

# Interacting with 3D Content on Stereoscopic Displays

Florian Daiber, Marco Speicher, Sven Gehring, Markus Löchtefeld, Antonio Krüger  
German Research Institute for Artificial Intelligence (DFKI)  
Campus D3 2, Saarbrücken, Germany  
{florian.daiber, marco.speicher, sven.gehring, markus.loechtefeld, krueger}@dfki.de

## ABSTRACT

Along with the number of pervasive displays in urban environments, recent advances in technology allow to display three-dimensional (3D) content on these displays. However, current input techniques for pervasive displays usually focus on interaction with two-dimensional (2D) data. To enable interaction with 3D content on pervasive displays, we need to adapt existing and create novel interaction techniques. In this paper we investigate remote interaction with 3D content on pervasive displays. We introduce and evaluate four 3D travel techniques that rely on well established interaction metaphors and either use a mobile device or depth tracking as spatial input. Our study on a large-scale stereoscopic display shows that the physical travel techniques (whole-body gestures) outperformed the virtual (mobile touch) techniques with respect to task performance time and error rate.

## Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: Input devices and strategies, Interaction styles

## Keywords

Spatial interaction; gestural interaction; mobile interaction; 3D travel; large displays; media facades

## 1. INTRODUCTION

Along with the increasing ubiquity of public displays, stereoscopic display technologies recently became available to the mass market. Although 3D interaction has been well-studied in a wide range of settings, interaction with stereoscopic content on large-scale public displays has rarely been covered. The most common modality to interact with the content of a public display is direct manipulation (e.g. touch input). However, such techniques are only partially applicable to 3D content due to parallax effects. Current approaches, as they are for example provided by virtual reality (VR) systems, usually consist of stereoscopic projection and external input

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

PerDis '14, June 03 - 04 2014, Copenhagen, Denmark

Copyright 2014 ACM 978-1-4503-2952-1/14/06...\$15.00.

<http://dx.doi.org/10.1145/2611009.2611024>



Figure 1: Gestural 3D navigation on a large stereoscopic display.

devices that are tracked within the potentially interactive environment. These are often expert systems with complex user interfaces and high instrumentation, which make them inappropriate for exploring stereoscopic content on pervasive displays in a public setting. With the recent rise of commodity input device that allow for pervasive tracking of the user (e.g. the Microsoft Kinect), novel promising approaches to address these issues became available.

In this work we evaluate gestural- and touch-based travel techniques to explore stereoscopic content using commodity hardware as spatial input devices. We propose four 3D travel techniques that rely on well established interaction metaphors. We evaluate and compare these techniques in a within subjects experiment, for which the participants were asked to perform a 3D virtual search task on a large-scale stereoscopic display. We expected that the (whole-body) physical interaction techniques are more effective than the (mobile) virtual techniques. To evaluate this we used task performance time and error rate as the main metrics. In addition, subjective feedback was collected. We expected that the physical demand is lower for the virtual techniques, which reflects the better overall workload of our virtual travel techniques. The results show that the virtual travel techniques outperform the physical techniques regarding physical demand and effort of the participants. However, the physical techniques were reported as less frustrating and less time-consuming.

## 2. RELATED WORK

Stereoscopic displays allow users to perceive 3D data in an intuitive and natural way [10]. However, the interaction with stereoscopic content is still a challenging task even in VR-based environments. The control of the virtual camera in 3D environments requires at least six degrees of freedom (DOF), which can be directly controlled by means of 6-DOF input devices using established metaphors, like the Scene-in-Hand, Eyeball-in-Hand and Flying Vehicle metaphors proposed by Ware and Osborne [11]. In the Scene-in-Hand technique, translations of the user's hand are directly linked to corresponding translations of the scene. In the Eyeball-in-Hand technique, the user's hand serves as a virtual camera. In a qualitative oriented study Ware and Osborne concluded that the different metaphors are best suited depending on a particular task. Steinicke et al. [10] discussed potentials and limitations for using multi-touch interfaces with multi-touch enabled devices to interact with stereoscopic content. Due to the accuracy provided by mobile multi-touch surfaces, they found them especially suited for fine-grained input.

In the last years several commodity input devices were introduced that can be used for 3D interaction. Boring et al. [2] presented three interaction concepts that provide remote control of a pointer on a display via scroll-, tilt- and move-gestures using a mobile phone. But their work only focussed on 2D interaction spaces. Liang et al. [9] investigated how mobile devices can be used as input for distant large 3D displays. In an exploratory study they asked participants to propose interactions for 3D tasks and apply their findings to a prototypical application for 3D object manipulation. Capin et al. [4] introduced a camera-based approach to navigate through virtual environments on mobile phones to estimate the device motion by analyzing the live input images from the camera of a mobile phone. Benzina et al. [1] investigated phone-based motion control for travel in VR and in particular the reduction of DOF and mapping between user action and mobile device. Kratz and Rohs [8] extended the virtual trackball metaphor to rear touch input. They evaluated their input techniques in a 3D rotation task. The navigation techniques investigated in this paper, in particular the tilt metaphor, were evaluated in an extended version of their approach. But instead of a simple rotation task, a 3D travel task was performed. The introduction of the Nintendo Wiimote brought spatial interaction not only to the mass market but also stimulated research in human computer interaction and 3D user interfaces (c.f. [12]). Since then many research projects have investigated spatial interaction using commodity hardware as an affordable tracking solution.

Interaction with public displays is an emerging research field. Jurmu et al. [7] explored mid-air gestural interactions with public displays. In a descriptive field study several interaction issues were identified namely problems with mid-air gestures, explicit vs. implicit interaction and opt-out from the interaction stream. Diniz et al. [5] discuss the emergence and need of principles and guidelines for the design of interaction spaces based on media facades as large public interactive spaces. As new architectural creations media facades are designed and built all over the world. These interactive facades are not restricted to 2D surfaces any more and thus new challenges for spatial interactions arise.

Based on related work, we designed four 3D travel techniques for the interaction with large stereoscopic displays. We evaluated them in a comparative user study using commodity hardware as spatial input. Our approach addresses the interaction dimension of Diniz et al. by explicit spatial interactions with large displays. In contrast to the more descriptive work of Jurmu et al. our paper complements public display interaction with a more quantitative approach.

## 3. 3D TRAVEL TECHNIQUES

Travel tasks are one of the most fundamental human tasks in our physical environment as well as universal interaction tasks in virtual environments (e.g. navigation in the world wide web, in layers of a spreadsheet or in a virtual world of a computer game). Travel is the task of performing the action that move the viewport from the current to a target location [3]. Once a goal is formulated, muscles are triggered by the brain to perform correct movements in order to achieve the goal. Turning a wheel, press a pedal or flipping a switch are examples for interfaces mapping various physical movements. In contrast to the real world in virtual environments simple physical motions are only effective in limited space and speed which results in more or less natural mappings of the actions.

Ware and Osborne [11] describe two fundamental different constraints for user interface metaphors: cognitive and physical constraint. According to this, the proposed travel techniques can be classified as physical or virtual. First, we present two physical techniques using a depth camera where the user physically changes the viewport with his whole body including both hands. Second, two virtual travel techniques are introduced where the user controls the virtual camera using a mobile device. Moreover, all travel techniques are active, i.e. users directly control viewport movement and orientation. In the following we describe the interaction metaphors regarding their classification and relate them to the corresponding classic travel techniques.

### 3.1 Physical: Bi-manual Grabbing (BMG)

The physical input technique *Bi-manual Grabbing (BMG)* is a manual manipulation technique that is inspired by the Grabbing-the-Air travel technique [3]. The user performs a grabbing gesture to initiate the travel interaction and moves his hand in order to move the viewport. The human hand is a remarkable device which is very useful to manipulate physical objects quickly, precisely and with little conscious attention. Therefore we have chosen this technique, which combines the Grabbing-the-Air technique (see Figure 2.2) and the Camera-in-Hand technique (see Figure 2.1). Grabbing-the-Air is performed with the non-dominant hand (NDH) for the translation (x,y,z) of the viewport. Camera-in-Hand uses the dominant-hand (DH) to orientate (yaw, pitch, roll) the scene viewport. In this scenario each of both hands controls 3-DOF respectively resulting in a simultaneous 6-DOF input. The bi-manual interaction approach allows an intuitive and flexible control, i.e. the user can look around while moving the viewport. The high sensibility of the tracking device can increase the user's effort for precise movements in small areas in contrast to travel larger distances. However, using this input technique might result in a high physical demand because both hands need to be held in the air while performing the interaction.

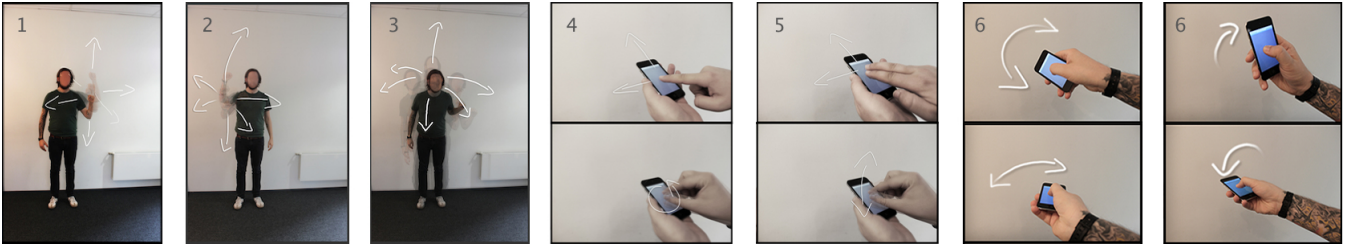


Figure 2: 3D travel techniques: Camera-in-Hand metaphor (1), Grabbing-the-Air metaphor (2), Whole-body tilt (3), Mobile one-finger pan and two-finger (4), rotate Mobile two-finger pan and pinch (5), Mobile tilt (6)

### 3.2 Physical: Whole-body Tilt & Grab (WTG)

The design of the *Whole-body Tilt & Grab (WTG)* technique is inspired by the control of a Segway vehicle. Leaning and bending the head and torso in combination with a grab gesture of the NDH results in moving the viewport continuously in the desired direction. This can be seen as a variant of the semi-automated steering technique. The movement speed can be controlled by the leaning angle (see Figure 2.3). The user controls the viewport’s orientation the same way as in the previously mentioned *BMG* technique, again based on the Camera-in-Hand technique by the user’s DH (see Figure 2.1). Due to the continuous movement, the user can focus on controlling the viewport’s orientation which results in a full 6-DOF interaction. Although this technique has a very high physical demand referable to the leaning and bending motions (particularly compared with the mobile variants), here the user can travel longer distances without high effort.

### 3.3 Virtual: Mobile Multi-touch (MMT)

The *Mobile Multi-touch (MMT)* technique is a combination of multi-touch gestures with two fingers for movement and one finger for camera orientation. The user moves the viewport in its horizontal and vertical axis by panning on the touchable surface of the mobile device with his two fingers and use the pinch gesture to move forward and backward (see Figure 2.5). In order to control the 3D camera orientation the user can pan with one finger in combination with the rotate gesture (see Figure 2.4). In summary the 6-DOF are composed by the mobile device panning gestures for translation (x,y), the pinch (z) gesture, the rotation (yaw, pitch) and the rotate (roll) gesture.

### 3.4 Virtual: Mobile Tilt & Touch (MTT)

The camera orientation control in the *Mobile Tilt & Touch (MTT)* technique is analogous to the orientation control in the *MMT* technique while the movement is controlled by using the gyroscope. Tilting the mobile device is comparable to the leaning and bending of the *WTG* technique. The interaction is toggled by the user’s finger touching the surface of the mobile device. Then the movement direction and speed can be directly controlled by changing the tilting angle of the mobile device (see Figure 2.6).

## 4. USER STUDY

We conducted an experimental study in order to evaluate our techniques in a 3D search task on a large stereoscopic display. We designed a parametrized search task that provides a flexible and easy to control experimental setup.

### 4.1 Participants

Ten participants (9 male, 1 female) from the university environment volunteered in the user study. The participants were aged between 20 and 35 years ( $M = 27.8, SD = 4.13$ ). They all owned a touch-enabled smart phone, but had no or only few experience with depth cameras, especially Microsoft Kinect. Only three of the participants had prior experience with stereoscopic 3D applications and 3D user interfaces. According to the participants’ self-assessment their experience level of 3D modeling ranged from 1 to 9 on a scale from 1 to 10 ( $M = 5.4, SD = 2.55$ ). Furthermore, the experience in computer graphics ranged from 2 to 9 ( $M = 5.9, SD = 2.47$ ). The participants received no monetary compensation for the participation in the study.

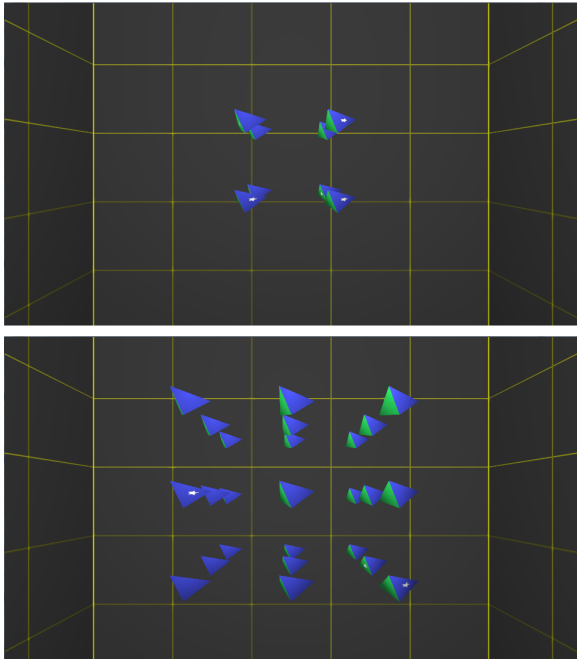
### 4.2 Task

We extended the experimental setup of Kratz and Rohs [8] who investigated 3D object rotation on a mobile device using a front and rear touch virtual trackball as well as tilt. According to previous work each of the four faces of one tetrahedron object was colored in a distinct color to allow the participants to remember the sides of the objects and give orientation in the scene. The experimenter was able to change each parameter, e.g. grid size or number of textured object faces, during the experiment remotely as well as starting and stopping the trials. The objects were not randomly chosen and the number of objects could be defined programmatically which enabled the experimenter to parametrize precisely the characteristics of the experiment. This approach provides a good control of the experiment conditions and thus even allows a reasonable way to compare future travel techniques.

Each travel task starts with an exploration task followed by a search task. We chose an introducing exploration task without an explicit goal for movement in order to browse the environment and obtain information about objects, orient the user to the world and build up spatial knowledge. Besides that, in this training phase the user is able to get familiar with the travel techniques. After that, the user is asked to perform the actual travel task, or more specifically a primed search task.

### 4.3 Design

The experiment had a  $4 \times 2 \times 2$  within subjects factorial design. Factors were interaction technique for navigation control (*BMG*, *WTG*, *MMT* and *MTT*), grid size (small:  $2 \times 2 \times 2$  and big:  $3 \times 3 \times 3$ ) as well as textured face count (easy:  $3 \pm 1$  and difficult:  $7 \pm 1$ ). The textured face count was



**Figure 3: Screenshots from the user’s perspective to the 3D scene consisting of a 2x2x2 grid of tetrahedrons (top) and a 3x3x3 grid (bottom).**

randomly chosen in a  $\pm 1$  range around 3 and 7 in order to prevent the participants from inferring the correct number of textured faces and forcing them to really count all textured faces in the scene presented to them. According to a Latin Square design, the order of our four chosen input techniques was counterbalanced, as well as the order of grid size and textured face count settings. The trials for each input technique were conducted in sequence followed by a short break of two minutes before starting a new trial sequence. Each setting results in a total of  $10 \times 4 \times 2 \times 2 = 160$  trials conducted.

#### 4.4 Procedure

After the participants filled out a short questionnaire to gather personal details (age, gender, etc.) and information about their level of experience with 3D graphics and computer science, they were placed in front of a  $5 \times 3$  meters projection wall at a distance of 2.5 meters during the trials. The freely navigable scene comprised a regular 3D grid of tetrahedrons where the goal of each trial was to count the number of object faces textured with a white star logo. Each task started with an exploration task in a scene with a  $3 \times 3 \times 1$  grid of tetrahedrons without any textured faces in order to let the participants focus on exploration of the world and acclimatization with the input technique. Only minimal instruction was given to the participants in how to use the input devices and thus perform the interaction techniques. Then they could test the techniques by travelling through the scene to get familiar with the device.

After the participants felt comfortable with the training task and decided to start the task the experimenter initiated a trial by determining grid size and number of textured faces

corresponding to Latin Square design. A trial was completed after the participant reported the number of found textured object faces to the experimenter. The trial completion time and the number of found textured faces were recorded for each trial. After each sequence of trials for each input technique, the participants were asked to subjectively rate the workload of the currently passed input technique using the NASA TLX [6] rating scale.

#### 4.5 Improvement to Existing Methodology

The experimental setup is well suited to evaluate 3D travel techniques. The subjects need to look at all faces of each object. In order to reach this goal they need to move and change the orientation to appropriate viewpoints. In contrast to previous work, we investigate a 3D travel task instead of a rotation task. Thus, we extended the setup with a 3D grid of tetrahedron objects. This allows a free and easy navigation control within a reasonable testbed environment. In previous work colored faces are introduced to support orientation. We extended this by adding light on top of the scene and placing the tetrahedron grid in center of a virtual cube with grid pattern textures on the inner walls. This extension of the setup is intended to amplify the user’s immersion.

#### 4.6 Apparatus

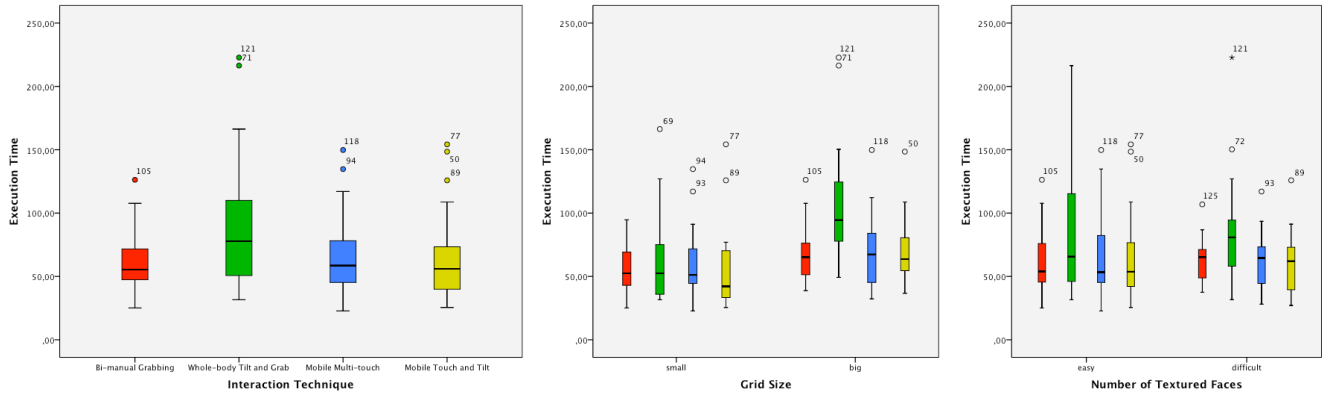
The system was an Intel Core i5 4x 3.20GHz with 8GB of RAM, and an NVIDIA GeForce GTX 660 Ti. The software running on Windows and was written in C++ and DirectX. A large back-projected wall for polarized stereoscopic display with a size of  $4.45 \times 2.8$  meters (diagonal: 5.26 meters) and a full HD resolution was used. User input was performed with a Microsoft Kinect depth camera and an Apple iPod Touch 4G for touch and orientation tracking. We used the Kinect for Windows SDK 1.7 and the associated Kinect for Windows Toolkit 1.7 for the skeleton tracking (head and hand positions) and for the reading of the hand states (grab gesture) from the interaction stream. On the iPod (iOS 6.1.3) input data from the touch-enabled surface and sensor data from the gyroscope was used to enable mobile gestures.

### 5. RESULTS

In the following we present the results of the experiment with respect to interaction technique (Bi-manual Grabbing: *BMG*; Whole-body Tilt & Grab: *WTG*; Mobile Multi-touch: *MMT*; Mobile Tilt & Touch: *MTT*), grid size (small or big) and number of textured faces (easy or complex) for the task completion time and error rate. Then additional subjective feedback of a NASA TLX test is reported. All figures use the same color scheme for the interaction techniques (*BMG*: red; *WTG*: green; *MMT*: blue; *MTT*: yellow).

#### 5.1 Task completion time

The results for task completion time are shown in Figure 4. The mean execution time for *BMG* was 62.24s,  $SD = 22.05$ , for *WTG* 85.08s,  $SD = 46.98$ , for *MMT* 66.88,  $SD = 28.46$  and for *MTT* 62.94,  $SD = 30.94$ . The mean completion time regarding grid size was 58.70s,  $SD = 29.25$  for small and 78.80s,  $SD = 36.39$  for big grid. The mean completion time regarding the number of textured faces was 69.35s,  $SD = 38.20$  for easy and 68.16,  $SD = 30.40$

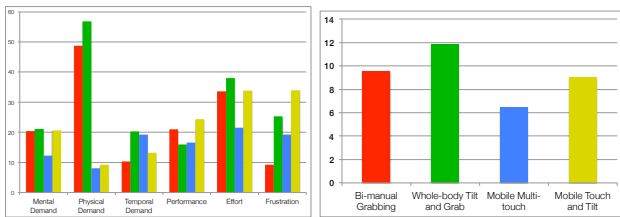


**Figure 4: Box plots of the execution time in seconds w.r.t. interaction technique (left). The box plots of execution time w.r.t. grid size (middle) and textured face count (right) are grouped by interaction technique. To see details, please zoom in.**

for difficult. An univariate ANOVA shows significant effect in the task completion time for interaction technique ( $F_{3,40} = 4.663, p < 0.05$ ) and grid size ( $F_{1,80} = 15.729, p < 0.05$ ), but no significant effect on the number of textured faces ( $F_{0,80} = .055, p = 0.815$ ). A Bonferroni pairwise comparison of interaction techniques shows a significant difference for *BMG* vs. *WTG*, *WTG* vs. *MMT* and *WTG* vs. *MTT* ( $p < 0.05$ ), but no significant difference could be found between the other techniques ( $p = 1.0$ ).

## 5.2 Error rate

The error rate, i.e. the ratio of the number of incorrect responses to the total number of responses, with respect to input method was 20.0% for *BMG*, 17.5% for *WTG*, 25.0% for *MMT* and 17.5% for *MTT*. In order to measure neutral error performance participants were not provided with feedback whether they counted the right number of textured surfaces. The responses of the physical techniques were closer to the actual numbers than the responses of the virtual techniques. This is reflected by the mean square error, i.e. the deviation of the reported count from the actual count, with 0.2 for *BMG* and 0.175 for *WTG*, and 0.475 for *MMT* and 0.4 for *MTT*.



**Figure 5: NASA TLX subscales (left) and overall workload (right). To see details, please zoom in.**

## 5.3 NASA TLX

The NASA TLX provided the following subjective results on the four input methods. Figure 5 (left) refers to results of the six NASA TLX subscales with respect to the four interaction techniques: Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Performance (OP),

Effort and Frustration (FR). The average overall workload of each interaction technique is 9.55 ( $SD = 2.75$ ) for *BMG* and 11.83 ( $SD = 2.85$ ) for *WTG*, and 6.46 ( $SD = 3.99$ ) for *MMT* and 9.00 ( $SD = 3.28$ ) for *MTT* (see Figure 5 right). In conclusion both physical input methods resulted in higher physical demand and effort but less frustration and temporal demand while the virtual input methods were performed with a very low physical demand and less effort but a high temporal demand and much more frustration.

## 6. DISCUSSION

*BMG* outperforms all other methods in task completion time and average in error rate. Regarding the remaining techniques, *MMT* performed well in task completion time but worst with respect to error rate. *WTG* has a low error rate (lowest mean square error), but bad task completion time. *MTT* was average for both metrics. In summary, the mobile interaction techniques performed average with respect to error rate, but evidently worse to mean square error. These high mean square errors lead to the conclusion that the physical techniques are superior compared to virtual techniques. The quantitative results of the experiment are in accordance with the subjective results of the NASA TLX. As expected the mid-air gestural input methods resulted in higher physical demand and effort which is a well known issue. But the good task performance time of these input methods explain the small temporal demand and low frustration level very well. Altogether, the task performance time and error rate of the mobile input methods are worse. This is also clearly reflected by the NASA TLX test that revealed low physical demand and effort but a high temporal demand and much more frustration. The *MTT* technique performs worse than the *MMT* technique regarding the task completion time but better in terms of error rate. This might be due to the fact that this technique is closely related to direct touch interaction that people already have adopted in their daily use. Although the *MTT* technique allows an effective interaction, applications that implement this technique need to be carefully designed in order to familiarize the user with the tilt interaction.

One of the most important advantages of our approach is the ease-of-use. Our goal was to keep the interaction very sim-

ple, intuitive and direct by guaranteeing that all parts in the 3D scene are easily reachable. In order to keep the frustration level at a minimum the user has free navigation control (i.e. no travel limitations in space and full control of the movement speed). Furthermore, the interaction techniques are cognitive friendly. This is because of the continuous movement, i.e. the whole environment can be reached from the current position to the desired location while the speed of camera movement gives an indication of the distance traveled. The NASA TLX resulted in low cognitive load, disregarding the physical demand of the physical techniques which is quite obviously compared to the virtual techniques and proves our design goal correct.

## 7. CONCLUSION AND OUTLOOK

In this paper we investigated travel techniques for mobile and gestural input. We proposed four 3D travel techniques that rely on well established interaction metaphors and either use a mobile device or depth tracking as spatial input. In order to evaluate our techniques we performed a user study on a large stereoscopic 3D display. Participants were asked to perform a 3D search task using all techniques in a within subjects design. Our results clearly show that the physical interaction techniques outperform the virtual technique. Although, the physical demand and effort was worse for the travel techniques based on the physical metaphor their overall performance was better.

The results of the study give implications for the design of intuitive 3D navigation techniques that might enable spatial interactions in public places. Interaction with public displays is often learned by observing other users [7]. Thus, pondering gestures might support the social learning aspect of public display interaction. This needs to be studied in the wild in future studies. One potential scenario is 3D gaming with stereoscopic public displays and media facades. The physical interaction techniques are very suitable candidates for such a scenario. The general drawback of physical demand and effort might even increase the complexity and thus the gaming experience. Another potential application is browsing rich content on public displays such as browsing a 3D movie (or music) database in an advertising scenario. The physical techniques might be inappropriate for exploring a movie database. On the other hand the remote control metaphor is a well known concept and thus the mobile interaction techniques might be better for this kind of applications. However, an appropriate mobile interaction technique needs to be carefully designed.

In future work we will investigate 3D travel with commodity tracking devices more in detail. Other interaction metaphors can be adapted to these input devices. We especially want to focus on the implications that such new interaction metaphors create in a public setting. In particular the physical travel metaphors need to be carefully tested in such settings and we therefore want to conduct an in-the-wild study to harvest an insight on audience responses to such techniques.

## 8. ACKNOWLEDGEMENTS

This research project is partially supported by the Deutsche Forschungsgemeinschaft (DFG KR3319/8-1).

## 9. REFERENCES

- [1] Benzina, A., Dey, A., Tonnis, M., and Klinker, G. Empirical Evaluation of Mapping Functions for Navigation in Virtual Environments Using Phones with Integrated Sensors. *International Journal of Innovative Computing, Information and Control (IJICIC)* 9, 12 (Dec. 2013), 4693–4709.
- [2] Boring, S., Jurmu, M., and Butz, A. Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays. In *Proceedings of the Annual Conference of the Australian Computer-Human Interaction, OZCHI '09*, ACM (2009), 161–168.
- [3] Bowman, D. A., Kruijff, E., LaViola, J. J., and Poupyrev, I. *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., 2004.
- [4] Capin, T., Haro, A., Setlur, V., and Wilkinson, S. Camera-based virtual environment interaction on mobile devices. In *Lecture Notes in Computer Science 4263*, 765 (2006), 773.
- [5] Diniz, N. V., Duarte, C. A., and Guimarães, N. M. Mapping interaction onto media facades. In *Proceedings of the International Symposium on Pervasive Displays, PerDis '12*, ACM (2012), 14:1–14:6.
- [6] Hart, S. G., and Stavenland, L. E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds. Elsevier, 1988, ch. 7, 139–183.
- [7] Jurmu, M., Ogawa, M., Boring, S., Riekk, J., and Tokuda, H. Waving to a touch interface: Descriptive field study of a multipurpose multimodal public display. In *Proceedings of the International Symposium on Pervasive Displays, PerDis '13*, ACM (2013), 7–12.
- [8] Kratz, S., and Rohs, M. Extending the virtual trackball metaphor to rear touch input. In *Proceedings of the Symposium on 3D User Interfaces, 3DUI '10*, IEEE (2010), 111–114.
- [9] Liang, H.-N., Williams, C., Semegen, M., Stuerzlinger, W., and Irani, P. An Investigation of Suitable Interactions for 3D Manipulation of Distant Objects through a Mobile Device. *International Journal of Innovative Computing, Information and Control (IJICIC)* 9, 12 (Dec. 2013), 4737–4752.
- [10] Steinicke, F., Hinrichs, K. H., Schöning, J., and Krüger, A. Multi-touching 3d data: Towards direct interaction in stereoscopic display environments coupled with mobile devices. In *Advanced Visual Interfaces (AVI) Workshop on Designing Multi-Touch Interaction Techniques for Coupled Public and Private Displays* (2008), 46–49.
- [11] Ware, C., and Osborne, S. Exploration and virtual camera control in virtual three dimensional environments. In *Proceedings of the Symposium on Interactive 3D Graphics, I3D '90*, ACM (1990), 175–183.
- [12] Wingrave, C., Williamson, B., Varcholik, P. D., Rose, J., Miller, A., Charbonneau, E., Bott, J., and LaViola, J. The wimote and beyond: Spatially convenient devices for 3d user interfaces. *IEEE Computer Graphics and Applications* 30, 2 (Mar. 2010), 71–85.