

# smARTbox – A Portable Setup for Intelligent Interactive Applications

Martin Fischbach<sup>1</sup>, Dennis Wiebusch<sup>1</sup>, Marc Erich Latoschik<sup>1</sup>,  
Gerd Bruder<sup>2</sup>, Frank Steinicke<sup>2</sup>

Human-Computer Interaction Group, University of Würzburg<sup>1</sup>  
Immersive Media Group, University of Würzburg<sup>2</sup>

## Abstract

This paper presents a semi-immersive, multimodal fish tank simulation realized using the smARTbox, an out-of-the-box platform for intelligent interactive applications. The smARTbox provides portability, stereoscopic visualization, marker-less user tracking and direct interscopic touch input. Off-the-shelf hardware is combined with a state-of-the-art simulation platform to assemble a powerful system environment, which facilitates direct (touch) and indirect (movement) interaction.

## 1 Introduction

Interactive graphics displays provide an innovative medium in promotional, artistic, or educational areas. Interaction and participation are central aspects. Diverse sensing and tracking technologies turn passive observers into active participants. Their actions change the behavior of the installation's content, from almost unobtrusive adaptations of art-like scenarios to obvious and clear interaction possibilities.

The smARTbox is build from off-the-shelf components. It realizes an out-of-the-box platform for semi-immersive interactive applications providing portability, stereoscopic visualization, marker-less user tracking and direct interscopic touch input. The smARTbox is powered by Simulator X, a software for intelligent interactive applications that provides state-of-the-art simulation software aspects as well as multimodal—speech and gesture— interaction paradigms (Latoschik & Tramberend 2011). The demonstration simulates an improved version of the virtual fish tank presented at the virtual reality international conference (VRIC) with a head-tracked stereoscopic view (Fischbach et al. 2012). Users can interact with a simulated liquid surface and a school of fish by touching and perturbing the pseudo watery surface.

## 2 Hardware Setup

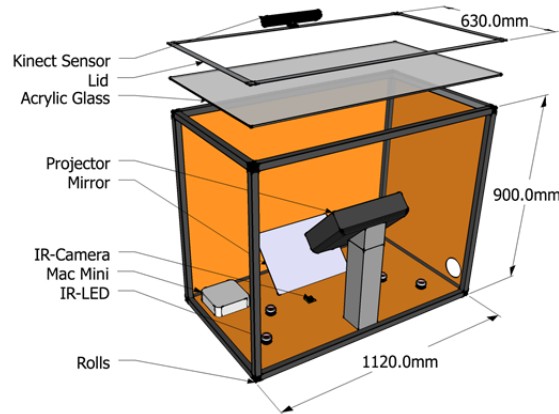


Figure 1: Hardware setup of the smARTbox.

A schematic of the smARTbox is shown in Figure 1. Enclosed in a  $63\text{cm} \times 112\text{cm} \times 90\text{cm}$  portable wooden box, the virtual environment shown in the demonstration is rendered on an nVidia Quadro 4000 graphics card, driven by an Intel Core i7 @3.40GHz processor with 8GB of main memory. An Optoma GT720 DLP-projector supports active stereoscopy and a resolution of  $1280 \times 720$  pixels at a refresh rate of 120 Hz. A wide-angle converter lens is used to back-project images at the screen via a mirror mounted at the bottom at a  $45^\circ$  angle. The hardware setup was constructed at a total cost of less than €2,500 during a students project.

### 2.1 Marker-less User Tracking

The smARTbox uses a Microsoft Kinect with the flexible action and articulated skeleton toolkit (FAAST) for real-time movement tracking (Suma et al. 2011, Tang 2011), e.g., to generate a head-coupled perspective projection. Multiple users can be tracked simultaneously, providing potential for interactive collaborative setups. The Kinect is a multi-sensor system that uses structured IR light patterns to reconstruct the distance of points in space (Freedman et al. 2008). The data provided by the Kinect sensor includes an RGB image with a resolution of  $1280 \times 1024$  pixels at a refresh rate of 15Hz with  $63^\circ$  horizontal and  $50^\circ$  vertical field of view (FOV), an IR image of  $1280 \times 1024$  pixels at 15Hz with a  $57^\circ$  horizontal and  $45^\circ$  vertical FOV, as well as a computed depth image of  $640 \times 480$  pixels at 30 Hz. The Kinect also supports a  $640 \times 480$  pixels reduced versions of RGB and IR images running at 30 Hz.

## 2.2 Direct Interscopic Touch Input

Multi-touch interaction has recently received considerable attention (Steinicke et al. 2011) due to the potential of comparably accurate input, and near-natural interaction with mono- and stereoscopic objects relative to display surfaces (Valkov et al. 2010 and 2011). The top of the smARTbox is modified to serve as a touch-sensitive input surface via the Rear-DI principle (Schöning et al. 2010). Six clusters of high-power IR LEDs illuminate the screen from the inside. Objects in contact with the surface reflect the IR light. This is sensed by a Point Grey Dragonfly®2 digital video camera with a wide-angle lens and infrared band-pass filter. Touch positions and gestures are analyzed by a modified version of the NUI Group's CCV software (Çetin et al. 2009).

## 3 Responsive Virtual Fish Tank

The simulation of a fish tank is a common theme in virtual environments (Mulder & Van Liere 2000, Ware et al. 1993). The responsive virtual fish tank demonstrated is reacting to user presence in multiple ways. The Kinect sensor tracks the user's head-position in order to render the stereoscopic three-dimensional scene. Touches are registered via the IR camera: When a contact with the display is established, a rippling effect is simulated and nearby fishes try to avoid the intruder (see Figure 2). Due to the sensors' oblique modes of operation, the user can freely interact with the application.



Figure 2: Responsive Fish Tank application: Overview (left) and user's perspective (right)

A general and distributed model of behavior inspired by Reynolds (1987) was implemented to simulate the fish-behavior. Simulator X fosters a semantic decoupling of components as well as of the simulated entities (here the fishes, the water, and the user). Entities form a global state, which is accessible by every component using a semantic abstraction layer. For example, the swarm simulation component is completely independent. Rendering, sensor, or interaction components can easily be exchanged without interfering with the simulation. Hence, the Simulator X based applications can drive a multitude of environments, ranging from standard desktop displays, to devices like the smARTbox with its touch-sensitive surface and multiscreen setups like CAVEs. Additionally, alternate species with different be-

haviors can easily be added to the fish tank without interfering with the rest of the simulation, e.g., by adding a new component.

## 4 Conclusion

The smARTbox has great potential for presentations and exhibitions since it is less expensive, easier to set up and transport than typical VR installations. The lack of obtrusive instrumentation improves the engagement with potential users, e.g. in exhibition scenarios. The utilized software platform Simulator X provides a state-of-the-art simulation engine with an integrated AI-core and several multimodal interaction components. Its scalable architecture provides an ideal platform, e.g., to experiment with alternate sensors, possibly enhancing accuracy or precision.

### Literature

- Çetin, G., Bedi, R. & Sandler, S. (Eds.) (2009). *Multi-touch technologies*. Natural User Interface (NUI).
- Fischbach, M., Latoschik, M. E., Bruder, G., Steinicke, F. (2012). smARTbox: Out-of-the-Box Technologies for Interactive Art and Exhibition. *Virtual Reality International Conference*. ACM.
- Freedman, B., Shpunt, A., Machline, M. & Arieli, Y. (2008). Depth mapping using projected patterns. WO Patent WO/2008/120217.
- Latoschik, M. E. & Tramberend, H. (2011). Simulator X: A Scalable and Concurrent Software Platform for Intelligent Realtime Interactive Systems. *Proceedings of the IEEE VR 2011*. IEEE.
- Mulder, J. D. & Van Liere, R. (2000). Enhancing fish tank VR. *Proceedings of the IEEE VR 2000*, pages 91–98. IEEE.
- Reynolds, C. W. (1987) Flocks, Herds, and Schools: A Distributed Behavioral Model. *Computer Graphics SIGGRAPH '87 Conference Proceedings*, 21(4), pages 25-34.
- Schöning, J., Hook, J., Bartindale, T., Schmidt, D., Oliver, P., Echtler, F., Motamedi, N., Brandl, P. & Zadow, U. (2010). Building interactive multi-touch surfaces. *Tabletops-Horizontal Interactive Displays*, pages 27–49.
- Steinicke, F., Benko, H., Daiber, F., Keefe, D. & de la Rivière, J. B. (2011). Touching the 3rd dimension (T3D). *Proceedings of the Conference extended Abstracts on Human Factors in Computing Systems*, pages 161–164. ACM.
- Suma, E. A., Lange, B., Rizzo, A., Krum, D. M. & Bolas, M. (2011). Faast: The flexible action and articulated skeleton toolkit. *Proceedings of the IEEE VR 2011*, pages 247–248, IEEE.
- Tang, M. (2011). Recognizing hand gestures with Microsoft's Kinect. *Technical report, Department of Electrical Engineering, Stanford University*.
- Valkov, D., Steinicke, F., Bruder, G., Hinrichs, K., Schöning, J., Daiber, F. & Krüger, A. (2010). Touching floating objects in projection-based virtual reality environments. *Proceedings of Joint Virtual Reality Conference*, pages 17–24.
- Ware, C., Arthur, K. & Booth, K. S. (1993). Fishtank virtual reality. *Proceedings of the Conference on Human Factors in Computing Systems*, pages 37–42. ACM.