use in expanding projects as it seems to be limited to a particular area. Where one user would like a one-to-one relationship when rotating the ball, another made an argument for continuous rotation after a point to avoid uncomfortable hand and wrist movements. Other participants commented on the ball sitting too loose, which made for some unwanted movements if they did not hold it in place. Other suggestions included adding buttons to the device to remove the need for a keyboard, a built-in multi-lens projector for increased space efficiency and in the future to focus on the ergonomics of the device. Moreover, several remote users commented that it was difficult to say anything for certain without physically being able to interact with the prototype.

These findings, where a majority of responses highlights the prototype's ease of use, suggest that VE-SpaND aligns well with people's mental map, the importance of which is argued by Burigat and Chittaro in terms of successfully navigating a 3D VE [1]. Moreover, although well-received by regular users and non-users alike, VE-SpaND seems to have had a more considerable impact on non-users' ability to navigate given their responses on how it makes more sense for them to use. Thus, current data show that a user-centric tangible user interface does make for a useful alternative for navigating 3D VEs. With that said, it bears to notice that these results might be different in a situation where all participants can interact with the prototype physically. Also, it's worth considering the prototype's inability to move in vertical lines and how this may have affected the outcome of this study.

Although efforts were made to design VE-SpaND according to frameworks and guidelines from the literature, there are still improvements to be made. I suggest that future work should explore how the design could better align with Ishii's double feedback loop to avoid unwanted effects such as the current lag [6]. Furthermore, I suggest studies should look into the ergonomics of the device and how users can operate such a device at various levels of attention, as some differences in use were observed, but this was not a focus area for this study [9]. Also, I suggest looking further into how this prototype works with- or could replace other input devices, given participants responses on using VE-SpaND in combination with a keyboard and how adding keys to it could replace the keyboard altogether. Lastly, further longitudinal studies revolved around whether the device affects a user's'ability to learn and navigate a 3D VE should be conducted to attain over-time usage data.

Acknowledgment

I would like to thank all of those who volunteered their valuable time to participate in my study, whom without I would not succeed. Moreover, thank you Thomas Vincent Saly for offering countless hours of technical assistance, problem-solving and creativity; Liv Karoline Torp Olsen for making sure I stayed on the correct path; and Lorna Macdonald for valuable insight and supervision throughout this study.

References

[1] S. Burigat and L. Chittaro, "Navigation in 3D virtual environments: Effects of user experience and location-pointing navigation aids," *Int. J. Hum.-Comput. Stud.*, vol. 65, no. 11, pp. 945–958, Nov. 2007, doi: 10.1016/j.ijhcs.2007.07.003.

[2] J. Lee, R. Post, and H. Ishii, "ZeroN: mid-airztangible interaction enabled by computer controlled magnetic levitation," in *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*, Santa Barbara, California, USA, 2011, pp. 327–336, doi: 10.1145/2047196.2047239.

[3] V. P. Nguyen, "Soft tangible interface for continuous control of virtual objects with pressure based input," M.S.E., Purdue University, United States -- Indiana, 2014.

[4] M. Leflar and A. Girouard, "Navigating in 3D space with a handheld flexible device," *Entertain. Comput.*, vol. 5, no. 4, pp. 205–209, Dec. 2014, doi: 10.1016/j.entcom.2014.07.002.

[5] B. Fanini, "A 3D Interface to Explore and Manipulate Multi-scale Virtual Scenes using the Leap Motion Controller," in *The Seventh International Conference on Advances in Computer-Human Interactions*, Barcelona, Spain, 2014, [Online]. Available: https://www.researchgate.net/publication/274832895_A_3D_Interface_to_Explore_and_Manipulate_Multi-scale_Virtual_Scenes_using_the_Leap_Motion_Controller.

[6] H. Ishii, "Tangible Bits: Beyond Pixels," in *Proceedings of the 2Nd International Conference on Tangible and Embedded Interaction*, New York, NY, USA, 2008, pp. xv–xxv, doi: 10.1145/1347390.1347392.

[7] R. A. Ruddle, S. J. Payne, and D. M. Jones, "Navigating buildings in 'desk-top' virtual environments: Experimental investigations using extended navigational experience," *J. Exp. Psychol. Appl.*, vol. 3, no. 2, pp. 143–159, 1997, doi: 10.1037/1076-898X.3.2.143.

[8] H. Ishii and B. Ullmer, "Tangible bits: towards seamless interfaces between people, bits and atoms," in *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '97*, Atlanta, Georgia, United States, 1997, pp. 234–241, doi: 10.1145/258549.258715.

[9] S. Bakker and K. Niemantsverdriet, "The Interaction-Attention Continuum: Considering Various Levels of Human Attention in Interaction Design," *Int. J. Des. Taipei*, vol. 10, no. 2, p. n/a, Aug. 2016.

Biography

Sigurd Sæther Sørensen is a Norwegian with a bachelors degree in international marketing and a year-study in information technology from the University of South-east Norway and is currently studying master of interaction design at the University of Queensland.