Mobile 3D Graphics and Virtual Reality Interaction

Wolfgang Hürst
Utrecht University
Department of Information and Computing Sciences
PO Box 80.089, 3508TB Utrecht, The Netherlands
huerst@uu.nl

Matthias Helder
Utrecht University
Department of Information and Computing Sciences
PO Box 80.089, 3508TB Utrecht, The Netherlands
mhelder@students.cs.uu.nl

ABSTRACT

Mobile devices such as smartphones and tablets offer great new possibilities for the creation of 3D games and virtual reality environments. However, interaction with objects in these virtual worlds is often difficult – for example due to the devices' small form factor. In this paper, we define different 3D visualization concepts and evaluate related interactions such as navigation and selection of objects. Detailed experiments with a smartphone and a tablet illustrate the advantages and disadvantages of the various 3D visualization concepts. Our results provide new insight with respect to interaction and highlight important aspects for the design of interactive virtual environments on mobile devices and related applications – especially for mobile 3D gaming.

Categories and Subject Descriptors

I.3.6 [**Computer Graphics**]: Methodology and techniques – *interaction techniques.*

General Terms

Design, Experimentation, Human Factors.

Keywords

Mobile 3D graphics, sensor input, virtual reality interaction.

1. INTRODUCTION

Today's tablets and mobile phones are able to display 3D graphics that allow for high-end gaming and realistic virtual reality applications. In addition, sensor data combined with the ability to freely move the device offers new possibilities for the creation of virtual 3D worlds. Navigation and interaction however remain a challenge due to the devices' small form factors and noise introduced by the sensors. In this paper, we define different visualization concepts for 3D data on mobiles and evaluate related interactions. Our experiments pinpoint important characteristics of each concept and provide valuable information for the development of 3D games on mobiles.

Contributions of our paper include: a formal introduction of different 3D visualization concepts for virtual worlds on mobile devices (Section 2), a detailed user study investigating basic interaction tasks and the consequences of the visualization concepts for such canonical interactions (Section 4), and an analysis of the results indicating useful applications for the different concepts and potential limitations (Section 5). Related and future work is discussed in Sections 3 and 6, respectively.

2. VISUALIZATION CONCEPTS

Standard visualization. Figure 1 illustrates the standard way to visualize 3D graphics: A 3D model of a virtual world is projected perspectively onto the 2D screen in the direction of the observer (who is assumed to sit in front of the center of the screen). The red arrows indicate the related coordinate systems for the virtual world (3D) and the screen space (2D). In the following, we refer to this kind of visualization as *standard 3D virtual world visualization*, or *standard visualization* for short.

Navigation in such virtual 3D worlds is usually realized by rotating the world in the opposite direction of where the user wants to go. For example, think about a flight simulator where pushing the left arrow key on the keyboard or moving a game controller to the left evokes a movement of the related virtual world to the right thus creating the illusion of flying to the left in this virtual world – as illustrated in Figure 2.

Because mobile phones and tablets usually lack a physical controller or keyboard, an onscreen controller is often used for navigation (cf. bottom left on the screen shown in the images in Figure 10 on the last page of this paper). Alternatively, we can use the accelerometer that is commonly integrated in high-end phones to map tilting movements of the device to opposite movements of the virtual world. For example, tilting the phone to the left results in movement of the virtual world to the right, thus creating the illusion of navigation to the left in the virtual world.

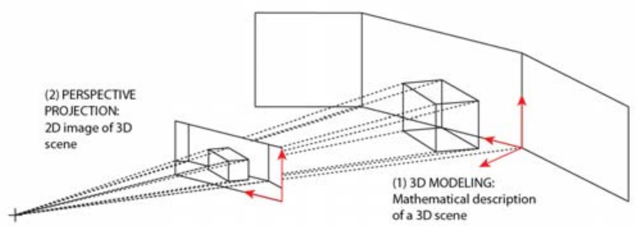


Figure 1. Standard 3D visualization by modeling and perspective projection of a 3D virtual world.

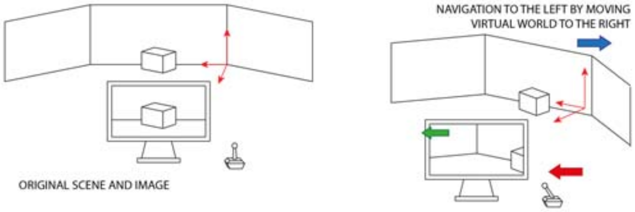


Figure 2. Standard navigation of 3D data on desktop PCs.

Short presentation, ACE'2011 – Lisbon, Portugal.

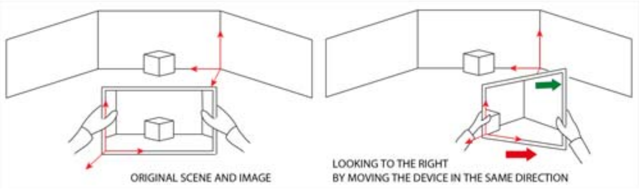


Figure 3. Fixed virtual world visualization by adaptive projection of the 3D model under consideration of the device's orientation.

Fixed world visualization. From sensors such as magnetometer (i.e. a digital compass) and accelerometer, we can get information about where a mobile device is pointed. Hence, on a mobile, we do not necessarily have to move the virtual world, for example to the left, to create the illusion of looking to the right. Instead, we can keep the virtual world fixed with respect to the real world and update the 2D visualization of it on the mobile's screen with respect to the device's movement and orientation – as illustrated in Figure 3. Again, the thin red arrows indicate the related coordinate systems. Notice that the device is now represented by a tridimensional coordinate system because its orientation with respect to the tridimensional virtual world needs to be specified in order to properly update the 2D graphics on the screen.

Hence, there is no need to move the virtual world in order to look around. Instead, moving the device in a specific direction in the real world creates the illusion of looking in exactly this direction in the virtual one as well. The device becomes a window into this virtual world. Because the 3D world does not move but stays fixed with respect to the real world, we refer to this concept as *fixed 3D virtual world visualization*, or *fixed world* for short.

Shoebox visualization. Considering the two concepts introduced so far, we have one where we assume a fixed device – e.g. the screen of a desktop PC – and create the illusion of movement in a virtual world by rotating the 3D graphics in the opposite direction. In the second concept, we assume a fixed virtual world and create the illusion of looking in a specific direction by moving the device there. Obviously, we can combine both approaches and create new experiences by allowing device and virtual world to move at the same time. Maybe the most interesting one is where we move both of them synchronously: The virtual world moves with the device as if it is "glued" to it. Updating the perspective view of the 3D world in the display with respect to the device's orientation and the (fixed) viewer position creates the effect of looking into a box – as illustrated in Figure 4. Hence we refer to this visualization in the following as *virtual 3D shoebox visualization* or *shoebox* for sort.

Further visualizations and actions in virtual worlds. Above, we defined different visualization concepts based on restrictions in terms of possible movements (e.g. by assuming a fixed screen for the standard approach, a fixed virtual world in the fixed world concept, and a fixation between the 3D graphic and the device in the shoebox approach). In general, we are basically dealing with three coordinate systems defined by the user/real world, the device, and the virtual world. Each of them offers six degrees of freedom, i.e. translation along all three axes and rotation around them. Hence, there are nearly endless options for other visualization concepts by introducing further restrictions or loosening some of the existing ones. For example, if we want to allow users to navigate in the fixed world concept, we obviously have to allow movement of the virtual world if we assume a fixed user position. Investigating such user navigation in these different concepts is part of our future work. In this paper, we restrict ourselves to studying different approaches for object navigation and selection. Being able to interact with objects in a virtual world, and to manipulate and influence them is an essential requirement for creating interactive 3D games. For example, one might want to *select* an object (e.g. to pick up a treasure in a treasure hunt game), *translate* objects (e.g. to move a piece in a 3D puzzle), *scale* or *rotate* them (e.g. to make a certain object fit into a related hole). While such interactions are well established and researched in traditional desktop-based virtual reality worlds, research in 3D interaction on mobile phones and tablets is still in its beginning. In the following, we are investigating the canonical interactions ***selection*** and ***navigation*** of objects with respect to the three visualization concepts introduced in this section.

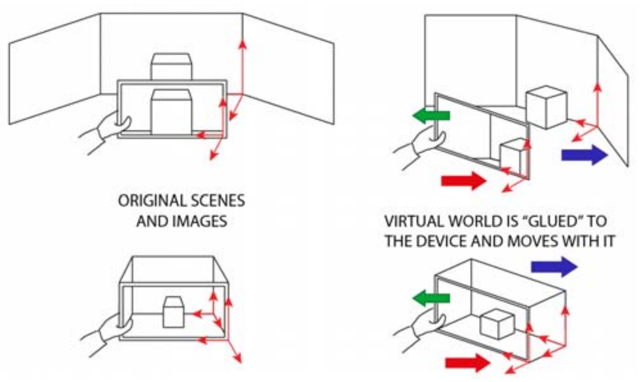


Figure 4. Shoebox visualization by adaptive projection of the 3D scene under ing the device's orientation. If the size of the scene matches the device's display – as in the lower case – the effect resembles looking into a box that is glued to the device.

3. BACKGROUND AND RELATED WORK

Work related to **virtual environments on handheld devices** dates back to the early 90s, especially the Chameleon prototype, a visualization and interaction concept for navigating an egocentric 3D information space [FZC93] [BF98] [Fit93]. The handheld prototype was compared against an equally sized static display, as well as a larger static display. It was concluded that depth perception on the small, yet moveable Chameleon concept is about as good as for a large display. Performance on the small static display, however, was much poorer. A multitude of studies extended the Chameleon concept, often supporting the results previously found. In [MWW06] and [Yee03], similar navigation and interaction tasks for 2D scenarios suggest the importance of spatial relationships for devices with limited screen sizes. More recent work includes concepts that are fully spatially aware in 3D space, and present a natural example of the metaphor where the screen of the mobile device acts as a window or peephole into virtual space [BKA05]. Another study shows the same concept can be applied to develop a fight simulation game named Mirage Money [GCCV08].

Little work exists that aims at specifying a clear **design space** for such virtual worlds on handhelds and related visualization concepts. For example, in a comparative study of Head-mounted Displays (HMD), Chameleon environments and so called CAVEs [BF98], a visually intuitive, however informal design space is proposed. It characterizes the different concepts for three types of virtual environments in terms of the relationships between the eyes and hands of the viewer, and the display. The common denominator for the three concepts – the feature that makes them suitable for structuring into a design space – is the "human-centric", also called "ego-centric" or "outside-in", viewpoint into the virtual environments. Game developers tend to name this the first-person view, as the graphical perspective is rendered from the viewpoint of the player's character. Other studies define visualizations of similar virtual environments by using intuitive metaphors, e.g. "window in the hand" and "scene in the hand" [Fit93], [WO90]. Here, the scene in the hand metaphor describes the scenario in which the viewpoint of the camera into the virtual world is static, and where the user manipulates, rotates and navigates the scene onscreen with the help of an input device. The window in the hand metaphor, sometimes also called "eye(ball) in the hand" [Han97] or "camera in the hand" [DBC02], specifies the mobile device as a physical window through which can be looked into a virtual world. Common examples that apply this concept are augmented reality applications, but also purely virtual variations like the previously mentioned Chameleon prototype.

Interaction with virtual worlds on mobile devices generally relies on sensor information. Examples for commonly used sensors include accelerometer, magnetometer, gyroscope, digital camera and sensors to measure light and temperature levels [HPSH00]. Previously, such sensors were often only available as separate external modules and lacked convenient access from within different software platforms. As described in the previous section, extending traditional and existing concepts with sensory input can create new interaction and visualization concepts. Some of these are specifically directed towards mobile devices, such as the Chameleon prototype, and augmented reality applications. Others tend to require complex setups with expensive hardware, such as the Cave Automatic Virtual Environment (better known by its recursive acronym CAVE) and fish tank concepts [BF98] [DJK+06]. In relation to mobile 3D games and interaction, most of them incorporate sensor usage relying on the magnetometer and/or accelerometer for input. From within the scientific field some early examples demonstrate possible applications; e.g. Tunnel Run and the afore mentioned Mirage Money [HPSH00] or use the tilting motion for navigational actions [Hod09] [Vis09], while others rely on data from a magnetometer or gyroscope to simulate the idea of a 360 degree world around the user [Zic09], or track a user's head or hand using a camera to interact with the device [Tar09].

4. INTERACTION STUDY

Motivation and goal. Normally buttons on a controller or keyboard are used for interaction with objects in a virtual environment on desktop PCs. Because of the small form factor, mobile phones and tablets lack such physical control elements but rely on touch screen interaction. Doing this with the new concepts introduced in Section 2 could have some inherent disadvantages. First, the position of objects in the virtual world and shoebox approach changes when the device is tilted and thus might be harder to target. See for example in Figure 5 how the position of the blue ball changes its location with respect to the observer's position when the device is tilted, whereas its absolute position within the scene says the same. Shaking hands and noise from the sensors used to adapt the 3D projection can introduce jitter and further instability. Tilting the device also changes the navigation direction from the perspective of the observer, as illustrated again in Figure 5. Hence, we need to study such interactions in order to get insight into what kind of interactions are possible in what kind of concept and thus what kind of applications we can create. In the following, we present a study investigating the canonical interactions *selection* and *navigation* in the shoebox and fixed world approach. More complex interactions such as *rotation* and *scaling* of objects are part of our future work (cf. Section 6).

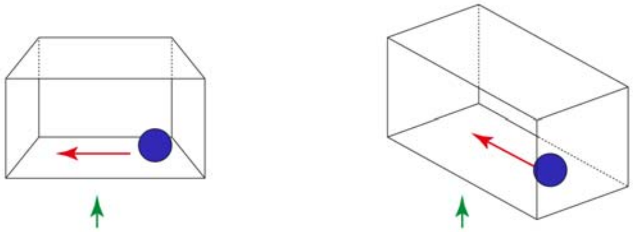


Figure 5. Potential interaction problem in the shoebox approach: when tilting, the position of the blue ball changes with respect to the observer's view point (green arrow), although it stays the same with respect to the rendered scene. Also notice that the from the observer's point of view, the ball now moves diagonal instead of straight (red arrow).

4.1 Concepts and Implementation

In the following, we discuss how to integrate object navigation into the previously introduced concepts and describe our implementation.

Navigation and interaction in standard approach. In standard 3D visualizations, object navigation is commonly implemented using the onscreen joystick. Selection is done by tapping on the touch screen. In order to be able to compare this approach with the shoebox and fixed world concepts, we created two scenes: one with a 180 degree view (because the shoebox does not allow 360 degree turns), and one with a full 360 degree view where the user is standing "inside of a box" – as illustrated in Figure 6. To eliminate influences of the graphics as much as possible, simple visuals (i.e. boxes for targets and obstacles, spheres for objects to navigate and select) and consistent coloring schemes have been used (red for targets, green for obstacles, blue for moving objects). A black and white checkerboard pattern was used for the environment in order to enable easy orientation while keeping influences of the graphics as low as possible.

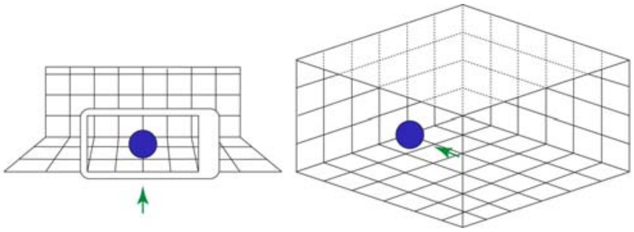


Figure 6. Scenes used in the evaluations (notice that for the cube in the fixed world approach only the floor and back walls are drawn for better visibility; during the tests, all faces were rendered with a checkerboard pattern). Likewise, the scenes in the test of the shoebox included a wall to the left and right. The green arrow illustrates the user's view point again.

Navigation and interaction in fixed world approach. The fixed world concept allows us to naturally look around in the virtual world by moving the device in the desired viewing direction. Hence, in this concept, moving objects is realized by coupling the object with the camera, i.e. the object stays in the center of the screen and is moved along with left and right movements of the device – as illustrated in Figure 7. For movements in depth, the onscreen joystick is used. Selection of objects is again implemented by tapping on the touch screen.

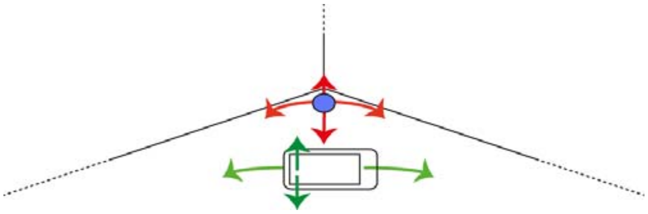


Figure 7. Navigation in fixed world approach: up/down movements of the on screen joystick (green up/down arrows) move the object further away/closer, left/right movements of the object are realized by moving the device in the respective direction (green left/right arrows).

Navigation and interaction in shoebox approach. By tilting the phone, the shoebox approach enables users to see a larger area of the attached virtual world compared to the standard visualization (cf. Fig. 8). In order to support also navigation and not just orientation, we allow movement of the attached virtual world parallel to the screen of the device by coupling a moving camera with the object that we want to move (which in turn is therefore always in the center of the screen) – as illustrated in Figure 9. Hence, in the actual implementation, navigation of objects is done via the onscreen joystick. As in the other concepts, objects can be selected by tapping at their position on the touch screen.

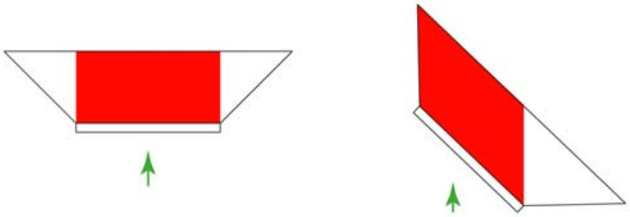


Figure 8. Visibility (red polygon) for the observer (green arrow) in the shoebox approach depending on different tilting angles of the device (top view for better illustration).

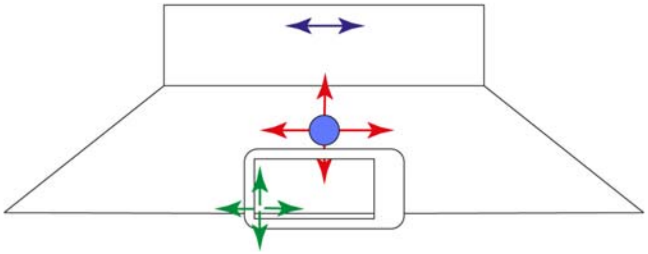


Figure 9. Navigation in the shoebox approach: Left/right movements of the on screen joystick (green arrows) result in related left/right movements of the object (red arrows) and scene (blue arrows). Up/down movements only move the object, but not the camera, so the scene stays fixed.

Figure 10, which can be found on the last page of this paper, illustrates the actual implementation on a Samsung Galaxy tablet (screen size 7”, resolution 1024x600 pixels) and a Motorola Droid/Milestone phone (screen size 3.7”, resolution 854x480 pixels) which were used in our evaluations. Both devices run under the Android OS version 2.1. Rending of the 3D graphics was implemented with the OpenGL ES 1.1 API for which both devices offer hardware support.

4.2 Experimental Setup

Given the potential problems mentioned in the beginning of this section, we put up the hypothesis that the advanced visualization concepts shoebox and fixed world are more engaging and appealing, but also harder to control when actual interaction takes place. In order to verify especially the second part of this statement, we set up an experiment evaluating how well people can solve typical interaction tasks with each of the concepts. Interaction in standard visualization served as a ground truth and reference.

Participants and study design. 24 subjects (20 male, 4 female, ages 18-25) participated in the experiments. We have purposely chosen this age range because we expect this to be a common target audience for 3D games and other popular 3D applications. None of the subjects was colorblind. We used a within subject design, i.e. each participant tested the shoebox and fixed world concept (each compared to the standard approach) on both devices. The order of devices and all concepts (shoebox vs. fixed word as well as standard vs. shoebox/fixed world) was counterbalanced across subjects. 18 participants had experience with a smartphone and touch screen interaction, and one of them had tablet experience. However, in the evaluation we could not observe any differences with respect to these levels of experience.

Tasks. Each participant was requested to fulfill three navigation tasks and one selection task in each of the setups. In the first one, users had to navigate the blue ball to the target – a randomly placed red box. Task two was similar, but green boxes were randomly placed in the scene as obstacles that users had to navigate around. In the third task, the obstacles were placed in a row and users had to navigate the ball in a slalom-like way around them to reach the target – a red box placed at the end of the row. The forth and final task required users to select a floating red sphere by clicking on the respective position on the touch screen. In order to create a random but comparable setting for all tests, we split the scenes in equally sized rectangular areas (16 for shoebox, 48 for fixed world related tests). Targets and obstacles were randomly placed within these areas under the condition that each square was used equally often, thus guaranteeing a comparable number of situations where targets were not visible in the initial position, so the scene had to be moved to see and reach it (cf. Figures 6 and 7 for the shoebox approach). In case of the selection task, we applied a similar approach for the vertical placement of the target. For the slalom task, obstacles were arranged in a vertical, horizontal, and diagonal layout (if seen from top). Notice that each of these tasks is designed to be completed by using two hands as previous studies have shown that for natural interaction two hands are preferred [Yee03], [Hen07]. Hence, one hand is targeted to hold the device steady, while the second can be used for interaction.

Procedure & data gathering. For later analysis and gathering of quantitative data, all interactions and sensor information was logged during the tests. For qualitative data analysis, users had to fill out a questionnaire at the end and give three keywords best describing each of the concepts. These were used to start an informal discussion and interview with the otherwise neutral observer. Task one (navigation w/o obstacles) and two (navigation w/ obstacles) each had to be done 12 times, task three (slalom) was done with three randomly selected setups, and task four (selection) 12 times. Results presented in the next section are averaged over participants and tasks unless stated otherwise. Overall, the experiments took between 20-30 min for each of the two devices (phone, tablet), i.e. 40-60 min in total per participant.

5. RESULTS AND DISCUSSION

Figure 11. Task completion times (in msec) with the shoebox approach for phone (top) and tablet (bottom).

Table 1. Accuracy indicators for shoebox tests

PHONE	**NAV. W/**	**NAV. W/O**	**SLALOM**	**SELECT.**
STANDARD	19.6 %	16.8 %	2.5	1.09
SHOEBOX	22.7 %	14.9 %	2.4	1.27
TABLET	**NAV. W/**	**NAV. W/O**	**SLALOM**	**SELECT.**
STANDARD	18.8 %	12.8 %	2.8	0.25
SHOEBOX	12.8 %	14.9 %	2.6	0.40

5.1 Shoebox versus standard visualization

Intuitively, we would assume that **navigation** with and without obstacles is more difficult with the shoebox approach than with the standard concept because of the potential problems discussed at the beginning of section 4. In order to verify this hypothesis, we compare two performance indicators: *time to solve a task* (measured in seconds) and *accuracy* (measured in percentage of the total distance with the ball nearby the target). The latter one is a measure of how easy users can hit a target, i.e. high, if users are having problems to directly hit it, and low otherwise. T-tests were performed to verify statistical significance. Results are illustrated in Figure 11 and Table 1. No statistically significant difference could be proven for values in case of the phone. Although only slightly worse for the shoebox case, the differences in time for the “w/o obstacles” task and accuracy for the “w/ obstacles” task were statistically significant ($p<0.0163$ and $p<0.0405$). Although this confirms our hypothesis to some degree, the rather small differences were unexpected. The shoebox approach performed surprising well compared to standard visualization. A likely explanation is that the advantages of the shoebox concept (faster visibility of the target and obstacles) compensated for the expected decrease in performance that led to the original hypothesis.

In the **slalom** task, users had to navigate in one of two possible zigzag tracks around a line of obstacles to reach a target placed after the last obstacle. Accuracy in this case was measured in terms of correct paths. We expected the shoebox concept to perform better in this case, because we assumed that tilting the device would enable users to better see the gap between two obstacles what could have a positive effect on navigation performance. However, using a t-test, statistical significance of the results presented in Figure 11 and Table 1 could not be proven, so our hypothesis turned out to be wrong. A closer look into the logged sensor data from the devices revealed a low average tilting angle and tilt deviation (= change in tilt angle) for this task. This can be explained by the fact that the participants knew the location of the target and obstacles (in contrast to the navigation with obstacles where objects were placed randomly). This awareness of the optimum navigation path suggests why the participants did not take advantage of the enhanced visualization.

For the **selection** task, we were not able to specify a clear hypothesis because on the one hand, with the shoebox concept, tilting the device enables one to see a larger area and thus should make it easier (and faster) to localize targets. On the other hand, the related change in the image (cf. Fig. 5) makes the scene less stable, so objects might become harder to hit (potentially resulting in a slower and less accurate performance). The results presented in Figure 11 and Table 1 seem to confirm these thoughts: using a t-test, no significant differences could be proven, thus suggesting that the advantages compensate for the obvious disadvantages of the shoebox approach. Accuracy was measured in number of failed first attempts to correctly tap a target.

On the negative side, our results for both navigation and selection tasks show that there are no improvements in performance when the shoebox approach is applied despite the fact that a better visibility would suggest so. On the positive side, the results also did not show any decrease in performance, despite the noise introduced by the concept. Hence, there does not seem to be much benefit in using this visualization technique for tasks where a good interaction performance should have a higher priority compared to “nice” visuals (such as in many serious applications). However, for games and other entertainment-related applications, interesting and engaging visuals are important – if they do not come at the cost of making the interaction more difficult to handle, which in turn does not seem to be the case.

Hence, **user feedback and interaction experience** are important issues that should be taken into account when discussing the practical usefulness of the introduced visualization concepts. However, our experiment was set up to evaluate performance indicators. We only evaluated canonical interactions. Also, we used rather simple visuals (basic shapes and few, simple colors) in order to minimize the possible influence of the rendered scenes. Hence, subjective user feedback about how appealing and engaging certain visualization concepts are should be treated with care. Also, the data was gathered based on subjective terms and in case of the standard approach referred to both scenes (cf. Fig. 6). Thus, they are not suitable for a statistical analysis. Nevertheless, the participant’s comments provide interesting and valuable insight into how the advanced visualization concept of the shoebox will be perceived in a concrete application.

Figure 12 illustrates one rating from the questionnaire that the participants had to fill out after the experiments. For the distribution of both navigation as well as selection tasks, we can see a trend towards a more positive rating for the shoebox approach (red bars), thus indicating that users appreciated the additional possibilities they had with this concept or at least liked and enjoyed the enhanced graphics. Ratings for the standard visualization (blue bars) were rather neutral and follow a Gaussian distribution. Typical remarks described the shoebox approach as “intuitive” and “fun”, whereas positive statements about the traditional approach often contained phrases such as “familiar” and “easy/clear”. Rather negative comments described the shoebox approach as “harder to control” and classified the standard visualization as “slow” and “limited”.

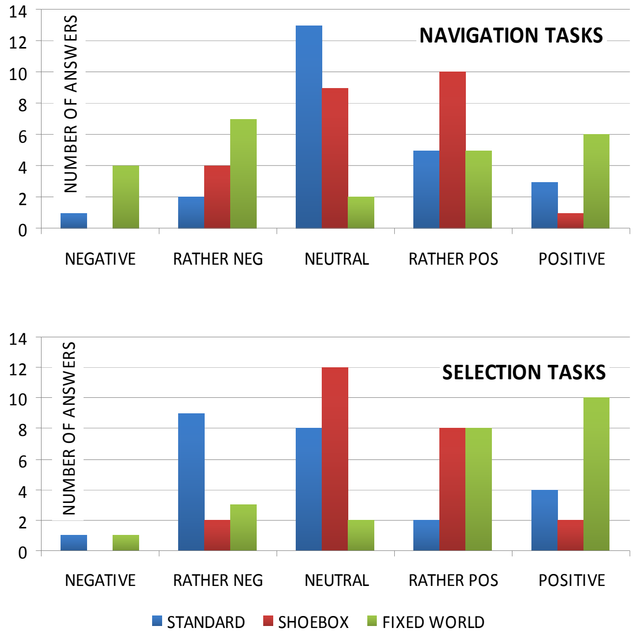


Figure 12. Questionnaire results.

5.2 Fixed world versus standard visualization

Figures 13 show the results of the performance indicators task completion time and accuracy (both measured similarly as in the shoebox experiment) for the comparison between fixed world and standard visualization in the navigation tasks. While using a t-test statistical significance of the differences in completion time could not be proven, **navigation with obstacles** showed a statistically significant difference in accuracy on the phone ($p<0.0146$) as well as the tablet ($p<0.0003$). Similarly, there are no significant differences in time for the **slalom task**, but the number of correct tracks was significantly worse for both phone and table in case of the fixed world visualization ($p<0.0026$ and $p<0.0191$). Our impression during the experiments was that participants were faster in finding a target (because they just had to move the device instead of operating the onscreen joystick) but actually hitting it was harder due to the introduced sensor noise. This subjective observation is confirmed by the lower accuracy values. It can also explain that there are no significant differences in task completion time because it suggests that a faster search time compensated for the latter disadvantage.

When considering the sensor data, several consecutive values were averaged in order to eliminate noise and sensor instabilities. Intensive filtering however can lead to a noticeable lag. Although we optimized our implementation in an informal pre-test to deal with this tradeoff, some instability of the visualization was noticeable. In addition, the compass – which was not needed for the shoebox approach but is necessary to realize the fixed world concept – sometimes adds an additional lag. These irregularities seem to be the major reason for the rather negative results for the **selection task**, where again no significant difference in timing could be observed, but accuracy (measured in failed first attempts to hit a target) decreased significantly on both phone and tablet ($p<0.0102$ and $p<0.0483$, t-test). Again, positive and negative characteristics seem to even each other out when it comes to task completion time: users were generally able to find a target faster because they just had to move the device. However, we often observed an “overshooting” of the target, requiring them to go back and do an exact positioning (which in turn was influenced negatively by the noise introduced by the small but sometimes noticeable lag due to sensor filtering).

These accuracy problems can also explain the rather distributed ratings in the subjective feedback from the participants for the fixed world concept in relation to the navigation tasks (cf. Fig. 12, green bars in top diagram). While some users made quite negative statements about it characterizing it as “hard-to-control” and “inaccurate”, others were rather positive and even expressed enthusiasm using phrases such as “fun”, “most immersive”, “innovative”, and “fast”. However, most of these positive statements were motivated by the experience with the selection task that was rated very positively (cf. Fig. 12, bottom diagram). Despite the lower accuracy, many users appreciated and enjoyed using this concept because of its intuitiveness for the given task. Most participants perceived moving the device around and using it like a window into the virtual world as very natural and intuitive for search tasks and exploring such worlds.

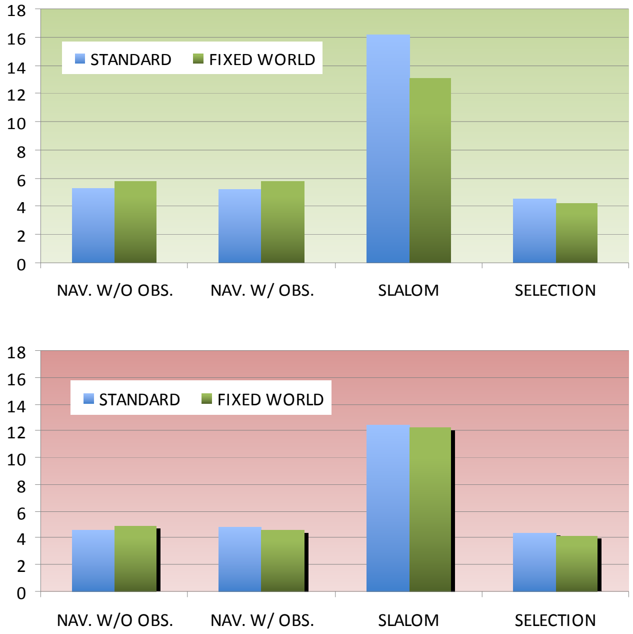


Figure 13. Task completion times (in msec) with the fixed world approach for phone (top) and tablet (bottom).

Table 2. Accuracy indicators for fixed world tests

PHONE	NAV. W/	NAV. W/O	SLALOM	SELECT.
STANDARD	19.5 %	19.4 %	2.3	0.09
FIXED W.	24.1 %	26.2 %	1.6	0.25
TABLET	**NAV. W/**	**NAV. W/O**	**SLALOM**	**SELECT.**
STANDARD	18.8 %	16.8 %	2.5	0.05
FIXED W.	21.4 %	22.6 %	2.0	0.10

5.3 Mobile phone versus tablet

The used phone and tablet feature different sensors which in turn can have an influence on the absolute values, Hence, a statistical comparison of the respective results can not be done, and the differences described below should rather be interpreted as general trends. The major difference between phone and tablet in relation to visualization concepts is of course the bigger screen size of the latter – which suggests that the interaction tasks become easier resulting in a better performance. In relation to interaction however, size and weight are also important issues to take into consideration. The first hypothesis – i.e. a better performance on the tablet – was confirmed for all tasks in our experiments. In particular, users were on average 9.5% faster on the tablet when using the standard visualization, 15% faster in case of the shoebox visualization, and 11% faster in case of the fixed world approach. Selection tasks had 60-77% less misses on the tablet, and for navigation and slalom tasks accuracy improved by 8-20% and 5-25%, respectively. Considering weight and size, we were quite surprised not to hear any complaints about the tablet being heavier and harder to hold – especially for the fixed world approach where the device has to be held up straight most of the time.

6. CONCLUSION AND FUTURE WORK

As described in Section 3, mobile devices such as phones and tablets offer exciting new possibilities for 3D graphics and virtual world visualizations because we can freely move them in space. Using sensor information, we can realize interesting visualization concepts such as the shoebox and fixed world visualization. However, these sensors are prone to noise. Also, the change of the view based on the orientation of the device adds motion to the objects what makes it questionable if these concepts can be used for applications that do not only require nice graphics but also rely on interaction with the created virtual worlds. Fortunately, our results could prove these negative expectations wrong: despite the existing problems, users were in general able to solve the given tasks successfully. Only slight decreases in performance could be observed, and for most of them a statistically significant difference could not be proven. On the negative side, our experiments also revealed obvious problems and not all interactions were perceived naturally. Sensor noise seems to be the most critical issue, especially for the fixed world approach that suffered from the lag of the compass. Better sensors, such as the gyroscope included in the iPhone, might reduce related problems, but we do not expect them to go away completely. Consequently, our results suggest rather limited usage of the advanced visualization concepts for serious applications that rely heavily on a fast, flexible, and reliable interaction. On the other hand, they offer huge potential for entertainment-related applications and 3D games where interesting and immersive visuals are very important and a slightly more difficult interaction can actually be an advantage by making the game more challenging and interesting. Based on our results, the shoebox visualization seems to be more suited for games that require intensive interaction, such as moving objects around. However, it is naturally restricted insofar as it does not allow for a 360-degree view – which is given per default very naturally by the fixed world approach. The latter in turn seems to offer high potential for games that rely more on exploring virtual worlds (e.g. search games where players have to find treasures and other objects such as keys, etc. in a tridimensional virtual fantasy world) rather than on high interactivity.

Despite the already mentioned tests with other target groups of different ages, the most obvious next steps for our future work include more complex interactions such as manipulating objects, for example, by scaling and rotating them. In addition, exploring different options on how users can navigate through those virtual worlds will offer new insight into what kind of games we can create with these visualization concepts and where the limitations are. Finally, using our idea of describing the different visualization concepts by three coordinate systems and fixations of degrees of freedom between the coordinate axes gives us lots of new opportunities to explore various further 3D visualization concepts on mobiles.

7. ACKNOWLEDGEMENTS

This work was partially supported by a Yahoo! Faculty Research Grant.

REFERENCES

[BKA05] Lynne Baillie, Harald Kunczier, and Hermann Anegg. Rolling, rotating and imagining in a virtual mobile world. In MobileHCI '05: Proceedings of the 7th international conference on Human computer interaction with mobile devices & services, pages 283-286, New York, NY, USA, 2005. ACM.

[BF98] Bill Buxton and George W. Fitzmaurice. Hmds, caves & chameleon: a human-centric analysis of interaction in virtual space. SIGGRAPH Comput. Graph., 32(4):69-74, 1998.

[DBC02] Joan De Boeck and Karin Coninx. Haptic Camera Manipulation: Extending the Camera in hand Metaphor. In Proceedings of Eurohaptics 2002, pages 36-40, Edinburgh, UK, July 2002.

[DJK+06] Cagatay Demiralp, Cullen D. Jackson, David B. Karelitz, Song Zhang, and David H. Laidlaw. Cave and fishtank virtual-reality displays: A qualitative and quantitative comparison. IEEE Transactions on Visualization and Computer Graphics, 12:323-330, May 2006.

[Fit93] George W. Fitzmaurice. Situated information spaces and spatially aware palmtop computers. Commun. ACM, 36:39-49, July 1993.

[FZC93] George W. Fitzmaurice, Shumin Zhai, and Mark H. Chignell. Virtual reality for palmtop computers. ACM Trans. Inf. Syst., 11(3):197-218, 1993.

[GCCV08] Paul Gilbertson, Paul Coulton, Fadi Chehimi, and Tamas Vajk. Using “tilt” as an interface to control “no-button” 3-d mobile games. Comput. Entertain., 6:38:1-38:13, November 2008.

[Han97] Chris Hand. A survey of 3d interaction techniques. Computer Graphics Forum, 16(5):269-281, 1997.

[Hen07] Anders Henrysson. Bringing augmented reality to mobile phones. Department of Science and Technology, Linköpings universitet, Norrköping, 2007. Diss.

[Hod09] Eli Hodapp. 'SurrounDEAD' 3GS Compass-Friendly Zombie Invasion. http://toucharcade.com/2009/07/14/surroundead-3gs-compass-friendly-zombie-invasion/, July 2009.

[HPSH00] Ken Hinckley, Jeff Pierce, Mike Sinclair, and Eric Horvitz. Sensing techniques for mobile interaction. In UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology, pages 91-100, New York, NY, USA, 2000. ACM.

[MWW06] Sumit Mehra, Peter Werkhoven, and Marcel Worring. Navigating on handheld displays: Dynamic versus static peephole navigation. ACM Trans. Comput.-Hum. Interact., 13(4):448-457, 2006.

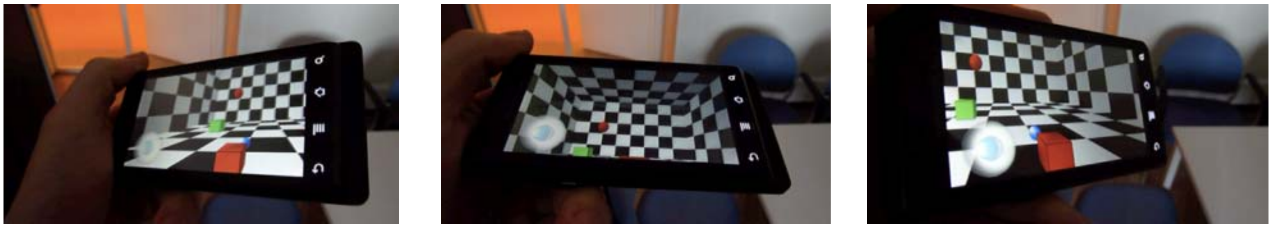
[Tar09] TarvainenProjects. Head Tracking for Smartphone VR Displays Using the Secondary Camera. http://www.youtube.com/watch?v=rczN3_G08kE, March 2009.

[Vis09] Filip Visnjic. 10 Creative Ways to Use the Accelerometer. http://www.creativeapplications.net/iphone/10-creative-ways-to-use-the-accelerometer-iphone/, January 2009.

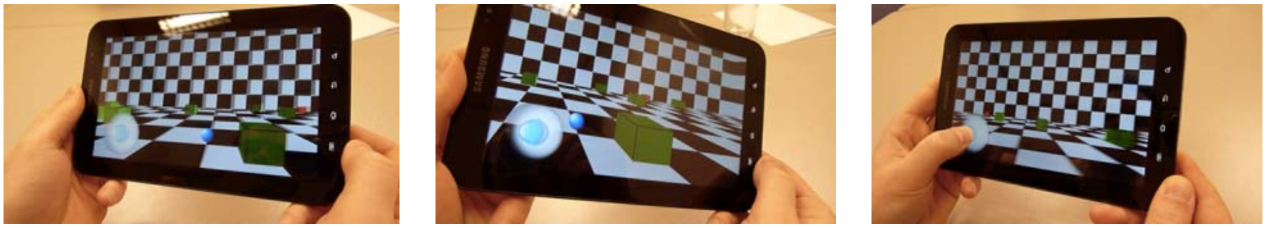
[WO90] Colin Ware and Steven Osborne. Exploration and virtual camera control in virtual three dimensional environments. In Proceedings of the 1990 symposium on Interactive 3D graphics, I3D '90, pages 175-183, New York, NY, USA, 1990. ACM.

[Yee03] Ka-Ping Yee. Peephole displays: pen interaction on spatially aware handheld computers. In CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems, pages 1-8, New York, NY, USA, 2003. ACM.

[Zic09] Ziconic. Aircoaster. http://www.ziconic.com/p/aircoaster.html, June 2009.



Shoebox approach (notice the effect of looking into a box when the device is tilted)



Shoebox; evaluation (left & center: w/obstacles – notice the change of visibility of the target when tilting; right: slalom task)

Shoebox; evaluation (selection task – notice that the target is only visible when the device is tilted; cf. Fig. 6)

Fixed world approach (notice that the red sphere stays at roughly the same position with respect to the real environment; slight changes in position are due to sensor noise and the bigger field of view of the scene)

Figure 10. Screenshots of our implementation with the scenes used in the evaluation: background = checkerboard pattern, moving object = blue ball, obstacles = green cubes, target for navigation and slalom tasks = red cube, target for selection tasks = red sphere. The onscreen joystick is visible in the lower left corner of the screen.