Interacting with Maps on Optical Head-Mounted Displays

David Rudi
ETH Zürich, Chair of Geoinformation Engineering
Stefano-Franscini-Platz 5
Zürich, Switzerland
davidrudi@ethz.ch

Ioannis Giannopoulos
ETH Zürich, Chair of Geoinformation Engineering
Stefano-Franscini-Platz 5
Zürich, Switzerland
igiannopoulos@ethz.ch

Peter Kiefer
ETH Zürich, Chair of Geoinformation Engineering
Stefano-Franscini-Platz 5
Zürich, Switzerland
pekiefer@ethz.ch

Christian Peier
ETH Zürich, Chair of Geoinformation Engineering
Stefano-Franscini-Platz 5
Zürich, Switzerland
chpeier@gmail.com

Martin Raubal
ETH Zürich, Chair of Geoinformation Engineering
Stefano-Franscini-Platz 5
Zürich, Switzerland
mraubal@ethz.ch

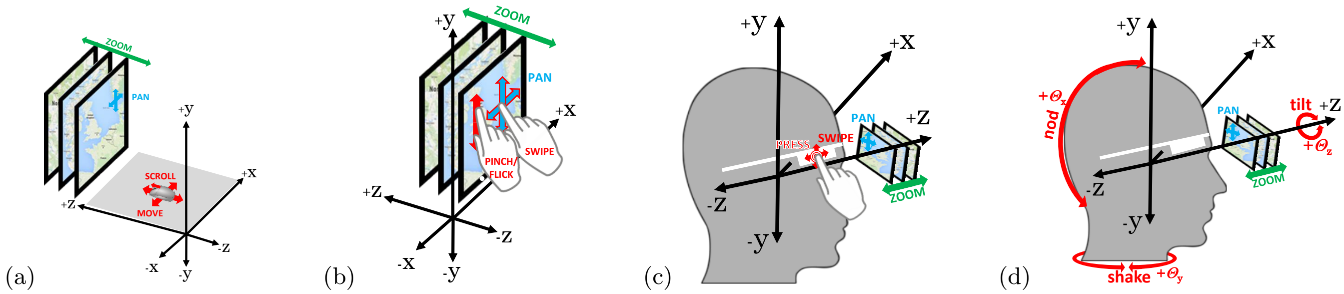


Figure 1: A depiction of how control inputs (e.g., moving the mouse along the x-axis to the left or right) correspond to map interactions (i.e., moving the map along the same axis/in the same direction) for: (a) mouse controls, (b) touch controls, (c) haptic controls on OHMD, (d) head controls on OHMD. (a) and (b) are adapted from [18].

ABSTRACT

This paper explores the design space for interacting with maps on Optical (See-Through) Head-Mounted Displays (OHMDs). The resulting interactions were evaluated in a comprehensive experiment involving 31 participants. More precisely, novel head-based interactions were compared with well-known haptic interactions on an OHMD regarding efficiency, effectiveness, user experience and perceived cognitive workload. The tasks involved navigating on maps by panning, zooming, and both panning and zooming. The results suggest that interaction methods exploiting congruent spatial relationships, i.e., mappings between the same axis in the control and display space, outperform others. In particular, the head-based interactions incorporating such mappings, significantly outperformed the haptic interactions for tasks involving panning, and combined tasks of panning and zooming.

SUI '16, October 15-16, 2016, Tokyo, Japan

DOI: http://dx.doi.org/10.1145/2983310.2985747

Keywords

Interaction Design; Optical Head-Mounted Display; Maps

1. INTRODUCTION

The market for wearables, such as OHMDs is rapidly growing, and constantly new interaction modalities are introduced. Nonetheless, many limitations still exist. A common limitation of OHMDs is a restricted surface for haptic inputs, resulting from amongst others, weight restrictions and ergonomic design guidelines. This limitation is especially cumbersome in situations where users wish to navigate in large information spaces [24]. For example, consider a surveyor who is taking measurements and simultaneously has to interact with a map of her surroundings. In that case, a limited interaction surface or even the use of hands is bothersome.

To overcome the limitations of OHMDs, some researchers extended the interaction modalities (i.e., how to interact) using the built-in OHMD-hardware by, e.g., adding gesture-recognition [17]. Others added hardware, such as eye trackers [13], while a number of researchers focused on extending the interaction methods (i.e., how to use the modalities) by, e.g., using the user's head-movement [26].

In this work, we focus on those two types of input which are already available on most OHMDs on the market: head-based gestures and touch interactions. We systematically explore the design space for interacting with these two types of input in large information spaces on OHMDs, such as maps. Both the head-based gestures and the touch interactions are performed in space. In particular, we argue that the congruency of the spatial movements (required for the interactions) with their mapping to the display space is decisive for their adequateness. Moreover, we introduce a novel design that allows for a better user performance while navigating on maps displayed on OHMDs using head-based gestures.

We developed a prototypical client for Google Glass supporting the Web Map Service (WMS) standard[1] and conducted a user study with 31 participants. Cartographic maps, amongst others, are documents containing a rich amount of information, requiring users to navigate using pan and zoom interactions [1, 24]. Considering that we live in a mobile information society, where people extensively use mobile maps and navigation services, there is a need to explore appropriate interaction methods with these [10]. This paper makes the following contributions concerning the navigation on maps using OHMDs:

- The exploration of the design space for the navigation on maps on OHMDs, following a method proposed by MacKenzie for interactions with desktops and tablets [18].
- A novel design for interactions with maps on OHMDs, based on head gestures with the aim to avoid unintentional interactions.
- The comparison of the novel head-based with the haptic interactions, with regard to efficiency (task completion time), effectiveness (error rate), user experience (UEQ), and perceived cognitive workload (raw NasaTLX) for tasks involving navigation on maps, i.e. panning, zooming and combinations thereof.

The remainder of this paper is structured as follows: sections 2 and 3 describe the existing research and evaluate the design space for map navigation on desktops and tablets, as well as OHMDs. Then, in section 4 we introduce our hypotheses and methodology. The results are presented in section 5 and discussed in section 6. Finally, the conclusions and future work are given in section 7.

2. RELATED WORK

This section reviews the interaction methods (i.e., input controls) for OHMDs that have been explored in previous research. Furthermore, section 2.2 explains the methodological approach proposed by [18] upon which we base our exploration of the design space in Section 3.

2.1 Interaction Methods for OHMDs

Head Mounted Displays (HMDs) are displays that are worn on the head and exist in many different variations [3]. They are usually categorized into optical see-through HMDs (OHMDs), video see-through HMDs or opaque ones. Information is either presented in front of one (monocular) or both eyes (binocular). The study in this paper is conducted using Google Glass [25], which is categorized as a monocular OHMD.

[1]A WMS is a web service providing (geographically referenced) map images, generated by a server using data from a Geographic Information System (GIS) database. See also http://www.opengeospatial.org/standards/wms

2.1.1 Haptic- and Other Interaction Methods

To circumvent the limited haptic interaction space, OHMDs were enhanced with external input devices, such as large external touch pads or a belt [19, 6].

Furthermore, several state-of-the art OHMDs provide alternatives to the haptic methods, such as external smartphones for controlling the OHMD remotely, or hand gesture recognition capabilities. The latter has been evaluated in detail [15] and extended by introducing input modalities that use different body parts for gesture recognition, such as the face [22] or feet [17].

Even though these methods provide a wide range of interaction possibilities, they all involve the use of extraneous body parts. This implies that they are very visible in public [26], and not suitable for users with disabilities [19] or in situations where users cannot involve the respective body parts in the task at hand. Some of these alternatives require the use of an additional device, which usually implies additional weight, costs and maintenance effort.

2.1.2 Hands-Free Interaction Methods

There has been extensive research on the use of voice as an interaction method for OHMDs, e.g., to perform game control tasks [26], to navigate between instructions [28] and to evaluate their use for people with disabilities [20]. However, it was found that voice commands have performance issues and are less favored interaction methods in public.

Head-based gestures were explored for simple interactions with mobile devices (e.g., scrolling [21]) or displays (e.g., point of view changes in games [5]). Concerning OHMDs, head-based gestures were for instance used to determine user activities (in combination with eye gestures) [12], for game control tasks [26], to select elements in physical spaces [27] and to evaluate smart spaces interactions [15]. Similar to voice commands, head-based gestures have the drawback of being visible in public and can result in fatigue.

Another approach to hands-free interaction with OHMDs is based on using gaze as an input method [11, 13]. While using gaze as an interaction method is promising, the currently available OHMDs can only detect blinks, but not fixations or saccades[2].

Regarding the previously described advantages and disadvantages of the various interaction methods, we decided to focus on the possibilities of using head-based gestures for interacting with OHMDs and compared our approach with the built-in capabilities of OHMDs using haptic inputs. A further reason for this decision is the fact that this type of interaction can also be used in scenarios where, e.g., the use of hands is not possible.

2.2 Spatial Control-Display Relationships

MacKenzie describes the *control-display relationship* as one that attributes to "how a controller property maps to a display property" (p. 75) and separates between *spatial*, *dynamic* and *physical relationships* (or "- mappings") [18]. *Spatial relationships* represent mappings from movements of the controller in space to movements on the display, e.g., the mapping from a mouse to a cursor. To systematically examine the relationships in more detail, he introduces (arbitrarily chosen) labels for the axes of a 3D Cartesian coor-

[2]The recently introduced HoloLens (https://www.microsoft.com/microsoft-hololens/en-us) is an exception, but has not yet surpassed the development state.

dinate space, i.e., the degrees of freedom (DOF) (see Figure 1). Table 1 visualizes these mappings, by showing how control inputs map to display outputs. For example, the first row shows that moving the mouse along the x-axis, moves the cursor (a map) along the same axis on a display.

spatially congruent mappings, i.e., a mapping between the same axes in the control and display space. These mappings pose natural relationships (as opposed to learned relationships) and thus result in a better user experience.

Cartographic maps can be considered 2D (or even 3D) representations of spatial phenomena [23]. They commonly follow the metaphor of viewing a certain extent of a space from a top-down perspective. To solve a task, a user is required to change this limited extent, either by moving it (panning) or by (de-) magnifying a specific part to the full extent (zooming). Both panning and zooming are considered to be the most important map interactions [8].

The spatial relationship between a mouse and a cursor easily translates to panning a map. A movement of the mouse along the x-axis (i.e., left or right) can be related to moving a map extent along the same axis (congruent), while moving the mouse along the z-axis (i.e., forwards or backwards) corresponds to moving the map along the y-axis (i.e., up or down) (see Figure 1). Zooming can be described as the movement along a virtual z-axis of a display, where each zoom-extent is a different layer. The mouse-wheel is commonly used as the control input when zooming and scrolling the wheel forwards or backwards along the z-axis results in movements along the same axis (see Figure 1 (a) and Table 1 ("Mouse")).

The mouse-display relationship can be extended to touch displays as depicted in Figure 1 (b) and Table 1 ("Touch"). That is, the controller is a user's touch input and the display a touch display, such as a tablet. Touch displays allow a fully congruent mapping for panning, i.e., moving the finger along the x-axis or y-axis (i.e., the z-axis, when the display is lying flat) results in a movement along the same axis for the map. As for zooming, the touch display's multi-touch capabilities allow users to flick or pinch with two fingers (moving them into opposite directions) to respectively zoom in or out. The control input by flicking or pinching takes place on a plane perpendicular to the z-axis (i.e., along the x-axis, the y-axis, or a combination of both).

Next to the ones described above, desktops and tablets have other hard controls, such as mouse-clicks and tapping. A thorough description of those and their spatial relationships can be found in [18]. For maps on OHMDs, a detailed evaluation of the spatial relationships is missing and is therefore evaluated in the next section for head-based gestures and touch interactions.

3. SPATIAL RELATIONSHIPS FOR MAPS ON OHMDS

Next to the interaction methods being explored in ongoing research, an analysis of the specifications of OHMDs on the market showed that they all support head-based inputs and most support touch inputs. For the studies conducted in this research Google Glass was used, which supports both. Hence, the spatial relationships that hold true for head and haptic gestures concerning panning and zooming on an OHMD are evaluated.

For pan interactions, movements of the map extent along the x- and y-axis need to be covered by two corresponding inputs of the controller. Zoom interactions require a control input for movements along the z-axis, while combinations of pan and zoom need a mapping for all axes.

Table 1: Visualization of how inputs along a control axis, map onto a display axis (adapted from [18]).

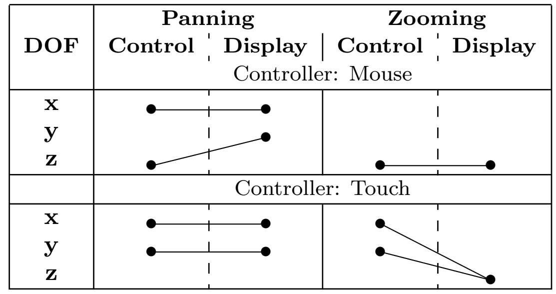
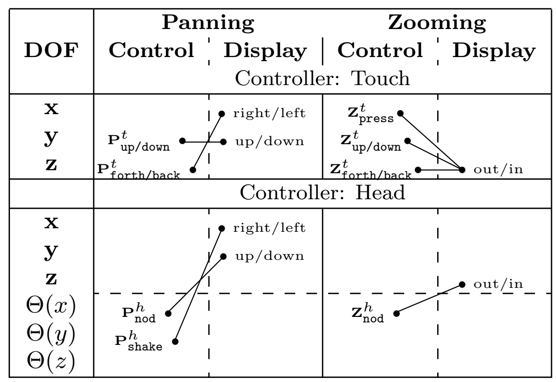


Table 2: An overview of the spatial relationships evaluated in this paper between touch and head control inputs with an OHMD for both pan and zoom.



Touch pads of OHMDs commonly provide 2.5 DOF for inputs. That is, swiping forwards or backwards along the z-axis, sliding upwards or downwards along the y-axis and tapping or pressing on the x-axis (see Figure 1 (c)). All of these inputs can be extended by considering the use of one or more fingers or (except for the x-axis) multi-touch gestures such as flicking or pinching.

The built-in gyroscope and accelerometers of OHMDs allow a distinction between six DOF for head-based gesture inputs. On the one hand, a user can move her head along the three axes (3 DOF), i.e., do "head slides" along the axes while standing still. On the other hand, a user can turn her head to the left or right around the y-axis (shake), look up or down around the x-axis (nod) and tilt it to the left or right around the z-axis (tilt) (see Figure 1 (d)).

3.1 Design Space

This section analyzes the design space. That is, spatial relationships are evaluated concerning potential map movements, for zoom-, pan- and mixed tasks. We justify which mappings were chosen to be included in the user study (see section 4.5) and provide an overview of the corresponding spatial relationships in Table 2.

For an easy and clear evaluation of the design space the following abbreviations are used: a large letter represents the map interaction, "**Z**" for zooming and "**P**" for panning; a small superscript letter represents the modality, "**h**" for head and "**t**" for touch; finally in subscript the modality is described, e.g., "**press**" for pressing the finger against the touch pad.

Panning, touch (P^t). The *x-axis* can be used for panning, either by considering subsequent taps (e.g., 1 tap = up, 2 taps = down) or different numbers of fingers that press against the touch pad (e.g., 1 finger = up, 2 fingers = down). However, simultaneously using up to 4 fingers (i.e., for 4 directions) would require a large interaction space, contradicting the common OHMD design restrictions. Additionally, our pilot study showed that excessive tapping led to displacements of the OHMD. Combinations of control inputs of pressing on the *x-axis* with swipes along the *y- or z-axis* were dismissed due to the anticipated higher workload resulting from the degree changes. Furthermore, the pilot study did not show any indication that swiping with a different number of fingers results in measurable performance differences. Finally, both flicking and pinching were abandoned, due to the limited space along the y-axis.

Consequently, the only remaining axes to be combined were the *y- and z-axis*, which were used with single finger inputs as $\mathbf{P}^t_{\texttt{up/down, forth/back}}$ (see Table 2 rows 2 and 3 under "Panning"). There is no consensus on the directionality of the input in the literature [14, 4], i.e., whether swiping forward on the touch pad of the OHMD corresponds to panning left or right on the map. In this work backward (forward) swipes are associated with pans to left (right) as suggested in [4].

Panning, head (P^h). The analysis of the design space for head-based gestures out-ruled the use of x, y and z, because "head slides" turned out to be difficult to perform and are difficult to differentiate from rotations (technically). While head "rotations" around all axes can be associated with either one of the pan movements, some variants were excluded based on the experiences from the pilot study. That is, "tilting" in general, "nodding" for left/right pans and "shaking" for up/down pans.

Hence, $\mathbf{P}^h_{\texttt{nod, shake}}$ (see Table 2 rows 7 and 8 under "Panning") was evaluated by congruently mapping pan directions according to the head movement, i.e., moving the head upwards (nodding) is mapped onto panning the map upwards. Note, that even though mapping "nodding" onto the y-axis does not appear congruent at first, it actually is, because a nodding user is looking up and down along an "invisible" y-axis.

Zooming, touch (Z^t). In contrast to panning, the *x-axis* can be used for zooming. While tapping was still excluded due to the inherent displacement of the OHMD, using different numbers of fingers that press and hold was a valid alternative. More precisely, based on our experience from the pilot study, $\mathbf{Z}^t_{\text{press}}$ (see Table 2 row 1 under "Zooming") was evaluated as the combination of pressing with two fingers to zoom-in and three to zoom-out.

Both the *y- and z-axis* can be used for zoom interactions. Regular swipes along these axes are denoted as $\mathbf{Z}^t_{\texttt{up/down}}$and $\mathbf{Z}^t_{\texttt{forth/back}}$ (see Table 2 rows 2 and 3 under "Zooming"). With regard to directionality, forward/backward swipes along the z-axis were interpreted as a zoom out/in interactions and upward/downward swipes as zooming out/in.

Zooming, head (Z^h). The axes x, y and z were excluded for zooming with the same argumentation as for panning. Furthermore, "tilting" and "shaking" were excluded for zooming, due to their lack of intuitiveness. Thus, $\mathbf{Z}^h_{\texttt{nod}}$ (see Table 2 row 7 under "Zooming")was evaluated and head-up (down) gestures were associated with zooming out (in).

Finally, combinations of haptic interactions and head-based gestures (e.g., swipe up = pan up and head left = pan left) for both panning and zooming were not evaluated, to avoid mode changes within a single interaction.

Panning and Zooming. Given the selected zoom and pan interactions, the design space for combinations thereof is large. However, combining $\mathbf{Z}^t_{\texttt{up/down}}$ or $\mathbf{Z}^t_{\texttt{forth/back}}$ with $\mathbf{P}^t_{\texttt{up/down, forth/back}}$, as well as $\mathbf{Z}^h_{\texttt{nod}}$ with $\mathbf{P}^h_{\texttt{nod, shake}}$ is not possible due to ambiguities.

Furthermore, the remaining combinations of $\mathbf{Z}^h_{\texttt{nod}}$ with $\mathbf{P}^t_{\texttt{up/down, forth/back}}$ and both $\mathbf{Z}^t_{\texttt{up/down}}$ and $\mathbf{Z}^t_{\texttt{forth/back}}$ with $\mathbf{P}^h_{\texttt{nod, shake}}$ were skipped in favor of comparing $\mathbf{P}^h_{\texttt{nod, shake}}$ with $\mathbf{P}^t_{\texttt{up/down, forth/back}}$ using the only zoom input that can be combined with both, i.e., $\mathbf{Z}^t_{\text{press}}$.

Thus, in this paper, the combinations ($\mathbf{P}^h_{\texttt{nod, shake}}$, $\mathbf{Z}^t_{\text{press}}$) and ($\mathbf{P}^t_{\texttt{up/down, forth/back}}$, $\mathbf{Z}^t_{\text{press}}$) are evaluated.

4. METHODOLOGY

In this section we first explain the overall study design, including the dependent and independent variables and present our hypotheses. Afterwards, we give an overview of the study, including information of the participants and the setup, i.e., the hardware and software. Next, we elaborate on how we realized the novel head-based gestures and explain the study procedure in detail. That is, the lessons learned from the pilot study, how we advanced for the main study, as well as a detailed explanation about the tasks involved.

4.1 Study Design and Hypotheses

A within subject design was employed for the experiment with the interaction method as the independent variable and effectiveness, efficiency, user experience and perceived cognitive workload as dependent variables.

More precisely, effectiveness was measured by the number of errors made. That is, each zoom or pan step that did not get closer to a target zoom level (i.e., the difference of the current zoom and the target zoom is increasing) or location (i.e., increasing euclidean distance) was counted as an error. However, errors did not accumulate for consecutive incorrect steps, i.e., moving farther away from the target twice was still counted as one error[3]. Efficiency was measured by the task completion time, while user experience as well as perceived cognitive workload were measured using standardized questionnaires (UEQ [16] and NASA TLX [9]).

With regard to the independent variables, users were given three different tasks:

- The first task involved zooming and had four conditions: $\mathbf{Z}^h_{\texttt{nod}}$, $\mathbf{Z}^t_{\text{press}}$, $\mathbf{Z}^t_{\texttt{up/down}}$ and $\mathbf{Z}^t_{\texttt{forth/back}}$.
- The second task required users to pan and had two conditions $\mathbf{P}^h_{\texttt{nod, shake}}$ and $\mathbf{P}^t_{\texttt{up/down, forth/back}}$.
- The final, third task was a combination of panning and zooming and had two conditions ($\mathbf{P}^h_{\texttt{nod, shake}}$, $\mathbf{Z}^t_{\text{press}}$) and ($\mathbf{P}^t_{\texttt{up/down, forth/back}}$, $\mathbf{Z}^t_{\text{press}}$).

[3]This was introduced to account for sensory imprecision.

To avoid a learning effect we used a counterbalanced design (balanced latin square) (see subsection 4.5.3 for details).

Considering the spatial mappings of the interaction methods and the assumption that congruent mappings between a control input and the display have a higher level of intuitiveness, the following **hypotheses** were made:

H1. For the panning task, the interaction method $\mathbf{P}^h_{\texttt{nod, shake}}$ performs best with regard to the given dependent variables.

H2. For the zooming task, the interaction method $\mathbf{Z}^t_{\texttt{forth/back}}$ performs best with regard to the given dependent variables.

H3. For the combination of panning and zooming, the interaction method ($\mathbf{P}^h_{\texttt{nod, shake}}$, $\mathbf{Z}^t_{\texttt{press}}$) performs best with regard to the given dependent variables.

4.2 Participants

In total 31 participants (18 males) were recruited for the experiment. The participants had a mean age of 30 years *(SD = 11)*. On average, the participants were using digital maps more than once per week and except for 5 participants, they had never used an OHMD before.

4.3 Setup

The room lighting was slightly dimmed to allow for a better contrast between the Google Glass display and the white wall participants were facing during task execution. A monitor was placed to the participants' right-hand side on which participants read questionnaires and instructions before and between the tasks.

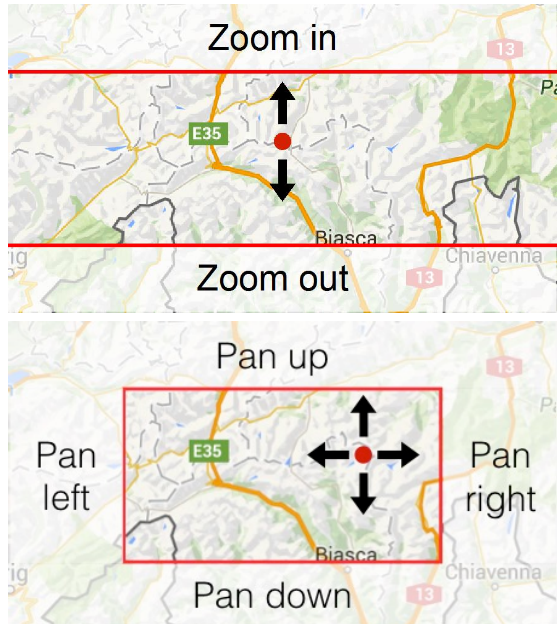
The hardware setup consisted of a Google Glass[4] (OS version: XE22) paired with a smartphone (Google Galaxy Nexus 2 (I9250), OS version: 4.2.1). The paired smartphone allowed us to observe participants remotely throughout the study and supplied the Google Glass with internet access.

The software setup included the WMS client, which loaded and displayed (study specific) maps and allowed participants to interact with them (see section 4.4 below). The application had a logging mechanism that stored each zoom or pan interaction made by the users with its time of occurrence.

The questionnaires and instructions were created using the Collector Application[5].

4.4 Implementation of Interaction Methods

Evaluating possible implementations for head-based gestures revealed that users had difficulties controlling their input. Users found it particularly difficult to identify the moment a head-based gesture results in a reaction from the application. To overcome these limitations, we introduced a novel approach for implementing head-based gestures. That is, for $\mathbf{Z}^h_{\texttt{nod}}$ the user was provided a dot representing her current head orientation, as well as two horizontal bars at the top and bottom area of the display (see Figure 2 (a)). Thus, whenever the user moved her head up or down (i.e., nods) the dot moved accordingly. As long as the dot was within the bars, the map did not move, i.e., zoom (or pan). However, as soon as the dot surpassed either one, the app started zooming. A single zoom was realized as (de)magnifying the current pane by a factor of two, every 400ms as determined in the pilot study. Moreover, if users moved the red dot to the far edge of the screen a zoom-in or -out happened every 200ms.



(a)

(b)

Figure 2: The user's current head orientation (red dot) in relation to the amount of rotation required for an interaction (red lines) was displayed for interaction methods (a) $\mathbf{Z}^h_{\texttt{nod}}$ and (b) $\mathbf{P}^h_{\texttt{nod, shake}}$. The arrows and text were not visible on the screen.[7]

$\mathbf{P}^h_{\texttt{nod, shake}}$ provided the user a dot and a rectangle (see Figure 2 (b)). As for $\mathbf{Z}^h_{\texttt{nod}}$, the dot moved according to the users nodding (up or down) or shaking (left or right) and panning began when the dot surpassed the borders of the rectangle. A pan interaction implied moving the pane by 1/8 of its overall width in the given head direction every 400ms. When the dot was moved to the edge of the pane, the panning took place every 200ms, too.

Touch interactions were identified by the number of units surpassed during a continuous swipe. A zoom occurred, when swipes surpassed 150 units (of 1366) for $\mathbf{Z}^t_{\texttt{forth/back}}$ and 30 units (of 187) for $\mathbf{Z}^t_{\texttt{up/down}}$. As for $\mathbf{P}^t_{\texttt{up/down, forth/back}}$, swipes in both directions exceeding 30 units were considered a pan. For $\mathbf{Z}^t_{\texttt{press}}$, a zoom occurred every 400ms.

4.5 Procedure

4.5.1 Pilot Study

A pilot study with three participants was conducted in order to ensure that the study procedure would be flawless, as well as to receive feedback on the interaction methods. As a result, we dismissed tapping as input control, due to displacement of the glasses. Furthermore, we chose to use

[4]The touch pad of Google Glass has a width of 8cm and a height of 1.1cm.

[5]http://www.survalyzer.ch

[7]Map data ©2016 Google

a single finger for swiping inputs, because there were no measurable differences between the number of fingers. After looking into the different head inputs, we chose the given head rotations to be most suitable as control inputs, due to the lower fatigue compared to actual movements and based on their technical feasibility. Finally, we chose the number of fingers for $\mathbf{Z}^t_{\texttt{press}}$ giving the interaction the highest stability and the step sizes (times and units) for zooming and panning, which the pilot study participants perceived to be the most convenient.

4.5.2 Main Study

While participants were sitting in the office, they received a brief introduction to the overall experiment and were asked to fill out a questionnaire assessing demographics and experience information (regarding OHMDs and digital maps). Google Glass and its mechanisms were explained and participants had time to adjust the prism to ensure that the edges of the screen were visible.

Each task was introduced in general and with an example instruction. A task consisted of several instructions. Participants were asked to execute instructions as fast as possible. Before solving a task with an interaction method, the method was explained, and a 30 second training phase was granted. A task was successful when participants completed all instructions, i.e., reaching a given location and zoom level. A task was aborted, if the time surpassed 180 seconds (for Task 1 and 2) or 240 seconds (for Task 3).

Each instruction requires users to repeat an interaction a different number of times. The frequency was chosen as to avoid fatigue, but to provide meaningful data. Furthermore, to avoid learning effects, the order of the instructions was changed between interaction methods. Finally, we chose the order of the tasks based on the number of axes involved.

After completing a task with an interaction method, the participant was asked to fill out the UEQ questionnaire [16] and the "Raw" Nasa TLX [9].

4.5.3 Tasks

Task 1 - *Zooming.* A total of 4 instructions had to be executed, performing a sequence of alternating zoom in and -out interactions starting and ending at zoom level 1. A typical instruction was of the form *"Please zoom in, until you see the label of the countries, for example, Switzerland"*.

Task 2 - *Panning.* In a total of 7 instructions, participants had to pan to a specific target location denoted by a white letter in a yellow pentagon. To know whether the correct map extent was reached, a small yellow rectangle was included. The target location was reached, when the pentagon was inside the rectangle (see Figure 3). A typical instruction was of the form *"Pan left, until location A is in the yellow box"*.

Task 3 - *Panning and Zooming.* Participants were asked to navigate to a specific target location by zooming and panning. This task included a total of 10 instructions, consisting of 3 zoom and 7 pan instructions. The procedure for those instructions was the same as for task 1 and 2, but included the yellow box for both panning and zooming. As an example, an instruction sequence of panning followed by zooming, was of the form *"Pan left, until the country label "CH" is in the yellow box"* followed by *"Zoom in, until location A is in the yellow box"*.

Figure 3: Map shown for the navigation tasks on the OHMD. In this example, the task is finished since the target position (A) is inside the yellow rectangle.

5. RESULTS

In this section we present the experimental results for the zooming, panning and their combination (see independent variables, section 4.1) with regard to efficiency, effectiveness, user experience (UX) and perceived cognitive workload (TLX).

5.1 Zooming

To evaluate the zooming task with four test conditions we applied the Friedman test to identify any statistical significance and the Wilcoxon signed-rank test for a pairwise comparison of the test conditions along with the Bonferroni correction [2], to weigh out the multiple comparison problem.

5.1.1 Effectiveness and Efficiency

Effectiveness. There was a significant difference over all interaction methods ($\chi^2 = 16.875$, $p < .05$). However, there were no significant differences between $\mathbf{Z}^h_{\texttt{nod}}$ and $\mathbf{Z}^t_{\texttt{up/down}}$, $\mathbf{Z}^h_{\texttt{nod}}$ and $\mathbf{Z}^t_{\texttt{press}}$, or $\mathbf{Z}^t_{\texttt{up/down}}$ and $\mathbf{Z}^t_{\texttt{press}}$. All other methods showed pairwise significant differences (see Table 3). The most effective zoom interaction method in terms of average number of errors was $\mathbf{Z}^t_{\texttt{forth/back}}$ with a mean of *1.13* errors *(SD = 1.76)*. The least effective one was $\mathbf{Z}^t_{\texttt{press}}$ with a mean of *4.65* errors *(SD = 5.096)* (see Figure 4).

Efficiency. There was a statistically significant difference over all interaction methods ($\chi^2 = 33.387$, $p < .05$). However, there were no significant differences between $\mathbf{Z}^t_{\texttt{forth/back}}$ and $\mathbf{Z}^t_{\texttt{up/down}}$ or between $\mathbf{Z}^h_{\texttt{nod}}$ and $\mathbf{Z}^t_{\texttt{press}}$ (after a Bonferroni correction). All other interaction methods showed pairwise significant differences (see Table 3). The most efficient zoom interaction in terms of average completion time was $\mathbf{Z}^t_{\texttt{up/down}}$ with a mean of *17.1* seconds *(SD = 6.9)*. The least efficient one was $\mathbf{Z}^t_{\texttt{press}}$ with a mean of *33.9* seconds *(SD = 18.2)*.

5.1.2 User Experience and Workload

UX. We will abbreviate the UX scales as follows: Attractiveness = **AT**, Perspicuity = **PE**, Efficiency = **EF**, Dependability = **DP**, Stimulation = **ST** and Novelty = **NV**. Furthermore, for the UX only the results for the UX-AT scale are described in detail, since it represents the overall impression of the method.

There was a significant difference over all interaction methods for all UX scales ($\chi^2 > 15$, $p < .001$). For the

Table 3: Inferential statistics on effectiveness (top) and efficiency (bottom) of the zoom methods. Z values in bold, when $p < .012$ (Bonferroni adjusted), i.e., indicating significant differences.

Method	$Z^h_{\texttt{nod}}$	$Z^t_{\texttt{forth/back}}$	$Z^t_{\texttt{up/down}}$
$Z^t_{\texttt{forth/back}}$	**-2.656**	-	-
$Z^t_{\texttt{up/down}}$	-0.186	**-3.225**	-
$Z^t_{\texttt{press}}$	-1.277	**-3.396**	-1.213
Method	$Z^h_{\texttt{nod}}$	$Z^t_{\texttt{forth/back}}$	$Z^t_{\texttt{up/down}}$
$Z^t_{\texttt{forth/back}}$	**-3.763**	-	-
$Z^t_{\texttt{up/down}}$	**-3.116**	-0.607	-
$Z^t_{\texttt{press}}$	-2.391	**-4.017**	**-4.507**

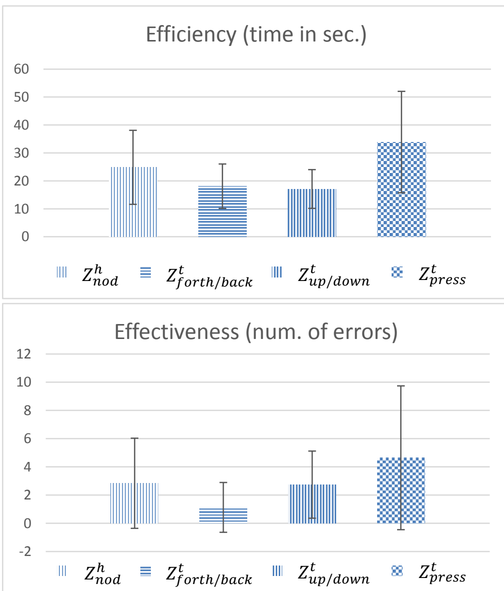


Figure 4: The efficiency (top) and effectiveness (bottom) of the zoom methods, measured as the average task completion time in seconds (SD as error bars).

UX-AT scale, there were no significant differences between $\mathbf{Z}^t_{\texttt{forth/back}}$ and $\mathbf{Z}^t_{\texttt{up/down}}$, as well as between $\mathbf{Z}^h_{\texttt{nod}}$ and $\mathbf{Z}^t_{\texttt{press}}$. All other methods showed a pairwise significant difference with regard to UX-AT (see Table 4). The highest overall UX-AT value was that of $\mathbf{Z}^t_{\texttt{forth/back}}$ with a value of *1.71 (SD = .75)*, while $\mathbf{Z}^h_{\texttt{nod}}$ had the lowest value with *.57 (SD = 1.24)* (see Figure 5).

TLX. We will abbreviate the TLX components as follows: Mental Demand = **MD**, Physical Demand = **PD**, Temporal Demand = **TD**, Performance = **OP**, Effort = **EF**, Frustration = **FR** and **OV** for the overall workload (sum of components). Furthermore, for the TLX only the results for the TLX-OV component are described in detail, since it represents the overall perceived workload.

There was a significant difference over all interaction methods for all components *($\chi^2 > 16$, $p < .001$)* except for TLX-TD *($\chi^2 = 5.581$, $p < .134$)* (see Table 5). However, for the TLX-OV there were no significant differences between $\mathbf{Z}^t_{\texttt{forth/back}}$ and $\mathbf{Z}^t_{\texttt{up/down}}$, or $\mathbf{Z}^h_{\texttt{nod}}$ and $\mathbf{Z}^t_{\texttt{press}}$. All other methods showed a pairwise significant difference with regard to TLX-OV. The highest TLX-OV value was that of $\mathbf{Z}^h_{\texttt{nod}}$ with a mean of 36 *(SD = 18.13)*, while $\mathbf{Z}^t_{\texttt{forth/back}}$ had the lowest value with a mean of 20.87 *(SD = 13.40)* (see Figure 6).

Table 4: The inferential statistics on the UX scales for the zoom methods. The Z values are given in bold, when $p < .012$ (Bonferroni adjusted), i.e., there was a significant difference.

Method	$Z^h_{\texttt{nod}}$		$Z^t_{\texttt{forth/back}}$		$Z^t_{\texttt{up/down}}$	
$Z^t_{\texttt{forth/back}}$	**-3.955**^{AT}	**-3.334**^{PE}	–	–	–	–
	-4.271^{EF}	**-4.159**^{DP}	–	–	–	–
	-2.054^{ST}	**-3.695**^{NV}	–	–	–	–
$Z^t_{\texttt{up/down}}$	**-3.150**^{AT}	-2.102^{PE}	-2.049^{AT}	-0.986^{PE}	–	–
	-3.451^{EF}	-1.830^{DP}	-1.250^{EF}	**-3.183**^{DP}	–	–
	-2.471^{ST}	**-3.667**^{NV}	-1.171^{ST}	-0.592^{NV}	–	–
$Z^t_{\texttt{press}}$	-0.125^{AT}	-0.626^{PE}	**-3.982**^{AT}	**-3.736**^{PE}	**-3.121**^{AT}	**-2.657**^{PE}
	-0.858^{EF}	-2.009^{DP}	**-3.791**^{EF}	**-4.630**^{DP}	**-3.631**^{EF}	**-3.355**^{DP}
	-2.398^{ST}	**-3.183**^{NV}	-1.760^{ST}	**-3.039**^{NV}	-1.498^{ST}	**-2.782**^{NV}

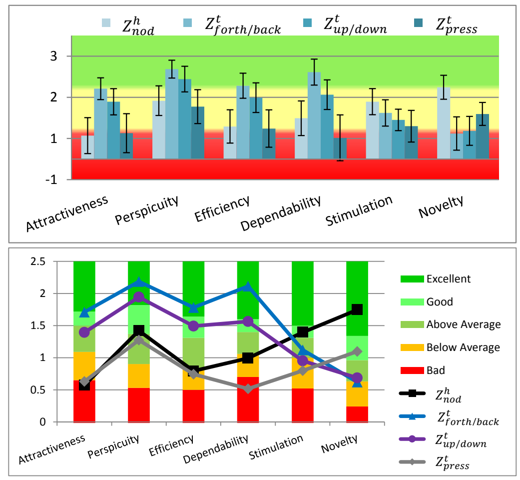


Figure 5: The UX scale (top) and benchmark (bottom) of the zoom methods.

Table 5: The inferential statistics on the TLX (top) and TLX-OV (bottom) for the zoom methods. The Z values are given in bold, when $p < .012$ (Bonferroni adjusted), i.e., there was a significant difference.

Method	$Z^h_{\texttt{nod}}$		$Z^t_{\texttt{forth/back}}$		$Z^t_{\texttt{up/down}}$	
$Z^t_{\texttt{forth/back}}$	**-3.451**^{MD}	**-4.008**^{PD}	–	–	–	–
	-0.792^{TD}	**-2.920**^{OP}	–	–	–	–
	-2.919^{EF}	**-3.312**^{FR}	–	–	–	–
$Z^t_{\texttt{up/down}}$	-2.483^{MD}	**-4.000**^{PD}	-1.654^{MD}	-0.525^{PD}	–	–
	-2.200^{TD}	-2.103^{OP}	-1.913^{TD}	-0.771^{OP}	–	–
	-2.607^{EF}	**-2.997**^{FR}	-0.824^{EF}	-0.343^{FR}	–	–
$Z^t_{\texttt{press}}$	-0.542^{MD}	**-3.384**^{PD}	**-3.055**^{MD}	-2.029^{PD}	-2.147^{MD}	-2.277^{PD}
	-0.635^{TD}	-1.152^{OP}	-0.128^{TD}	**-3.515**^{OP}	-2.284^{TD}	**-3.604**^{OP}
	-0.308^{EF}	-0.653^{FR}	**-3.193**^{EF}	**-3.143**^{FR}	**-2.779**^{EF}	**-3.144**^{FR}

Method	$Z^h_{\texttt{nod}}$	$Z^t_{\texttt{forth/back}}$	$Z^t_{\texttt{up/down}}$
$Z^t_{\texttt{forth/back}}$	**-4.087**	-	-
$Z^t_{\texttt{up/down}}$	**-3.940**	-0.154	-
$Z^t_{\texttt{press}}$	-0.588	**-3.685**	**-3.632**

5.2 Panning

For panning we applied the Wilcoxon signed-rank test, since we had only two test conditions.

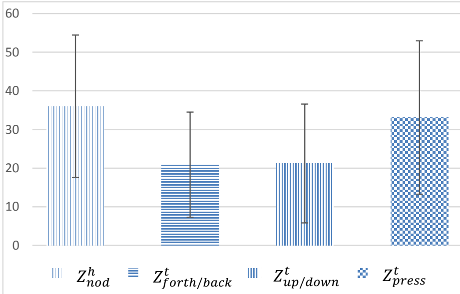


Figure 6: The results of the TLX-OV for the zoom methods (SD as error bars).

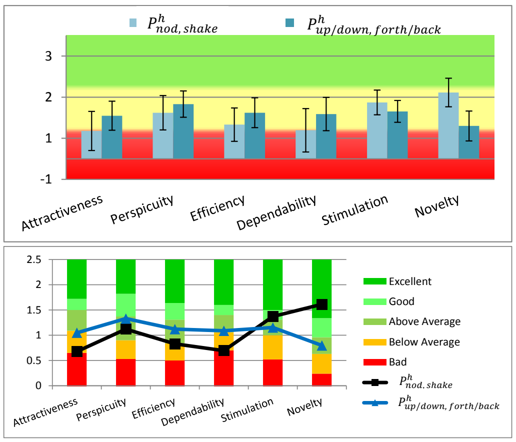


Figure 7: The UX scales (top) and benchmark (bottom) of the pan methods.

5.2.1 *Effectiveness and Efficiency*

Effectiveness. There was a significant difference between the interaction methods *($Z = -2.582$, $p < .05$)*. The more effective pan interaction method was $\mathbf{P}^h_{\texttt{nod, shake}}$ with a mean of *12.6* errors *(SD = 18.37)*, while $\mathbf{P}^t_{\texttt{up/down, forth/back}}$ resulted in a mean of *25* errors *(SD = 20.69)*.

Efficiency. There was no significant difference between the interaction methods *($Z = -1.920$, $p = .055$)*. $\mathbf{P}^t_{\texttt{up/down, forth/back}}$ had a mean completion time of *56* seconds *(SD = 28.2)*, while $\mathbf{P}^h_{\texttt{nod, shake}}$ resulted in a mean of *72.3* seconds *(SD = 39.2)*.

5.2.2 *User Experience and Workload*

UX. There were no significant differences over the two interaction methods for all UX scales except for UX-NV *($Z = -3.240$, $p < .05$)*. The higher UX-AT scale value was that of the interaction method $\mathbf{P}^t_{\texttt{up/down, forth/back}}$ with a value of *1.05 (SD = 1.00)*, while $\mathbf{P}^t_{\texttt{up/down, forth/back}}$ had the lower value with *0.68 (SD = 1.35)* (see Figure 7).

TLX. There were no significant differences over the two interaction methods for all TLX components except for TLX-PD *($Z = -2.078$, $p < .05$)*. The higher TLX-OV component value was that of the interaction method $\mathbf{P}^h_{\texttt{nod, shake}}$ with a mean of 39.48 *(SD = 22.55)*, while $\mathbf{P}^t_{\texttt{up/down, forth/back}}$ had the lower value with a mean of 35.10 *(SD = 18.95)*.

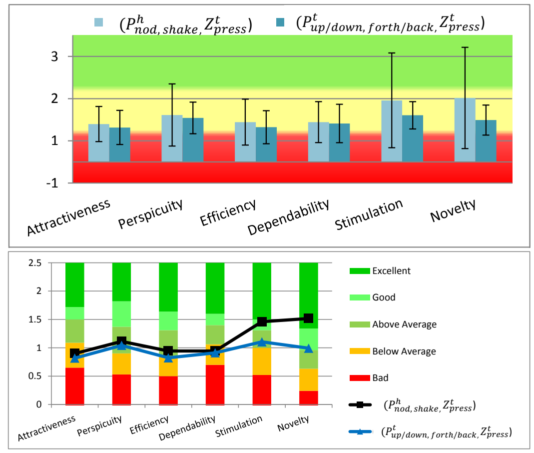


Figure 8: The UX scales (top) and benchmark (bottom) of the pan and zoom methods.

5.3 Panning and Zooming

For the combined tasks of panning and zooming, we applied the Wilcoxon signed-rank test, because there were only two test conditions.

5.3.1 *Effectiveness and Efficiency*

Effectiveness. There was a significant difference between the interaction methods *($Z = -3.832$, $p < .05$)*. ($\mathbf{P}^h_{\texttt{nod, shake}}$, $\mathbf{Z}^t_{\texttt{press}}$) had a mean of *12.50* errors *(SD = 13.74)*, while ($\mathbf{P}^t_{\texttt{up/down, forth/back}}$, $\mathbf{Z}^t_{\texttt{press}}$) had a mean of *26* errors *(SD = 28.50)*.

Efficiency. There was no significant difference between the interaction methods *($Z = -.921$, $p = .357$)*. ($\mathbf{P}^t_{\texttt{up/down, forth/back}}$, $\mathbf{Z}^t_{\texttt{press}}$) had a mean completion time of *94.18* seconds *(SD = 42.48)*, while ($\mathbf{P}^h_{\texttt{nod, shake}}$, $\mathbf{Z}^t_{\texttt{press}}$) had a mean of *104.50* seconds *(SD = 52.77)*.

5.3.2 *User Experience and Workload*

UX. There was a significant difference over the two interaction methods for the UX scales UX-ST *($Z = -2.225$, $p < .05$)* and UX-NV *($Z = -3.023$, $p < .05$)*. The higher UX-AT scale was that of the interaction method ($\mathbf{P}^h_{\texttt{nod, shake}}$, $\mathbf{Z}^t_{\texttt{press}}$) with a value of *0.90 (SD = 1.37)*, while ($\mathbf{P}^t_{\texttt{up/down, forth/back}}$, $\mathbf{Z}^t_{\texttt{press}}$) had the lower value with *0.82 (SD = 1.15)* (see Figure 8).

TLX. There was no significant difference over the two interaction methods for all TLX components. The higher TLX-OV component value was that of the interaction method ($\mathbf{P}^t_{\texttt{up/down, forth/back}}$, $\mathbf{Z}^t_{\texttt{press}}$) with a mean of 37.74 *(SD = 18.38)*, while ($\mathbf{P}^h_{\texttt{nod, shake}}$, $\mathbf{Z}^t_{\texttt{press}}$) had the lower score with a mean of 37.13 *(SD = 19.97)*.

6. DISCUSSION

Hypotheses **H1** and **H2** were confirmed, while **H3** was partially confirmed. First, the results confirmed that $\mathbf{P}^h_{\texttt{nod, shake}}$ was significantly more effective than $\mathbf{P}^t_{\texttt{up/down, forth/back}}$, while there were no significant differences concerning efficiency, user experience and perceived cognitive workload (**H1**). Second, $\mathbf{Z}^t_{\texttt{forth/back}}$ was the significantly most effective zoom interaction method. Furthermore, except for when being compared to $\mathbf{Z}^t_{\texttt{up/down}}$, it was the significantly most efficient interaction, had the significantly highest user experience and lowest perceived cognitive workload. Finally, ($\mathbf{P}^h_{\texttt{nod, shake}}$, $\mathbf{Z}^t_{\texttt{press}}$) was significantly more effective, but there was no significant difference compared to ($\mathbf{P}^t_{\texttt{up/down, forth/back}}$, $\mathbf{Z}^t_{\texttt{press}}$) otherwise.

Thus, the results reflected the assumption that interaction methods with congruent spatial mappings are able to outperform others independent of the interaction modality. That is, no matter whether head-based or haptic interactions are used.

The results also indicate that familiarity with the input method correlates with performance. That is, the more common haptic input methods performed significantly better than the less common head-based gestures. This can be observed when comparing $\mathbf{Z}^t_{\texttt{up/down}}$ with $\mathbf{Z}^h_{\texttt{nod}}$, where $\mathbf{Z}^t_{\texttt{up/down}}$ performed significantly better for all DVs and both can be considered as input modalities along the y-axis.

We observed high SD values in our data, which can be explained by extreme outliers. These were not excluded from evaluation, in order to truly reflect difficulties with the interaction methods. In particular, we identified two possible causes for the deviations. First, some participants appear to have had particular difficulties with a specific zoom method resulting in bad efficiency and effectiveness values. For example, one participant had a completion time of 81.56 seconds and 15 errors for $\mathbf{Z}^h_{\texttt{nod}}$. Considering the average completion time of 24.84 seconds and average number of errors of 2.84, this demonstrates the extremity of these outliers. We observed similar effects for $\mathbf{Z}^t_{\texttt{press}}$. The second reason for these outliers is our abort-condition, i.e., we stopped the experiment after 180 seconds, e.g., when participants claimed or we realized they are not able to complete the experiment, because they "got lost" on the map or simply could not cope with the interaction method. For example, we had 3 such observations for the efficiency measurements of $\mathbf{P}^h_{\texttt{nod, shake}}$ and 1 for $\mathbf{P}^t_{\texttt{up/down, forth/back}}$.

Finally, such an extensive comparison has many possible sources for confounding variables. However, we narrowed those down by analyzing the design space, conducting a pilot study and choosing the same parameters for the interaction methods whenever possible. It should be noted, that the aim of this paper was not to find the optimal settings for either one of the inputs, but to explore the potential of the novel head-based approaches.

Limitations As explained in section 3.1, we did not implement and evaluate all possible interaction methods that were identified in the design space, amongst others, because they contradicted with the assumption of congruent mappings. Furthermore, the explored design space was limited to basic gestures, such as swipes and refrained from evaluating gesture patterns on the touch pad.

Furthermore, we only tested the setup with one OHMD, i.e. Google Glass, but we investigated the specifications of different OHMDs and found that our input modalities were provided by the majority of them. Even though some do not provide a touch pad for interaction input, they all allowed for inputs based on head-based gestures.

We decided to have a 30 second introductory session with each new interaction method and further chose to have 5 seconds between each of the instructions. These intervals could have been chosen differently, but their usefulness was determined during the pilot study. We did not receive any feedback indicating the need for additional time.

For zooming and panning, we predefined the movement direction (e.g. "zoom in") and required the participants to reflect on their actions, by asking them to perform an action until they see a specific label (or it is centered within the yellow rectangle). This approach was chosen for several reasons. First, restricting the direction and destination allowed for comparability. By letting all participants, for example, zoom in until they see the label "Paris", we ensure that they perform the same actions. Additionally, we know how many actions they should perform to reach the label. Thus, the results for task completion time and error rate can be compared. Second, asking participants to look for a particular label ensures a more realistic setup, where a user is actually looking for a particular piece of information.

The study was conducted under lab conditions and the results for user experience and perceived cognitive workload might have been different if the experiment were conducted in an outdoor environment, as discussed in [26]. We refrained from letting participants stand, because the combination with the arm movements required for haptic inputs might have led to earlier physical exhaustion and lower user experience ratings. However, efficiency and effectiveness results would not have been influenced, because, for the given interaction methods (using a touch pad and head movements), they are less dependent on the surrounding environment.

The choice of the step sizes for the zoom interactions is a potential source for improvement, especially for $\mathbf{Z}^h_{\texttt{nod}}$ and $\mathbf{Z}^t_{\texttt{press}}$. Those values were chosen based on the pilot study, but alternative solutions are thinkable, such as an adjustable zoom speed.

Finally, choosing the sequence of the tasks based on the number of axes involved, might result in order effects, since some of the interaction methods for panning were already used for zooming.

7. CONCLUSIONS AND FUTURE WORK

This paper explored the design space for map interaction methods on OHMDs. After a thorough analysis, novel head-based and common haptic interaction methods were selected and compared with regard to panning, zooming, as well as the combination of panning and zooming on maps.

In a study with 31 participants, the selected interaction methods were evaluated concerning effectiveness, efficiency, user experience and perceived cognitive workload. We were able to demonstrate that congruent mappings outperform others. More precisely, for panning interactions, head-based gestures were competitive, since they were able to congruently map control inputs onto the display. As for the zooming interactions, even though our approach was outperformed by the more common haptic interactions, it still showed potential for improvement.

In consequence, the head-based interaction methods introduced in this paper are suitable in scenarios where users want to navigate on an OHMD, but cannot use their hands or other bodily extremities except the head, e.g., due to disabilities or because their current activity does not allow for it. An example could be a plumber looking at a plan of the pipes in a house while fixing one. Many more examples can be thought of, e.g., the use by airplane engineers, car mechanics, etc.

Based on this study's results, evaluating the combinations of $\mathbf{P}^h_{\texttt{nod, shake}}$ with $\mathbf{Z}^t_{\texttt{forth/back}}$ and $\mathbf{Z}^t_{\texttt{up/down}}$, as well as

$\mathbf{P}^t_{\texttt{up/down, forth/back}}$ with $\mathbf{Z}^h_{\texttt{nod}}$, would give us further valuable insights on the impact of congruent spatial mappings on the performance for interacting with OHMDs. Furthermore, to fully evaluate the potential of our novel head-based interaction methods, it is necessary to conduct a longitudinal study.

Finally, the possibility to integrate Eye Tracking capabilities [13] into OHMDs, holds the potential for many new interaction methods. For example, extending the mixed interaction methods with the approach introduced by [7], i.e. using markers to emphasize regions on a display that were previously consulted for a given task, might allow for enhanced user orientation and more coherent user experience.

8. REFERENCES

[1] B. B. Bederson, J. D. Hollan, K. Perlin, J. Meyer, D. Bacon, and G. Furnas. Pad++: A zoomable graphical sketchpad for exploring alternate interface physics. *Journal of Visual Languages & Computing*, 7(1):3–32, March 1996.

[2] J. M. Bland and D. G. Altman. Multiple significance tests: the bonferroni method. *Bmj*, 310(6973):170, 1995.

[3] O. Cakmakci and J. Rolland. Head-worn displays: A review. *J. Display Technol.*, 2(3):199–216, 2006.

[4] J. Chen and R. W. Proctor. Up or down: Directional stimulus-response compatibility and natural scrolling. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, pages 1381–1385. SAGE Publications, 2012.

[5] S. Christian, J. Alves, A. Ferreira, D. Jesus, R. Freitas, and N. Vieira. Volcano salvation: Interaction through gesture and head tracking. In *CHI*, pages 297–300. ACM, 2014.

[6] D. Dobbelstein, P. Hock, and E. Rukzio. Belt: An unobtrusive touch input device for head-worn displays. In *CHI*, pages 2135–2138. ACM, 2015.

[7] I. Giannopoulos, P. Kiefer, and M. Raubal. Geogazemarks: Providing gaze history for the orientation on small display maps. In *ICMI*, pages 165–172. ACM, 2012.

[8] M. Harrower and B. Sheesley. Designing better map interfaces: A framework for panning and zooming. *Transactions in GIS*, 9(2):77–89, 2005.

[9] S. G. Hart and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human Mental Workload*, pages 139–183. North-Holland, 1988.

[10] S. C. Hirtle and M. Raubal. Many to many mobile maps. In *Cognitive and Linguistic Aspects of Geographic Space*, pages 141–157. Springer, 2013.

[11] H. Hua, X. Hu, and C. Gao. A high-resolution optical see-through head-mounted display with eyetracking capability. *Opt. Express*, 21(25):30993–30998, 2013.

[12] S. Ishimaru, K. Kunze, K. Kise, J. Weppner, A. Dengel, P. Lukowicz, and A. Bulling. In the blink of an eye: Combining head motion and eye blink frequency for activity recognition with google glass. In *AH*, pages 15:1–15:4. ACM, 2014.

[13] S. Jalaliniya, D. Mardanbegi, I. Sintos, and D. G. Garcia. Eyedroid: An open source mobile gaze tracker on android for eyewear computers. In *PETMEI*, pages 873–879. ACM, 2015.

[14] J. A. Johnson. A comparison of user interfaces for panning on a touch-controlled display. In *CHI*, pages 218–225. ACM, 1995.

[15] B. Kollee, S. Kratz, and A. Dunnigan. Exploring gestural interaction in smart spaces using head mounted devices with ego-centric sensing. In *SUI*, pages 40–49, 2014.

[16] B. Laugwitz, T. Held, and M. Schrepp. Construction and evaluation of a user experience questionnaire. In *HCI and Usability for Education and Work*, pages 63–76. Springer Berlin Heidelberg, 2008.

[17] Z. Lv, A. Halawani, S. Feng, S. Ur Réhman, and H. Li. Touch-less interactive augmented reality game on vision-based wearable device. *Personal Ubiquitous Comput.*, 19(3-4):551–567, 2015.

[18] I. S. MacKenzie. *Human-computer interaction: An empirical research perspective*. Newnes, 2012.

[19] M. Malu and L. Findlater. Personalized, wearable control of a head-mounted display for users with upper body motor impairments. In *CHI*, pages 221–230. ACM, 2015.

[20] R. McNaney, J. Vines, D. Roggen, M. Balaam, P. Zhang, I. Poliakov, and P. Olivier. Exploring the acceptability of google glass as an everyday assistive device for people with parkinson's. In *CHI*, pages 2551–2554. ACM, 2014.

[21] M. F. Roig-Maimó, J. Varona Gómez, and C. Manresa-Yee. Face me! head-tracker interface evaluation on mobile devices. In *CHI*, pages 1573–1578. ACM, 2015.

[22] M. Serrano, B. M. Ens, and P. P. Irani. Exploring the use of hand-to-face input for interacting with head-worn displays. In *CHI*, pages 3181–3190. ACM, 2014.

[23] T. A. Slocum, R. B. McMaster, F. C. Kessler, and H. H. Howard. *Thematic cartography and geovisualization*. Pearson Prentice Hall Upper Saddle River, NJ, 2009.

[24] M. Spindler and R. Dachselt. Paperlens: Advanced magic lens interaction above the tabletop. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, pages 7:1–7:1. ACM, 2009.

[25] T. Starner. Project glass: An extension of the self. *Pervasive Computing, IEEE*, 12(2):14–16, 2013.

[26] Y.-C. Tung, C.-Y. Hsu, H.-Y. Wang, S. Chyou, J.-W. Lin, P.-J. Wu, A. Valstar, and M. Y. Chen. User-defined game input for smart glasses in public space. In *CHI*, pages 3327–3336. ACM, 2015.

[27] B. Zhang, Y.-H. Chen, C. Tuna, A. Dave, Y. Li, E. Lee, and B. Hartmann. Hobs: Head orientation-based selection in physical spaces. In *SUI*, pages 17–25. ACM, 2014.

[28] X. S. Zheng, C. Foucault, P. Matos da Silva, S. Dasari, T. Yang, and S. Goose. Eye-wearable technology for machine maintenance: Effects of display position and hands-free operation. In *CHI*, pages 2125–2134. ACM, 2015.