Interacting with large maps using HMDs in VR settings

Ioannis Giannopoulos
University of Patras
Rio, 26504, Greece
igiannopou@ceid.upatras.gr

Andreas Komninos
University of Patras
Rio, 26504, Greece
akomninos@ceid.upatras.gr

John Garofalakis
University of Patras
Rio, 26504, Greece
garofala@ceid.upatras.gr

MobileHCI '17, September 04-07, 2017, Vienna, Austria

ACM ISBN 978-1-4503-5075-4/17/09.
https://doi.org/10.1145/3098279.3122148.

Abstract

Location based services are a common application scenario in mobile and ubiquitous computing use. A typical issue with cartographic applications in this domain is the limited size of the displayed map, which makes interaction and visualization a difficult problem to solve. With the increasing popularity of head mounted displays for VR and AR systems, an opportunity is presented for map-based applications to overcome the limitation of the small display size, as the user's information visualization space can extend to his entire surroundings. In this paper we present a preliminary investigation into how interaction with such very large display maps can take place, using a virtual reality headset as the sole input and interaction method.

Author Keywords

Augmented Reality; Virtual Reality; Map-based applications; Digital Maps; Interaction

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

Introduction

Digital maps used on desktop computers in geographical information systems or web-based

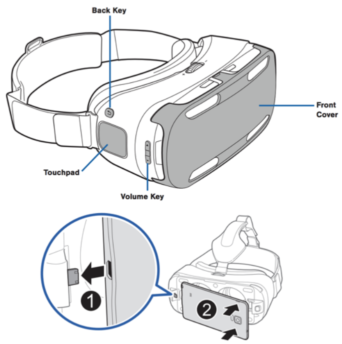

Figure 1: The Samsung Gear VR accessory, showing the placement of the input controls and smartphone.

Figure 2: The fade-in of the panning controls in the delay with feedback interface (V3)

location services have been studied for their usability for considerable time [1, 3, 13]. In addition to interaction with traditional methods (e.g. keyboard and mouse), non-standard interaction (e.g. gestural, and multimodal input) has also been the focus of study, in an attempt to improve the usability and interactivity of maps [4, 12]. Mobile maps require careful design to ensure the usability, mostly because of the limited screen area available and the richness of the information that needs to be displayed [7]. Some of the common problems during interaction with mobile maps arise from the limited display area, which results in the need for frequent interaction and manipulation of the displayed information. Mobile maps need to be frequently zoomed and panned to gain a thorough spatial understanding [8, 11]. Another frequent problem is the display and selection of markers denoting points of interest (POIs), which are often difficult to select by tapping due to their small size, or due to overlap when POIs are clustered together [5]. The small display size also adds the problem of how to visualize off-screen POIs and help the user understand the spatial relationships between on and off-viewport information [2].

By pairing mobile displays with sensors like GPS, accelerometers, magnetometers and inertial gyroscopes, an application can be made aware of the user's position and direction of view. Thus it can provide the user with information that is spatially relevant to their field of vision. This concept of augmenting reality is typically implemented with the superimposition of digital spatial information on camera video feeds, on the screens of mobile devices. These devices could thus be used as a “magic lens”, through which the user can experience augmented views of the real world surrounding them [9]. Of course, this method of interaction requires the user to hold their devices at eye-level, which may be tiring during prolonged use, and also does not overcome the problem of the limited size of the information space presented to the user, as the digital information is only visible through the narrow keyhole of the mobile device screen. Augmented reality environments have been found to be tied with regard to usability criteria compared to standard mobile maps and still have several issues to be resolved [6]. In recent years, interest and advances in wearable display technology including augmented reality and virtual reality headsets, have resulted in freeing up the user from the constraints of the mobile screen, by extending the available information space to the entire world around the user. The use of head mounted displays (HMDs) increases the space upon which digital information can be displayed to the entire field of view of the user, thus minimizing the need for frequent interaction (e.g., panning) and also results in more natural interaction that is consistent with everyday experiences (i.e. users simply have to turn their heads towards the location where the information they need might be found).

As such, the use of AR or VR headsets to explore large information spaces, like maps, can offer a more natural and comfortable way to afford users a better spatial understanding of information. However, the use of such headsets requires a different approach to designing interaction and input, compared to the touch-based mobile screens, or keyboard and mouse used when exploring maps on large screen desktop computers. The only relevant paper covering use of maps with an HMD is by [10], however in this paper the authors evaluate a limited set of gestural design for map

Figure 3: The progress bar on zoom controls & POIs (hover and then select) in the delay with feedback interface (V3)

control, focusing solely on zooming and panning. Our paper therefore focuses on addressing the issue of controlling very large maps displayed via headset technologies, by exploring how different gestural or limited hand-input modalities can be mapped onto actions related to the control of digital maps, as well as the information objects presented in these.

Designing interactions with maps on head mounted displays

To explore the interaction with maps in a HMD use context, we proceeded with the design of an application using the Samsung Gear VR headset. This is a virtual reality accessory in which a smartphone is inserted at the front of the accessory, effectively making the device screen act as the display. By splitting the screen in two segments, on which images are rendered at a slightly different angle and viewed over a separate lens, the system successfully emulates stereoscopic vision that affords 3D vision effects like depth perception. The smartphone's embedded sensors are used as input to the applications running on the smartphone, allowing the device to accurately know where the user is looking and therefore adapt the displayed views to provide an immersive virtual reality experience. The system also allows for further input options via a dedicated touchpad surface on the right side of the HMD, as well as a programmable "back" button and volume buttons near the touchpad. We opted for this system because of its very low cost and thus likely availability to most users.

For the purposes of our experiment we developed an application for Android using the Unity IDE, which renders a very large map area in front of the user. The map area is dynamically populated from the OpenStreetMaps API. The application can also display points of interest as markers on the map. The map extends beyond the user's lateral field of view, as shown in Figure 4.

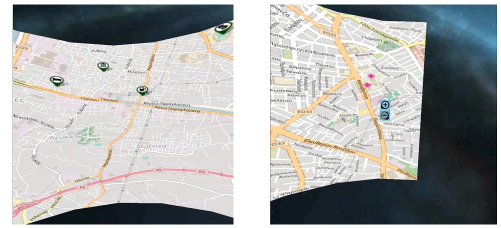

Figure 4: The map application demonstrating the size of the map, which extends beyond the user's lateral FOV.

We then proceeded to design the interaction with the map itself, by considering the core functions that a digital map should offer and how they might be mapped to the input modalities afforded by the Gear VR device. We therefore considered two main interaction techniques: A hybrid technique incorporating touch and button controls with natural head movement, and an all-gestural (virtual) technique consisting solely of head movements. In all cases the user is presented with a red circular reticle at the center of her FOV, that acts as a "pointer". By placing the reticle over the various map controls (gazing), the user is able to then perform input actions with these controls as will be described below.

The map actions that we designed for are as follows (Table 1): Panning the map, zooming (in and out) of the map, hovering over a POI to obtain a tooltip and selecting a POI (a common action used to bring up more detailed information about it). Panning, zooming and selecting are actions available on all desktop and mobile maps, while hovering is an action available on desktop maps and augmented reality spatial applications. For panning we note here that a user is able to pan horizontally, vertically

	Delay	Feedback
T-H		
Virtual 1		
Virtual 2	X	
Virtual 3	X	X

Table 1. The four designed control methods and their use of delays and feedback

or diagonally by gazing the four edges or corners of the map.

	Panning	Zooming	Hover	Select
Tactile hybrid	Tap & hold + head movement	Double tap or double click on back button	Gaze on POI	Tap on POI
Virtual	Gaze on map edges	Gaze on zoom buttons	Gaze on POI	Continue to gaze on POI after hover

Table 2: The input modalities in the TH and Virtual input methods

We considered the concept of delays between preparing to perform an action and issuing the relevant command. On tactile controls (e.g. a mouse, or the Gear VR buttons), a user can position their finger on the control element in preparation to perform an action, without actually performing it (e.g. using a light touch). This allows users to change their mind before committing the input command. For the gestural interface we also considered the option of introducing a short delay between gazing and the registering of a command, in order to afford the users the ability to change their minds or avoid unintended interactions.

Feedback during interactions is also an aspect which we considered as important. When using a touch (or tactile) interaction, primary feedback is immediately available to the user that they have provided an input command correctly (e.g., a user can feel their fingers resting on a button before it is pressed and when it is clicked – the same applies also to touch areas). Further feedback is provided when the system performs an observable action, therefore indicating that the input command was successfully registered. On purely gestural interfaces however, this primary feedback is normally lacking as there is no way that a user can know that the input command has been performed, until they observe some effect taking place on the interface. Hence we designed a mechanism to provide this feedback to users by displaying visual cues that an input command was being registered prior to being enacted. This was implemented by providing some feedback to users, in the form of gradually fading in panning visual indicators or using a loading bar on the zoom icons and POIs. As a result, we ended up with four method designs that incorporated the following interaction elements (Table 2):

- Tactile hybrid vs. gestural only control
- Delay or immediate action (for gestural only)
- Delay feedback or no feedback (for gestural only)

To elucidate on the design, the key features of each interface, we provide the following implementation details. For the Tactile Hybrid (TH) interface we did not implement any delays or feedback on the interface, since the primary feedback is achievable via the tactile sensation and the user is able to postpone a primed action until ready to commit to it, by resting their hands on the controls. For V1, there is no delay in an action when the reticle is moved into a control area of the display. As such, feedback about the impending input is pointless to implement, since this occurs immediately. Interfaces V2 and V3 are similar with the exception that in V2, no feedback on the impending action is provided (after a short delay where nothing changes in the interface, the input command is enacted). In V3, visual feedback of the time elapsed during the delay is provided (see Fig. 2& 3).

Selection task
Please find and select the POIs labelled "Shop 1", "University 2" and "Bank 1".
Zoom Task
Please zoom in until the map displays the label "Gulf of Corinth".
Panning task
Follow the railway line in a western direction, until you find the POI labelled "St. Andrew's railway station".
Combination task
Starting at the center of the map (George Square) and moving along the traffic direction on Corinth Street, select the marker positioned on the first intersection with a street whose traffic direction is towards the right of Corinth Street.

Table 3: Examples of the experiment task scenarios

The application delays in V2 and V3 are set empirically to 750ms to provide a reasonably quick response that offers some opportunity for the user to reconsider an action. Additionally for the zoom buttons in V1-3 we implement an artificial delay of 500ms in case the user gaze persists on the control area, to prevent the uncontrollable continuous zoom in/out (simulating thus the time elapsing between successive double-taps/clicks used for zooming in the TH interface).

Finally, we added a logging mechanism in our application which captures the timestamped interactions of the user with the application (scrolling, selection, hovering, zooming). In this late-breaking results paper, we report only on the subjective feedback reported by participants, as the quantitative data is pending analysis.

Experiment

We recruited 25 participants (6 female), all computer science students aged between 18-30. Fourteen participants reported some experience with VR HMDs, mostly by trying them out in electronics stores. We proceeded to perform a laboratory experiment with the above interfaces, where we attempted to evaluate participants' performance using all four interfaces and the map control options (panning, zooming, hovering, selecting) in each. For this purpose, we randomly assigned an interface order to each participant, and with each interface the participant was asked to perform four tasks (see Table 3) – One task related to identifying and selecting POIs (all other controls disabled), one to panning (starting from a given location and following map features, e.g. roads or railway lines to find another location, zoom controls disabled), one to zooming (continuously zoom in/out until a map feature is visible, panning disabled) and a combination task which included zooming, panning and identifying/selecting a POI. The order of the first three tasks was random, and the combination task was always last. The participants were given some time to freely interact and familiarize themselves with the interfaces prior to a set of tasks and began when they reported they felt ready. After the completion of a set of tasks with each interface, the participants filled in an electronic version of the NASA-TLX questionnaire on a tablet device and continued to the next set of tasks with another interface. At the end of the experiment, participants were given a final set of subjective questions to respond to. The experiment took place in an office, with the participants sat down in an armchair for safety. Participants did not receive compensation for their time and participation took approximately 30 minutes.

Experiment Results

Following the completion of tasks with each method, we issued a NASA-TLX questionnaire to each participant (scaled 0-100 in increments of 5 units). Data from the NASA-TLX questionnaire did not exhibit any outliers so there was no need to filter any cases out of the analysis (Figure 5). All results that follow are reported using appropriate parametric or non-parametric tests, based on the normality of the data distribution, which is examined with the Shapiro-Wilk test.

Mental demand, Performance, Effort and Frustration
Friedman tests revealed no statistically significant differences in any of these factors (Mental demand $\chi^2_{(25)}=0.816$, $p>0.05$; Performance $\chi^2_{(25)}=3.211$, $p>0.05$; Effort $\chi^2_{(25)}=7.188$, $p>0.05$; Frustration $\chi^2_{(25)}=1.942$, $p>0.05$).

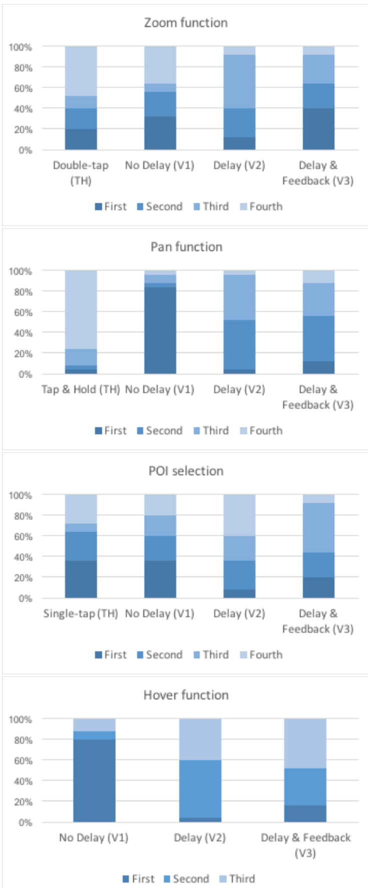

Figure 5: Results of the participant ordering of control methods

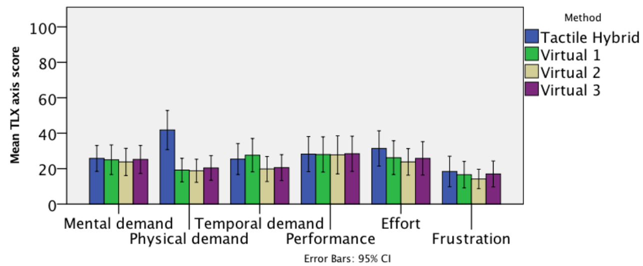

Figure 6: Results of the participant ordering of control methods

Physical demand

In terms of physical demand, a Friedman test reveals a statistically significant difference across the methods ($\chi^2_{(25)}$=22.415, $p<0.01$). Post-hoc Bonferroni-corrected Wilcoxon signed rank tests, reveal statistically significant differences between Tactile Hybrid and all Virtual methods (TH-V1 Z=-3.599 $p<0.01$, TH-V2 Z=-3.650 $p<0.01$, TH-V3 Z=-3.534 $p<0.01$), showing that the Tactile Hybrid method was perceived as the most tiring method due to the combination of manual and head motor control required to control the map.

Temporal demand

A Friedman test reveals statistically significant differences in the temporal demand perceived by participants ($\chi^2_{(25)}$=10.728, $p<0.05$). Post-hoc Bonferroni-corrected Wilcoxon signed rank tests, reveal statistically significant differences only between methods Virtual 1 and Virtual 2 (Z=-2.730, $p<0.01$). This was a surprising result because we expected that the delay introduced in V2 might cause participants to take longer to complete tasks, however this did not seem to be the case (we suspect that the pending analysis of our quantitative data might reveal the cause to be more actions taken in the V1 condition, due to the need for correcting inadvertent actions).

Appropriateness of the map size

At the end of the session, we asked participants for their opinion about the size of the displayed map area on a 5-point Likert scale (1=very small, 3=about right, 5=very big). Participants were asked to imagine that they might have this service available to them as an application during daily use while navigating an unfamiliar city. 88% of participants responded that they felt that the map was about the right size for this purpose, while the remaining 12% felt that the map could be classified as "big".

Subjective preference of individual map controls

We further asked participants to reflect on the map control options used in each method, and to rank the different control options by assigning their preferred order of preference for each control option (Figure 6). First, we asked them about their opinion of the zoom functions. For this, most participants (40%) placed the "Delay & Feedback" (V3) option in 1st place of preference, followed by the "No Delay" (V1) option (32% of participants). With regard to the hover function (displaying a tooltip prior to selection), 80% of participants placed the "No Delay" (V1) option as their 1st preference, followed by "Delay & Feedback" (V3) which 16% of participants preferred most. With regard to selecting a marker (POI) on the map, participants were tied between placing the "No Delay" (V1) and "Single Tap" (TH) option as their 1st preference (36%). However, the "Single Tap" option was rated 2nd by 28% of participants while the "No Delay" option was rated 2nd by 24%, so the "Single Tap" is considered as marginally preferable. Finally, with regard to panning, the “No Delay” (V1) option was rated as most preferred by 84% of participants, followed by “Delay & Feeback” (V3) (12%). Participants liked the combination of tap & hold combined with the head gesture option the least.

General User feedback
We also asked users’ feedback on what types of application they thought the use of large maps would be appropriate for. Twelve participants provided some ideas. Four participants stated that they felt this application would be good for educational purposes (e.g. teaching children geography, or to help students explore the spatial distribution of information on a subject). Three participants felt this application could replace mobile and desktop maps in services of any kind. One further participant indicated that this application would be useful for tourism applications only. Finally, two participants felt this application would benefit users with special needs (limited motor control) and also one user stated the large map would be good for gaming purposes.

As a last question, we asked participants to provide feedback on improvements or changes they would like to make to the application and received responses from seven participants. From this feedback, most comments reflected the operation of the interface (e.g. where controls should be placed or different combinations of control methods, as indicated in the order of preference in the previous questions). We highlight however the comments of two participants, one of whom felt that the control methods could also benefit from spoken instructions (voice command recognition) and a further participant who believed that the application could be supplemented by hand/arm gesture recognition.

Discussion and further work
The outcome of this subjective evaluation highlights some interesting findings. First, the combination of hand and head input appears to cause tiredness to the users, which is understandable as the tactile area is not visible and thus the arm needs to be constantly positioned appropriately in order to maintain the ability to quickly find the input areas (touchpad & buttons). Additionally, the introduced delays between actions in the virtual conditions do not seem to cause any significant overall delay or frustration in the users’ ability to complete tasks. We hypothesize that this is because the delay affords users the ability to prevent and correct inadvertent input, thus requiring less actions to correct it. Finally, the breakdown of user preference by individual map control rather than the entire input reveals that delay, visual feedback and simple touch gestures can play a role in designing a better control method for large AR/VR maps, though each technique has to be carefully applied to individual controls and not over the entire interface. Our analysis of the quantitative data gathered during the experiment might shed more light into these tentative findings. We also hope to re-design the control interface according to these findings and evaluate it in a laboratory and field setting, using AR instead of VR technology. Further work should also investigate participants’ performance while stationary and while mobile (or in situations like public transport), which may introduce problems during the hand or head gesture controls and highlight issues in resolving input uncertainty.

References

1. Balciunas, Andrius. 2013. User-Driven Usability Assessment of Internet Maps. In *Proceedings of the 26th international cartographic conference*.

2. Stefano Burigat, Luca Chittaro, and Silvia Gabrielli. 2006. Visualizing Locations of Off-screen Objects on Mobile Devices: A Comparative Evaluation of Three Approaches. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services* (MobileHCI ’06), 239–246. https://doi.org/10.1145/1152215.1152266
3. Arzu Çöltekin, Benedikt Heil, Simone Garlandini, and Sara Irina Fabrikant. 2009. Evaluating the Effectiveness of Interactive Map Interface Designs: A Case Study Integrating Usability Metrics with Eye-Movement Analysis. *Cartography and Geographic Information Science* 36, 1: 5–17. https://doi.org/10.1559/152304009787340197
4. E. Dubois, Ph. Truillet, and C. Bach. 2007. Evaluating Advanced Interaction Techniques for Navigating Google Earth. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...But Not As We Know It - Volume 2* (BCS-HCI ’07), 31–34.
5. Haosheng Huang and Georg Gartner. 2012. A Technical Survey on Decluttering of Icons in Online Map-Based Mashups. In *Online Maps with APIs and WebServices*, Michael P. Peterson (ed.). Springer Berlin Heidelberg, 157–175.
6. Manousos Kamilakis, Damianos Gavalas, and Christos Zaroliagis. 2016. Mobile User Experience in Augmented Reality vs. Maps Interfaces: A Case Study in Public Transportation. In *Augmented Reality, Virtual Reality, and Computer Graphics*, 388–396. https://doi.org/10.1007/978-3-319-40621-3_27
7. Kuparinnen, Liisa, Silvennoinen, Johanna, and Isomäki, Hannakaisa. 2013. Introducing Usability Heuristics for Mobile Map Applications. In *Proceedings of the 26th International Cartographic Conference (ICC 2013)*.
8. Rosemarijn Looije, Guido M. te Brake, and Mark A. Neerincx. 2007. Usability Engineering for Mobile Maps. In *Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems and the 1st International Symposium on Computer Human Interaction in Mobile Technology* (Mobility ’07), 532–539. https://doi.org/10.1145/1378063.1378150
9. Thomas Olsson, Tuula Kärkkäinen, Else Lagerstam, and Leena Ventä-Olkkonen. 2012. User evaluation of mobile augmented reality scenarios. *Journal of Ambient Intelligence and Smart Environments* 4, 1: 29–47. https://doi.org/10.3233/AIS-2011-0127
10. David Rudi, Ioannis Giannopoulos, Peter Kiefer, Christian Peier, and Martin Raubal. 2016. Interacting with Maps on Optical Head-Mounted Displays. In *Proceedings of the 2016 Symposium on Spatial User Interaction* (SUI ’16), 3–12. https://doi.org/10.1145/2983310.2985747
11. Vidya Setlur, Cynthia Kuo, and Peter Mikelsons. 2010. Towards Designing Better Map Interfaces for the Mobile: Experiences from Example. In *Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application* (COM.Geo ’10), 31:1–31:4. https://doi.org/10.1145/1823854.1823890
12. Edward Tse, Chia Shen, Saul Greenberg, and Clifton Forlines. 2006. Enabling Interaction with Single User Applications Through Speech and Gestures on a Multi-user Tabletop. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (AVI ’06), 336–343. https://doi.org/10.1145/1133265.1133336
13. Manlai You, Chun-wen Chen, and Lin, Hsuan. 2009. A usability evaluation of navigation modes in interactive maps. In *Proceedings of the 2009 Conference of the International Association of Societies of Design Research*.