# Hands-Free Navigation in Immersive Environments for the Evaluation of the Effectiveness of Indoor Navigation Systems

Volker Settgast*†, Marcel Lancelle⋆, Dietmar Bauer‡, Dieter Fellner*†⋄

*Fraunhofer Austria Research GmbH,
Visual Computing
Inffeldgasse 16c
8010 Graz, Austria
Tel.: +43 (0)316 / 873 5410
E-Mail: volker.settgast@fraunhofer.at

†Institute of Computer Graphics
and Knowledge Visualization
TU Graz, Austria

‡AIT Austrian Institute of Technology,
Mobility Department
DTS, Austria

⋆Fraunhofer IDM@NTU,
Nanyang Technological University
Singapore

⋄Fraunhofer Institute for
Computer Graphics Research
Darmstadt

**Abstract:** While navigation systems for cars are in widespread use, only recently, indoor navigation systems based on smartphone apps became technically feasible. Hence tools in order to plan and evaluate particular designs of information provision are needed. Since tests in real infrastructures are costly and environmental conditions cannot be held constant, one must resort to virtual infrastructures. In this paper we present a hands-free navigation in such virtual worlds using the Microsoft Kinect in our four-sided Definitely Affordable Virtual Environment (DAVE). We designed and implemented navigation controls using the user's gestures and postures as the input to the controls. The installation of expensive and bulky hardware like treadmills is avoided while still giving the user a good impression of the distance she has travelled in virtual space. An advantage in comparison to approaches using head mounted augmented reality is that the DAVE allows the users to interact with their smartphone. Thus the effects of different indoor navigation systems can be evaluated already in the planning phase using the resulting system.

**Keywords:** Navigation systems, smartphone app, route guidance, virtual environments

## 1 Introduction

Transportation hubs such as airports, train stations and other junctions of mass public transport become increasingly complex, potentially posing obstacles for the orientation of passengers. In particular, older people as well as people with mobility restrictions rely on the timely and effective provision of information in order not to get lost or at least incur unnecessary delays. They also have special requirements that typically are not in the main focus of route guidance.

Increasingly, this information provision is accomplished using indoor navigation systems realized using smartphones to complement the static guidance systems. A testimony in this respect is the inclusion of indoor navigation in Google Maps 6.0 [Vol12]. While the use of smart phones for indoor navigation is technically feasible today there are no tools at present to test prior to implementation whether the suggested system is appropriate or to compare two different systems with respect to their effects.

In the planning phase virtual infrastructures are necessary to this end through which the test persons have to navigate in order to measure the effects of different navigation aids. In order to realize such immersive environments there are two options: Kretz et al. used telepresence conveyed via a head mounted display [KHR$^+$11]. The test person walked normally and her motion was fed back to the virtual environment. In such a setting smartphone usage cannot be incorporated realistically. The second option is to use a CAVE environment which attains the immersion via shutter glasses inside a room onto whose walls the virtual infrastructure is projected upon. Interaction with smart phones works as in reality as the smartphone can be carried and the shutter glasses are no obstacle for natural vision of the smartphone in this environment. However, in this setting the question arises how the test persons navigate in the model, as the scene is a lot larger than the CAVE and natural walking is limited to a short distance. Thus, navigation is often accomplished using a joystick or a similar pointing device. However, such a handheld device is not desirable when test persons need to interact with the smartphone app or when in a travel setting they potentially carry a suitcase, a ticket or other items.

It is known that the degree of immersion increases with the level of physical effort necessary to navigate [SUS95]. Therefore, in this paper a method using the Microsoft Kinect for the navigation based on simple and intuitive walking related gestures is described. To perform the gestures, the test persons walk on the spot and hence allow natural interaction with both the static as well as the mobile guidance systems in a realistic fashion. This allows the evaluation of effectiveness of the indoor navigation system in the planning phase. We decided to implement a navigation control with walking in place as a gesture for moving forward in the virtual world. We hope to create physical stress similar to walking in real life.

## 2 Related Work

Many traveling techniques and navigation methods have been developed for immersive virtual environments. Obvious options like pointing based methods, often in combination with a joystick or a wand, were already described over two decades ago. Some of them are summarized by Ware and Osborne [WO90]. More recent overviews categorize existing travel methods like the ones from Bowman [BKH97], [BKLP04] and Mine [Min95].

Mechanical locomotion interfaces can be used, such as 1D or 2D treadmills or large hollow rotating spheres the user walks in. Among others, Iwata et al. developed several innovative locomotion interfaces like treadmills or moving tiles [IYFN05]. CyberSphere [FRE03] and Cyberwalk [STU07] are special platforms that allow the user to walk within a virtual environment. While the former system uses the rotating sphere, the latter one employs balls which are actuated by a belt on a turntable. However, these locomotion systems require a huge mechanical effort and in practice are often more difficult to use than one would expect. Eventually they still cannot reduce simulator sickness problems because the real physical and virtual visually perceived accelerations do not match.

Bourdot et al. use a stateless approach where different zones for the user's position are used for special behavior of the navigation [BDA99]. Leeb et al. describe a simple VR navigation with a brain computer interface by measuring neural impulses with an electroencephalogram [LSFP07]. However, the setup time is long, navigation is very limited and lengthy and physical motion is also strongly restricted.

LaViola et al. describe multiple hands-free techniques for multi scale ground based walk navigation and also address the problem of a missing back wall in a CAVE. By amplifying the mapping of the user's orientation, 360 degree views become possible [LFKZ01]. They also introduce a pair of special slippers for the navigation task. Their setup uses a magnetic tracking system and the user is required to wear a belt for the tracking of the waist in addition to head tracking.

Adamo-Villani et al. developed and evaluated a travel interface using stepping on a dance mat [AVD07]. Beckhaus et al. also used a dance mat and developed a chair interface for traveling [BBH05]. This hands-free method of navigation is not very natural and the user is distracted too much by choosing the correct floor button for traveling actions. These techniques do not cover the experiences of a physical movement with the possibility to precisely estimate the travelled distance.

Recent works describe experiments with the Microsoft Kinect as Natural User Interface for CAVEs. For example Jung et al. use the Kinect in combination with a Nintendo Wii controller for traveling in virtual worlds [JKS11]. Other techniques use hand and arm gestures for navigation and traveling tasks.

None of these approaches realize a true hands-free navigation that is intuitive, allows interaction with a smartphone and the carrying of suitcases and the like while at the same time allowing to estimate the travelled distance realistically. Therefore a new approach has been developed which is presented in the following sections.

## 3 The Setup

### 3.1 The DAVE

The Definitely Affordable Virtual Environment (DAVE) is an immersive projection room with three side walls and one floor projection [FHH03, LSF08]. The projection screens are 3.3 meters wide and 2.7 meters high (see Figure 1). Stereoscopic shutter glasses are used, similar to the ones also known from 3D TV sets or 3D cinemas. In addition, an optical head tracking system allows a correct parallax and an undistorted view for the main user. The user can simply walk around an object to see it from all sides. A big advantage compared to most head mounted displays is the very wide field of view. Such a CAVE provides the most visually convincing immersive experience.

However, natural walking is very limited due to the small room size and no haptic feedback is available. In order to explore a larger 3D world, navigation or so called travel techniques as mentioned in the last section are necessary. By mostly using standard hard-

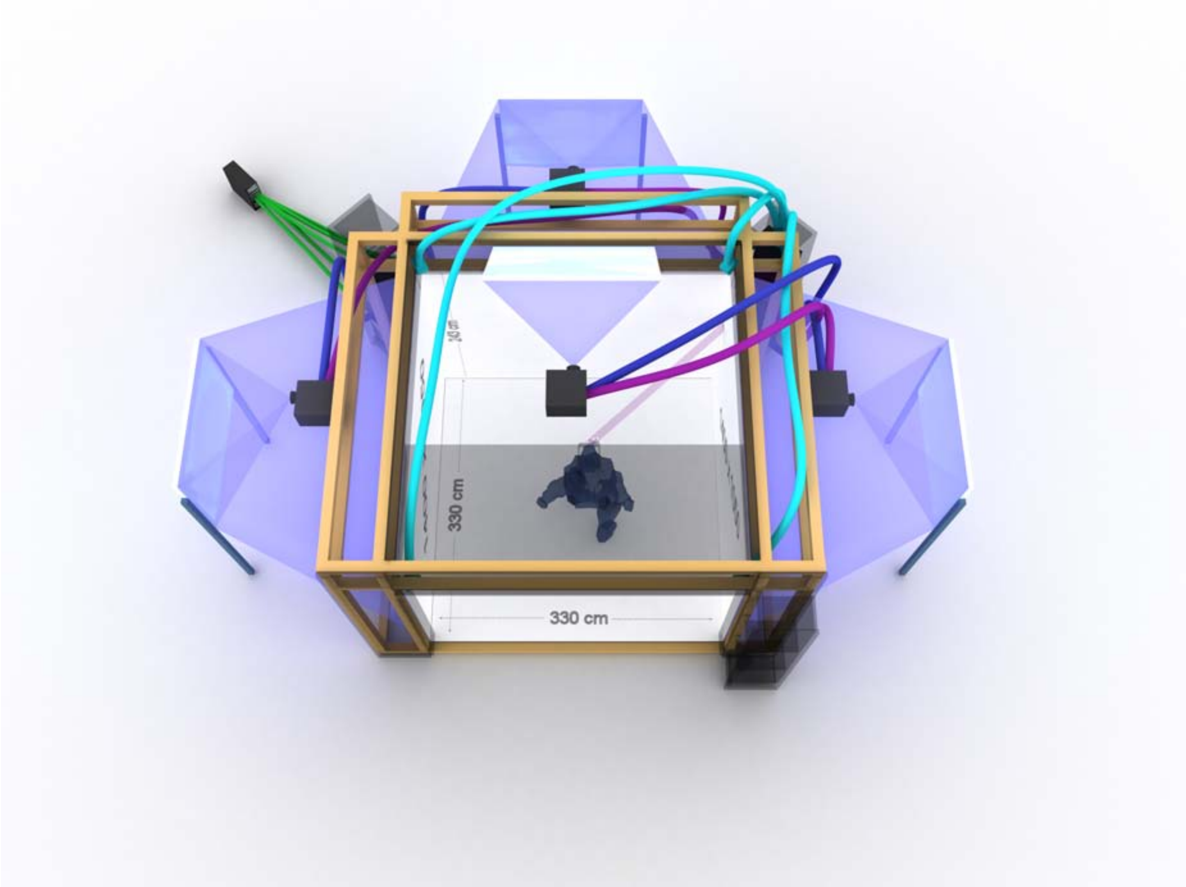


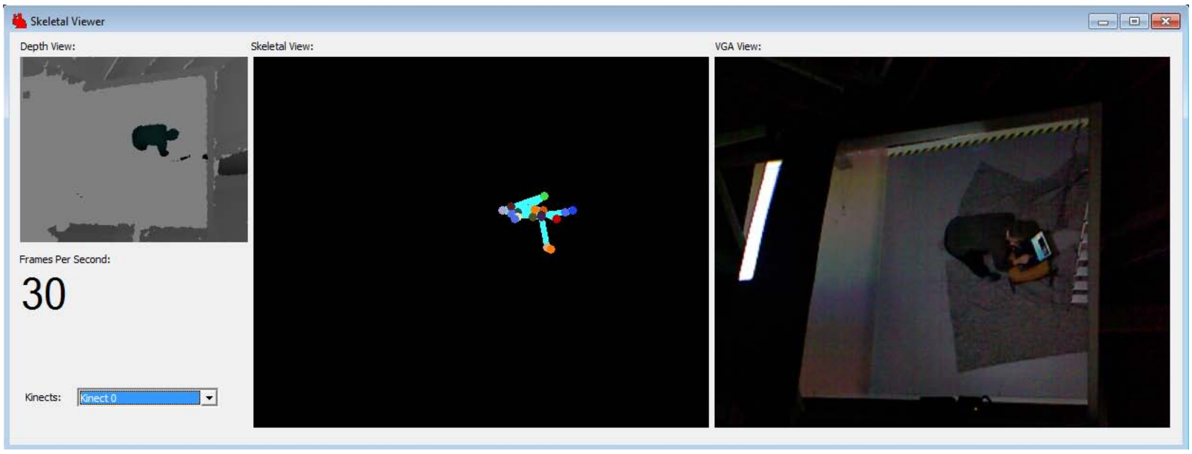
Figure 1: This is the *Definitely Affordable Virtual Environment* (DAVE). Large mirrors are used to minimize the required space for the setup. Schematically: tracking and cameras (turquois), projector synchronization (magenta), video signal (blue) and network (green).

ware components the system as well as its upgrades over time are cost effective. Large mirrors are used to fold the light paths from the projectors to the screens in order to minimize the necessary room size. In order to track objects or the user's head, multiple reflective markers are rigidly attached to objects or the glasses. These passive markers are detected by multiple cameras and their position is computed by triangulation. Since all markers look the same identification is only possible with heuristic estimations or a fixed constellation of markers that we call target. At least three markers are necessary to compute all six degrees of freedom of an object. In our current setup the cameras are attached above the screen and the limited field of view also restricts the possible tracking volume. It is mainly used and optimized for head and hand tracking and the performance is limited for objects close to the floor such as for foot tracking. In addition, the rather low power infrared lighting setup prevents marker detection during fast motions. With one target attached to the glasses the system determines the position and orientation of the main users head. From this information it is possible to estimate the position of the users eyes. A dynamic asymmetric view frustum is used to provide undistorted stereoscopic imagery to the main user.

### 3.2 Installation of the Kinect Sensor

The installation of the Kinect Sensor was difficult because of the spatial restrictions in the DAVE. Three positions are plausible:

- At the ceiling, similar to the floor projection using a mirror
- At the back opening of the DAVE
- Above the front wall



Frames Per Second:
30

Kinects: Kinect 0

Figure 2: Kinect placed at the ceiling of the DAVE

All of those positions were tested for suitability. We evaluated the resulting data of the Kinect with a test application SkeletalViewer which is part of Microsofts software development kit to visualize recognized skeletons. The application also shows the camera image and depth map. The position at the ceiling, next to floor projector means for the Kinect to look through a mirror. In principle that is not a problem, but the viewing angle is much too steep for the recognition of the user standing in the DAVE (see Figure 2). The test application was not able to reconstruct a skeleton for that angle.

Placing the Kinect at the back opening of the DAVE gives much better results for the recognition of skeletons. Users are visible from the back and the software is assuming to see them from the front. This is not an issue for the navigation control. The setup has two disadvantages. First, observers at the outside of the DAVE might occlude the test user. For user studies this would be a problem due to the spatial restrictions of the current DAVE location. Second, the recognition of arm gestures is very limited because the body of the test user occludes the arms in many situations.

In the third configuration the Kinect is placed on top of the front wall (see Figure 3). The viewing angle to the user is again rather steep, but still acceptable for the recognition of the skeleton. The test user is completely visible only if he stands about 1.5 m away from the front wall. If she moves closer, the feet and legs are outside of the viewing frustum of the Kinect Sensor. The person is never occluded by observers and also her arms are visible as long as she is facing the front wall. We choose this position for the Kinect as the most suitable.

Figure 3: Kinect placed above front wall

Even although both the Kinect Sensor as well as the optical tracking system work with infrared light, both systems work at the same time. With the chosen position, all sensors do not directly see the light sources of the other system respectively, so that they do not interfere.

# 4 Implementation

## 4.1 Kinect Application

For accessing and controlling the Kinect we use the Microsoft software development kit (SDK). To acquire the skeleton data from the Kinect and use it in our DAVE applications, we modified the test application SkeletalViewer. The application shows the camera output and depth image and also any detected skeletons. The recognized skeleton data consists of twenty 3D points for twenty joints of the human body. Arms and legs are divided into three segments, the head is one segment and the hip consists of two segments. With our modification the joints data is made available over a standard TCP/IP socket to the DAVE controller application. This is done analogous to the optical tracking system.

## 4.2 Calibration of the Kinect

The test application is working within the coordinate system of the Kinect. 3D data is sent without any modification. To match the coordinate system of the DAVE, we define a transformation matrix composed by two small rotations ( $< 45°$) and a translation. Scaling is not necessary as both coordinate systems are based on meters. To calibrate the Kinect we use the already calibrated optical tracking system of the DAVE. With the known position of the head in both systems it is easy to calculate the required transformation. The user has to stand in the middle of the DAVE with his arms spread to form a T-shape. Now the head position and the midpoint of the feet represent the up vector of the Kinect coordinate system and the elbow joints can be used to calculate a vector pointing to the right. Those two vectors are used to calculate the pitch and yaw rotation. The latter is necessary because the Kinect is not placed at the exact center of the front wall. Finally, a translation offset is calculated using the rotated head joint of the Kinect data and the head position of the optical tracking. While other methods can lead to a much more precise registration, this method is already sufficient for our purposes.

### 4.3 Test Data for Movement

To determine the correct gesture for forward movement we recorded some test data of people walking in place in the DAVE. The test person was told to walk slowly for 30 seconds and then walk faster for another 30 seconds. The joints positions were recorded and analyzed for obvious repeating patterns. The positions of the feet turned out to be too noisy to be usable. The knees positions show a more favorable pattern. Figure 4 shows the up/down motion (blue) of the knees and the forward/backward motion (green). It also shows the same data for the ankle (red and turquoise). Even for a simple sign function of the knee position relative to the mean, the resulting patterns are distinct (yellow and pink).

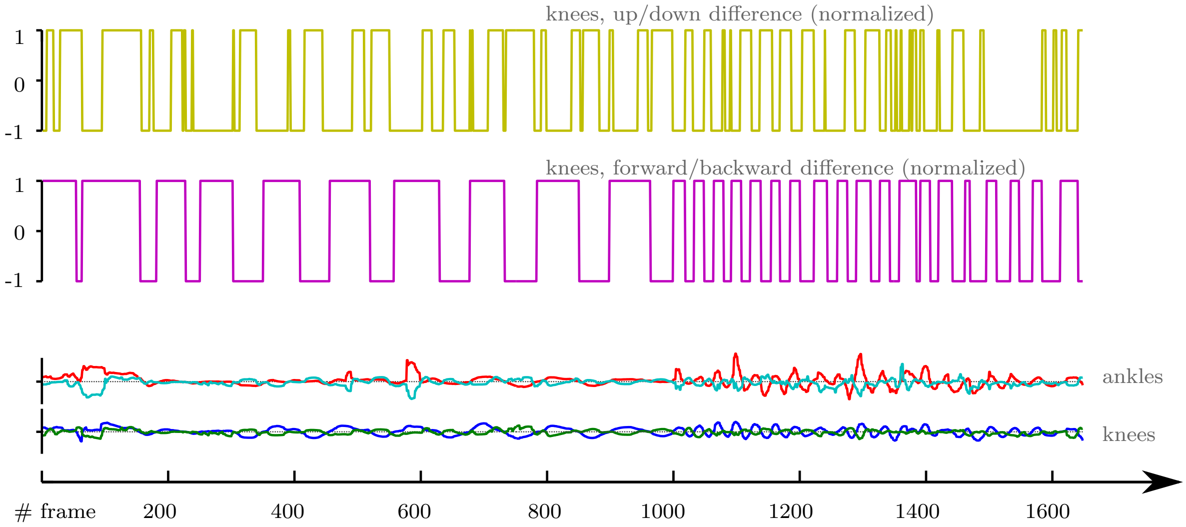


Figure 4: Selected test data of a person walking in place, recorded by the Kinect.

### 4.4 Navigation Control Module

The navigation control module consists of two functions:

1. Rotation, centered at the test person
2. Moving forward

The rotation is realized using a combination of the shoulder joints and the hip joints. Shoulder orientation weights twice the hip orientation. If the person rotates his shoulders to the left, the navigator starts rotating the world to the right. By also rotating the lower body to the left, the speed of rotation to the right is increased. This technique allows the user to freely look around without influencing navigation. Visual feedback is given in the form of an arrow on the floor pointing towards the recognized direction. As the Kinect can only recognize the skeleton of a person if the legs and arms are not occluded, the maximal allowed rotation angle is less than 45 degree. If the person is facing one of the side walls, self-occlusion prevents the Kinect to accurately determine the correct rotation angle.

To move forward in the virtual world, the user has to physically move his legs up and down. The sign change of the local distance of the knees in the direction of traveling is used to trigger impulses of movement (pink plot in Fig. 4). The height of the feet influences the amount of force for the impulse, simulating the fact that a tall person makes larger steps. To prevent a jerky movement, the forward motion does not stop abruptly but simulates a simple kind of inertia as long as the legs are moving. Otherwise, a damping actively slows down the forward motion.

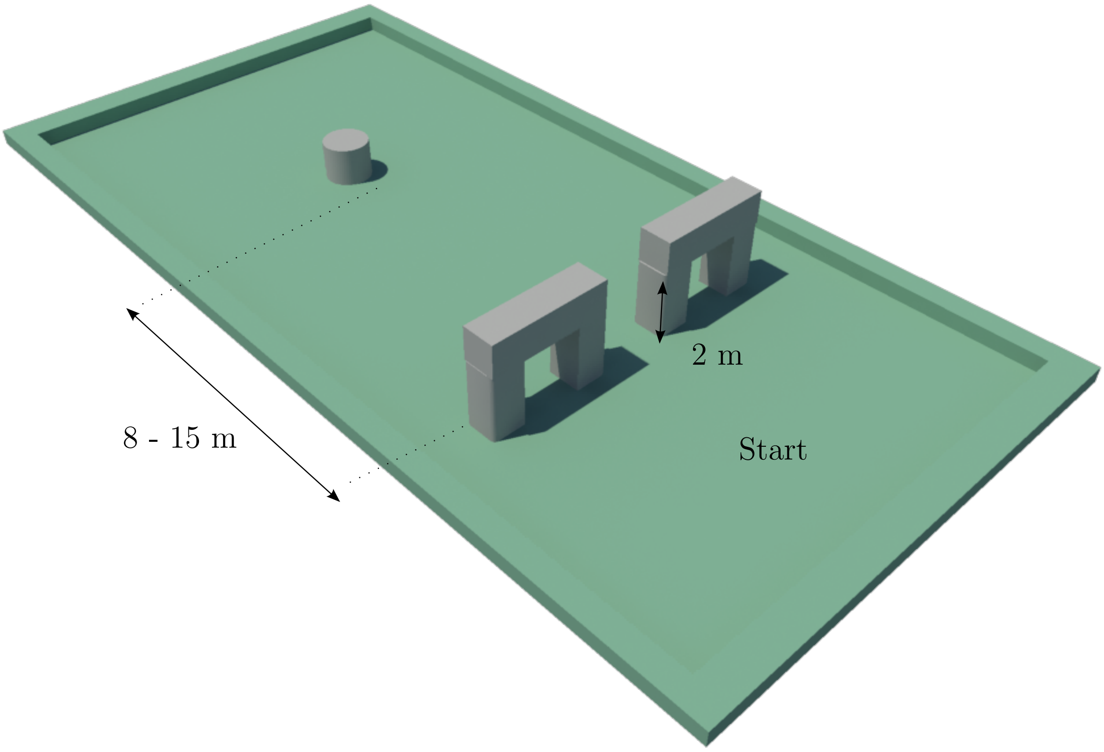


Figure 5: Simple test scenario for the pilot study. The distance of the round column is adjustable between 8 and 15 meters.

# 5 Pilot-Study

A simple scenario was created to test the navigation control module. The scenario consists of a large planar area with two archways and a column (see Figure 5). The distance of the column is adjustable between 8 and 15 meters. 14 users (4 female, 10 male) attended the pilot study. Before walking through the virtual world the test persons were asked to estimate a distance of 6 meters in the real world only by vision. The results were kept secret in order to avoid influences of the estimation in the virtual world.

The task in the DAVE was to walk through one of the archways, turn around at the column and walk back through the other archway. Seven persons used the Kinect navigation method and the other seven a pointing device. So half of the test persons had to move their legs to move forward and the other half just pressed a button. Afterwards the test persons were asked two questions:

- Did you have difficulties navigating through the archway?
- How long is the distance from the archway to the column?

In the result the estimation of the real world distance was too low: the mediate estimation was 5.19 meters with a standard deviation of 0.699. The average error was 15.6. All of the test persons had no problem navigating through the archway. The estimation of the distance to the column is shown in Table 1.

| Person | Navigation | actual distance (m) | estimated distance (m) | Difference (m) | Real world estim. of 6 m |
|---|---|---|---|---|---|
| 1 | Kinect | 15 | 12.0 | 3.0 | 4.40 |
| 2 | Kinect | 15 | 12.0 | 3.0 | 4.72 |
| 3 | Kinect | 10 | 10.0 | 0.0 | 4.50 |
| 4 | Kinect | 15 | 10.0 | 5.0 | 5.15 |
| 5 | Kinect | 15 | 7.0 | 8.0 | 5.60 |
| 6 | Kinect | 15 | 20.0 | 5.0 | 4.65 |
| 7 | Kinect | 15 | 17.5 | -2.5 | 5.15 |
| 8 | Joystick | 10 | 7.0 | 3.0 | 5.32 |
| 9 | Joystick | 15 | 13.0 | 2.0 | 4.80 |
| 10 | Joystick | 15 | 13.0 | 2.0 | 6.25 |
| 11 | Joystick | 15 | 12.0 | 3.0 | 5.10 |
| 12 | Joystick | 10 | 8.0 | 2.0 | 4.40 |
| 13 | Joystick | 10 | 13.0 | -3.0 | 6.48 |
| 14 | Joystick | 15 | 12.0 | 3.0 | 6.18 |
| standard deviation, total: 2.86, Kinect: 3.47; Joystick: 2.14 | | | | | |

Table 1: Estimated distance in meters in the virtual world.

The average error in estimating the distance is 23.1%. For the joystick navigation only it is 20.9% and for the Kinect navigation only it is 25.2%. The Kinect navigation had no relevant effect on the ability to estimate distances. Probably the test scenario was too small for the walking movement to. Most of the test persons estimated the distance by looking at the geometry and not by the walking time or physical stress. A test with a longer travel distance should be used in further studies. Nevertheless the first pilot study shows that in general the navigation task can be accomplished without holding any controller device or using the arms or hands for traveling.

## 6 Conclusion

In this paper an approach for hands-free navigation in an immersive environment has been described. Using the Microsoft Kinect in our four-sided DAVE, we designed and implemented navigation and movement controls using the user's gestures and postures. Compared to other solutions, the installation of the Kinect sensor is inexpensive and can be realized in limited space. As our technique is completely vision-based, the user does not have to learn how to use a new device. This is a more intuitive interaction and more related to real walking than commonly used approaches based on special input devices (3D joystick, cyber gloves, etc.), which compromise the user's VR presence.

The approach is already implemented and tested with respect to the perceived realism. The perception of distances and walking time will be the focus of further investigations in this respect. This navigation builds the cornerstone for the development of a test lab that can be used in order to evaluate the effectiveness of smartphone based indoor navigation systems already in the planning phase.

## Acknowledgements

This research is partly funded via the project IMITATE funded by the BMVIT under the research initiative IV2SPlus which is gratefully acknowledged. This research was partially done for Fraunhofer IDM@NTU, which is funded by the National Research Foundation (NRF) and managed through the multi-agency Interactive & Digital Media Programme Office (IDMPO) hosted by the Media Development Authority of Singapore (MDA). The authors would also like to thank Torsten Ullrich for his support.

## References

[AVD07] Nicoletta Adamo-Villani and Jones David. Travel in Smile: a study of two immersive motion control techniques. *Proceedings of 2007 International Association for Development of the Information Society (IADIS) International Conference on Computer Graphics and Visualization*, 1:43–49, 2007.

[BBH05] Steffi Beckhaus, Kristopher J. Blom, and Matthias Haringer. Intuitive, Hands-free Travel Interfaces for Virtual Environments. *Proceedings of the 2005 IEEE Conference on Virtual Reality, Workshop on New Directions in 3D User Interfaces*, 1:57–60, 2005.

[BDA99] Patrick Bourdot, Martin Dromigny, and Laurent Arnal. Virtual navigation fully controlled by head tracking. *Proceedings of the 7th International Scientific Workshop on Virtual Reality and Prototyping*, 7:1–9, 1999.

[BKH97] Doug A. Bowman, David Koller, and Larry F. Hodges. Travel in Immersive Virtual Environments: An Evaluation of Viewpoint Motion Control Techniques. *Proceedings of the 1997 Virtual Reality Annual International Symposium*, 5(7):45–52, 1997.

[BKLP04] Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, and Ivan Poupyrev. *3D User Interfaces: Theory and Practice.* Addison-Wesley Professional, 2004.

[FHH03] D. W. Fellner, S. Havemann, and A. Hopp. Dave – eine neue technologie zur preiswerten und hochqualitativen immersiven 3d-darstellung. In R. Möller, editor, *Proc. 8. Workshop: Sichtsysteme – Visualisierung in der Simulationstechnik*, pages 77–87, Bremen, nov 2003. Shaker Verlag.

[FRE03] Kiran J. Fernandes, Vinesh Raja, and Julian Eyre. Cybersphere: the fully immersive spherical projection system. *Communications of the ACM*, 46(9):141 – 146, 2003.

[IYFN05] Hiroo Iwata, Hiroaki Yano, Hiroyuki Fukushima, and Haruo Noma. CirculaFloor. *IEEE Computer Graphics and Applications*, 25(1):64 – 67, 2005.

[JKS11] T. Jung, S. Krohn, and P. Schmidt. Ein Natural User Interface zur Interaktion in einem CAVE Automatic Virtual Environment basierend auf optischem Tracking. In *Workshop 3D-NordOst*, pages 93–102, 2011.

[KHR+11] T. Kretz, S. Hengst, V. Roca, A. Pérez Arias, S. Friedberger, and U.D. Hanebeck. Calibrating dynamic pedestrian route choice with an extended range telepresence system. In *2011 IEEE International Conference on Computer Vision Workshops*, pages 166–172, 2011. First IEEE Workshop on Modeling, Simulation and Visual Analysis of Large Crowds, 6-13 November 2011, Barcelona, Spain.

[LFKZ01] Joseph J. Jr. LaViola, Daniel Acevedo Feliz, Daniel F. Keefe, and Robert C. Zeleznik. Hands-free multi-scale navigation in virtual environments. *Proceedings of the 2001 Symposium on Interactive 3D Graphics*, 1:9–15, 2001.

[LSF08] Marcel Lancelle, Volker Settgast, and Dieter W. Fellner. Definitely affordable virtual environment. *Virtual Reality Conference*, 1(2):1–1, 2008.

[LSFP07] Robert Leeb, Volker Settgast, Dieter Fellner, and Gert Pfurtscheller. Self-paced exploration of the Austrian National Library through thought. In *International Journal of Bioelectromagnetism*, volume 9, 2007.

[Min95] Mark R. Mine. Virtual Environment Interaction Techniques. *University of North Carolina at Chapel Hill, Technical Report TR95-018*, 18:1–18, 1995.

[STU07] Martin Schwaiger, Thomas Thümmel, and Heinz Ulbrich. Cyberwalk: Implementation of a Ball Bearing Platform for Humans. *Lecture Notes in Computer Science*, 4551:926–935, 2007.

[SUS95] Mel Slater, Martin Usoh, and Anthony Steed. Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Trans. Comput.-Hum. Interact.*, 2(3):201–219, September 1995.

[Vol12] Joseph Volpe. Google maps 6.0 hits android, adds indoor navigation for retail and transit. `http://www.engadget.com/2011/11/29/google-maps-6-0-hits-android-adds-indoor-navigation-for-retail/`, 5 2012.

[WO90] Colin Ware and Steven Osborne. Exploration and virtual camera control in virtual three dimensional environments. *Proceedings of the 1990 Symposium on Interactive 3D Graphics*, 1:175 – 183, 1990.