Terrain Modelling with a Pen & Touch Tablet and Mid-Air Gestures in Virtual Reality

Fabrice Matulic
Preferred Networks Inc.
fmatulic@preferred.jp

Daniel Vogel
School of Computer Science, University of Waterloo
dvogel@uwaterloo.ca

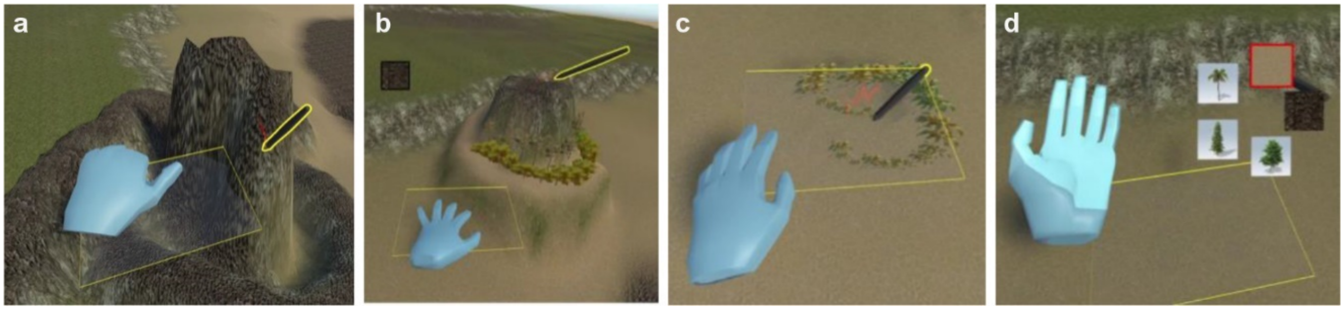

Figure 1: Terrain editing in VR with bimanual pen+touch+mid-air interaction: Sculpting mountains with the pen while moving the terrain plane with the other hand using mid-air gestures (a) and dragging on the tablet (b); Adding trees with the pen on the tablet and controlling their height using pressure (c); Invoking a mid-air marking menu by turning the palm of the other hand up and selecting menu items with the pen (d).

ABSTRACT

World building or terrain modelling is an essential task when designing games, natural simulations or artistic creations involving virtual 3D landscapes. To support this task, we propose a virtual reality (VR) system based on a pen and touch tablet used in a sitting position (desktop VR) such that both hands are free to interact in an asymmetric way (pen hand + other bare hand). We present and compare several techniques to perform navigation, sculpting and menu operations using the two hands, which interact on and above the tablet surface, i.e. using the pen, touch and mid-air input spaces. A qualitative evaluation with 16 participants confirms the entertaining nature and practical benefits of our system. The study further underlines the complementarity of the different modalities and identifies the promising—and as of yet underexplored—combination of bimanual touch + pen + mid-air interaction in desktop VR.

CCS CONCEPTS

• **Virtual reality**; • **Interaction techniques**;

KEYWORDS

pen and touch interaction, VR modelling

CHI '22 Extended Abstracts, April 29–May 05, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9156-6/22/04…$15.00
https://doi.org/10.1145/3491101.3519867

ACM Reference Format:
Fabrice Matulic and Daniel Vogel. 2022. Terrain Modelling with a Pen & Touch Tablet and Mid-Air Gestures in Virtual Reality. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22 Extended Abstracts), April 29–May 05, 2022, New Orleans, LA, USA*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3491101.3519867

1 INTRODUCTION

3D open-world games and sandbox creation environments enjoy wide popularity. To design these virtual worlds, tools that efficiently—and for gamers also playfully—support terrain modelling are required. While most world editors are designed for traditional mouse and keyboard computer environments, the immersive nature of those worlds and the proliferation of virtual reality (VR) devices make it possible to consider more intuitive and engaging alternatives. The popular 3D game developing platforms Unity [28] and Unreal Engine [29] as well as sandbox games such as Minecraft [30] and Horizon Worlds [18] include tools to create virtual worlds directly in VR. Interaction relies on standard controllers or mid-air hand gestures for input, which is generally sufficient for coarse editing, but may become tedious when working at a finer level of granularity, in particular when the user wants to precisely "sketch" features into the landscape. Sketching or drawing is best performed with a pen and systems using pens and tablets in VR have been proposed to support such tasks in mobile and standing contexts [2, 6, 8]. Although they do not target world building and terrain editing, these systems show how the pen can be used for precise input both *on* the tablet and *above* it, in mid-air. In these contexts, the tablet can be moved around to situate input in the 3D space, but because the device is held in one hand, contact/surface input is mostly limited to the pen. As a result, the two hands cannot be used together for touch input and thus no bimanual pen+touch synergies are possible to enhance the interaction vocabulary like in stationary desktop environments [3, 12, 16, 27]. VR systems using fixed horizontal touch surfaces exist, but either their mid-air interaction capabilities are limited, or they do not use pens. Furthermore, none of those previous systems, mobile or stationary, support terrain modelling using asymmetric bimanual pen and barehand interaction.

In this work, we propose a novel desktop VR system based on a pen+touch tablet and a set of techniques involving both hands interacting on the tablet surface as well as above it, to support terrain sculpting. We consider different combinations of the pen-holding hand and the other hand interacting both in the same space (surface or mid-air) or each in a different space. To explore the potential of these different hand and input space combinations, we create several alternative methods for common operations such as navigation, invoking menus, changing the relief and topography of the scene and inserting terrain details. We conduct a qualitative evaluation to compare the different versions of our techniques and based on participant feedback we attempt to derive more general insights about bimanual pen+barehand input for pen+touch+mid-air interaction in VR that hopefully can inform the design of other applications in different domains.

2 RELATED WORK

As immersive environments, virtual reality systems lend themselves to creative design in 3D space such as modelling and sketching. In most consumer-focused VR applications, the user is standing and interacts with virtual content using handheld controllers or their bare hands in mid-air, but tablets and pens have also been considered for surface-bound and precise input: TabletInVR is a touch-based tablet system with a set of multitouch gestures to perform modelling operations on 3D cubes, where the physical position and orientation of the tablet can also be used to define extrusion directions and to cut through virtual volumes [22]. SymbiosisSketch [2] and VRSketchIn [6] are two tablet and pen-based VR systems that allow the user to define virtual surfaces in 3D space mapped to the tablet display for precise sketching on a physical surface. They also support unconstrained mid-air input for freehand sketching in the VR space. In these systems, the tablet is used in mid-air and requires a hand to hold it. In our setting the tablet is placed on a desk and thus allows both hands to participate directly in 2D and 3D input.

VR systems for large fixed horizontal surfaces, where the space above them is available for visualisation of 3D content as well as mid-air input using both hands, have also been proposed. Mockup Builder is a stereoscopic tabletop system for 3D modelling with asymmetric bimanual touch and mid-air interactions [1]. Shapes can be sketched and extruded with the dominant hand (DH) while the non-dominant hand (NDH) invokes menus, moves objects and sets constraints for extrusions. The stereoscopic nature of the system and the fact that only the index finger and the thumb of both hands are tracked limit the types of mid-air interactions and likely reduce the sense of immersion compared to modern VR environments. Furthermore, those bimanual gestures are mostly independent or sequential. Gesslein et al. developed a spreadsheet application for pen-based tablets in VR, where the augmented space above and around the device can show data in 3D as well as layered menus [9]. The application is very pen-centric and the NDH is only used to sporadically perform air clicks with the index finger.

With regard to world building, commercial 3D games and game engines with world-creation features such as Unity [28] and Unreal Engine [29] include VR editors operated using standard VR controllers. These tools focus on element selection and placement and do not include dedicated terrain sculpting functionality like in their 2D desktop counterparts. Other research prototypes to build scenes [14] and create road networks for the simulation of traffic scenarios [7] use tracked hands, but similarly their interfaces are designed for picking, placing and arranging elements in a 3D scene, not for terrain sculpting. DIY World Builder is a system which supports terrain editing and texturing in a room-scale environment that includes audio, wind and haptic feedback [25]. Operations are performed with a wand and a touchscreen worn on the forearm, a configuration that precludes bimanual touch interaction. VRBox is an augmented sandbox system that allows users to form landscapes using real sand [8]. Terrain properties can be modified by bringing floating spheres over the area of the sandbox to edit. While DIY World Builder and VRBox both target terrain editing in VR, these systems are more focused on the enhanced immersive experience using various forms of haptic feedback. Such setups require additional sensors and equipment, which take up a large amount of space. Our approach concentrates solely on input and utilises a different interaction environment in a fixed desktop setting with a pen+touch tablet. We take inspiration from previous work on (non VR) binanual pen+touch input, which has shown that the two modalities can be used together to create powerful hybrid combinations, in particular for quickly switching the pen mode using NDH postures [3, 12, 16, 27]. We consider the extension of that interaction space to mid-air and VR to design tools for terrain editing that make use of two-hand and cross-space interactions.

3 APPLICATION

Our prototype is based on Unity's terrain editor with which many game developers are familiar with. The editor allows the user to locally raise or lower the terrain level by pointing and dragging with the mouse on the 3D scene. Details such as textures, trees and grass can be added by painting them on the landscape surface. We adapt these basic tools for a 3D VR interface used with pen+touch+mid-air input.

A popular framework adopted by asymmetric bimanual interfaces, and especially pen+touch systems, is Guiard's kinematic chain model [10]. This model identifies a division of labour between the dominant hand (DH) and the non-dominant hand (NDH), in which the latter sets the frame of reference. For instance, the NDH positions and adjusts the canvas for the DH to execute the main task (sketching or editing). We largely adopt this principle when designing our techniques for our augmented pen+touch+mid-air input space, but depart from it when considering simultaneous editing and navigation.

We explore different alternatives of techniques in that input space to perform three types of basic operations: navigation, sculpting and menu invocation. Furthermore, we include a method to reset the scene as well as palm rejection to avoid unwanted touch

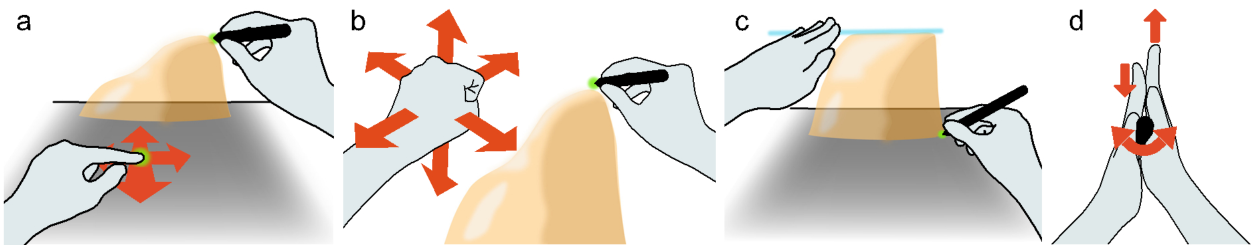

Figure 2: Editing interactions: (a) Sculpting the terrain with the pen while moving the camera by dragging on the tablet with the NDH; (b) Sculpting the terrain with the pen while moving the camera using mid-air grab-and-release NDH gestures; (c) Sculpting the terrain on the tablet with the pen and the height of the terrain defined by the NDH in mid-air; (d) Rubbing the pen between the hands to reset the scene.

input from the DH. The system and all techniques are shown in the accompanying video.

3.1 Pen and Hand Representation

In our implementation, the NDH is rendered as a virtual hand model, while for the DH only the pen is shown (Figure 1). When illustrating interactions in this paper, we also represent the DH holding the pen for increased clarity. The tablet's touch input space is shown as a yellow contour in VR to minimise occlusion with the content. The pen is equipped with a barrel button to engage input in mid-air. In VR, the virtual pen model "glows" when the button is pressed to provide visual feedback to the user.

3.2 Palm Rejection

Our prototype supports palm rejection, which detects and discards unwanted touch input when the palm of the pen-holding hand contacts the surface. We use the position and orientation of the pen to reject touches occurring within a rectangle covering the whole hand and arm and whose orientation aligns with the pen, a simplified version of Vogel et al's occlusion model [24].

3.3 Reset Gesture

We include a gesture to delete all content and reset the application as an example of an action that is rarely performed but has a high penalty if accidentally triggered. We choose a bimanual gesture that requires both hands to manipulate the pen: The palms are brought together in a "praying pose" with the pen clasped in-between. Then, the pen is spun back and forth by rapidly rubbing the hands together (Figure 2d). Detection examines a 2-second window of accumulated pen rotations about its longitudinal axis, and executes the delete/reset action when a threshold is exceeded. To avoid false positives, a threshold that can only be achieved with two hands is chosen.

3.4 Terrain Editing

Terrain features can be added with the pen interacting on the tablet or above it, in mid-air. When in mid-air, pressing the pen button modifies the terrain at the pen location in 3D space according to the currently selected pattern (shown in the top left corner of the viewport) (Figure 1b). We support two types of terrain patterns: land (2 types) and trees (3 types). When a land pattern is selected (using a mid-air marking menu described below), mountains and valleys are formed, with the elevation of the relief corresponding to the height of the pen. Trees are added on the terrain surface at the projection point of the pen direction. The pen button can be maintained pressed while the pen is waved over the terrain like a magic wand to continuously sculpt the topography.

For more precise editing, the pen can be used directly on the tablet, i.e. the user can place trees or make fine adjustments to the landscape by tapping and tracing on the surface. The height of a terrain feature can be controlled in two ways: If the NDH is in mid-air, it is matched to the hand height (Figure 2c). If the hand is resting on the tablet or is out of tracking range, the pressure of the pen on the tablet is used to modulate land height or tree size (Figure 1c).

3.5 Navigation

Like editing, navigation can also be performed using touch or mid-air input. When using touch, the user can move the camera to navigate through the scene by dragging with fingers on the surface. One finger pans, two rotate, and three move the camera vertically (dragging upwards/downwards on the tablet moves the camera up/down). We use simple absolute movement mappings without any inertia, i.e. the effect is like moving a mouse cursor with a trackpad. Alternatively, navigation can be performed with the NDH in mid-air. Grabbing (clenching the fist) engages navigation mode. The camera can then be moved and rotated according to the hand motion. Releasing the grab disengages the mode. The user can navigate with repeated grabbing and pulling, similarly to Coomer et al's point-tugging technique [5], but using the bare hand instead of a controller. Since degrees of freedom are separated when navigating with touch while the NDH in mid-air controls all DoF at once, we add the possibility to constrain mid-air movements to panning only, by extending the thumb. We use the same gain function translating dragging and mid-air motions to camera movements so that the two navigation methods can be fairly compared.

3.6 Editing While Navigating

The user can edit the terrain while moving the camera using any combination of the editing and navigation methods, i.e. maintaining the pen button pressed in mid-air or tracing on the tablet while grabbing-pulling in mid-air or dragging on the tablet with the NDH

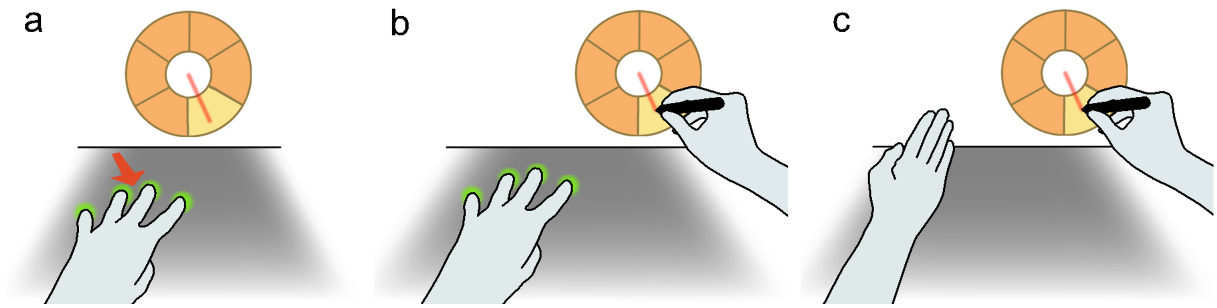

Figure 3: Different methods to invoke the marking menu with the NDH and select menu items: (a) A four-finger touch invokes the menu in the middle of the screen, a menu item is selected by dragging; (b) A four-finger touch invokes the menu at the pen location, a menu item is selected with the pen in mid-air; (c) Turning the NDH with the palm facing upwards invokes the menu at the pen location, a menu item is selected with the pen in mid-air.

(Figure 2a and b). This technique may violate the kinematic chain model if the NDH performs the main movement to sculpt the terrain (see video), but it allows to rapidly add terrain elements over large areas of the scene. We believe this novel way to edit and sculpt landscapes in VR is a particularly powerful example of barehand + pen input, which we surmise would be more difficult to achieve with two controllers.

3.7 Marking Menu

The current terrain editing pattern can be changed by invoking a marking menu. Here as well we support several techniques. The first method requires touching the tablet with four fingers of the NDH to display the menu in the middle of the screen, then dragging towards the desired item and releasing the fingers to select (Figure 3a). The cursor is anchored to the rightmost finger (if the NDH is the left hand) so that gesture relaxation [26] is possible by dragging only with that finger after triggering the menu with four fingers. The second method (which requires a manual switch to be enabled) follows a bimanual NDH trigger + pen marking action sequence. A four-finger touch with the NDH causes the menu to appear at the tip of the pen (Figure 3b). An alternative invocation method in mid-air is to turn the NDH palm facing up (Figure 3c). The pen is then moved inside the desired menu item and the fingers of the NDH released or the palm turned down to confirm the selection. While bimanual marking menus have been previously proposed, they were for non-VR multitouch interfaces and mostly to support multi-level menus [13] or higher integration with tracing tasks [11, 19], so we believe these techniques for VR marking menus to be novel.

4 APPARATUS

Our prototype is implemented in Unity running on a Windows PC with a GeForce GTX 1080 Ti GPU. Our system uses an HTC Vive Pro for the VR display and a Wacom Intuos Pro L with an active area of 311×216 mm for the tablet. An Optitrack motion capture system precisely tracks the headset, tablet, and pen in the 3D space. The required reflective markers are mounted on custom 3D printed mounts. The Wacom tablet's sensors only detect pen barrel button presses when the pen is in contact with the tablet. To achieve mid-air button press detection, we repurpose the internal hardware of a Bluetooth remote shutter and attach it with a resin sheath on the pen barrel. The entire mount affixed to the pen adds respectively 27 g and 8.6 cm to its original weight of 18 g and length of 15.5 cm. Since it is not possible for our motion capture system to track non-rigid bodies like hands, we use a Leap Motion sensor fixed to the front of the headset to track the NDH. While it would have been desirable for the DH to be tracked as well to be rendered in VR, the Leap Motion is not able to robustly track a hand holding a pen.

5 USER STUDY

We conducted a qualitative user evaluation to assess the usability and suitability of our techniques. The main goal of this investigative study was to obtain high-level, spontaneous feedback about the use of bimanual hybrid gestures combining surface and mid-air input rather than formally comparing our prototype with standard terrain editors like Unity's terrain tools. Therefore, we designed our experiment as a walkthrough-style exploration, where participants tried the application and its interactions and gave oral feedback about their experience.

We recruited 16 volunteers from within our institution: 13 males, 3 females, mean age 38 (SD=6.9). 9 had previous experience with VR and 9 were frequent or occasional users of styli on tablets. None had ever used a pen in VR or for mid-air input.

5.1 Procedure

Before starting the study, the participant adjusted the height of their chair and the desk to their liking. We then proceeded to demonstrate the reset gesture as it was more difficult to explain once the participant was wearing the headset. Next, the participant donned the headset and made the necessary adjustments for it to fit comfortably and the VR scene to appear clearly. The gestures and techniques were then explained and participants asked to try them one by one. Afterwards, the participant was free to explore the interactions to further familiarise themselves with them. At all stages, the participant was encouraged to share their impressions, but a short interview after the experience formed the primary source for the recorded feedback. During the interview, the participant was asked for specific preferences between the different alternative methods they tried. In all cases, participants were asked to ignore tracking issues as much as possible and assume hand detection was robust when judging the techniques.

Each session took approximately 35 minutes, and the participant was offered a selection of snacks as a thank-you for their participation.

5.2 Results

Overall, participants enjoyed their experience as they considered the application playful and its potential use for game design and world building "à la Minecraft" appealing. They did not encounter major difficulties when performing the gestures, even though most were unfamiliar.

5.2.1 Navigation. Navigating with multitouch drag gestures felt like using a large trackpad and therefore was the preferred locomotion method for 8 participants. The DoF separation further allowed movements to be more controlled. 4 people preferred mid-air grabbing, as they found they had more freedom to place the hand and move it anywhere, compared to the tablet, where input was only possible in a limited surface area. Mid-air grabbing allowed faster navigation, but quick movements also caused motion sickness for 4 participants and repeated clutching was fatiguing. 4 participants stated they liked to have both navigation options, depending on how fast or carefully they wanted to navigate. One person, in particular, wanted to pan horizontally using the tablet, but found moving the camera vertically with a horizontal dragging movement less intuitive, and thus preferred moving their fist up and down for that action as it matched the direction of the camera motion.

5.2.2 Terrain Editing. 5 participants found that editing the terrain with the pen on the tablet was more precise and particularly useful to add detailed features like trees. Given the choice between using pen pressure or the NDH to control land height or tree size, 7 participants preferred pressure, as it allowed them to use only one hand, whereas 5 said they would rather use their NDH. The rest stated they would edit in mid-air only if both height and ground position needed to be controlled.

When experimenting with concurrent navigation and editing, most participants fixed the pen position, while moving the NDH. This is not surprising considering the motor coordination demands of asymmetric bimanual interaction [23]. For coarse mountain sculpting over large areas, mid-air navigation was more flexible, but when creating linear features, such as lines of trees or ridges at a constant elevation, the tablet offered more control.

5.2.3 Marking Menu. Regarding the marking menu, 7 participants preferred invoking and marking with four fingers on the surface, the easiness and convenience of using only one hand given as the main reason. 5 people liked the bimanual techniques better, with 4 participants favouring the mid-air palm up gesture to invoke the menu and 1 preferring tapping with four fingers. 2 participants said they would like to use touch with the NDH only for navigation. The remaining 4 people found all three methods equally acceptable. 2 participants commented that they would rather have the menu shown near their NDH (i.e. similar to wrist menus) rather than in the middle of the screen or the pen tip.

5.2.4 Reset Gesture. With regard to the delete-all/reset technique, opinions were evenly split. Half of the participants said the gesture was easy to remember and a good choice for this action. Some participants commented that it is not often used and difficult to trigger unintentionally. Yet the other half of the participants expressed difficulty when rapidly spinning the pen between their hands because of the thick resin attached on one side of the barrel. This problem is a side-effect of our prototype pen button, and would likely not occur with a perfectly round pen.

6 DISCUSSION

Overall, our observations and the feedback we received from our participants tend to show the complementarity of the two hands and the two interaction spaces. The tablet and mid-air space above it are useful for both navigation and editing, depending on precision and range needs. There is a slight preference for the tablet for navigation with the NDH as control is similar to a trackpad and fatigue is reduced compared to gesturing in mid-air, because the hand can rest on the surface. Pen input, however, felt appropriate and comfortable both on and above the tablet, although our study may not have been sufficiently long to examine potential fatigue issues with the pen.

Our sitting desktop condition limits the interaction space to the tablet on the desk and to roughly arms reach in mid-air, but our NDH-based navigation techniques allow the user to efficiently move in the scene so that these restrictions can be overcome. For example, the camera can be quickly repositioned using successive short drags on the tablet (i.e. with clutching) to continue sculpting in areas that are initially beyond reach. This is similar to how artists work with 2D compositions on paper [10] or pen and touch devices [12], where the NDH can make precise adjustments to the canvas between pen strokes. The use of the NDH for touch-based trackpad-like navigation while the pen edits in mid-air is particularly powerful and felt intuitive to most participants, even though the NDH may be more active than the pen during those interactions. As an input space in general, we observe that NDH touch + mid-air DH input is a combination of bimanual interaction that is still very much underexplored, and we see this type of cross-space hybrid input as particularly promising and worthy of future research, including in mobile contexts (e.g. phone touch+pen).

Regarding the marking menu, we believe touch-based invocation is a compelling alternative to commonly used wrist-attached menus triggered on palm up, as touching the tablet with a specific number of fingers involves little hand movement if the hand is already resting on the surface. We also believe that the pen with its pointed tip is an apt instrument for mid-air selection compared to controllers [20] and perhaps even compared to fingers, similar to 2D touch interfaces. We did not include classic touch-operated UIs on the virtual tablet surface, floating panels or palm menus with widgets operated by mid-air NDH taps, but of course such elements of more standard interfaces can also be used. However, for touch UIs appearing on the tablet surface, the user would be required to look down, which may not be desirable if most content is above the tablet. As with many gestural interfaces, interactions using gestures that need to be learned can exist as expert shortcuts in addition to more discoverable widget-based UIs.

7 LIMITATIONS AND EXTENSIONS

Our sensing environment relied on a motion-capture system for precise pen tracking and a more limited hand-tracking sensor fixed to the headset. This setup would not be practical for a regular consumer-oriented pen+touch+mid-air system, which needs to be both mobile and track robustly. However, we are confident that technical solutions will emerge to robustly track both hands and pen, including the DH gripping the pen. Recent approaches for hand pose estimation using deep learning have shown considerable promise in this area already [4, 15].

Our system uses a tablet to precisely detect pen and touch input, but when in VR and the hands and pen are tracked in 3D space, detecting contact with a hard surface is theoretically also possible using that tracking system so that any horizontal surface such as a desk or a table could potentially be used for pen and touch input. This would allow surface input to not be limited to the small touch-sensitive area of the tablet. However, with current camera-based tracking systems, detection is less precise and occlusion-prone compared to the dedicated touch sensors of a tablet. Future solutions might include highly precise and specially designed pens for input on arbitrary surfaces like Flashpen [21], while touch detection could be aided by wrist-worn motion sensors [17].

8 CONCLUSION

We presented a terrain editor for desk-based pen+touch+mid-air input in VR featuring hybrid bimanual interactions on and above the surface. We outlined different possible combinations of the bare NDH with the pen-holding DH to realise basic navigation, terrain editing and menu invocation operations. We evaluated the usability and suitability of the techniques in a qualitative study with 16 participants. Their feedback highlighted the complementarity of the two spaces to flexibly support precise constrained 2D input on the surface and fast unconstrained 3D interaction in mid-air with both hands.

Our goal with this work was both to support terrain editing using novel interactions as well as to initiate an investigation of the pen+touch+mid-air design space in VR and we hope this space will be further explored in other scenarios to identify compelling examples of hybrid cross-space interactions.

REFERENCES

[1] Bruno R De Araùjo, Géry Casiez, and Joaquim A Jorge. 2012. Mockup Builder: Direct 3D Modeling on and above the Surface in a Continuous Interaction Space. In *Proceedings of Graphics Interface 2012* (GI '12), 173–180.

[2] Rahul Arora, Rubaiat Habib Kazi, Tovi Grossman, George Fitzmaurice, and Karan Singh. 2018. SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–15. Retrieved from https://doi.org/10.1145/3173574.3173759

[3] Peter Brandl, Clifton Forlines, Daniel Wigdor, Michael Haller, and Chia Shen. 2008. Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces. *Proceedings of the working conference on Advanced visual interfaces*, 154–161. https://doi.org/10.1145/1385569.1385595

[4] Theocharis Chatzis, Andreas Stergioulas, Dimitrios Konstantinidis, Kosmas Dimitropoulos, and Petros Daras. 2020. A Comprehensive Study on Deep Learning-Based 3D Hand Pose Estimation Methods. *Applied Sciences* 10, 19. https://doi.org/10.3390/app10196850

[5] Noah Coomer, Sadler Bullard, William Clinton, and Betsy Williams-Sanders. 2018. Evaluating the Effects of Four VR Locomotion Methods: Joystick, Arm-Cycling, Point-Tugging, and Teleporting. In *Proceedings of the 15th ACM Symposium on Applied Perception* (SAP '18). https://doi.org/10.1145/3225153.3225175

[6] Tobias Drey, Jan Gugenheimer, Julian Karlbauer, Maximilian Milo, and Enrico Rukzio. 2020. VRSketchIn: Exploring the Design Space of Pen and Tablet Interaction for 3D Sketching in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376628

[7] Sevinc Eroglu, Frederic Stefan, Alain Chevalier, Daniel Roettger, Daniel Zielasko, Torsten W Kuhlen, and Benjamin Weyers. 2021. Design and Evaluation of a Free-Hand VR-based Authoring Environment for Automated Vehicle Testing. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, 1–10. https://doi.org/10.1109/VR50410.2021.00020

[8] Thomas Fröhlich, Dmitry Alexandrovsky, Timo Stabbert, Tanja Döring, and Rainer Malaka. 2018. VRBox: A Virtual Reality Augmented Sandbox for Immersive Playfulness, Creativity and Exploration. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play* (CHI PLAY '18), 153–162. https://doi.org/10.1145/3242671.3242697

[9] Travis Gesslein, Verena Biener, Philipp Gagel, Daniel Schneider, Per Ola Kristensson, Eyal Ofek, Michel Pahud, and Jens Grubert. 2020. Pen-based Interaction with Spreadsheets in Mobile Virtual Reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 361–373. https://doi.org/10.1109/ISMAR50242.2020.00063

[10] Yves Guiard. 1987. Asymmetric division of labor in human skilled bimanual action: the kinematic chain as a model. *Journal of Motor Behavior* 19, 4: 486–517.

[11] François Guimbretière and Chau Nguyen. 2012. Bimanual Marking Menu for near Surface Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 825–828. Retrieved from https://doi.org/10.1145/2207676.2208521

[12] Ken Hinckley, Koji Yatani, Michel Pahud, Nicole Coddington, Jenny Rodenhouse, Andy Wilson, Hrvoje Benko, and Bill Buxton. 2010. Pen + touch = new tools. *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, 27–36. https://doi.org/10.1145/1866029.1866036

[13] Kenrick Kin, Björn Hartmann, and Maneesh Agrawala. 2011. Two-Handed Marking Menus for Multitouch Devices. *ACM Trans. Comput.-Hum. Interact.* 18, 3. https://doi.org/10.1145/1993060.1993066

[14] Josen Daniel O De Leon, Romelio P Tavas, Rodolfo A Aranzanso, and Rowel O Atienza. 2016. Genesys: A Virtual Reality scene builder. In *2016 IEEE Region 10 Conference (TENCON)*, 3708–3711. https://doi.org/10.1109/TENCON.2016.7848751

[15] Shaowei Liu, Hanwen Jiang, Jiarui Xu, Sifei Liu, and Xiaolong Wang. 2021. Semi-supervised 3d hand-object poses estimation with interactions in time. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14687–14697.

[16] Fabrice Matulic and Moira Norrie. 2013. Pen and Touch Gestural Environment for Document Editing on Interactive Tabletops. *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces*, 41–50.

[17] Manuel Meier, Paul Streli, Andreas Fender, and Christian Holz. 2021. TaplD: Rapid Touch Interaction in Virtual Reality using Wearable Sensing. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, 519–528. https://doi.org/10.1109/VR50410.2021.00076

[18] Meta. 2021. Horizon Worlds. Retrieved from https://www.oculus.com/facebook-horizon/

[19] Daniel L Odell, Richard C Davis, Andrew Smith, and Paul K Wright. 2004. Toolglasses, Marking Menus, and Hotkeys: A Comparison of One and Two-Handed Command Selection Techniques. In *Proceedings of Graphics Interface 2004* (GI '04), 17–24.

[20] Duc-Minh Pham and Wolfgang Stuerzlinger. 2019. Is the Pen Mightier than the Controller? A Comparison of Input Devices for Selection in Virtual and Augmented Reality. In *25th ACM Symposium on Virtual Reality Software and Technology* (VRST '19). https://doi.org/10.1145/3359996.3364264

[21] Hugo Romat, Andreas Fender, Manuel Meier, and Christian Holz. 2021. Flashpen: A High-Fidelity and High-Precision Multi-Surface Pen for Virtual Reality. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, 306–315. https://doi.org/10.1109/VR50410.2021.00053

[22] Hemant Bhaskar Surale, Aakar Gupta, Mark Hancock, and Daniel Vogel. 2019. TabletInVR: Exploring the Design Space for Using a Multi-Touch Tablet in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–13. Retrieved from https://doi.org/10.1145/3290605.3300243

[23] Stephan P Swinnen. 2002. Intermanual coordination: From behavioural principles to neural-network interactions. *Nature Reviews Neuroscience* 3, 5: 348–359. https://doi.org/10.1038/nrn807

[24] Daniel Vogel, Matthew Cudmore, Géry Casiez, Ravin Balakrishnan, and Liam Keliher. 2009. Hand Occlusion with Tablet-Sized Direct Pen Input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 557–566. Retrieved from https://doi.org/10.1145/1518701.1518787

[25] Jia Wang, Owen Leach, and Robert W Lindeman. 2013. DIY World Builder: An immersive level-editing system. In *2013 IEEE Symposium on 3D User Interfaces (3DUI)*, 195–196. https://doi.org/10.1109/3DUI.2013.6550245

[26] Mike Wu, Shen Chia, Kathy Ryall, Clifton Forlines, and Ravin Balakrishnan. 2006. Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, 185–192. https://doi.org/10.1109/tabletop.2006.19

[27] Robert Zeleznik, Andrew Bragdon, Ferdi Adeputra, and Hsu-Sheng Ko. 2010. Hands-on math: a page-based multi-touch and pen desktop for technical work and problem solving. *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, 17–26. https://doi.org/10.1145/1866029.1866035

[28] Unity. https://unity3d.com

[29] Unreal Engine VR Mode. https://docs.unrealengine.com/4.27/en-US/BuildingWorlds/VRMode

[30] Minecraft in VR. https://www.minecraft.net/en-us/vr