Look & Turn: One-handed and Expressive Menu Interaction by Gaze and Arm Turns in VR

Katharina Reiter
katharina.reiter1@stud.sbg.ac.at
Paris London Universität Salzburg
Salzburg, Austria

Ken Pfeuffer
ken@cs.au.dk
Aarhus University
Aarhus, Denmark
Bundeswehr University Munich
Munich, Germany

Augusto Esteves
augusto.esteves@tecnico.ulisboa.pt
ITI / LARSyS, Instituto Superior
Técnico, University of Lisbon
Lisbon, Portugal

Tim Mittermeier
tim.mittermeier@unibw.de
Bundeswehr University Munich
Munich, Germany

Florian Alt
florian.alt@unibw.de
Bundeswehr University Munich
Munich, Germany

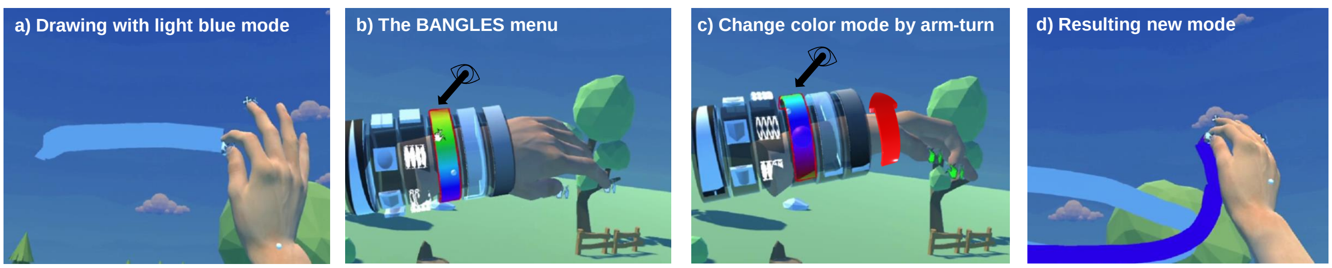

Figure 1: We explore design and interaction of handheld menu for freehand gestural Virtual Reality (VR). Specifically, how to hold and interact with the menu via just a single hand and gaze – keeping the other hand free for drawing. Example: The user looks at a menu sub-group (b), and then rotates the arm to adjust a parameter (c), to then immediately continue drawing with the dominant hand (d).

ABSTRACT

A user's free hands provide an intuitive platform to position and design virtual menu interfaces. We explore how the hands and eyes can be integrated in the design of hand-attached menus. We synthesise past work from the literature and derive a design space that crosses properties of menu systems with an hand and eye input vocabulary. From this, we devise three menu systems that are based on the novel concept of Look & Turn: gaze indicates menu selection, and rotational turn of the wrist navigates menu and manipulates continuous parameters. Each technique allows users to interact with the hand-attached menu using the same hand, while keeping the other hand free for drawing. Based on a VR prototype that combines eye-tracking and glove-based finger tracking, we discuss first insights on technical and human factors of the promising interaction concept.

ETRA '22, June 8–11, 2022, Seattle, WA, USA

ACM ISBN 978-1-4503-9252-5/22/06…$15.00
https://doi.org/10.1145/3517031.3529233

CCS CONCEPTS

• **Human-centered computing → Interaction techniques**; *Virtual reality*; Interaction design.

KEYWORDS

Multimodal interaction, virtual reality, VR, gaze interaction, handheld menus, on-body menus, extended reality, XR

ACM Reference Format:
Katharina Reiter, Ken Pfeuffer, Augusto Esteves, Tim Mittermeier, and Florian Alt. 2022. Look & Turn: One-handed and Expressive Menu Interaction by Gaze and Arm Turns in VR. In *2022 Symposium on Eye Tracking Research and Applications (ETRA '22), June 8–11, 2022, Seattle, WA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3517031.3529233

1 INTRODUCTION

As spatial interfaces grow in popularity through Virtual and Augmented Reality (VR/AR), exploring novel ways in which users can interact with these systems is important [Billinghurst et al. 2015]. Gestural interfaces controlled by free hands of the user are promising for natural interaction, as currently available in headsets (E.g., Meta Quest 2 or Microsoft Hololens 2). However, it also introduces challenges. Prolonged use of gestures and hand pointing can lead to a plethora of user interaction problems [Norman 2010] such as physical fatigue [Jang et al. 2017], and robust tracking is often hindered when both hands are used close to each other such as in the interaction with handheld menus.

This paper explores how a multi-modal approach combining gestures and gaze can facilitate interaction with freehand-anchored menus in VR. Multi-modal input is a popular approach where one modality tends to compensate for the limitations of another [Nizam et al. 2018]. For example, by offloading strenuous [Hincapié-Ramos et al. 2014] and occasionally hard to track [Shah et al. 2012] manual pointing to the user's gaze and reserving the former for short and discrete gestures that confirm the user's intent (e.g., pinch to select [Bowman and Wingrave 2001]). Interaction concepts such as *Gaze + Pinch* [Pfeuffer et al. 2017] and *Pinpointing* [Kytö et al. 2018] exploit the strength of our eyes and hands in concert. For hand-anchored menus, Pfeuffer et al. employed gaze and gesture to easily select items without tracking issues as the hands do not overlap in space [Pfeuffer et al. 2017, 2020]. However, the gaze input limits the menu expressiveness, as only a few options can be supported as of accuracy (compared to hand pointing), and gaze does not allow for fine-grained manipulation of continuous parameters. Further, it feels inefficient to still need both hands to interact with the menu, when already using gaze to substitute one hand's pointing task.

We propose a novel interaction concept called *Look & Turn*. It is based on 1) gaze pointing for implicit acquisition of menu groups and items, and 2) manual arm-turn and pinch gesture to navigate the menu and manipulate parameters. The concept affords two advantages. First, it is one-handed – the same hand that holds the menu executes turn and pinch gesture for manipulation, leaving the other hand completely free for other tasks. Second, it is highly expressive, as multiple menu levels and continuous parameter control can be supported. To investigate the concept from ground up, we first distilled a design space of menu design parameters and gaze + hand input vocabulary. Based on this, we designed three novel interaction techniques that specifically exploit the novel concept.

One example is the BANGLES technique, as illustrated in Figure 1. Here the user draws with their dominant hand as common in current VR design tools (a), but can change drawing parameters in a novel way. Users hold up their non-dominant hand and gaze at a submenu in form of one of the rings (b). By holding a pinch gesture and turning the arm, the user manipulates a colour parameter (c). If the desired setting is found, the user continues drawing operations with their dominant hand (d). This provides the advantage that the drawing hand can be kept at a user-preferred position, and can immediately start after menu access. In our paper, we describe two more techniques using the Look & Turn principle, and assess initial user feedback in an informal evaluation.

In sum, we contribute (1) A design space of the interaction possibilities between manual and gaze controls and wrist and hand menu placements, (2) three interaction techniques that derive from our design space and illustrate several novel interactions between the input modalities (i.e., gaze pointing and dwell, pinches, and wrist rotations), and (3) a user evaluation that provides preliminary feedback on the feasibility and trade-offs of the prototypes above.

2 RELATED WORK

Gestural input and direct manipulation are natural means of interaction in VR [Hand 1997]. Users' hands can be tracked using RGB [Wang et al. 2020] or depth sensing cameras [Sharp et al. 2015], or worn gloves [Mapes and Moshell 1995; Pierce et al. 1999; Weimer and Ganapathy 1989]. The latter tends to be favoured when haptic feedback is desired [Perret and Vander Poorten 2018], or when user manipulations are expected to cause occlusions to optical tracking [Shah et al. 2012]. Examples include the VPL DataGlove [Jacoby and Ellis 1992], where different menus can be called upon by extending a different number of fingers; or [Bowman and Wingrave 2001]'s Pinch Gloves, where menu options are displayed in users' fingers and are selected via pinching the appropriate finger and the thumb. Pinches between the little finger and thumb change the options displayed in the remaining three fingers; while both hands are used to enable two-level menus. [Piekarski and Thomas 2001] proposed an approach to the latter using multi-level menus: touching the palm of the hand with any finger returns the user to the main menu.

Gaze input is popular for interaction, as it is innately hands-free and is not reliant on explicit user commands [Pfeuffer et al. 2021], and can enhance UI elements from buttons to virtual keyboards [Rivu et al. 2020b, 2019] – an interaction style known as "non-command-based" [Jacob 2003]. In VR, gaze pointing has been found to be faster than traditional pointing – particularly for distant objects – and most users report preferring this approach [Tanriverdi and Jacob 2000]. As a consequence, several research efforts have investigated the use of gaze input for virtual menus. For example, gaze-interactive menus in face-to-face conversations on people [Rivu et al. 2020a], or see-through palettes to apply modes to 3D objects [Mardanbegi et al. 2019]. [Pfeuffer et al. 2020] compared five menu selection techniques in VR. Direct grab gesture was the fastest, easiest, and preferred technique, but required the most physical exertion; while a purely hands-free approach using gaze-based dwell time was the second fastest approach, but also most fatiguing on the eyes. [Esteves et al. 2017] provides an example in AR, where a spatial-based menu used smooth pursuits eye movements to select menu options and access the menu's sub-levels – while looking away and back would return the user to the top or main menu.

Multimodal input for selection and manipulation in virtual environments is long established. Early work by [Bolt 1980] introduced how the combination of hand pointing and speech can render virtual object manipulation highly natural and expressive. Recent work by Sidenmark et al. explored the use of gaze and head input to enable interaction with radial menus in AR, where the alignment of both pointing modalities enables selection of menu commands [Sidenmark et al. 2021]. Closely related is the gaze and gesture interface, that has been explored for desktop and environmental interaction where hand gestures allow expressive manipulation of the gaze-selected objects [Chatterjee et al. 2015; Hales et al. 2013]. In VR, [Pfeuffer et al. 2017] explored the Gaze + Pinch concept for immersive 3D manipulations. In particular, introducing the Virtual Eye-gaze Interaction Armband (VEIA) where a pinch gesture selects the item being gazed at in a menu located in the user's forearm or palm. These approaches are popular as one modality compensates for another's limitations [Nizam et al. 2018]. Our work provides a deeper look into hand-attached menus that users operate via combined gaze pointing with wrist and hand gesture.

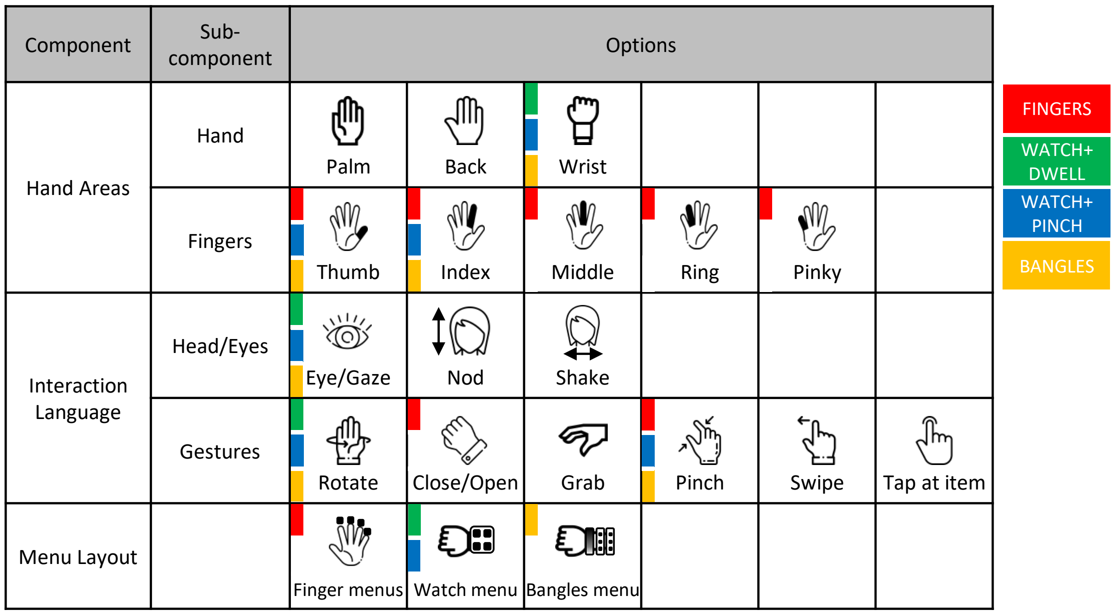

Figure 2: Design space of interaction possibilities for hand and gaze based menus (rows describe dimensions, columns the various instances). It shows the many possibilities of creating gaze interfaces that emerge when creating arm-attached menus. It can be used as an overview, and as well to draw out new interaction techniques. For example, we derived four novel techniques that each provide a unique instantiation of the design space (as color coded).

3 DESIGN SPACE

A novel design space is proposed, intersecting menu UI design properties with the user's input modalities of gaze and hand gestures (see figure 2). The goal is to inform a VR interaction designer about the possibilities to create advanced interaction techniques given such input possibilities. All parts of the design space are informed by prior work. To focus our work, the design space currently supports the following properties:

- *Menus with a hierarchy of two levels* (i.e. top-level menu, and detail-level submenus). This is particularly useful for design applications such as Google Tiltbrush, Microsoft Paint, or Adobe Photoshop, e.g., when first selecting the colour palette submenu and then selecting a colour in it, but in principle can generalise to any two-level menu type.
- *Menu systems spatially fixed at areas of the non-dominant hand*; to take advantage of the user's proprioceptive senses in the body, as humans are intuitively aware of the movement and location of their body parts [Mine et al. 1997].
- *Asymmetric division of labour between the hands.* This adheres the division of labour as advocated by Guiard [Guiard 1987], where the non-dominant hand plays a supportive role for the other hand, a principle implemented across many design tools (e.g., Google Tiltbrush), and research efforts [Pfeuffer et al. 2020]. Here we explore a novel potential advantage for the user: using the non-dominant hand (and gaze input) to both access and interact with the menu. This leaves the other (dominant) hand completely focused on drawing tasks and remain at a user-preferred position.

We now describe each design space dimension. The matrix representation allows the designer to iterate over each component and mark the cells to be used, as done with the prototypes presented in this paper (coloured boxes).

3.1 Hand areas

For hand areas, the hand including forearm or wrist is listed first. E.g., [Pfeuffer et al. 2017]'s VEIA menu is located on the user's forearm. The hand itself can be divided into its front (palm) and back, where each side can become useful for a distinct interaction metaphor such as opening and closing a menu [Bowman and Wingrave 2001]. Next, the fingers are named, starting with the thumb, and followed by the index, middle, ring and little finger. [Bowman and Wingrave 2001]'s TULIP menu locates menu options on the users' fingers. Starting with hand areas can be beneficial since not all hardware can address the individual fingers.

3.2 Interaction language

An important high-level consideration is how one integrates the user's hands and eye gaze into the menu interaction. How one can invoke UI commands and selections can be defined by an interaction language. We describe example primitives divided into each modality, i.e. gestures triggered by the hand and head/eyes/gaze.

Head/Eyes. The user can trigger a reaction from the system through eye and head based input mechanisms. Eye/gaze can be used to select elements on which the user has focused, through implementing a dwell time, eliminating the need for explicit commands [Jacob 2003; Tanriverdi and Jacob 2000]. Other work has suggested, aligning, shaking or nodding the head [Esteves et al. 2017; Sidenmark and Gellersen 2019].

Gestures. Gestures must be 1) executed without problems, 2) recognisable and distinguishable, and 3) chosen for the right context [Jacoby and Ellis 1992]. Thus, we derive six gesture types. The user can rotate her wrist, e.g., to open a menu or to get to another view. [Bowman et al. 2002] use this gesture to interact with internal menu items. Opening and closing the fist can also be used to confirm an interaction. With the closed fist, objects can also be grabbed and moved [Chatterjee et al. 2015; Jacoby and Ellis 1992]. But selecting virtual objects can also be done with individual fingers, e.g., pressing the thumb and a finger to accomplish a pinch gesture. This creates four possible pinch gestures on each hand to exploit for interaction design [Bowman and Wingrave 2001; Piekarski and Thomas 2001; Pierce et al. 1999]. Other gestures that can be made with one or more fingers is swiping and tapping, similar to touch-enabled screens. [Chatterjee et al. 2015] use a swipe gesture to choose between menu items, and [Buchmann et al. 2004] use tap gestures on menu items to adjust AR information.

3.3 Menu layout

As with the TULIP menus, an intuitive menu activation is to open the hand, and use fingers for options in a submenu. This works when the amount of options matches the amount of fingers. Next, the Watch and Bangles menu layouts are related to "grid menus" that are spatially fixed to the user's hand or arm. Menus within the hand's location have been designed for controller-based interfaces, where menu is attached to the controller (e.g., Google Tiltbrush menus, or with [Pfeuffer et al. 2020]'s gaze-based handheld menus). This can be intuitive as it is similar to holding a mobile phone in one's hand. Several prior works have exploited gaze for hands-free interaction with smartwatch interfaces [Esteves et al. 2015; Wang and Grossman 2020] – this metaphor can be explored in VR, and potentially expanded to multiple virtual watch-like interfaces on the user's arm to support control of multiple menu levels.

4 GAZE & TURN INTERACTION TECHNIQUES

We now describe example interaction techniques that emerge from the design space. The first can be considered as a baseline as a fully manual technique. The remaining three exploit the *Look & Turn* principle. An overview about how the techniques map to the dimensions of the design space is provided in Figure 2. Our interactions are designed for a drawing scenario that involves many possible menu items (e.g., colours, brushes and shapes), as inspired by the literature [Billinghurst et al. 1997; Bowman and Wingrave 2001; Pfeuffer et al. 2020]. We use four menu groups: colour, material, shapes and thickness. Unity3d is used to develop the applications on a Vive Pro Eye head-mounted display. Eye tracking is done with Tobii XR SDK, hand tracking with the VRfree gloves from sensoryx [Sensoryx AG 2019].

All of the examples share similarities and differences in how atomic tasks are accomplished by the user. These include *Pre-selection*, where users can "preliminary" select a sub-menu by pointing at it with their eyes. Pre-selection triggers visual feedback that communicates to the user that a full selection can commence [Sidenmark and Gellersen 2019; Sidenmark et al. 2021]. Next, *selection confirmation*, to avoid the Midas Touch problem of eye-tracked interfaces [Jacoby and Ellis 1992]. We include pinch gestures (touch of both thumb and another finger) and dwell time, where users look at a target for a specified time. Lastly, *menu navigation*, to navigate from a high-level menu to a sub-menu (e.g., colour palette) to select a particular mode (e.g., colour), by an explicit gesture (here the *Hand-turn* gesture) or implicitly through the UI design when the entire menu is directly visible.

FINGERS. This technique is a modification of the TULIP menus [Bowman and Wingrave 2001], where each finger represents a menu group that can be rapidly accessed by pinch gestures. To select a menu item, first the menu group is selected in the main menu with a pinch gesture with the respective finger (figure 3 - step 1). The menu items of the sub-menu are then mapped to the fingers and can be selected with a pinch gesture. To return from a sub-menu to the main menu, a "spider-man gesture" is used where middle and ring finger are brought to the palm (step 2). Figure 3a) shows the main menu with the menu elements on the fingertips. In b) we are already in the sub-menu for colour selection. Here a pinch gesture is performed with the middle finger, which selects the colour blue. Visual feedback is enabled by the border around the menu element turns green. c) shows an example "spider-man gesture".

WATCH+DWELL. This menu design embraces a layout that is located at the user's wrist. It allows users to rapidly engage the hand and simply select by their eyes only. The top level of the hierarchy is displayed on the outside of the wrist, the second layer on the inside. All menu elements are selected with gaze (figure 3 - step 1). A dwell time of 1s based on [Penkar et al. 2012; Riegler et al. 2020] is used for confirmation (step 2). A circular progress indicator is used, i.e., a circle above the targeted element, which fills up as long as the user keeps looking. To navigate between the menu levels, a turn of the wrist is used (step 3). Figure 3a) shows the main menu with the red circle filling up ahead of the colour menu. So the user has gazed at that element and is now waiting for the required dwell time. In b) the circle is full, so the dwell time has been reached and the menu item is selected. The same applies to the sub-menus. In c) you can see the material sub-menu. The full red circle over one of the menu items indicates that the cloud texture has been selected.

WATCH+PINCH. This technique is like WATCH+DWELL, but uses a pinch gesture instead of dwell time for selection. Pinch can be executed quickly and allows users to visually inspect a target without invoking selections, but potentially adds another substep. As visual feedback, a pinch icon can be seen above the sphere (Figure 3a)) to indicate that this element is gazed at and what has to be done next. Small cubes with the same symbol are also added to the two required fingertips. In b) the icons turned red, you can see that the fingers are getting closer. In c) the icon above the menu item is hidden and the symbols on the fingertips are green. This indicates that the sphere shape has been selected.

BANGLES. To reduce wrist rotations, this menu type uses rings that each represents menu groups and that are directly accessed without a need to navigate menu hierarchies. Gaze is used for the pre-selection (figure 3 - step 1), followed by a pinch gesture for confirmation (step 2). The dynamic sub groups options are then

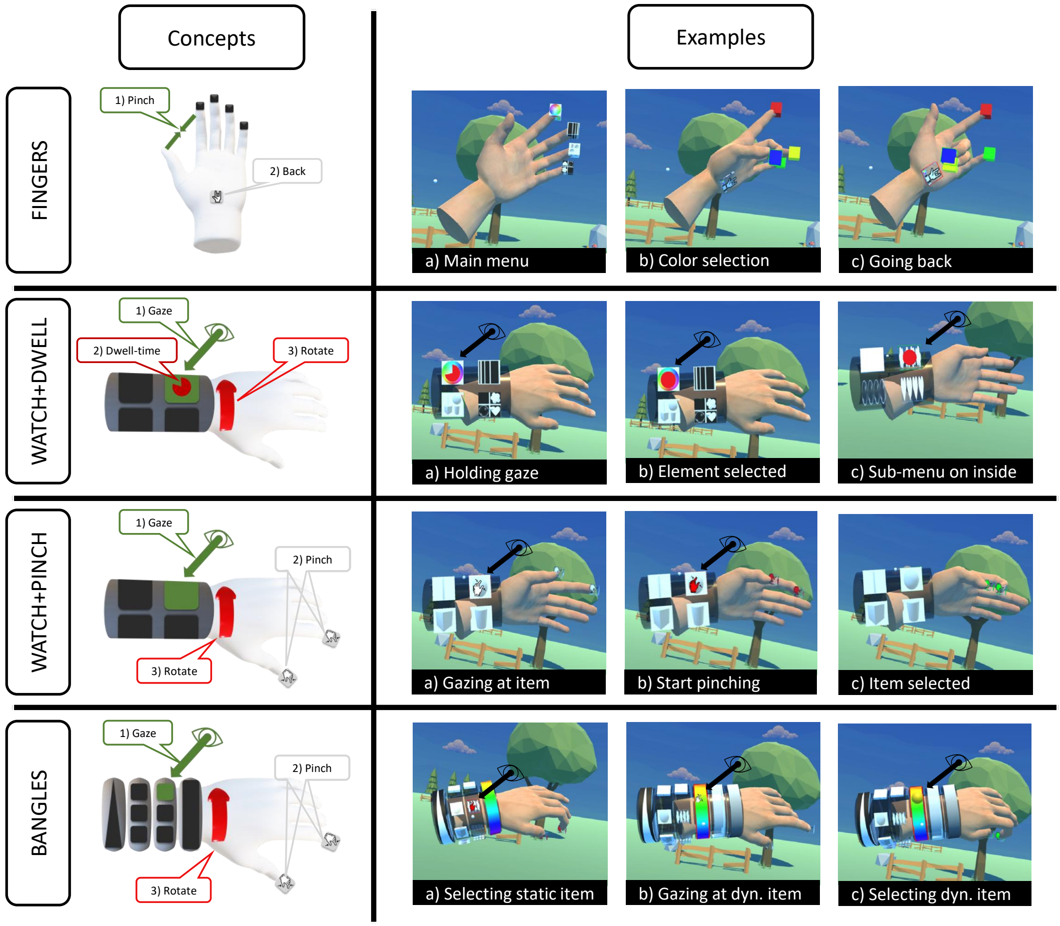

Figure 3: Overview over the explored interaction techniques

selected by a wrist rotation (step 3) – as long as the pinch gesture persists. Figure 3a) shows an example of a static menu item with a set of options. b) shows the second version with three colour rings, including accommodate "Saturation" and "Value" of the HSV colour space. The Hue-ring is selected with gaze, which can be seen by the red border and pinch icon. In c) a pinch gesture has been carried out and instead of the pinch symbol, a coloured ball can now be seen above the still gaze-targeted colour ring. If the user now turns her wrist while holding the pinch gesture, the colour of this ball changes depending on this ring. As soon as the user stops the pinch gesture, the colour of the ball is selected.

5 EVALUATION

We conducted an informal evaluation where users experienced the different menu types. 7 users (6 male, 1 female) aged 39.7 years (SD=12.28) of mixed background (students, non-CS employees) participated. On a scale between 1 (no experience) to 5 (expert), users rated themselves with low experience with VR (M=1.29, SD=0.49). The procedure included: 1) introduction and calibrating of the hardware; 2) learning the use of the drawing operations; and 3) performing a drawing task with each menu. In the task, users were drawing clouds and spheres with different colours and textures, which implied frequent use of the menu. The order of the menu systems has been counterbalanced. The study took about 40 minutes, and we collected verbal and written user feedback through a questionnaire after each condition on usability. The questions in the forms are based on the System Usability Scale (SUS) and NASA Task Load Index (TLX).

In FINGERS, the gestures were clear to most of the participants and some even found the gestures visibly good and had fun executing them. But it was hard for some to bend the two required fingers far enough for the back gesture to be recognised. Being able to use all fingers was noted several times as a positive feature of the technique.

With WATCH+DWELL, it was mentioned multiple times that one would find it good not to need any further interaction and not having to do anything but look. But there are also critical comments regarding the more time-consuming dwell time. In addition to the suggestion of a shortened dwell time, it was also noted that one would pick something wrong relatively quickly if you aren't looking away fast enough. If one wants to take a moment to look at the available menu options, it can be annoying that the selection of a menu item starts immediately after gazing at it.

With WATCH+PINCH, one study participant noted preferring the pinch version over the dwell time version, since she does not have to wait as long as compared to the dwell time. Two challenges emerged from the user tests. First, the visual pinch symbols attracted the user's gaze. When looking at them, they lost focus on the targeted menu item. This problem was amplified with poor hand tracking, as the user cannot rely on the haptic feedback of their fingertips and instead looked at the virtual hands to press the virtual fingertips together to invoke pinch.

Half of participants agreed that BANGLES can be described as easy to use, while others found it difficult to choose a colour here and needed further instructions on how to use the colour ring. The calibration caused difficulties, as the user has to keep the gaze focused on the ring while holding the pinch gesture and rotating the wrist. A poor hand/eye calibration of the used hardware strongly affected the usability. Some participants were however positive of the advanced menu options. It was appreciated that there is just one hierarchical level, since this leads to a better overview. The appearance was also highlighted positively since it would look uniform.

6 DISCUSSION AND CONCLUSION

In this paper we explored Look & Turn interaction, by a design space that exploits various user input modalities and menu configurations for parametric control in VR. We conceived three novel menu techniques, each based on *Look & Turn*, that allow users to interact with multi-level menu hierarchies. While our semi-informal evaluation with a glove-based hand tracking system provided useful early insights, several limitations need to be addressed before the techniques presented can become a realistic alternative to current menu interfaces.

The use of arm turns for menu navigation was viewed positively by users. This is likely due to allowing for more coarse input than the back gesture in FINGERS, which can require a more precise gesture. Taken together with the fact that it enables users to cycle through menus directly, the arm rotation was ultimately perceived as "less strenuous". Further research is needed to explore how this arm rotation can be accomplished eyes-free – a relevant aspect for experienced or mobile users. It has several advantages, e.g., users do not have to take their eyes off the drawn object. We observed in the study that the gestures were always performed with eye contact. Therefore, the question arises whether users would take advantage of eyes-free input – further studies should take into account long-term usage and users' eye-hand coordination. Another venue for future work expands our design space towards the context of mobile use, via Augmented or eXtended reality (AR, XR). Parametric control during the use of these technologies could also benefit from the techniques developed, e.g., adjusting the volume of media playback or scrolling through a list of songs.

In our work we have focused on the design space and three specific multimodal techniques that exploit fine-grained hand and finger tracking coupled with gaze. This can be further explored, e.g., through optimising the amount of sub-tasks required to perform menu commands. The BANGLES menu for example, could potentially even be used without a pinch gesture. The act of focusing on the arm-attached ring UI could represent the trigger. Furthermore, it is important to perform systematic evaluations of the techniques to assess the overall user performance and experience.

REFERENCES

Mark Billinghurst, Sisinio Baldis, Lydia Matheson, and Mark Philips. 1997. 3D palette. In *Proceedings of the ACM symposium on Virtual reality software and technology - VRST '97*, Daniel Thalmann, Steve Feiner, and Gurminder Singh (Eds.). ACM Press, New York, New York, USA, 155–156. https://doi.org/10.1145/261135.261163

Mark Billinghurst, Adrian Clark, and Gun Lee. 2015. A survey of augmented reality. (2015).

Richard A. Bolt. 1980. "Put-That-There": Voice and Gesture at the Graphics Interface. In *Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques* (Seattle, Washington, USA) *(SIGGRAPH '80)*. Association for Computing Machinery, New York, NY, USA, 262–270. https://doi.org/10.1145/800250.807503

D. A. Bowman and C. A. Wingrave. 2001. Design and evaluation of menu systems for immersive virtual environments. In *IEEE virtual reality 2001*, Haruo Takemura (Ed.). IEEE Computer Soc, Los Alamitos, Calif., 149–156. https://doi.org/10.1109/VR.2001.913781

D. A. Bowman, C. A. Wingrave, J. M. Campbell, V. Q. Ly, and C. J. Rhoton. 2002. Novel Uses of Pinch Gloves™ for Virtual Environment Interaction Techniques. *Virtual Reality* 6, 3 (2002), 122–129. https://doi.org/10.1007/s100550200013

Volkert Buchmann, Stephen Violich, Mark Billinghurst, and Andy Cockburn. 2004. FingARtips – Gesture Based Direct Manipulation in Augmented Reality. In *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia - GRAPHITE '04*, Yong Tsui Lee, Alan Chalmers, Hock Soon Seah, and Stephen N. Spencer (Eds.). ACM Press, New York, New York, USA, 212–221. https://doi.org/10.1145/988834.988871

Ishan Chatterjee, Robert Xiao, and Chris Harrison. 2015. Gaze+Gesture: Expressive, Precise and Targeted Free-Space Interactions. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, Zhengyou Zhang, Phil Cohen, Dan Bohus, Radu Horaud, and Helen Meng (Eds.). ACM, New York, NY, USA, 131–138. https://doi.org/10.1145/2818346.2820752

Augusto Esteves, Eduardo Velloso, Andreas Bulling, and Hans Gellersen. 2015. Orbits: Gaze Interaction for Smart Watches Using Smooth Pursuit Eye Movements. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 457–466. https://doi.org/10.1145/2807442.2807499

Augusto Esteves, David Verweij, Liza Suraiya, Rasel Islam, Youryang Lee, and Ian Oakley. 2017. SmoothMoves: Smooth Pursuits Head Movements for Augmented Reality. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 167–178. https://doi.org/10.1145/3126594.3126616

Yves Guiard. 1987. Asymetric division of labor in human skilled bimanual action: the kinematic chain as model" Journal. 19 (01 1987).

Jeremy Hales, David Rozado, and Diako Mardanbegi. 2013. Interacting with objects in the environment by gaze and hand gestures. In *Proceedings of the 3rd international workshop on pervasive eye tracking and mobile eye-based interaction*. 1–9.

Chris Hand. 1997. A Survey of 3D Interaction Techniques. *Computer Graphics Forum* 16, 5 (1997), 269–281. https://doi.org/10.1111/1467-8659.00194

Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1063–1072. https://doi.org/10.1145/2556288.2557130

Robert J.K Jacob. 2003. Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-Command Interfaces. Washington, D.C., 1–58.

Richard H. Jacoby and Stephen R. Ellis. 1992. Using virtual menus in a virtual environment. In *Visual Data Interpretation (SPIE Proceedings)*, Joanna R. Alexander (Ed.). SPIE, 39–48. https://doi.org/10.1117/12.59654

Sujin Jang, Wolfgang Stuerzlinger, Satyajit Ambike, and Karthik Ramani. 2017. Modeling Cumulative Arm Fatigue in Mid-Air Interaction Based on Perceived Exertion and Kinetics of Arm Motion. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3328–3339. https://doi.org/10.1145/3025453.3025523

Mikko Kytö, Barrett Ens, Thammathip Piumsomboon, Gun A. Lee, and Mark Billinghurst. 2018. *Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173655

Daniel P. Mapes and J. Michael Moshell. 1995. A Two-Handed Interface for Object Manipulation in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 4, 4 (1995), 403–416. https://doi.org/10.1162/pres.1995.4.4.403

Diako Mardanbegi, Benedikt Mayer, Ken Pfeuffer, Shahram Jalaliniya, Hans Gellersen, and Alexander Perzl. 2019. Eyeseethrough: Unifying tool selection and application in virtual environments. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 474–483.

Mark R. Mine, Frederick P. Brooks, and Carlo H. Sequin. 1997. Moving Objects In Space: Exploiting Proprioception In Virtual-Environment Interaction. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques - SIGGRAPH '97*, G. Scott Owen, Turner Whitted, and Barbara Mones-Hattal (Eds.). ACM Press, New York, New York, USA, 19–26. https://doi.org/10.1145/258734.258747

SS Muhammad Nizam, Rimaniza Zainal Abidin, Nurhazarifah Che Hashim, Meng Chun Lam, Haslina Arshad, and NAA Majid. 2018. A review of multimodal interaction technique in augmented reality environment. *Int. J. Adv. Sci. Eng. Inf. Technol* 8, 4-2 (2018), 8–4.

Donald A. Norman. 2010. Natural User Interfaces Are Not Natural. *Interactions* 17, 3 (may 2010), 6–10. https://doi.org/10.1145/1744161.1744163

Abdul Moiz Penkar, Christof Lutteroth, and Gerald Weber. 2012. Designing for the eye: Design parameters for dwell in gaze interaction. In *Proceedings of the 24th Australian Computer-Human Interaction Conference on - OzCHI '12*, Vivienne Farrell, Graham Farrell, Caslon Chua, Weidong Huang, Raj Vasa, and Clinton Woodward (Eds.). ACM Press, New York, New York, USA, 479–488. https://doi.org/10.1145/2414536.2414609

J. Perret and E. Vander Poorten. 2018. Touching Virtual Reality: A Review of Haptic Gloves. In *ACTUATOR 2018; 16th International Conference on New Actuators*. 1–5.

Ken Pfeuffer, Yasmeen Abdrabou, Augusto Esteves, Radiah Rivu, Yomna Abdelrahman, Stefanie Meitner, Amr Saadi, and Florian Alt. 2021. ARtention: A design space for gaze-adaptive user interfaces in augmented reality. *Computers & Graphics* 95 (2021), 1–12. https://doi.org/10.1016/j.cag.2021.01.001

Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze + pinch interaction in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*, Adalberto L. Simeone and A. Special Interest Group on Computer-HumanC.M. Interaction (Eds.). ACM, [Place of publication not identified], 99–108. https://doi.org/10.1145/3131277.3132180

Ken Pfeuffer, Lukas Mecke, Sarah Delgado Rodriguez, Mariam Hassib, Hannah Maier, and Florian Alt. 2020. Empirical Evaluation of Gaze-enhanced Menus in Virtual Reality. In *26th ACM Symposium on Virtual Reality Software and Technology (ACM Digital Library)*, Robert J. Teather (Ed.). Association for Computing Machinery, New York,NY,United States, 1–11. https://doi.org/10.1145/3385956.3418962

W. Piekarski and B. H. Thomas. 2001. Tinmith-Metro: new outdoor techniques for creating city models with an augmented reality wearable computer. In *Proceedings Fifth International Symposium on Wearable Computers*. IEEE Comput. Soc, 31–38. https://doi.org/10.1109/ISWC.2001.962093

Jeffrey S. Pierce, Brian C. Stearns, and Randy Pausch. 1999. Voodoo Dolls: Seamless Interaction at Multiple Scales in Virtual Environments. In *Proceedings of the 1999 symposium on Interactive 3D graphics - SI3D '99*, Jarek Rossignac, Jessica Hodgins, and James D. Foley (Eds.). ACM Press, New York, New York, USA, 141–145. https://doi.org/10.1145/300523.300540

Andreas Riegler, Bilal Aksoy, Andreas Riener, and Clemens Holzmann. 09212020. Gaze-based Interaction with Windshield Displays for Automated Driving: Impact of Dwell Time and Feedback Design on Task Performance and Subjective Workload. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, New York, NY, USA, 151–160. https://doi.org/10.1145/3409120.3410654

Radiah Rivu, Yasmeen Abdrabou, Ken Pfeuffer, Augusto Esteves, Stefanie Meitner, and Florian Alt. 2020a. StARe: Gaze-Assisted Face-to-Face Communication in Augmented Reality. In *ACM Symposium on Eye Tracking Research and Applications* (Stuttgart, Germany) *(ETRA '20 Adjunct)*. Association for Computing Machinery, New York, NY, USA, Article 14, 5 pages. https://doi.org/10.1145/3379157.3388930

Radiah Rivu, Yasmeen Abdrabou, Ken Pfeuffer, Mariam Hassib, and Florian Alt. 2020b. Gaze'N'Touch: Enhancing Text Selection on Mobile Devices Using Gaze. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382802

Sheikh Rivu, Yasmeen Abdrabou, Thomas Mayer, Ken Pfeuffer, and Florian Alt. 2019. GazeButton: Enhancing Buttons with Eye Gaze Interactions. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications* (Denver, Colorado) *(ETRA '19)*. Association for Computing Machinery, New York, NY, USA, Article 73, 7 pages. https://doi.org/10.1145/3317956.3318154

Sensoryx AG. 2019. *VRFree by sensoryx*. https://www.sensoryx.com/

Manisah Mohd Shah, Haslina Arshad, and Riza Sulaiman. 2012. Occlusion in augmented reality. In *2012 8th International Conference on Information Science and Digital Content Technology (ICIDT2012)*, Vol. 2. IEEE, 372–378.

Toby Sharp, Cem Keskin, Duncan Robertson, Jonathan Taylor, Jamie Shotton, David Kim, Christoph Rhemann, Ido Leichter, Alon Vinnikov, Yichen Wei, Daniel Freedman, Pushmeet Kohli, Eyal Krupka, Andrew Fitzgibbon, and Shahram Izadi. 2015. Accurate, Robust, and Flexible Real-Time Hand Tracking. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3633–3642. https://doi.org/10.1145/2702123.2702179

Ludwig Sidenmark and Hans Gellersen. 2019. Eye&Head: Synergetic Eye and Head Movement for Gaze Pointing and Selection. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 1161–1174. https://doi.org/10.1145/3332165.3347921

Ludwig Sidenmark, Dominic Potts, Bill Bapisch, and Hans Gellersen. 05062021. Radi-Eye: Hands-Free Radial Interfaces for 3D Interaction using Gaze-Activated Head-Crossing. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yoshifumi Kitamura, Aaron Quigley, Katherine Isbister, Takeo Igarashi, Pernille Bjørn, and Steven Drucker (Eds.). ACM, New York, NY, USA, 1–11. https://doi.org/10.1145/3411764.3445697

Vildan Tanriverdi and Robert J. K. Jacob. 2000. Interacting with eye movements in virtual environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, Thea Turner (Ed.). ACM, New York, NY, 265–272. https://doi.org/10.1145/332040.332443

Bryan Wang and Tovi Grossman. 2020. *BlyncSync: Enabling Multimodal Smartwatch Gestures with Synchronous Touch and Blink*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376132

Jiayi Wang, Franziska Mueller, Florian Bernard, Suzanne Sorli, Oleksandr Sotnychenko, Neng Qian, Miguel A. Otaduy, Dan Casas, and Christian Theobalt. 2020. RGB2Hands: Real-Time Tracking of 3D Hand Interactions from Monocular RGB Video. *ACM Trans. Graph.* 39, 6, Article 218 (nov 2020), 16 pages. https://doi.org/10.1145/3414685.3417852

D. Weimer and S. K. Ganapathy. 1989. A synthetic visual environment with hand gesturing and voice input. *ACM SIGCHI Bulletin* 20, SI (1989), 235–240. https://doi.org/10.1145/67450.67495