User Experience of 3D Map Navigation – Bare-Hand Interaction or Touchable Device?

Robert Tscharn, Philipp Schaper, Jan Sauerstein, Sarah Steinke, Sebastian Stiersdorfer, Carsten Scheller, Huy Tan Huynh

Psychologische Ergonomie, Julius-Maximilians-Universität Würzburg

Abstract

New tracking technologies allow users to interact with digital systems in a more naturalistic way, while touchable interfaces represent a more precise but also less direct interaction. In this study, example devices for each of these interaction styles were compared in terms of their suitability for free navigation in a 3D Map. Efficiency, subjective attractiveness ratings and joy-of-use based on displayed emotions in differently complex tasks were measured. We focused on users without prior experience with both respective devices, allowing insights on first impressions and unbiased interaction. Our results showed higher joy-of-use and better efficiency in an easy navigation task for bare-hand interaction. Interestingly efficiency also improved for more complex tasks, after participants first worked with a touchable interface.

1 Introduction

In the last years, several technological developments paved the way for the applicability of naturalistic interaction styles with digital systems. The general shift away from classical user interfaces to a natural collaboration between computers and human beings was particularly facilitated by the upcoming of powerful sensor technologies. Markerless tracking of human movements in free space (e.g., Leap Motion Controller, Kinect v2) allows for input and information manipulation that mirrors our daily behavior with our social and physical environment. For example, gestures as lowering or raising a hand to emphasize verbal speech are a fundamental part of human-to-human communication. Providing systems with the ability to recognize these gestures promises direct and intuitive human-computer interaction (Rautaray & Agrawal, 2015). Therefore, more and more domains apply so-called bare-hand interaction and several approaches exist on its introduction to daily contexts. E.g., in the automobile context, this interaction style promises to be an innovative and useful feature for

Veröffentlicht durch die Gesellschaft für Informatik e.V. 2016 in
W. Prinz, J. Borchers, M. Ziefle (Hrsg.):
Mensch und Computer 2016 – Tagungsband, 4. - 7. September 2016, Aachen.

http://dx.doi.org/10.18420/muc2016-mci-0167

controlling in-car systems with reduced cognitive load and less driver distraction (e.g., Ohn-Bar, Tran, & Trivedi, 2012). Another domain for bare-hand gestures is 3D-modelling. Here, virtual objects can be manipulated and rotated in a very natural way compared to screen-based interactions (e.g., Dangeti, Chen, & Zheng, 2016). A growing market for immersive virtual reality hardware adds another field of application and bare-hand interactions are seen as a candidate for successful interaction with virtual or mixed realities' objects (e.g., Colaço et al., 2013). Still, the trend to add physical experience while interacting with digital content results in a growing interest in physical and touchable interaction. Here, digital content and its abstract objects are sought to be physically graspable to make use of the whole variety of humans' perceptions. Also, while classical interaction is on retreat the focus of new interactions is more and more on providing a positive user experience. The question is now, which facets of new and hand-free interaction are to be regarded as advantageous over classical interaction styles.

In fact, certain tasks have been identified, where classical interaction is advantageous to bare-hand interaction. For example, Bachmann, Weichert, and Rinkenauer (2014) showed that contact-free interactions with lower accuracy rates are outperformed by traditional pointing devices as the mouse device for pointing tasks. Physical interaction is often more precise and during specific tasks like 3D-modelling precision is a crucial aspect. But there are also tasks, where accuracy has not the highest priority, for example, while freely exploring a large scale virtual reality in contrast to following an exact path (Nabiyouni, Laha, & Bowman, 2014). Additionally, while navigating in the free space of a three-dimensional virtual reality we would not expect to get haptic feedback from this free space.

The underlying question of the present study is therefore, under which circumstances a less accurate, but direct and touchless bare-hand interaction is superior to a highly accurate, but indirect interaction style with a physical device while navigating in a 3D Map (Google Earth). We investigated, if the user experience – taking into account standard measurements as well as basic emotions analysis - would be influenced by the input device. Unlike prior studies which focused on prototypes and usability of complex (two-handed) gestures for interaction with Google Earth (Boulos et al., 2011; Stannus, Rolf, Lucieer, & Chinthammit, 2011; Stellmach, Jüttner, Nywelt, Schneider, & Dachselt, 2012), we looked at simple one-handed interaction with an established without the need to instruct participants about the interaction. Also, we analyzed carry-over effects between both styles to get insights of the users' application of experience with one interaction style to the other one. The contribution of this work is therefore a) the direct comparison of an interaction style based on bare-hand gestures with an interaction style based on a physical device and b) the investigation of the application of insights from one to the other interaction style.

2 Background

Interaction technologies

Wanderley, Kelner, Costa, and Teichrieb (2006) distinguish three interfaces to augment users' interaction with digital content. First, *Tangible User Interfaces* (TUI) use physical objects, tools or surfaces to represent and manipulate data. This provides users with a natural way of perceiving or manipulating digital content. Second, *Bare-Hand Interfaces* use hand positions, movements or gestures as the primary input for controlling the system in real time. Here, no direct contact between user and system is required and movement recognition can be achieved over distance. This also allows for very natural and usable interaction styles. Third, *Traditional Virtual Reality Interfaces* are used in form of gloves, joysticks, or carpets with sensors to interact with a system. These devices are used to indirectly interact with the system. In our study, we compared two of these interaction styles in the context of free space navigation: the *Leap Motion Controller* as a bare-hand interface and the *Space Navigator®* as a touchable, traditional virtual reality interface.

The Leap Motion™ Controller (www.leapmotion.com) is a new technology based on two monochromatic, infrared depth-sensors. It allows to detect and track hand or finger positions with a sampling rate of 300Hz in an area between 3cm and 50cm above the controller with a view span of about 150 degrees. Besides the position and orientation of one or more fingers it can also recognize tools as, for example, a pen. Despite the high accuracy for some finger constellations, recognition rates drop as soon as fingers are crossed or located very near to each other or the palm rotation exceeds 80 degrees (Nabiyouni et al., 2014). In sum, the main advantage of this device is the possibility of bare-hand interaction. On the other hand, the reliability of the Leap Motion Controller is high only under certain conditions.

The Space Navigator® is a 3D mouse developed and distributed by 3Dconnexion™ (http://www.3dconnexion.de). It is similar to a static joystick but holds six degrees of freedom (three types of translation: x-, y-, and z-axis; three types of rotation: pitch, roll, and yaw). Its sensors recognize changes of 4 micrometers. A first application in the 1970s was controlling robotics with six degrees of freedom. Nowadays, it is often used in the areas of robotics, CAD and in assisted 3D-modeling. The main advantage of this device is the high accuracy and reliability. Still it is a static and fixed device.

Classical and physical interfaces as a mouse or the Space Navigator provide an indirect interaction between user and computer since the user has to transform visual information depicted on a screen to physical movement of a controller. Also, predicting changes prior to a controller movement are prone to transformation errors and demand a high cognitive effort. In contrast, devices as the Leap Motion Controller or Kinect v2 provide a more direct interaction style. Still, these technologies are not accurate enough for precise object manipulation. Previously, comparison between bare-hand and classical interaction styles have been subject to research for standard devices as mouse controllers (e.g., Dangeti et al., 2016), but not devices as the Space Navigator. Because of its six degrees of freedom, this device allows for more complex interaction styles compared to a mouse based interaction with only two degrees of freedom and represents therefore a suitable comparison with the Leap Motion Controller.

Free Space Navigation

Navigating in a 3D virtual space (such as Google Earth) is a normal task nowadays. To manipulate the viewpoint of the camera while navigating, six different parameters are to be modified: 1) up/down, 2) front/back, 3) left/right, 4) roll, 5) yaw, and 6) pitch (an extra parameter for a field of view scale factor is ignored here). Traditional 2D input devices with only two degrees of freedom are not sufficient for this task without further adjustments to the interaction style. Only devices with six degrees of freedom can directly match these six parameters and represent an ideal interaction style for this task.

Different metaphors for traveling in a three-dimensional space have been suggested by Ware and Osborne (1990) that can be used for interaction design. The term "metaphor" is here used in the way of an "internal model" that represents and explains the behavior of the system in the user's mind. The first metaphor, *"eyeball in hand"*, assumes that the user places and moves an eyeball in the scene with the input device. Here, the user only sees what the eyeball is turned to in the fixed scene. Moving the eyeball in the scene changes the view on the screen. Using the second metaphor, "*scene in hand*", the perspective is static, but the scene can be moved around and rotated in front of the view. This results in a direct linkage between the position or rotation of the input device and the position or rotation of the scene. In the third metaphor, "*flying vehicle control*", a virtual vehicle is controlled by a device. In contrast to the "eyeball in hand" metaphor, the user "steers" this vehicle including direction and velocity with his input device. The evaluation of Ware and Osborne (1990) of these metaphors revealed that – from the perspective of the users - free navigation was best represented by the "*flying vehicle control*" metaphor. Halin, Humbert, and Bettenfeld (2015) also successfully implemented the "*flying vehicle control*" metaphor (called "*métaphore de l'avion*" in their work). On the other side, Nabiyouni et al. (2014) found that for free space navigation an interaction style based on the *"eyeball in hand"* metaphor resulted in better task times compared with an interaction style based on the "*flying vehicle control*" metaphor. One difference between both studies is the input device. Six degrees of freedom were provided in both studies. But Ware and Osborne used a control device for the tests of all metaphors presuming that the user has "something in his hand", either an eyeball, the scene or a physical controlling device. Nabiyouni et al. on the other hand implemented the metaphors using the Leap Motion Controller. Also, Ware and Osborne opted for a qualitative methodology using semi-structured interviews and content analysis. Nabiyoumi et al. mainly relied on task times to identify the more appropriate metaphor. Thus, an open question is how basing interaction styles on different metaphors is perceived by the user and if a positive user experience can be better achieved by a physical and touchable device like the Space Navigator or by the bare-hand interaction of the Leap Motion Controller.

User experience and instant emotions

Considering good user experience as one crucial factor for successful interaction design is unquestioned. Besides standardized questionnaires as, for example, the AttrakDiff (Hassenzahl, Burmester, & Koller, 2003), analyzing emotions during interacting with digital systems has been suggested to measure the quality of the users' experience: User-friendly devices are likely to trigger more positive emotions as satisfaction and the "fun to use" is manifested in emotions (Mahlke, Minge, & Thüring, 2006). Several approaches exist to measure these emotions as, for example, questionnaires, speech recognition, or physiological data (Terzis, Moridis, & Economides, 2013). In this study, we focused on facial expressions as markers for good user experience. One reliable instrument in this context is the FaceReader™ (Zaman & Shrimpton-Smith, 2006). Especially the "fun of use", defined by Desmet (2003) as "the fun one experiences from owning or using a product" was of particular interest in the present study. The FaceReader analyses recordings of participants' faces for six different emotions. It distinguishes between happy, angry, sad, surprised, scared, and disgusted plus a neutral state. It follows the Facial Action Coding System (FACS) based on Ekman and Friesen's theory that basic emotions are linked to facial models (Ekman & Friesen, 1977). It has been proven to be a valid and reliable instrument for instant emotions and fun of use (Zaman & Shrimpton-Smith, 2006).

Research question

The underlying research question for this study was to compare two interaction styles for free space 3D navigation: the Leap Motion Controller provides a bare-hand, direct interaction style and the Space Navigator as a touchable, indirect interaction style. It was predicted that the Leap Motion Controller would be experienced as a joyful, but not very precise interaction device. The Space Navigator on the other side was expected to allow for a very precise navigation in space without a high joy-of-use.

First, the difference in joy of use should be indicated by more happy emotions during the interaction with the Leap Motion Controller, based on the FaceReader data. In addition the Leap Motion Controller was expected to be perceived as more attractive, due to its bare-hand navigation capabilities, this should be reflected in higher attractiveness ratings in the AttrakDiff questionnaire (Hassenzahl et al., 2003). Second, different tasks demands should affect efficiency measures, based on the task time for different task complexities. Users should be more efficient with the Leap Motion Controller for a simple task, which does not require high precision, such as simple navigation tasks over short distances. For long distance navigation tasks, which require more complex input, and for tasks requiring high precision, the efficiency for the Space Navigator should be higher. Third we expected to find an carry-over effect between the interaction with both input devices: since both interaction styles are based on six degrees of freedom, user should perform better when they had used the other device in a previous task.

3 Methods

3.1 Participants

A total of 20 students of the Julius-Maximilians-Universität of Würzburg volunteered to participate in the study (13 males, average age 21.7 years). Due to the focus of the study we only allowed volunteers, who had never used either of the input devices. One participant was excluded from the analysis as the overall handling time for one of the input devices was more than three standard deviations above the group mean.

3.2 Design

A balanced within-subjects design was used. Half of the participants first worked with the Space Navigator, while the other half started with the Leap Motion Controller. Participants were randomly assigned to their starting device. Four different use cases were used for the study with different complexities in the context of 3D Navigation. Thus, participants completed two blocks of four use cases using another interaction style for each block. Dependent variables were for each use case and device the task time and emotion analysis of the FaceReader. Also, each participant rated both devices on a AttrakDiff questionnaire (Hassenzahl et al., 2003).

3.3 Procedure and materials

After the informed consent, participants were seated in front of a monitor and instructed about the general task, navigating within the google earth software, and using the supplied input device. The current input device (Space Navigator or the Leap Motion Controller) was the only available hardware for the use cases. The participants saw a short video about the current input device before the use cases were presented. Participants were set at a predetermined starting location in Google Earth before each use case. To eliminate geological knowledge as a possible confound the target locations were labeled within the 3D Map.

We used four different use cases: First and second, two navigation tasks over long distances, moving from the statue of liberty in New York to the center of Paris as indicated by the software itself, and subsequently to the library of the University of Würzburg. Third, a straight, short distance navigation task, crossing the Golden Gate Bridge in San Francisco, was used. Fourth, a rotation navigation task, circuiting the statue of liberty, was assigned. The time used between start and arrival was measured for each use case. For the use cases with a set destination a minimum height above ground had to be reached to consider the use case passed.

After working on all four use cases with one input device, participants were asked to fill out the AttrakDiff questionnaire (Hassenzahl et al., 2003). Participants were filmed during the first two minutes with each input device. The videos were subsequently analyzed using the Face Reader 6.0 software. After the first block of use cases with the first device, participants repeated this procedure for the second device.

4 Results

To test for differences in overall efficiency we used the sum of all times needed to finish each of the four use cases for each input device. Here, the Space Navigator ($M = 169.68$ seconds, $SD = 51.00$) indicated higher overall efficiency compared to the Leap Motion Controller ($M = 229.68$ seconds, $SD = 68.45$; $t(18) = 3.46$, $p = .003$, $d = .99$).

Subsequently we first used a t-test for depended variables to specifically compare the handling time for the straight, short distance navigation task. We found significantly faster handling times for the Leap Motion Controller (*M* = 35.58 seconds, *SD* = 22.12) compared to the Space Navigator (*M* = 53.32 seconds, *SD* = 24.95; *t*(18) = -2.72, *p* = .014, *d* = .75). To compare the overall efficiency for the long distance and high precision tasks, as well as for the interaction between device and presentation order we conducted a 2 × 2 mixed ANOVA for the aggregated handling times for these three use cases. Presentation order was used as between-subject factor (Leap Motion Controller first, Space Navigator first) and input device as within-subject factor (Leap Motion Controller, Space Navigator). The analysis revealed a significant main effect of input device, *F*(1, 17) = 55.95, $p < .001$, $\eta_p^2 = .77$, and a significant interaction, *F*(1, 17) = 13.36, *p* = .002, $\eta_p^2 = .44$, but no significant effect of presentation order *F*(1, 17) = 3.37, *p* = .084, $\eta_p^2 = .17$ (Figure 1).

A planned contrast for the Leap Motion Controller showed significantly longer handling times if participants worked with the Leap Motion Controller first (*M* = 231.44 seconds, *SD* = 58.28), compared to using the Space Navigator first (*M* = 160.50 seconds, *SD* = 35.15; *t*(17) = 3.25, *p* = .005, *d* = 1.50.

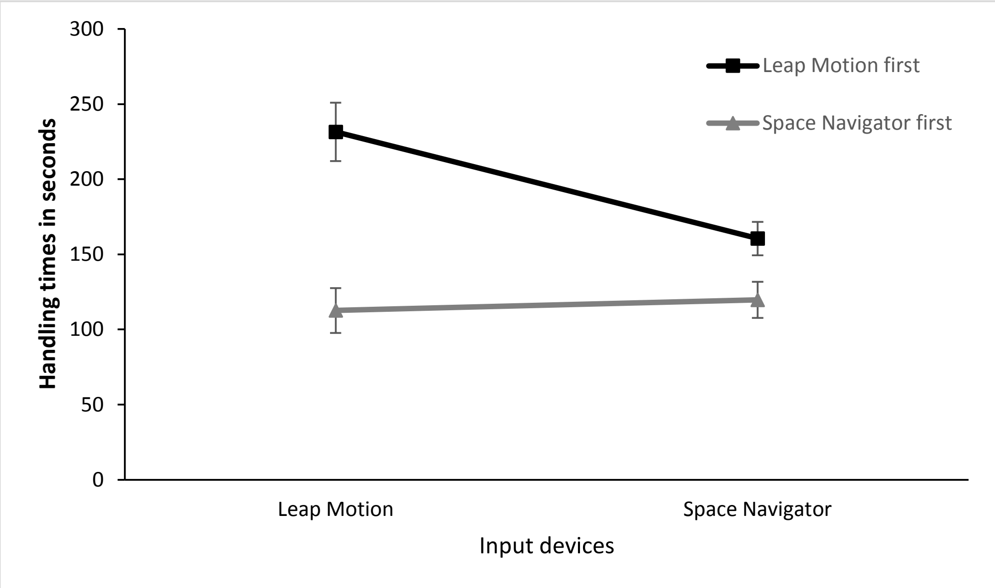

Figure 1. Aggregated reaction times for use cases 1, 2 and 4 separated by presentation order. Error bars represent SE.

For the five subscales of the AttrakDiff (see table 1), we used Bonferroni-corrected depended t-tests. We found a significantly better evaluation of the attractiveness of the Leap Motion Controller (*M* = 3.98, *SD* = .43) compared to the Space Navigator (*M* = 3.69, *SD* = .28; *p* = .045). The input devices did not differ significantly, with descriptively higher values for the Space Navigator on the other four scales.

	Pragmatic Quality	Hedonic Quality	Hedonic Identity	Hedonic Stimulation	Attractiveness
Leap Motion Controller	3.90 (.34)	4.04 (.29)	4.33 (.49)	3.76 (.50)	3.98 (.43)
Space Navigator	4.09 (.31)	4.16 (.31)	4.41 (.43)	3.91 (.45)	3.69 (.28)

Note: Standard deviations are in parentheses

Table 1: Results for the AttrakDiff questionnaire

The recorded video-data was analyzed with the Face Reader 6.1 software. We calculated the average of each of the six detectable emotions plus the neutral state (see table 2). Two participants had to be excluded from this analysis as the software could not read their expression during more than 40% of at least one of their videos. Due to the low occurrence of emotional states beside neutral and happy (all other below five percent of the time), we excluded these states from further analysis. We found descriptively more happiness for the first interaction with the Leap Motion Controller (M = 28.14, SD = 20.30) compared to the Space Navigator (M = 17.08, SD = 15.58), but no significant difference between both devices $t(17) = -2.00$, $p = .062$, $d = .47$. Similarly the neutral state occurred descriptively less during the interaction with the Leap Motion Controller (M = 61.23, SD = 20.24) compared to the Space Navigator (M = 71.17, SD = 15.58), without reaching significance $t(17) = 2.00$, $p = .062$, $d = .47$. A subsequent power analysis for the happy state revealed that the power to find an effect was below .5 given the sample size of this study.

	Happy	Sad	Angry	Surprised	Scared	Disgusted	Neutral
Leap Motion Controller	28.14 (20.30)	.43 (.47)	2.06 (2.27)	4.53 (5.18)	.31 (.49)	3.31 (5.85)	61.23 (20.24)
Space Navigator	17.08 (15.58)	.93 (1.59)	2.18 (2.45)	4.78 (6.15)	.19 (.25)	2.67 (3.15)	72.17 (15.58)

Note: Standard deviations are in parentheses

Table 2: Percentage of emotions shown during first interaction with the respective device

5 Discussion

By using navigation tasks in a 3D map we compared the performance and user experience of bare-hand interaction (Leap Motion Controller) to a physical, touchable input device (Space Navigator). The perceived attractiveness of the Leap Motion Controller was higher than for the Space Navigator, indicating that interacting with the device as a whole is perceived more positively in the context of 3D navigation. This is contrasted with the significantly higher overall efficiency of the Space Navigator for the sum of all use cases as well as the descriptively higher ratio of happy emotions during the interaction. Even though this difference was not significant, most likely due to insufficient power, descriptively it is still supportive for our claim of higher joy-of-use for bare-hand interaction in 3D navigation. For the less complex short distance navigation task we found significantly shorter handling times for the Leap Motion Controller compared to the Space Navigator, while the opposite was true for the combined handling times for the other use cases. In addition we also found the predicted interaction between the input device and the order in which they were used within our study. After using the Space Navigator, which presumably allowed for a more controlled exploration with a high accuracy, participants were significantly faster when using the Leap Motion Controller compared to the participants who used it first. Thus, users might obtain valuable knowledge about this kind of task by previously using a precise indirect controller as the Space Navigator. This knowledge can be used later for a different interaction style: the bare-hand interaction of the Leap Motion Controller. This shows that bare-hand interaction might be not the best way to learn free space navigation. Rather, users should start with an indirect, but precise device before operating with the bare-hand interaction. The implications for instructional design are interesting, since further studies could further investigate this finding and provide recommendations for learning environments focusing on navigation in 3D maps.

6 Conclusion

Within the scope of our study we conclude that the direct, bare-hand interaction for free space navigation is perceived as more attractive and more fun to use. In addition, it is more efficient in less complex tasks. The navigation following the touchable but indirect interaction style has been shown to be more efficient for complex tasks, but most notably we found indications, that the interaction with the touchable input device strongly affects the later use of the bare-hand interaction. As most users are not used to interaction with more than two degrees of freedom, the Space Navigator might provide an easier and more controlled introduction to this sort of interaction, demonstrated in this study by significantly higher efficiency in later interactions with the Leap Motion Controller. In summary, this study contributes to a growing literature body on bare-hand interactions by systematically comparing it to an indirect, but more precise and touchable baseline technology and providing implications for a better match between interaction technology and task-complexity.

References

Bachmann, D., Weichert, F., & Rinkenauer, G. (2014). Evaluation of the leap motion controller as a new contact-free pointing device. Sensors, 15(1), 214-233.

Boulos, M. N. K., Blanchard, B. J., Walker, C., Montero, J., Tripathy, A., & Gutierrez-Osuna, R. (2011). Web GIS in practice X: a Microsoft Kinect natural user interface for Google Earth navigation. International journal of health geographics, 10(1), 1.

Colaço, A., Kirmani, A., Yang, H. S., Gong, N.-W., Schmandt, C., & Goyal, V. K. (2013). Mime: compact, low power 3D gesture sensing for interaction with head mounted displays. In Proceedings of the 26th annual ACM symposium on User interface software and technology, 227-236. ACM

Dangeti, S., Chen, Y. V., & Zheng, C. (2016). Comparing bare-hand-in-air Gesture and Object-in-hand Tangible User Interaction for Navigation of 3D Objects in Modeling. In Proceedings of the TEI'16, 417-421. ACM

Desmet, P. (2003). Measuring emotion: Development and application of an instrument to measure emotional responses to products. In Blythe, M.A., Monk, A.F., Overbeeke, K., & Wright, P.C. (eds.), Funology: From Usability to Enjoyment (pp. 111-123). Springer.

Ekman, P., & Friesen, W. (1977). Manualfor the Facial Action Coding System: Palo Alto: Consulting Psychologists Press.

Halin, G., Humbert, P., & Bettenfeld, V. (2015). Navigation main libre dans une reconstitution urbaine 3D: application à la valorisation de plans reliefs. Paper presented at the 27th Conférence Francophone sur l'Interaction Homme-Machine.

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In Mensch & Computer 2003 (pp. 187-196). Springer.

Mahlke, S., Minge, M., & Thüring, M. (2006). Measuring multiple components of emotions in interactive contexts. In CHI'06 Extended Abstracts, 1061-1066. ACM.

Nabiyouni, M., Laha, B., & Bowman, D. A. (2014). Poster: Designing effective travel techniques with bare-hand interaction. In Proceedings of the Symposium on 3D User Interfaces, 139-140. IEEE.

Ohn-Bar, E., Tran, C., & Trivedi, M. (2012). Hand gesture-based visual user interface for infotainment. In Proceedings of AutomotiveUI '12, 111-115. ACM.

Rautaray, S. S., & Agrawal, A. (2015). Vision based hand gesture recognition for human computer interaction: a survey. Artificial Intelligence Review, 43(1), 1-54.

Stannus, S., Rolf, D., Lucieer, A., & Chinthammit, W. (2011). Gestural navigation in google earth. In Proceedings of the 23rd Australian Computer-Human Interaction Conference, 269-272. ACM.

Stellmach, S., Jüttner, M., Nywelt, C., Schneider, J., & Dachselt, R. (2012). Investigating Freehand Pan and Zoom. In Mensch und Computer, 303-312. Oldebourg Verlag, München.

Terzis, V., Moridis, C. N., & Economides, A. A. (2013). Measuring instant emotions based on facial expressions during computer-based assessment. Personal and ubiquitous computing, 17(1), 43-52.

Wanderley, I., Kelner, J., Costa, N., & Teichrieb, V. (2006). A survey of interaction in mixed reality systems. Paper presented at the Symposium on Virtual Reality, 2006.

Ware, C., & Osborne, S. (1990). Exploration and virtual camera control in virtual three dimensional environments. In Proceedings of SIGGRAPH Computer Graphics, 175-183. ACM.

Zaman, B., & Shrimpton-Smith, T. (2006). The FaceReader: Measuring instant fun of use. In Proceedings of the NordiChi '06, 457-460. ACM.