2019-03

Augmented Reality Map Navigation with Freehand Gestures

Satriadi, Kadek Ananta; Ens, Barrett; Cordeil, Maxime; Czauderna, Tobias; Willett, Wesley J.; Jenny, Bernhard

Augmented Reality Map Navigation with Freehand Gestures

Kadek Ananta Satriadi*
Monash University

Barrett Ens†
Monash University

Maxime Cordeil‡
Monash University

Tobias Czauderna§
Monash University

Wesley Willett¶
University of Calgary
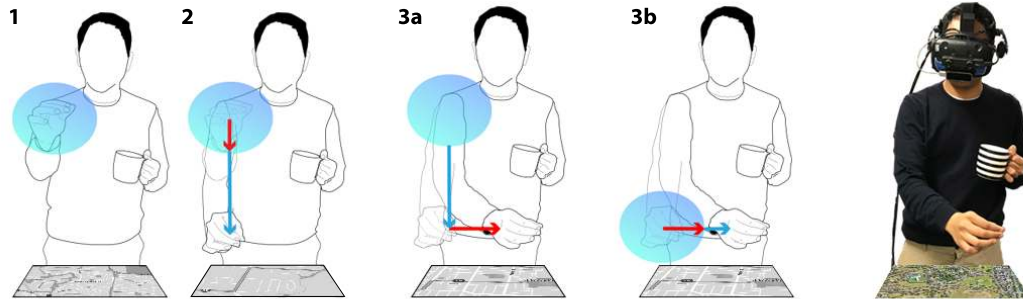
Bernhard Jenny‖
Monash University

Figure 1: An example use of hybrid interaction: (1) a pinch gesture starts navigation, followed by (2) vertical hand movement to zoom, transitioning from *Indirect Grab* (red line) to *Joystick* (blue line) input mappings. Two variations are (3a) *DiveZoom*, where horizontal movement results in simultaneous zoom and pan, and (3b) *TerraceZoom* where horizontal movement relocates the *Indirect Grab* region (blue ellipse), to begin a new pan action. (Right) Study setup using video see-through AR.

ABSTRACT

Freehand gesture interaction has long been proposed as a 'natural' input method for Augmented Reality (AR) applications, yet has been little explored for intensive applications like multiscale navigation. In multiscale navigation, such as digital map navigation, pan and zoom are the predominant interactions. A position-based input mapping (e.g. grabbing metaphor) is intuitive for such interactions, but is prone to arm fatigue. This work focuses on improving digital map navigation in AR with mid-air hand gestures, using a horizontal intangible map display. First, we conducted a user study to explore the effects of handedness (unimanual and bimanual) and input mapping (position-based and rate-based). From these findings we designed *DiveZoom* and *TerraceZoom*, two novel hybrid techniques that smoothly transition between position- and rate-based mappings. A second user study evaluated these designs. Our results indicate that the introduced input-mapping transitions can reduce perceived arm fatigue with limited impact on performance.

Index Terms: Human-centered computing— Mixed / augmented reality; Human-centered computing—Gestural input; Human-centered computing—Empirical studies in interaction design

1 INTRODUCTION

The recent commercialisation of wearable Augmented Reality (AR) hardware provides new opportunities to support human data understanding in applications such as immersive map browsing. However, there are currently no standard interaction methods for wearable AR, especially for navigation-intensive applications such as map

---

*e-mail: kadek.satriadi@monash.edu

†e-mail:barrett.ens@monash.edu

‡e-mail::max.cordeil@monash.edu

§e-mail:tobias.czauderna@monash.edu

¶e-mail:wesley.willett@ucalgary.ca

‖e-mail:bernie.jenny@monash.edu

browsing. The predominant interactions for digital map browsing are panning and zooming, so it is essential to develop efficient multiscale navigation methods for this platform.

This work focuses on improving multiscale navigation in the context of horizontally situated interactive maps, viewed using wearable AR (Figure 1). With the aim of extending these developments in future to collaborative interaction with 3D maps, we begin this exploration with single-user interaction with 2D maps, using a wide-FoV, video see-through AR platform [38, 57]. While in-air gestures have been explored for large wall displays [53, 65, 69] and mobile devices [28, 36], these previous developments use large arm motions or small finger movements that do not necessarily transfer directly to wearable AR and horizontal maps.

Freehand gesture input has become a common interaction method for wearable AR displays (e.g. Microsoft HoloLens [50], Meta 2 [48], Project North Star [41]), since it is intuitive [12], and does not require users to carry cumbersome controllers. However, mid-air gestures have known limitations, mainly limited precision with direct input on intangible surfaces [20,40,66], and arm fatigue with extended use [30, 31]. This work aims to address these limitations with a hybrid input mapping approach that combines intuitive position-based input with clutch-free rate-based input.

To improve our understanding of position- and rate-based input mappings in this context, we first conduct a comparative study of these input methods (respectively named *Indirect Grab* and *Joystick*) with two tasks: panning, and integrated pan and zoom. We then evaluate our two hybrid techniques against modifications of these baseline methods, with a long-duration, integrated pan-and-zoom task. Our results show that the introduced hybrid methods provide a good compromise in performance, comfort, and ease of learning between the use of either method alone. In particular, we found that a fully-integrated pan and zoom approach (*DiveZoom*) outperforms a constrained approach (*TerraceZoom*), and was preferred by study participants.

The contributions of this paper include 1) a comprehensive analysis of position- and rate-based input for freehand multiscale navigation, 2) the introduction of two novel hybrid techniques that combine the benefits of position- and rate-based input mappings, and 3) an evaluation that reveals a good compromise achieved by this hybrid

approach. In particular, the *DiveZoom* technique reduces perceived arm fatigue by 39.9% over the easy-to-learn *Indirect Grab* baseline without reducing the performance.

2 RELATED WORK

This work builds on prior research on navigation in large workspaces and multiscale navigation, in particular techniques that use in-air gestures.

2.1 Approaches for Interaction for Large Workspaces

In computer applications, workspace is independent of display size. If a display is very large, or smaller than the workspace, then efficient navigation is required.

Researchers have found that spatial interaction, including finger [46], arm [17], body [5], and head [23] movements, has advantages when navigating large display spaces. Similarly, *proxemics*, spatial relations between people and display devices, have been explored for pan and zoom navigation [4, 34].

Several approaches use hybrid interactions that combine multiple input modes or strategies. For instance, Vogel [69] uses mid-air pointing and grabbing gestures to combine absolute pointing with relative cursor movement on a large display. Other techniques combine multiple input modes for traversing large distances versus navigating within local regions [47, 52].

One closely related example of hybrid input is RubberEdge [16], which combines position- and rate-based input to reduce clutching on a laptop trackpad. Previous work has also explored such hybrid input in virtual environments [11] and haptic force-feedback devices [22]. These works used hybrid position- and rate-based input mapping for pointing, whereas we apply it to navigation with mid-air gesture input. However, we expect the key principles to carry over to navigation, as multiscale navigation is akin to pointing with one's view [26].

2.2 Multiscale Navigation

Multiscale navigation involves both panning and zooming, which allows viewing the workspace at different scales and eases navigation over long distances.

Several techniques aim to reduce the required effort for navigation: CycloPan and CycloZoom [45, 53] use continuous, clutch-free gestures for panning and zooming. Pinch-to-Zoom-Plus [2] reduces panning and clutching by providing control of the gain transfer function. Transient gestures [3] support rapid, temporary zooming to and from a local context. Negulescu et al. [54] similarly explore bimanual techniques to allow rapid, controlled zooming.

Other strategies attempt to eliminate navigation altogether: Polyzoom [35] uses hierarchies of multiple displays at different zoom levels to eliminate navigation effort. SpaceTokens [49] allow users to create widgets on a map's edge that save favourite locations.

Recent work has explored combinations of bimanual gesture [74] and gaze [56] that allow zooming with the non-dominant hand to be combined with other input commands made by the dominant hand.

2.3 Pan and Zoom with Mid-Air Gestures

Researchers have explored map navigation using spatial device input [58, 63], around-device gestures [28, 36], hand and head gestures for AR [59, 62], and mid-air input with large displays [1, 9, 43, 53, 61, 68]. We focus on mid-air gesture use with AR, a platform where freehand input is particularly desirable. However, most previous work on mid-air gesture navigation has considered large displays.

For instance, Gunslinger [43] uses a hip-mounted hand tracker to detect hands-down gestures, combined freely with direct touch input on the display. Tscharn et al. [68] found that, compared to mouse interaction, gesture input is more joyful, but less efficient. Nancel et al. [53] found that mid-air pan and zoom is more effective when

degrees of freedom are constrained. Additional works have explored gaze input along with gestures for pan and zoom input [27, 64].

Similar to techniques explored in our work are the joystick metaphor used for pan and zoom interaction by Stellmach and Dachselt [65], and the VolGrab technique of Shun et al. [61], which uses a 3D, indirect clutching metaphor.

Some related works exist in virtual reality [21, 51, 67]. Most of them focus on locomotion in virtual environment. These works differ from our application context of exocentric map navigation where content is moving relative to static user position.

3 FREE HAND MULTISCALE NAVIGATION

This research explores free hand gesture input, which allows users to control AR applications with their bare hands, eliminating the need to hold a physical device. More specifically, we use *mid-air* gestures, which are performed within the wearable device's sensor range in front of the user. Mid-air gestures are often used in AR and other spatial interaction contexts because of their 'intuitive' nature. Users can, for instance, select an object of interest by pointing [6] or manipulate a virtual object by 'grasping' it and orienting it as desired [57]. Our aim is to develop efficient map navigation techniques using mid-air gestures.

In this section we introduce the design factors relevant to multiscale navigation using mid-air hand gesture interaction. First, we describe our envisioned context for these developments.

3.1 Application Context

This work differs from most prior work on mid-air interaction in the context of the target technology and application setup. Whereas much previous work on hands free multiscale navigation took place in the context of vertical maps on large wall displays, our investigation centres on horizontal maps shown with head-worn AR displays.

Figure 2 illustrates our motivation to position maps horizontally. A horizontal layout allows for an intimate collaborative setting (see Conclusion and Future Work, below), where users can gather around a map in a face-to-face discussion [13, 60, 71]. Unlike large wall displays, a horizontal orientation allows users to reorient their viewpoint by walking around a floor or table layout. Furthermore, when 3D terrain or structures are shown, a horizontal layout reveals these details in their correct orientation. This requirement is particularly important for geospatial data exploration where the third dimension can be used to show geospatial phenomena, such as atmospheric data [29], ground water modeling [70], city modeling [24], and spatio-temporal data visualisation [37].

Usage of mid-air gestures in this context may differ from vertical displays, which tempts users to raise their arms toward the display. By placing a horizontal display at table-height, where a user's hand can remain between the shoulder and the waist [30], we can reduce overall arm rotation to increase the interaction duration before significant muscle fatigue is perceived [19].

Figure 2: Application context of our study. A horizontal map is suitable for exploratory geospatial analysis tasks because it allows for a relaxed arm position and supports collaborative 3D map tasks.

3.2 Design Factors

This discussion builds on previous design space discussions from Hinckley et al. [31], Nancel et al. [53], and Jones et al. [36], which cover primary factors for mid-air interaction.

3.2.1 Input Mapping

Two primary mappings of input space to control space are position-based and rate-based mappings. Position-based mapping maps the position of the input mechanism (the tracked hand for mid-air gestures) to the control space (currently, pan and/or zoom commands). Rate-based mapping maps force or displacement (the latter in our case) to the control space.

Early 3D interaction research [75] suggests that position-based input is preferable for position sensors, while rate-based mapping works best with a force-sensing (isometric) control device. Rate-based input is well-suited to elastic feedback [16, 75]. It is also used in other contexts such as device-tilt control, but is shown to be less accurate than position-based input [55].

With large workspaces such as maps, and with limited input range due to arm-reach, position-based input requires clutching. Clutching can be reduced by introducing a suitable control-display gain transfer function [52]. Such a function (commonly known as cursor acceleration) is frequently used with cursor positioning to allow both rapid movement and high precision at high and low velocities respectively.

Another consideration is the relative orientation of the input space (the physical space where free hand input occurs) to the display space (the map surface). Prior work [72] shows that a mismatch between input and display space can negatively impact performance.

3.2.2 Fatigue

Mid-air interaction is prone to arm fatigue [31] and users holding their arms in the air for an extended length of time leads to the well-known 'gorilla-arm effect' [8]. The adoption of micro-gestures is a promising solution, but awaits improved reliability in sensing technology [42]. Meanwhile, arm fatigue can be reduced by limiting the duration and range of arm movements; by allowing relative rather than absolute motions; and by facilitating a lowered arm posture with a bent elbow [30].

3.2.3 Handedness and Integration

Jacob and Sibert [33] propose that pan and zoom are integrally-related operations, and can be mapped to a 3D control space. When used together they are perceived by users as a single operation, similar to Buxton's concept of 'chunking' [14], where multiple parameters are controlled simultaneously within a single gesture.

Nancel et al. [53] take a bimanual input approach by dividing the pan and zoom commands to the dominant and non-dominant hands, respectively, while keeping the operations integrated by simultaneous input. The result aligned well with the known theory that distributing degrees of freedom over two hands yields better control over pan and zoom parallelism [10, 15, 25]. Stellmach and Dachselt [65], conversely, found unimanual input to be faster than bimanual in a map navigation task.

4 STUDY 1: INPUT MAPPING AND HANDEDNESS

To guide our design of hybrid techniques, we first need to refine and gain understanding of the underlying position- and rate-based input mappings with mid-air gestures. We conducted a controlled study to evaluate the impact of *Input Mapping*, and *Handedness* on a map navigation task.

While *position-based*[1] mapping is more intuitive [75], it requires *clutching* which may cause *fatigue*. The *rate-based* input is less intuitive, but allows long pan and zoom actions without any clutching.

---

[1]Hereafter we use *italics* to denote reference to the above design factors.

Although this work focuses on navigation, it is not the only operation needed for map use. To allow other commands to be seamlessly *integrated* along with pan and zoom [14], we aim to develop a *unimanual* technique (see Hybrid Input Mapping Design below) so that other operations can be performed by the second hand. Whereas several previous studies have found superior performance with *bimanual* interaction [10, 15, 25], we investigate whether this holds true in the current context.

4.1 Interaction Techniques

For this study, we designed four techniques that combine position- and rate-based input mappings with both unimanual and bimanual input (Figure 3).

4.1.1 Indirect Grab

Our position-based input method, termed *Indirect Grab*, mimics direct manipulation, where one simply grabs and moves the map using a pinch gesture. However, we instead use *indirect input* (Figure 3) to limit occlusion of the map by the user's hands and reduce *fatigue* by allowing users to keep their arms closer to their body. This technique is akin to several previous implementations [36, 53, 61] that use movement along the map plane for panning, and orthogonal movement for zooming. However, these works all used a vertical *display space*, whereas ours is horizontal. This results in a vertical *control space* for zoom, which may reduce *fatigue* due to a lower and more natural arm pose.

4.1.2 Joystick

We name our rate-based method *Joystick* after its namesake metaphor, where the rate of pan and zoom actions is given by the displacement from a starting pinch position (Figure 3). This method also uses *indirect input*, but allows continuous pan or zoom actions without clutching.
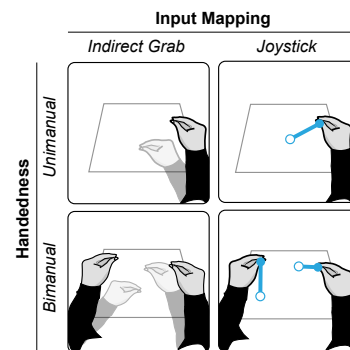

Figure 3: Study 1 input techniques.

4.1.3 Unimanual vs Bimanual

We implemented *Unimanual* and *Bimanual* variations of both *Input Mapping* methods. The *Unimanual* method maps the 3D position of the dominant hand to pan and zoom. The *Bimanual* method maps the dominant hand to the 2D panning plane, and the non-dominant hand to the orthogonal zoom dimension, following the predominant division of labour for bimanual pan and zoom interaction [10, 14, 25, 53]. All variations support *integrated* pan and zoom, since in all cases these actions can be performed simultaneously.

Another study found that bimanual and unimanual gesture techniques perform comparably in an AR scaling task [18]. However, their approach uses a 3DUI which is different from our study context.

4.1.4 Control-Display Gain Transfer and Visual Feedback

For both input mappings, we use a sigmoid gain transfer function [2] from [52]. The slope of the function provides a smooth transition from low to high velocity movement. We adjusted the parameters by conducting pilot studies until we found the parameters that are suitable for most of the users. Slow hand movement is mapped 1:1 to panning for *Indirect Grab* while fast hand movements allow rapid panning. Zoom gains are tuned to maintain a comfortable feel alongside pan. For internal consistency, we tuned the parameters of *Joystick* to match the performance of *Indirect Grab*. As with parameter tuning, we performed pilot studies to balance performance for both input mappings for short and long distances. We display simple visual feedback to indicate detected pinches for both mappings, as well as to show the displacement for the *Joystick* techniques.

4.2 Apparatus

We assembled a *video see-through AR* setup consisting of an HTC Vive equipped with a Zed Mini RGB stereo depth camera to allow HD (60 FPS @ 720p, 110° FoV) scene capture and correct occlusion. A Leap Motion mounted onto the headset provided hand tracking (Figure 1, right). We used a desktop PC with an Intel i7 processor, 32GB of RAM, and an Nvidia GeForce GTX 1080, ensuring a frame rate of 60–100 FPS. The size of the map viewport was 1 m × 1 m. We used a simple repeating grid instead of a map in this study.

4.3 Measures and Analysis

We measured performance efficiency with *Completion Time*. For each technique, participants rated perceived *Arm Fatigue* using Borg's RPE (Rating of Perceived Exertion) [7] measure. A post-study questionnaire was used to collect general comments and technique preference rankings.

We analysed completion time using repeated measures ANOVA. First, we checked for normality using Shapiro–Wilk tests. For non-normal completion time, we performed analyses on normally distributed log-transformed data. Post-hoc analyses were conducted using paired sample t-tests with Bonferroni p-value adjustments.

In Part 2 Results, both the raw and log-transformed completion time did not pass the normality test. Thus, we used Aligned Rank Transform (ART) technique [73] that allows non-parametric factorial analysis using repeated measures ANOVA. In this case, a Wilcoxon signed-ranks test [32] was performed for the post-hoc analysis. The questionnaire data were analysed using the Friedman test with a post-hoc Wilcoxon signed-ranks test.

4.4 Participants and Procedure

We recruited 16 participants (13 male, 3 female), all right-handed. Participants were recruited from the authors' host institution. Ten participants were in the 25 to 34 age group, 5 were in the 18 to 24 age group, and 1 was in the 35 to 44 age group. Five participants considered themselves highly familiar with VR/AR, and 3 had never experienced VR/AR before the study.

The study is broken down into two parts, containing a panning task, and an integrated pan and zoom task. The task in part 2 is modelled after the multiscale navigation task of Guiard et al. [26], also followed by Nancel et al. [53]. We added an additional panning task to evaluate each technique with extended panning motions (e.g. when tracing a path). The panning task was performed first, as pilot testing showed it was easier to learn, and served as additional training before learning the more difficult integrated task. The first part of the study lasted for 20–30 minutes and the second part lasted for 50–60 minutes with a break in between. For each interaction technique in each part, participants began with ample training and finished by completing the questionnaires.
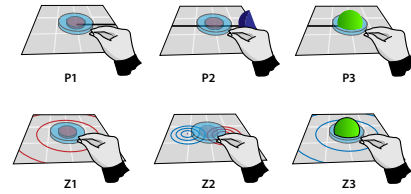
Figure 4: Top row: In the panning task, participants follow the black line (P1) until the blue target becomes visible (P2), then align the target with the centre of the map (P3). Bottom row: Zooming task steps consist of (Z1) starting point, (Z2) zoom out until target becomes visible, (Z3) zoom in to target and align it with the map centre. The size of the inner and outer circles are 10 cm and 30 cm respectively.

4.5 Part 1: Panning Task

For the initial *Panning task*, zooming was disabled, as we were interested to see the performance of the two input mappings without zoom integration. As such, participants only used the unimanual tasks in this part.

4.5.1 Task and Design

Participants were required to pan the map from the start location to the target location. To begin a trial, the participant was required to move their hand to a neutral starting position, marked by a cube. The trial ended when the participant successfully aligned the target sphere to a static cylinder at the map centre and confirmed by releasing their pinch (Figure 4, top row). The direction of the target was indicated by a straight line. We varied the task orientation and distance across four compass directions[3] and two target distances (0.5 m for *Close* and 2 m for *Far*).

We used a within-participants design with 3 factors:

- *Input Mapping* ∈ { *Indirect Grab*, *Joystick* }
- *Distance* ∈ { *Close*, *Far* }
- *Direction* ∈ { *N, S, E, W* }

Within each *Input Mapping*, 5 blocks of trials were presented, with each combination of *Distance* and *Direction* presented in random order. The order of *Input Mapping* was counter-balanced across subjects using a Latin square design. Each participant completed 2 × 2 × 4 × 5 = 80 trials for a total of 1280 trials.

4.6 Part 1 Results

4.6.1 Completion Time

For time analysis of this part and those that follow, we removed outliers caused by problems such as lost hand tracking or participants breaking mid-trial to relieve hand fatigue. Here we omited 24 trials (1.88%) with SD > 3. The ANOVA test did not show a significant main effect of *Input Mapping* on *Completion Time*. However, a main effect was found for *Distance* ($F_{1,14} = 767.226$, $p < .001$, $\eta^2 = .98$) indicating that the two levels of difficulty were set up correctly. An interaction effect was also found for *Direction* × *Distance* ($F_{3,42} = 12.317$, $p < .001$, $\eta^2 = .47$). Post-hoc test showed that, for the *Far* distance, panning to target in E and W directions is significantly faster ($p < .001$) than N and S directions (Figure 5, left).

The test also revealed interactions between *Input Mapping* and *Distance*. Although there is an indication that *Indirect Grab* is faster than *Joystick* with Close trials and slightly slower with *Far* ones, this difference is not significant.

4.6.2 Arm Fatigue

We analysed Borg's RPE using the Wilcoxon signed-rank test. The result indicated that *Indirect Grab* is more tiring than *Joystick* ($Z = 3.32$, $p < .001$).
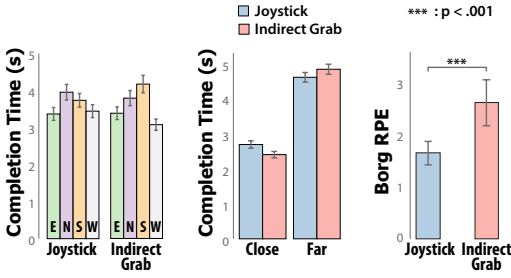
Figure 5: Study 1, Part 1 (left-to-right): Time by input mapping × direction and distance. Exertion by input mapping. Error bars in all of bar charts in this paper represent 1 SD.

4.6.3 Preference and subjective measure

The preference for this task is close with 56% of participants preferring *Indirect Grab* over *Joystick*. Subjective measures did not show any significant difference.

4.7 Part 2: Integrated Pan and Zoom

As we are interested in the combined effects of *Input Mapping* and *Handedness*, all four interaction techniques were employed in this task.

4.7.1 Task and Design

This task is adapted from Guiard's multiscale navigation task with two sets of abstract circles representing target and origin [26]. Participants are required to zoom out from the start location until the target becomes visible in the map frame, then zoom in and pan to the target location. The target is then selected and confirmed as in Part 1 (Figure 4, bottom row). In this task, target distances were set to 3 and 10 m.

We used a within-participant design with 4 factors:

- *Input Mapping* ∈ { *Indirect Grab*, *Joystick* }
- *Handedness* ∈ { *Unimanual*, *Bimanual* }
- *Distance* ∈ {*Close*, *Far* }
- *Direction* ∈ { *N, S, E, W* }

Within each technique (*Input Mapping × Handedness*) participants completed 3 blocks of trials, with each combination of *Distance* and *Direction* presented in random order. This yielded in $2 \times 2 \times 2 \times 4 = 96$ trials per participant, for a total of 1536 trials. Techniques were counterbalanced as in Part 1.

4.8 Part 2 Results

4.8.1 Completion Time

We removed 27 outliers (0.2% of total number of trials) in this analysis. We found main effects on the following factors: *Handedness* ($F_{1,14} = 6.488$, p < .05, $\eta^2 = .32$), *Input Mapping* ($F_{1,14} = 8.397$, p < .05, $\eta^2 = .38$), *Distance* ($F_{1,14} = 351.863$, p < .001, $\eta^2 = .96$), and *Direction* ($F_{3,42} = 4.355$, p < .005, $\eta^2 = .23$). The analysis showed that *Unimanual* (mean 9.31 s, SD 5.09 s) is faster than *Bimanual* (mean 10.13 s, SD 5.34 s), p < .001, and *Indirect Grab* (mean 8.80 s, SD 4.76 s) is faster than *Joystick* (mean 10.65 s, SD 5.51s), p < .05 (Figure 6, centre). These two results are confirmed by the absence of an interaction effect in *Handedness × Input Mapping*. The pairwise comparison of *Input Mapping × Direction* showed that for *Indirect Grab*, N and S directions are slower than E, N direction is slower than W. There was no significant difference in target directions with *Joystick* mapping (Figure 6, left).
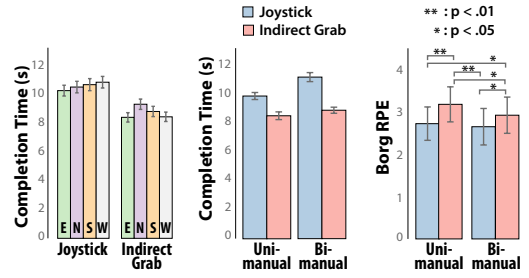

Figure 6: Study 1, Part 2 (left-to-right): Time by input mapping × direction and handedness. Exertion by input mapping × handedness.

4.8.2 Arm Fatigue

The Friedman test showed a significant effect in Borg's RPE ($\chi^2 = 23.64$, p < .001). The post-hoc analysis showed that the *Unimanual Indirect Grab* technique caused more exertion compared to all other techniques. *Bimanual Indirect Grab* is less tiring than *Unimanual Indirect Grab*, but it still has a higher RPE value than *Unimanual* and *Bimanual Joystick* (Figure 6).

4.8.3 Preference Ranking

We found that the user preference is not conclusive with the percentage of first rank as follows: *Unimanual Joystick* (31%), *Unimanual Indirect Grab* (25%), *Bimanual Joystick* (19%), and *Bimanual Indirect Grab* (25%).

4.9 Summary of Results

We found lower perceived fatigue for *Joystick*. This is likely due to the reduced amount of hand movement required to perform the task. An inspection of total hand movement for task 1 revealed that the *Joystick* input mapping covered 40.3% of total hand movement distance (normalised as percentage of total for all techniques). The data inspection also indicated that the average hand speed for *Indirect Grab* (mean 0.41 m/s, SD 0.37 m/s) is higher than *Joystick* (0.18 m/s, SD 0.21 m/s), which could potentially contribute to the result. One participant related fatigue with the amount of clutching required by the *Indirect Grab* technique when performing task 2 (P15: "*It takes a lot of pinches to complete the missions*"). In term of performance, the *Indirect Grab* mapping is faster than *Joystick* for the integrated Pan and Zoom task but not for Panning alone.

In an interesting contrast to previous studies [10, 15, 25], the bimanual techniques performed worse than the unimanual techniques. In part, this can be attributed to the fact that the majority of these studies used devices with constrained degrees of freedom and/or passive haptic feedback. Nancel et al. [53] showed that such constraints improve performance. Contrary to our result, the freehand mid-air bimanual technique by Nancel et al. [53] outperformed the unimanual version. We attribute this to requiring the participants to hold a mouse, and use its button for clutching, whereas our technique used a more natural hands-free pinching gesture. The difference may also be due in part to our horizontal layout which allows for a more relaxed arm posture than the wall display. Jones et al. [36] also found that around-device, unimanual input performed comparably to a touchscreen-based method.

We assume differences in direction are primarily due to ergonomic factors. In general, N and S are slower than E and W because panning N or S requires arm motions toward and away from the body. This requires the whole arm to be moved, while E and W panning can be done with the elbow fixed in place. This result is supported by participants' comments (P7: "*The forward/backward movement feels really frustrating as the distance panned is the same as the distance made by the hand but you are more limited because of your arm length compared to left or right movements, ...*", P6: "*Depth panning was not good enough*", P8: "*Going backwards gesture*
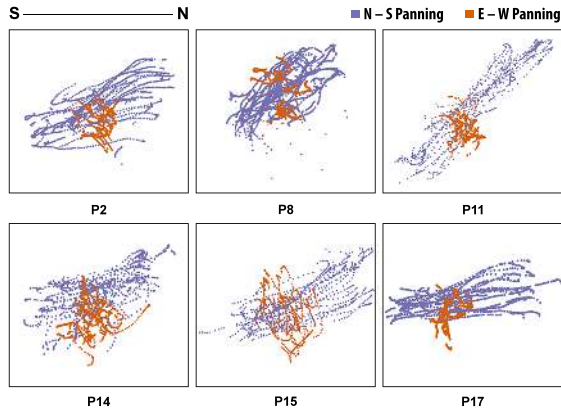
Figure 7: Side view of a some participants' hand position during *Panning task* with *Indirect Grab* technique. It can be seen that panning in *N* and *S* directions (blue) tend to follow oblique trajectories.

wasn't natural to me", P9: "*I feel panning tasks with targets at backward side are harder*").

5 HYBRID INPUT MAPPING DESIGN

In this section, we introduce two novel techniques that use hybrid input mappings. These techniques aim to maximise performance and minimise exertion by supporting both navigation over long distances and local control, with small gestures performed around the hand's comfort area. Our discussion begins with our strategy for pan and zoom integration, followed by the primary challenge of how to smoothly transition between input mappings.

5.1 Pan and Zoom Integration

A fundamental consideration is the degree of integration between pan and zoom. The most intensive multiscale navigation occurs when zooming in to a desired target [10]; in this zooming phase, pan and zoom integration allows *steering* towards the target. On the other hand, integrated zooming is not desirable during panning in regional exploration, because accidental zooming — which is likely to occur with in-air gestures — would interfere with the panning task. This is particularly so for in-air gestures when there is no separation between degrees of freedom [53].

When analysing hand movements in Part 1 of Study 1 (where the task was to pan, but not zoom) we indeed found that participants moved their hand also in the vertical direction when panning, particularly when panning in the N–S direction (Figure 7).

Given these considerations, we modified the *Indirect Grab* and *Joystick* techniques by differentiating between pan only (*Pan*) and integrated (*Zoom+Pan*) actions. We do this by assigning the hand's current movement vector to the appropriate region as illustrated in Figure 8. The *Zoom+Pan* region is represented as two axis-aligned mirrored cones with their apex at the current palm position and apex angles set to $90°$. The axis direction is always perpendicular to the forearm to support extension and rotation of the lower arm, thereby curtailing elbow movement and reducing fatigue [30, 44]. When the hand is inside the *Zoom+Pan* region, panning and zooming may be done simultaneously. When the hand is outside the *Zoom+Pan* region, only panning is possible.

Aside from these changes, the basic technique implementations are the same as in Study 1, including the control-display gain transfer function parameters.

5.2 *Indirect Grab*-to-*Joystick* Input Mapping Transition

In addition to switching between *Pan* and *Zoom+Pan* modes, our hybrid techniques support *Input Mapping* transitions between *Indirect Grab* and *Joystick*. Our aim is to provide transitions that are seamless and easy to learn.
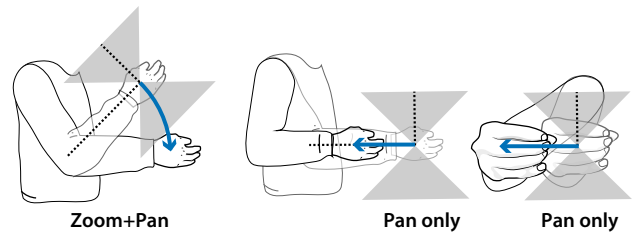

Figure 8: Integration of pan and zoom with two cones perpendicular to the forearm: hand movement inside the cones results in zooming (left), movement outside the cones results in N-S (centre) and E-W (right) panning. Blue arrows indicate hand movements.

Each time the user begins a pan or zoom action, an invisible ellipsoid region is placed around the user's initial palm centre (Figure 9). This ellipsoid defines the boundary between the *Indirect Grab* and *Joystick* input mapping regions. Inside the ellipsoid, the user can perform quick and small *Indirect Grab* movements, allowing precise local panning and zooming. Moving the hand outside of this ellipsoid causes a transition to *Joystick* mode. To make a seamless transition, the initial speed for the *Joystick* mode is set to the exit velocity of the hand from the ellipsoid. For *Indirect Grab* movements within the ellipsoid, differentiation between *Pan* and *Zoom+Pan* actions are determined by the two axis-aligned cones.

We use a tri-axial ellipsoid to provide ergonomic support, based on our analysis and user observations from Study 1. As a reminder, we found that participants were less effective navigating in the *N* and *S* directions than *E* and *W*, due to extraneous arm movement.

To reduce the required elbow movements, we compress the ellipsoid on the N–S axis to lower the escape distance threshold. In addition, we combine two semi-ellipsoids, such that the semi-axis from the centre towards the camera (in *S*-direction) is shorter than the semi-axis in *N*-direction. We selected the following semi-axis lengths: vertical: 2 cm, E–W: 4 cm, centre–N: 3 cm, centre–S: 2 cm.

We scaled the control-display gain transfer function in proportion to the ellipsoid dimensions, so that *Indirect Grab* is more sensitive to gestures along the shorter axes.
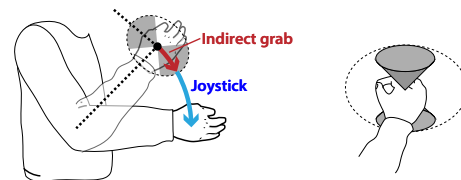

Figure 9: Transition between *Indirect Grab* (inside ellipsoid) and *Joystick* (outside ellipsoid).

5.3 Hybrid Techniques: *DiveZoom* and *TerraceZoom*

Based on the above approach, we designed two hybrid technique variations combining *Indirect Grab* and *Joystick* that we call *DiveZoom* and *TerraceZoom*. Based on our participant observations in Study 1, we noticed that several participants had difficulty simultaneously controlling pan and zoom with the fully integrated *Joystick* technique when steering toward a target. We therefore do not include fully integrated *Joystick* navigation in our hybrid techniques. Further pilot testing determined that the steering is easier when panning uses position-based *Indirect Grab* rather than rate-based *Joystick*.

Our two hybrid variations use *Indirect Grab* inside the ellipsoid, and distinguish between pan and zoom using the two axis-aligned cones. Figure 10 illustrates the state machine: inside the ellipsoid the *Indirect Grab* transitions between *Pan* and *Zoom+Pan* are identical for both hybrid techniques. Both techniques also transition to *Joystick* panning when the hand leaves the ellipsoid in approximately horizontal direction (i.e. the hand was not inside a conical

Zoom+Pan region before leaving the ellipsoid). However, the two methods differ when the hand leaves the ellipsoid in the vertical direction.

DiveZoom[4] combines *Joystick* for zooming with *Indirect Grab* for panning. An example sequence of gestures is illustrated in Figure 1. After the hand has left the ellipsoid (Figure 1-2) and a *Joystick* action has started, the hand moves horizontally (Figure 1-3a), which adds *Indirect Grab* panning to the mix, allowing for simultaneous rapid zooming and relatively short-distance panning. Note that once simultaneous pan and zoom is initiated, pan only mode can only be initiated by re-entering the ellipsoid, as shown in the state machine diagram in Figure 10 (or by releasing and re-initiating the pinch gesture).

TerraceZoom[5] only allows *Joystick* zooming, but relocates the ellipsoid without re-pinching. This is illustrated in Figure 1-3b: when the hand moves horizontally, *Joystick* zooming stops and the ellipsoid is moved to the current hand position. This causes the hand to enter the *Pan Only* zone inside the ellipsoid, allowing quick panning adjustments when zooming in to a target. In Figure 1-3b, the hand then leaves the ellipsoid and transitions to *Joystick* panning.
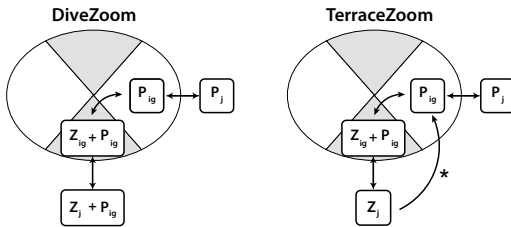

Figure 10: *DiveZoom* and *TerraceZoom* state machine diagrams. The * on *TerraceZoom* indicates the ellipsoid relocation. P_{ig}: pan in *Indirect Grab*, P_j: pan in *Joystick*, Z_{ig}: zoom in *Indirect Grab*, Z_j: zoom in *Joystick*.

6 STUDY 2: HYBRID INPUT MAPPING

We conducted a user study to evaluate the hybrid *DiveZoom* and *TerraceZoom* techniques. As baselines, we compare these against modifications of the unimanual *Indirect Grab* and *Joystick* techniques. We wanted to evaluate these with a more externally-valid setting than the first study, so we designed a longer task to reflect real map usage with a combination of zooming and panning.

We use the same setup as in the previous study, but show a satellite image map instead of an abstract grid.

6.1 Task

To encourage a combination of panning and zooming, each task consists of a series of pan and zoom segments. The start viewport is represented as a square on Figure 11. The task begins with a zoom segment, which requires zooming out to find the target (round point on the schema), align on the target and zoom in to reveal the route. This simulates a situation where a user zooms out to find a desired location, then zooms in to explore the region closely.

Panning segments are similar to the panning task in Study 1, where participants must trace the route, except the routes now contain direction changes. Zooming remains enabled, but to encourage panning the route is hidden when the user zooms out. Once a pan or zoom destination target is reached, a dwelling time of 500 ms confirms the target hit. Overshoots occur when the participant fails to maintain the hit position (within 30 cm radius of the target point) within the dwelling time.

---

[4]The name *DiveZoom* is inspired by skydiving, where the direction of free-falling can be steered within limits.

[5]*TerraceZoom* is akin to climbing down a terraced slope, where one can pause after each step and walk horizontally.

Each task consists of three alternating zoom and pan segments as illustrated in Figure 11. The distance, direction (N, E, S or W), and route direction changes (0–3) were randomised. For consistency, the total pan (5 m) and total zoom distances (3 km) within each task were normalised across all tasks.

6.2 Design and Procedure

We evaluated the four *Interaction Techniques* described in the previous section (*Indirect Grab*, *Joystick*, *DiveZoom*, *TerraceZoom*). Each participant completed a set of 16 pre-constructed *Tasks*, 4 with each technique. We used a Latin square to balance the presentation order of the *Interaction Techniques* to reduce learning effects. Training was given for each technique, and each was followed by questionnaires. The 4 *Interaction Techniques* × 4 *Tasks* design yielded 16 trials for each participant and 192 trials, with 576 segments of pan and zoom each, across 12 participants. The study ran for roughly 60 minutes.
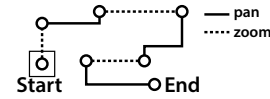

Figure 11: Composite navigation task.

Task measures were *Completion Time*, *Panning Overshoots*, *Arm Fatigue* (Borg), and subjective preference rankings. In addition, we asked participants to rate the ease of learning.

6.3 Participants

We recruited 12 participants (8 male, 4 female), between the ages of 21 to 43 (mean 28.50, SD 5.82). One participant was ambidextrous and chose their left hand to perform the tasks, while all others were right-handed. Ten participants were highly familiar with panning and zooming navigation. Four participants were highly experienced with VR/AR, six had little experience, and one had no experience.

6.4 Results

We used the same analysis method as in the first study, which consisted of the combination of parametric and non-parametric tests.

6.4.1 Completion Time

Four outliers (2.1% of total number of trials) were removed. The analysis found significant differences between techniques on *Completion Time* ($F_{3,33} = 7.861$, $p < .001$, $\eta^2 = .42$), as seen in Figure 12. Post-hoc tests revealed that *Indirect Grab* (mean 69.72 s, SD 25.11 s) is faster than *TerraceZoom* (mean 80.58 s, SD 23.31 s), and *Joystick* (mean 89.67 s, SD 42.65 s), but is not faster than *DiveZoom* (mean 77.20 s, SD 30.31 s). *DiveZoom* is significantly faster than *Joystick*.

6.4.2 Panning Overshoots

We also found a main effect of technique on *Overshoots* ($\chi^2 = 16.183$, $p < .05$). The pairwise comparison (Figure 12, right) result shows that *DiveZoom* caused significantly more overshoots than *Joystick* ($p < .01$) and *Indirect Grab* ($p < .05$). Interestingly, the
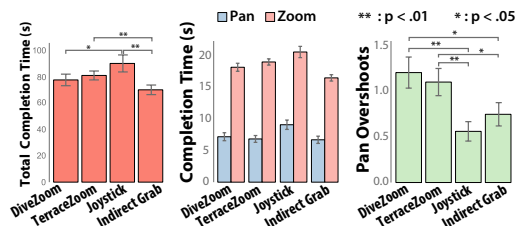

Figure 12: Study 2 mean total completion time, mean pan/zoom segment time (task time = approx. $3 \times (pan + zoom)$), and overshoots.

TerraceZoom technique resulted in more overshoots than *Joystick* (p < .01), but there is no significant difference with *Indirect Grab*.

6.4.3 Arm Fatigue

The analysis of Borg's RPE data showed a main effect of *Technique* ($\chi^2 = 15.61$, p < .005). Pairwise comparisons reveal several differences as seen in Figure 12, left. From the result, we can see that *DiveZoom* caused less *Arm Fatigue* than *TerraceZoom* (p < .01) and *Indirect Grab* (p < .01). As expected, *Indirect Grab* is more fatiguing than *Joystick* (p < .05). No significant difference was found between *DiveZoom* and *Joystick*. However, it was found that *TerraceZoom* technique is more fatiguing than *Joystick* (p < .05)

From further analysis of hand tracking data, we found that overall fatigue reflects the amount of movement required for each technique. Normalised as a percentage of the total for all techniques combined, relative hand travel distances are 20.4% for *DiveZoom*, 24.6% for *TerraceZoom*, 20.1% for *Joystick*, and 34.8% for *Indirect Grab*.

6.4.4 Ease of Learning and Preference

Ease of Learning was rated on a 5-point Likert scale (1 lowest, 5 easiest). The analysis indicates differences ($\chi^2 = 17.48$, p < .001). Pairwise comparisons show that *Indirect Grab* is easier to learn than all other techniques, and *Joystick* is significantly worse than *DiveZoom* and *Indirect Grab* (Figure 13, middle). In term of user preferences, the *Indirect Grab* and *DiveZoom* techniques are the most favorable (Figure 13, right).
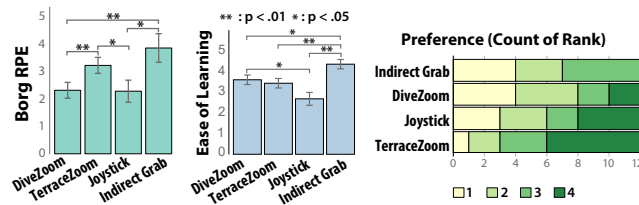

Figure 13: Study 2. Exertion, Ease-of-learning, and preferences.

7 DISCUSSION

Overall, both hybrid techniques show a compromise in performance and fatigue between *Joystick* and *Indirect Grab*. As in the Study 1, *Indirect Grab* was substantially faster than *Joystick*, while *Joystick* had lower fatigue. The *Completion Time* for the hybrid methods falls midway between both baselines, with *DiveZoom* nearing *Indirect Grab* (Figure 12). Meanwhile, for *Fatigue*, *TerraceZoom* falls midway between both baselines, and *DiveZoom* has roughly the same lower value as Joystick (Figure 13). *DiveZoom* provides the better compromise of the two hybrids, with similar task performance to the speedy *Indirect Grab* baseline and fatigue roughly equal to the less-tiresome *Joystick* method.

The hybrid techniques had the highest number of *Overshoots*, possibly due to higher panning velocities, which are set on exiting the transition boundary. This indicates that further tuning of the control-display gain transfer function parameters is required to provide better control. This result contrasts with the findings of Nancel et al. [53] who found that overshoots correlate with task completion time. We also attributed the overshoot to the absence of ellipsoid indicator that caused participants to accidentally escape *Indirect Grab* volume.

Interesting aspects of the four techniques are revealed by the plots in Figure 14, which show example zoom segments from Study 2. The initial zooming out phase appears similar for all techniques except for *Indirect Grab*, in which the many clutches can be seen. When zooming back in, the extra time required for the terraced steering is visible for *TerraceZoom*, and the speed-energy consumption trade-off is apparent in *Indirect Grab*.

Study 2 preference rankings indicate that *Indirect Grab* and *DiveZoom* are the most preferable among the techniques. Participant
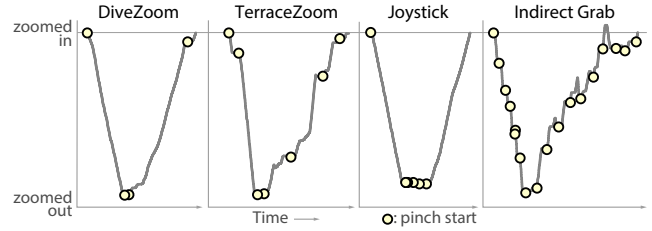

Figure 14: Examples of zoom level from one participant. The circles indicate pinch events.

comments reinforce this finding. The panning integration when zooming in is seen as an advantage of the *DiveZoom* technique (P2: "*... Having the option to pan when zoomed in felt a lot better, because it gave me options and it was mostly seamless...*"), which *TerraceZoom* does not have (P9: "*... Zooming in is difficult because you want to also pan by 'flying in' as zooming in.*").

Several comments reinforce the relative trade-offs of the *Indirect Grab* and *Joystick* techniques: *Indirect Grab* was seen as easy to use, but tiring (P10: "*Easiest technique so far, but probably the most strenuous on the long run. Still I loved how easy I could control my movements.*"). Most participants found *Joystick* difficult to learn, but several participants managed to use it enjoyably (P9: "*Zooming in and 'flying' is fun. Changing to pan when zoomed in enough requires some mental reminder of the direction.*"). These comments give us confidence that hybrid input mapping is a good approach for overcoming the limitations of both methods to provide efficient and enjoyable interaction for free-hand navigation input for maps AR.

Lastly, we acknowledge two main limitations of this study. The current hand tracking technology could be affecting the interaction, especially when pinching detection is unstable causing issues in input mapping transition (P3:"*triggering joystick took some effort for certain directions ...*"). We spent a substantial amount of effort to tune the transfer function parameters for these studies. However, the fact remains that a different choice of transfer function could affect the results.

8 CONCLUSION AND FUTURE WORK

We introduce two novel techniques to investigate the benefits of hybrid input mapping for in-air gestures. These techniques are explored in the context of multiscale map navigation in AR. These techniques seamlessly integrate position- and rate-based input mappings to minimise arm motion, with an ergonomic arm posture. Overall, we find the methods provide a good compromise in performance, fatigue and learnability between both baseline input mappings. Although the hybrid techniques are not faster than the position-based baseline, the *DiveZoom* technique provides comparable performance, but with reduced arm fatigue for long exploratory tasks. These results also reinforce prior findings that integrating pan with zooming action performs better than a constrained approach [10, 26].

As our controlled studies used abstract navigation tasks, we would like to further examine this hybrid approach when used with real map navigation, to identify opportunities for further improvement. Having achieved our goal of creating a robust, one-handed input method, our next step is to explore how multiscale navigation can be further integrated with additional commands. Such a bimanual approach combining selection with action has led to intriguing designs in other domains [39, 56, 74], and may provide similarly interesting possibilities by combining map navigation with common map operations. Further opportunities lie in exploring how such interaction methods can support effective collaboration.

REFERENCES

[1] V. K. Adhikarla, P. Wozniak, A. Barsi, D. Singhal, P. T. Kovács, and T. Balogh. Freehand interaction with large-scale 3d map data. In *3DTV-*

Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2014, pp. 1–4. IEEE, 2014.

[2] J. Avery, M. Choi, D. Vogel, and E. Lank. Pinch-to-zoom-plus: An Enhanced Pinch-to-zoom That Reduces Clutching and Panning. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pp. 595–604. ACM, New York, NY, USA, 2014.

[3] J. Avery, S. Malacria, M. Nancel, G. Casiez, and E. Lank. Introducing Transient Gestures to Improve Pan and Zoom on Touch Surfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 25:1–25:8. ACM, New York, NY, USA, 2018.

[4] S. K. Badam, F. Amini, N. Elmqvist, and P. Irani. Supporting visual exploration for multiple users in large display environments. In *Visual Analytics Science and Technology (VAST), 2016 IEEE Conference on*, pp. 1–10. IEEE, 2016.

[5] R. Ball, C. North, and D. A. Bowman. Move to improve: promoting physical navigation to increase user performance with large displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 191–200. ACM, New York, NY, USA, 2007.

[6] R. A. Bolt. "put-that-there": Voice and gesture at the graphics interface. In *Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '80, pp. 262–270. ACM, New York, NY, USA, 1980.

[7] G. Borg. *Borg's perceived exertion and pain scales*. Human Kinetics, Champaign, IL, USA, 1998.

[8] S. Boring, M. Jurmu, and A. Butz. Scroll, Tilt or Move It: Using Mobile Phones to Continuously Control Pointers on Large Public Displays. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group*, OZCHI '09, pp. 161–168. ACM, New York, NY, USA, 2009.

[9] M. N. K. Boulos, B. J. Blanchard, C. Walker, J. Montero, A. Tripathy, and R. Gutierrez-osuna. Web GIS in practice X: a Microsoft Kinect natural user interface for Google Earth navigation. *International Journal of Health Geographic*, 10(1):45:1–14, 2011.

[10] F. Bourgeois and Y. Guiard. Multiscale Pointing: Facilitating Pan-zoom Coordination. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '02, pp. 758–759. ACM, New York, NY, USA, 2002.

[11] D. A. Bowman and L. F. Hodges. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, pp. 35–ff. ACM, New York, NY, USA, 1997.

[12] D. A. Bowman, R. P. McMahan, and E. D. Ragan. Questioning naturalism in 3d user interfaces. *Communications of the ACM*, 55(9):78–88, 2012.

[13] S. Butscher, S. Hubenschmid, J. Müller, J. Fuchs, and H. Reiterer. Clusters, trends, and outliers: How immersive technologies can facilitate the collaborative analysis of multidimensional data. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 90. ACM, 2018.

[14] B. Buxton. Chunking and phrasing and the design of human-computer dialogues. In *Proceedings of the IFIP World Computer Congress*, pp. 475–480, 1986.

[15] W. Buxton and B. Myers. A Study in Two-Handed Input. *ACM SIGCHI Bulletin*, 17(4):321–326, 1986.

[16] G. Casiez, D. Vogel, Q. Pan, and C. Chaillou. RubberEdge: Reducing Clutching by Combining Position and Rate Control with Elastic Feedback. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, UIST '07, pp. 129–138. ACM, New York, NY, USA, 2007.

[17] J. Cauchard, M. Löchtefeld, M. Fraser, A. Krüger, and S. Subramanian. M+pspaces: Virtual workspaces in the spatially-aware mobile environment. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '12, pp. 171–180. ACM, New York, NY, USA, 2012.

[18] N. Chaconas and T. Höllerer. An evaluation of bimanual gestures on the microsoft hololens. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1–8, March 2018.

[19] D. B. Chaffin, G. Andersson, B. J. Martin, et al. *Occupational biome-*

chanics. Wiley New York, 1999.

[20] L.-W. Chan, H.-S. Kao, M. Y. Chen, M.-S. Lee, J. Hsu, and Y.-P. Hung. Touching the void: direct-touch interaction for intangible displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pp. 2625–2634. ACM, New York, NY, USA, 2010.

[21] R. Codd-Downey and W. Stuerzlinger. LeapLook: A Free-hand Gestural Travel Technique Using the Leap Motion Finger Tracker. In *Proceedings of the 2Nd ACM Symposium on Spatial User Interaction*, SUI '14, pp. 153–153. ACM, New York, NY, USA, 2014.

[22] L. Dominjon, A. Lecuyer, J.-M. Burkhardt, G. Andrade-Barroso, and S. Richir. The "bubble" technique: interacting with large virtual environments using haptic devices with limited workspace. In *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics Conference*, pp. 639–640. IEEE, 2005.

[23] B. M. Ens, R. Finnegan, and P. P. Irani. The personal cockpit: A spatial interface for effective task switching on head-worn displays. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pp. 3171–3180. ACM, New York, NY, USA, 2014.

[24] T. Glander and J. Döllner. Abstract representations for interactive visualization of virtual 3d city models. *Computers, Environment and Urban Systems*, 33(5):375–387, 2009.

[25] Y. Guiard. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behavior*, 19(4):486–517, 1987.

[26] Y. Guiard and M. Beaudouin-Lafon. Target acquisition in multiscale electronic worlds. *International Journal of Human Computer Studies*, 61(6):875–905, 2004.

[27] D. W. Hansen, H. H. Skovsgaard, J. P. Hansen, and E. Møllenbach. Noise tolerant selection by gaze-controlled pan and zoom in 3d. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, pp. 205–212. ACM, New York, NY, USA, 2008.

[28] K. Hasan, D. Ahlström, and P. P. Irani. Comparing Direct Off-Screen Pointing, Peephole, and Flick & Pinch Interaction for Map Navigation. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, SUI '15, pp. 99–102. ACM, New York, NY, USA, 2015.

[29] C. Helbig, H.-S. Bauer, K. Rink, V. Wulfmeyer, M. Frank, and O. Kolditz. Concept and workflow for 3d visualization of atmospheric data in a virtual reality environment for analytical approaches. *Environmental earth sciences*, 72(10):3767–3780, 2014.

[30] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pp. 1063–1072. ACM, New York, NY, USA, 2014.

[31] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell. A survey of design issues in spatial input. In *Proceedings of the 7th annual ACM symposium on User interface software and technology*, pp. 213–222. ACM, 1994.

[32] M. Hollander and D. A. Wolfe. *Nonparametric Statistical Methods*. Wiley-Interscience, New York, NY, USA, 1999.

[33] R. J. K. Jacob and L. E. Sibert. The Perceptual Structure of Multidimensional Input Device Selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '92, pp. 211–218. ACM, New York, NY, USA, 1992.

[34] M. R. Jakobsen, Y. S. Haile, S. Knudsen, and K. Hornbæk. Information visualization and proxemics: design opportunities and empirical findings. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2386–2395, 2013.

[35] W. Javed, S. Ghani, and N. Elmqvist. Polyzoom: Multiscale and Multifocus Exploration in 2D Visual Spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pp. 287–296. ACM, New York, NY, USA, 2012.

[36] B. Jones, R. Sodhi, D. Forsyth, B. Bailey, and G. Maciocci. Around Device Interaction for Multiscale Navigation. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '12, pp. 83–92. ACM, New York, NY, USA, 2012.

[37] T. Kapler and W. Wright. Geotime information visualization. *Information visualization*, 4(2):136–146, 2005.

[38] E. Kruijff, J. Orlosky, N. Kishishita, C. Trepkowski, and K. Kiyokawa. The influence of label design on search performance and noticeability in wide field of view augmented reality displays. *IEEE Transactions on Visualization and Computer Graphics*, 2018.

[39] B. Laha and D. A. Bowman. Volume cracker: a bimanual 3d interaction technique for analysis of raw volumetric data. In *Proceedings of the 1st symposium on Spatial user interaction*, pp. 61–68. ACM, 2013.

[40] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.

[41] Leap Motion. Project North Star. http://blog.leapmotion.com/northstar/, 2018.

[42] J. Lien, N. Gillian, M. E. Karagozler, P. Amihood, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev. Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar. *ACM Transactions on Graphics*, 35(4):1–19, 2016.

[43] M. Liu, M. Nancel, and D. Vogel. Gunslinger: Subtle arms-down mid-air interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, pp. 63–71. ACM, New York, NY, USA, 2015.

[44] P. Lubos, G. Bruder, O. Ariza, and F. Steinicke. Touching the Sphere: Leveraging Joint-Centered Kinespheres for Spatial User Interaction. In *Proceedings of the 2016 Symposium on Spatial User Interaction*, SUI '16, pp. 13–22. ACM, New York, NY, USA, 2016.

[45] S. Malacria, E. Lecolinet, and Y. Guiard. Clutch-Free Panning and Integrated Pan-Zoom Control on Touch-Sensitive Surfaces: The CycloStar Approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pp. 2615–2624. ACM, New York, NY, USA, 2010.

[46] S. Malik, A. Ranjan, and R. Balakrishnan. Interacting with large displays from a distance with vision-tracked multi-finger gestural input. In *Proceedings of the 18th annual ACM symposium on User interface software and technology*, pp. 43–52. ACM, New York, NY, USA, 2005.

[47] D. C. McCallum and P. Irani. Arc-pad: absolute+ relative cursor positioning for large displays with a mobile touchscreen. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pp. 153–156. ACM, New York, NY, USA, 2009.

[48] Meta Company. Meta 2. https://www.metavision.com/, 2016.

[49] D. Miau and S. Feiner. SpaceTokens: Interactive Map Widgets for Location-centric Interactions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 248:1–248:12. ACM, New York, NY, USA, 2018.

[50] Microsoft. HoloLens. https://www.microsoft.com/en-US/hololens, 2016.

[51] M. Nabiyouni, B. Laha, and D. A. Bowman. Poster: Designing effective travel techniques with bare-hand interaction. In *3D User Interfaces (3DUI), 2014 IEEE Symposium on*, pp. 139–140. IEEE, 2014.

[52] M. Nancel, O. Chapuis, E. Pietriga, X.-D. Yang, P. P. Irani, and M. Beaudouin-Lafon. High-precision Pointing on Large Wall Displays Using Small Handheld Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pp. 831–840. ACM, New York, NY, USA, 2013.

[53] M. Nancel, J. Wagner, E. Pietriga, O. Chapuis, and W. Mackay. Mid-air Pan-and-Zoom on Wall-sized Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 177–186. ACM, New York, NY, USA, 2011.

[54] M. Negulescu, J. Ruiz, and E. Lank. ZoomPointing Revisited: Supporting Mixed-resolution Gesturing on Interactive Surfaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '11, pp. 150–153. ACM, New York, NY, USA, 2011.

[55] I. Oakley and S. O'Modhrain. Tilt to scroll: Evaluating a motion based vibrotactile mobile interface. In *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics Conference*, pp. 40–49. IEEE, 2005.

[56] K. Pfeuffer, J. Alexander, and H. Gellersen. Partially-indirect Bimanual Input with Gaze, Pen, and Touch for Pan, Zoom, and Ink Interaction. In *Proceedings of the 2016 CHI Conference on Human Factors in*

Computing Systems, CHI '16, pp. 2845–2856. ACM, New York, NY, USA, 2016.

[57] T. Piumsomboon, D. Altimira, H. Kim, A. Clark, G. Lee, and M. Billinghurst. Grasp-Shell vs gesture-speech: A comparison of direct and indirect natural interaction techniques in augmented reality. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 73–82. IEEE, 2014.

[58] M. Rohs, J. Schöning, M. Raubal, G. Essl, and A. Krüger. Map navigation with mobile devices: virtual versus physical movement with and without visual context. In *Proceedings of the 9th International Conference on Multimodal interfaces*, pp. 146–153. ACM, New York, NY, USA, 2007.

[59] D. Rudi, I. Giannopoulos, P. Kiefer, C. Peier, and M. Raubal. Interacting with Maps on Optical Head-Mounted Displays. In *Proceedings of the 2016 Symposium on Spatial User Interaction*, SUI '16, pp. 3–12. ACM, New York, NY, USA, 2016.

[60] V. Rusnak, C. Appert, O. Chapuis, and E. Pietriga. Designing coherent gesture sets for multi-scale navigation on tabletops. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 142. ACM, 2018.

[61] S. Sekiguchi and T. Komuro. VolGrab: Realizing 3D View Navigation by Aerial Hand Gestures. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, pp. 2064–2069. ACM, New York, NY, USA, 2017.

[62] M. Serrano, B. M. Ens, and P. P. Irani. Exploring the use of hand-to-face input for interacting with head-worn displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pp. 3181–3190. ACM, New York, NY, USA, 2014.

[63] M. Spindler, M. Schuessler, M. Martsch, and R. Dachselt. Pinch-drag-flick vs. Spatial Input: Rethinking Zoom & Pan on Mobile Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1113–1122. ACM, New York, NY, USA, 2014.

[64] S. Stellmach and R. Dachselt. Investigating gaze-supported multimodal pan and zoom. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, pp. 357–360. ACM, New York, NY, USA, 2012.

[65] S. Stellmach, M. Jüttner, C. Nywelt, J. Schneider, and R. Dachselt. Investigating Freehand Pan and Zoom. In *Proceedings of Mensch & Computer*, pp. 303–312. Oldenbourg Verlag, München, 2012.

[66] Z. Szalavári and M. Gervautz. The Personal Interaction Panel–a Two-Handed Interface for Augmented Reality. *Computer Graphics forum*, 16(3):335–346, 1997.

[67] M. Tomberlin, L. Tahai, and K. Pietroszek. Gauntlet: Travel technique for immersive environments using non-dominant hand. In *2017 IEEE Virtual Reality (VR)*, pp. 299–300, March 2017.

[68] R. Tscharn, P. Schaper, J. Sauerstein, S. Steinke, S. Stiersdorfer, C. Scheller, and H. T. Huynh. User Experience of 3D Map Navigation - Bare-Hand Interaction or Touchable Device? In *Mensch und Computer 2016 - Tagungsband*. Gesellschaft für Informatik e.V., Aachen, Germany, 2016.

[69] D. Vogel and R. Balakrishnan. Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*, UIST '05, pp. 33–42. ACM, New York, NY, USA, 2005.

[70] M. Walther, L. Bilke, J.-O. Delfs, T. Graf, J. Grundmann, O. Kolditz, and R. Liedl. Assessing the saltwater remediation potential of a three-dimensional, heterogeneous, coastal aquifer system. *Environmental Earth Sciences*, 72(10):3827–3837, 2014.

[71] F. Wiehr, F. Daiber, F. Kosmalla, and A. Krüger. Artopos: augmented reality terrain map visualization for collaborative route planning. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pp. 1047–1050. ACM, 2017.

[72] D. Wigdor, C. Shen, C. Forlines, and R. Balakrishnan. Effects of display position and control space orientation on user preference and performance. In *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*, p. 309. ACM Press, New York, New York, USA, 2006.

[73] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned

rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 143–146. ACM, 2011.

[74] H. Xia, K. Hinckley, M. Pahud, X. Tu, and B. Buxton. WritLarge: Ink Unleashed by Unified Scope, Action, & Zoom. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pp. 3227–3240. ACM, New York, NY, USA, 2017.

[75] S. Zhai and P. Milgram. Human performance evaluation of manipulation schemes in virtual environments. In *Proceedings of IEEE Virtual Reality Annual International Symposium*, pp. 155–161. IEEE, 1993.