An Immaterial Pseudo-3D Display with 3D Interaction

Stephen DiVerdi[a], Alex Olwal[b], Ismo Rakkolainen[c & d], Tobias Höllerer[a],

[a] University of California at Santa Barbara, Santa Barbara, CA 93106, USA

[b] KTH, 100 44 Stockholm, Sweden

[c] FogScreen Inc., 00180 Helsinki, Finland

[d] Tampere University of Technology, 33101 Tampere, Finland

ABSTRACT

We present a novel walk-through pseudo-3D display, which enables 3D interaction and interesting possibilities for advanced user interface designs. Our work is based on the patented FogScreen, an "immaterial" indoor 2D projection screen which enables high-quality projected images in free space. We extend the basic 2D FogScreen setup with stereoscopic imagery and two-sidedness, in addition to the use of head tracking to provide correct perspective 3D rendering for a single user. We also add support for 2D and 3D interaction for multiple users with the objects on the screen, via a number of wireless input technologies that let us experiment with interaction with or without encumbering devices. We evaluate the usability of these interaction techniques by observing non-expert use in real settings to quantify the effects they have on 3D perception. The result is a wall-sized, immaterial pseudo-3D display that enables engaging 3D visuals with intriguing 3D interaction.

Introduction

Many techniques have been developed to create the impression of a 3D image floating in mid-air. These technologies all attempt to artificially recreate the depth cues we naturally perceive when viewing a real 3D object. For example, stereoscopic imaging simulates binocular disparity cues by presenting slightly different

DiVerdi, S., Olwal, A., Rakkolainen, I., and Höllerer, T. An Immaterial Pseudo-3D Display System with 3D Interaction. In *Three-Dimensional Television: Capture, Transmission, and Display*. Springer, 2008, pp. 505-528. ISBN 978-3-540-72531-2. Editors: Haldun M. Ozaktas and Levent Onural.

images of the same scene to the left and right eyes, which is interpreted by the brain as a single 3D image. Virtual reality applications tend to track the user's head and render different views of the 3D object depending on where the user is in relation to the object, to simulate motion parallax. 3D applications in general simulate realistic imagery with perspective and complex shading algorithms to create the impression that the virtual object is seamlessly integrated with the 3D scene.

We have applied these simulated depth cues to a novel immaterial display technology [1], creating an engaging new way to view 3D imagery. In addition, multiple unencumbered users can naturally manipulate the objects floating in mid-air. This paper summarizes our work on the interactive, immaterial pseudo-3D display and its implications on perception of 3D content.

Our novel walk-through pseudo-3D display and interaction system is based on the patented FogScreen, an "immaterial" indoor 2D projection screen [2,3,4], which enables high-quality projected images in free space. We have extended the basic 2D FogScreen setup in three major ways. First, we use head-tracking to provide motion parallax and correct perspective rendering for a single user. Second, we support multiple types of stereo vision technology for binocular disparity cues. Third, we take advantage of the two-sided nature the FogScreen to render the front and back views of the 3D content on the two sides of it, so that the user can cross the screen to see the content from the back.

Commonly, users want to interact with the displayed objects, not just view them. The ability for an observer to walk completely through the FogScreen makes it appropriate for enabling new viewing and interaction possibilities. While two traditional displays mounted back-to-back could present a similar pseudo-3D display to multiple viewers, the opaque displays would prohibit viewers from effectively collaborating across the two displays by obscuring users' views of each other, blocking or distorting speech, and making it difficult to pass physical objects across the display. This permeability of the FogScreen is important not only for imaging and visualization, but also to provide an additional perceptive cue of the virtual 3D content's integration into the physical environment.

Our interactive, dual-sided, wall-sized system allows a single user to view and manipulate objects floating in mid-air from any angle, and to reach and walk through them. Two individual, but coordinated, images are projected onto opposite sides of a thin film of dry fog, and an integrated 3D tracking system allows users on both sides to interact with the content, while the non-intrusive and immaterial display makes it possible to freely pass physical objects between users or move through the shared workspace. Our system opens up possibilities for a wide range of collaborative applications where face-to-face interaction and maximum use of screen estate is desirable, as well as the maintenance of individual views for different users.

We first discuss related work in Section 2. The basic FogScreen is described in Section 3, and its pseudo-3D extension is explained in Section 4. In Section 5, we discuss the interaction and input technologies, and Section 6 describes our interactive demonstration applications. Section 7 presents the results and evaluation of our pseudo-3D display and interaction, and finally we present future work and conclusions in Sections 8 and 9, respectively.

Advanced Displays

A fundamental goal for all 3D displays is to create an illusion of depth, such that the user perceives a full 3D scene that seems to float in mid-air. It can be generated in a variety of ways by artificially recreating the effects of depth cues from natural viewing. We briefly discuss the variety of such displays here.

Stereoscopic displays [5] provide slightly different images for the left and right eye, creating the appearance of 3D objects that float in front of or behind the screen. The viewing area for correct perspective is restricted, and user-worn glasses are required. Autostereoscopic displays [6] require no special glasses for stereoscopic viewing, but the correct viewing area and resolution are typically somewhat limited. The viewer's 3D position can be tracked, allowing the rendered images to be modified according to the user's perspective. This expands the viewing area and enables the user to experience parallax through head-motion.

Traditional augmented [7] and virtual reality often use head-worn, tracked displays [8] which draw virtual images directly in front of the user's eyes. These setups typically only provide a private image which cannot be seen without cumbersome user-worn equipment. Artifacts such as misregistration and lag detract from the sense of presence and may cause eye-strain, headache, and other discomforts.

Volumetric displays create a 3D image within a volume. The objects can be viewed from arbitrary viewpoints with proper eye accommodation. Unfortunately, existing displays create their 3D imagery in a fairly small enclosed volume that the viewer cannot enter. They also have problems with image transparency when parts of an image that are normally occluded are seen through a foreground object.

Many research projects investigate large displays with user tracking as interactive surfaces. It has proven advantageous to use a screen material that supports rear-projection and tracking from behind the display such that occlusion can be minimized. The HoloWall [9], MetaDESK [10] and Perceptive Workbench [11] use a diffusion screen for rear-projection, while IR illumination enables IR cameras to track objects near the surface of the screen. Projection screens like the dnp Holo Screen [12] and HoloClear [13] consist of a transparent acrylic plate that is coated with a holographic film, such that it only diffuses light projected from a 30-35º angle. These transparent displays show only projected objects, are single-sided and not penetrable. Touchlight [14] uses such a screen and allows the gestures of the users to be tracked through the screen. Hirakawa and Koike [15] combine user tracking with a transparent 2D screen for a projection-based optical see-through AR system, whereas ASTOR [16] achieves autostereoscopic AR with 3D imagery using a holographic optical element. A serious limitation of these setups, however, is their inherent single-sidedness, which limit collaboration.

There have been several displays using water, smoke or fog, with an early example presented by the Ornamental Fountain dating back to the end of the 19th century [17]. More recently, water screen shows such as Water Dome [18], Aquatique Show [19] and Disney's Fantasmic [20], spray sheets of freely flowing or high-velocity water to create impressive displays for large audiences. The magnitude and wetness of these screens, as well as their large water consumption, make them impractical for indoor applications, as well as preclude the viewers from comfortably passing through them and seeing clear images from short distances.

Many types of fog projection systems [21,22] have been used for art and entertainment purposes, but the rapid dispersion of the fog seriously limits the fidelity of projected images. The dispersion is caused by turbulence and friction in the fog flow, which disrupts the desired smooth planar surface, causing projected points of light to streak into lines. This streaking causes severe distortion of the image from off-axis viewing angles.

Closely related to the FogScreen is the Heliodisplay [23], a medium-sized (22"-42" diagonal) immaterial rear-projection display based on the same principle as the patented FogScreen. It harvests humidity in the air by condensing it into water, which is then broken into fog. An IR laser can be used to illuminate a finger while a camera provides unobtrusive finger tracking. However, considering its single-sidedness, smaller format (compared to the 100 inches of the FogScreen) and tabletop setup, it is not a suitable basis for the kind of walk-through human-scale interactive display we are pursuing.

The "Immaterial" FogScreen

To achieve an immaterial display, we base our system on the FogScreen [2,3,4], which uses fog as a projection surface to create an image that floats in mid-air (see Figure 1). If people walk through the FogScreen, the image will instantly re-form behind them. It allows projection of interactive content, such as images or videos, to appear floating in free space. It also enables creation of special effects like walking through a brick wall or writing fiery characters in thin air. FogScreens are currently used for special effects at various high-profile venues, events and trade shows. Entertainment is one big application area, including performing arts [24], but the screens are increasingly used for other applications as well.

The FogScreen employs a patented method for forming a physically penetrable 2D particle display. The basic principle (see Figure 2a) is the use of a large non-turbulent airflow to protect a flow of dry fog particles inside it from turbulence. The outer airflow may become slightly turbulent, but the inner fog layer remains flat and smooth, enabling high-quality projections. Ordinary tap water is broken into fine fog droplets and trapped inside this non-turbulent airflow. The resulting stable sheet of fog enables projections on a screen that is dry and feels like slightly cool air. The light from a standard projector is scattered through this sheet of fog, creating a rear-projection image.

Fig. 1. The FogScreen can create fully opaque or very translucent high-quality images in mid-air. It can provide high visual detail.

The FogScreen works much like an ordinary screen in terms of projection properties. Light from a projector is scattered by the fog, creating an image that floats in mid-air. However, not all the light is scattered, so a bright projector is needed. A 5000 ANSI lumens projector is usually sufficient for lit environments such as trade shows, if the background is dark. In dark rooms dimmer projectors will suffice.

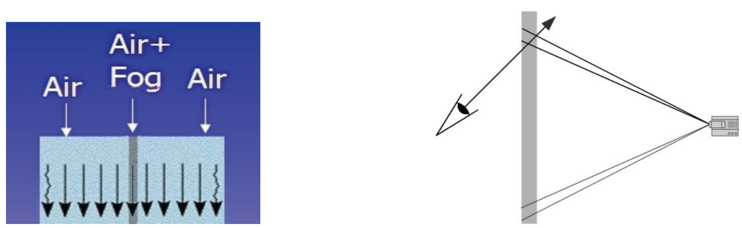

Fig. 2. *left to right*: (a) The principle of the FogScreen. (b) As the FogScreen image plane (grey cross-section) is not infinitesimally thin, the pixels may mix with the neighboring ones when viewed or projected at steep angles.

The image can be viewed in most off-axis viewing directions, although an on-axis viewing direction towards the projector yields an optimal image. The field-of-view is currently up to 120º, depending on flow quality and environmental issues, such as air flow and characteristics of the projected imagery. When viewed from a very steep angle, the thickness of the fog causes adjacent pixels to blur into one another, reducing image quality from the sides (see Figure 2b).

The FogScreen characteristics require some considerations in order to design the best possible content and take advantage of the screen as a novel media space [24]. Projectors with SVGA resolution (800×600 pixels) are adequate due to the currently limited fidelity of the FogScreen. Tiny details like small text may be hard to view and use, so it is recommended to design large buttons and objects. The effective image resolution is optimal towards the top of the screen and deteriorates somewhat with increasing distance from the fog dispenser. Most natural imagery looks good on the screen, and color and contrast are vividly preserved.

We conducted a test on the effect of projection angle with a NEC WT-610 [25] ultra-short throw distance projector, which can create a 100” diagonal image from as close as 25 inches. The projection angle over the screen area then varies between 40-70º. Because of the thickness of the fog, the images become quite blurry compared to conventional projector setups. Only very large, uniform objects and text remained identifiable and legible. For the remainder of this work we used ordinary projectors.

The opacity of FogScreen depends on many parameters such as fog density, projector and image brightness, and the background of the viewing area. Depending on what imagery is projected onto the FogScreen, a variety of different effects can be achieved. When showing a normal photograph it acts much like a traditional projector screen. However, if the image is mostly black except for a few objects, the black areas appear transparent and create the effect of virtual objects floating in space in front of the user. If a full-screen image of a virtual environment is displayed without text or abstract imagery, it creates more of a portal effect, giving the impression of looking through a window into another world.

3D and Pseudo-3D Display Technologies

While fundamentally a 2D display technology, the basic FogScreen can be extended to become a pseudo-3D display, via dual-sided rendering, head-tracked rendering, and stereoscopic imaging.

The screen affords a 100” diagonal image in the center of a large open viewing area that is limited only by available space and coverage of a 3D position tracker (see Figure 3). By tracking a single viewer’s head, using correlated projectors on each side and adjusting the projected 3D graphics rendering accordingly, we create a pseudo-3D display. This makes the 3D effect more convincing by showing the 3D object from the appropriate angle. The viewer can see objects floating in mid-air from both sides and freely walk around and through them to examine the scene from almost any angle. The eye cannot correctly focus at a real 3D point within the image, but an impression of depth is still achieved due to other monocular cues, most notably motion parallax. Also, stereoscopic imaging techniques can be used with the FogScreen. These techniques were all informally evaluated by six different viewers to get an idea of 3D perception performance.

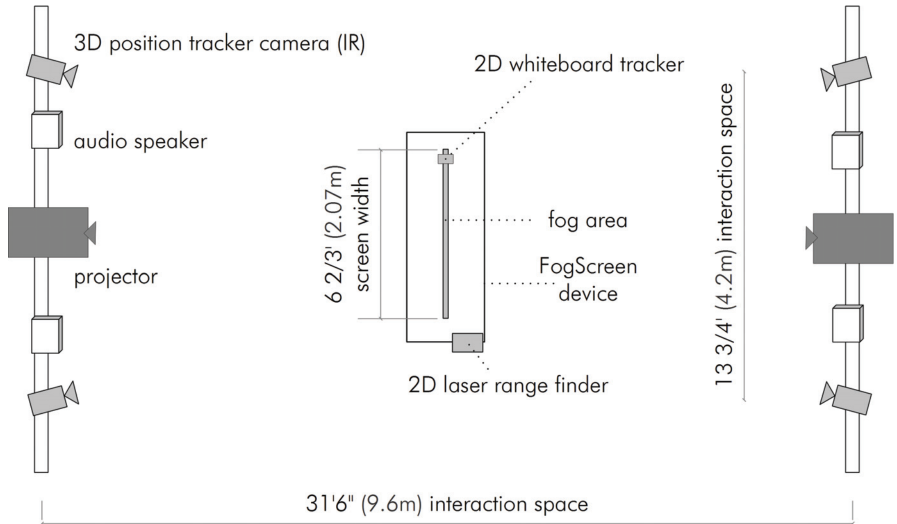

Fig. 3. Our dual-sided prototype system setup consists of a FogScreen with a 100" diagonal screen, created by two SVGA projectors mounted 3 m above the floor. Interactivity is added via a whiteboard tracker or a laser range finder for 2D position tracking, and/or via four infrared cameras around the work-space for 3D position tracking.

4.1. Stereoscopic projection

We experimented with a variety of passive and active stereoscopic rendering techniques on our display.

Passive stereoscopy with linear polarized glasses and filters [1] is possible, as a thin fog layer accurately preserves polarization of rear-projected light. We used standard polarization filters and glasses for our experiments. Cross-talk between the left and right images is comparable with that resulting from the use of a standard silvered screen. Polarization requires two projectors, which raises the setup cost. The computer must also be able to drive two separate projectors for a single-sided display, requiring four different views being rendered for dual-sided polarized stereo.

Passive stereoscopy with red-cyan colored glasses [1] also worked well since the FogScreen maintains proper image colors. Red-cyan stereoscopy only requires a single projector, making the system less expensive and complex than polarized stereoscopy, but the effect is limited to monochromatic imagery.

Since the FogScreen preserves image colors, the Infitec [26] passive stereo system could also be used, but so far we did not have one available for testing.

The last passive stereoscopy technique we tried was ChromaDepth [27], which color-codes an image by depth value, with red pixels being nearest to the camera, followed by orange, yellow, green and finally blue in the background (see Figure 7). A pair of diffraction grating glasses shift colors so red areas appear near the user, while blue appears far away. The main advantage of this technique is that if the users are not wearing the special glasses, they still see a single coherent image, instead of two superimposed views as with red-cyan or polarized stereo. However, the tradeoff is that ChromaDepth is more of an ad hoc technique for creating binocular depth cues and does not actually simulate eye separation and focal length of the user's visual system, resulting in less effective 3D perception.

For active stereoscopy, we used a DepthQ 3D projector [28] with shutter glasses. While this projector model is quite affordable and may serve as an example for the ongoing reduction in costs for active stereo systems, this was still overall the most expensive option we explored, and the resulting image brightness was lower than that of cheaper passive stereo solutions. Initial results indicate that the quickly changing turbulence pattern of the fog surface over time causes a subtle difference between the left and right images of an active stereo projection, disrupting accurate separation and making it slightly more difficult than on a standard silvered screen to see a clear stereoscopic image. This problem will be partially solved as the screen quality will improve in the future.

4.2. 3D head tracking

Motion parallax, achieved by tracking the user's head position, is a strong monoscopic depth cue. Most any tracking technology suitable for virtual or augmented reality work could be used with our system. For this work, we employed a WorldViz Precision Position Tracker (PPT) wide-area 3D optical tracker [29] for head tracking. Our 3D tracking is explained in more detail in Section 5.2.

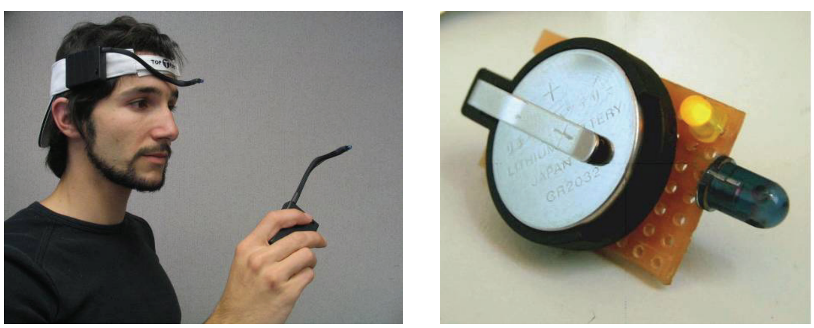

Fig. 4. *left to right:* (a) Prototype headset with an IR LED for viewer's 3D tracking, and a hand-held IR pointer for 3D interaction. (b) Custom-made 2×3 cm miniature version of the PPT marker.

We use an active LED marker on a headset for the head's 3D position tracking (see Figure 4a). The marker could also be custom-made into a miniature version for stereoscopic glasses. Figure 4b shows the current results of our miniaturization efforts. Our system works correctly for a single viewer, similar to virtual rooms and immersive workbenches [30].

4.3. Dual-sided projection

To accentuate the sensation that these virtual objects actually exist in the physical world, the dual-sided capabilities of the FogScreen are used to show both the front and back of the objects, so that viewing the scene from opposite sides will present a consistent perception.

Fig. 5. The FogScreen allows two independent images to be projected on each side of the screen, such that opposite sides of a 3D object can be rendered on the screen for a pseudo-3D effect. These photographs of a static two-sided scene illustrate how the back-projected image completely overshadows the simultaneous front-projection (which finds its way through the screen to the ground and back wall behind it).

Very little of the projected light actually reflects from the fog layer back towards the projecting source. Therefore, the image is predominantly visible for a viewer on the opposite side of a projector (viewing a rear-projected image). A front-projected image is extremely faint. This feature enables us to simultaneously project different images on the two sides of the FogScreen with the back-projected image completely dominating the view. The faint front-projection image means that there will be slight ghosting in high-contrast regions, but, in our experience, the cross-talk is acceptable, and, in fact, negligible apart from the case where very bright imagery on one side coincides with very dark regions on the other. In the cases where this cannot be avoided, dynamic photometric correction between the front and back projectors based on screen content could alleviate the effect.

More interestingly, two coordinated views of a 3D object can be shown on each side of the screen. For example, an application that displays a 3D object, such as a modeling and animation package, could show both sides of the object on the two sides of the FogScreen, creating a more convincing sense of presence of the virtual object in the physical environment. Figure 5 illustrates the dual-sided screen with the example of a cartoon shark seen from the front and back.

Our system is also well suited for multiple users. Different users that wish to view the 3D scene do not even need to be situated on the same side of the screen, as is the case with traditional display technologies. With conventional displays, users must crowd inside a small viewing. A typical tabletop system allows users to spread out around the display, but each user will see the data from a different orientation, some upside-down, hindering collaboration. Dual-sided rendering on the FogScreen allows the same layout to be presented on both sides, but with text and images properly oriented for viewers who can spread out on either side of the large display.

Pseudo-3D Experiments

While the image quality is not perfect from off-axis viewing angles, the system works reliably and produces an appealing and intriguing human-scale reach- and walk-through pseudo-3D display. The spectator can view floating objects from the front and back and freely walk around and through them and, with head-tracking, see the scene from any angle (see Figures 6 and 7). As the projection plane is 2D, the eye does not accommodate to the correct distance, but even without using stereoscopic effects the 3D nature of objects is emphasized by giving the impression of floating in free space. The 3D objects look fairly natural when viewed on-axis. As the viewing direction is moved to the side, the image will start to degrade, and finally becomes unusable when viewed more than about 60º off-axis.

Fig. 6. A gray, non-textured teapot on the pseudo-3D FogScreen as seen from the viewer.

Users who tried the head-tracking system took an initial period to become accustomed to the interaction, but found that the effect was convincing in making it seem as though a 3D object floated in the space of the screen. We had similarly encouraging results when we demonstrated the dual-sided rendering – users would often naturally walk around or through the screen to see the full scene and they could easily get a better idea of the entire 3D volume without the need for interaction devices.

To test the stereoscopic techniques, we showed our users a number of 3D images, including stereo photographs, random dot stereograms and rendered 3D images of simple geometric objects [31]. Overall, we found that the FogScreen is suitable for stereoscopic rendering, with an impressive 3D effect. The stereoscopic effect was dominant, and it became difficult to estimate where the screen plane lies.

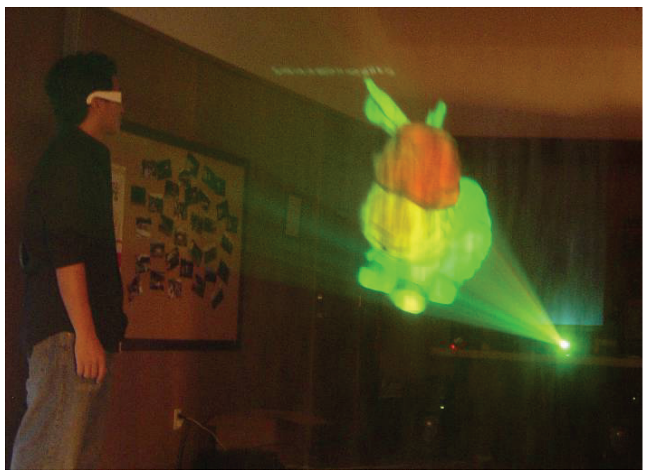

Fig. 7. The pseudo-3D FogScreen, displaying the Stanford Bunny [31] in mid-air, here with ChromaDepth [27] stereoscopic imaging.

In particular, polarized stereo provided the clearest 3D perception effect, based on qualitative assessments of many users for a variety of 3D scenes. The downside to the required two projectors is that the stereo effect demands careful calibration of the two images. The different projection angles for corresponding pixels also created some minor loss of edge sharpness at depth discontinuities, due to pixel smearing. Red-cyan stereo also performed well, but inherently it can not reproduce as vivid colors as polarization. ChromaDepth had the weakest 3D effect, but this is not surprising as it gives a poorer impression when viewed on a regular display as well.

Active frame-sequential stereo did not perform quite as well as we had hoped, as bright light from the projector easily disrupted the infrared stereo sync. This can be alleviated by using a more powerful IR emitter and actually be avoided in most cases if the projector is placed high enough so that the FogScreen device occludes it, except in the near proximity of the screen. If infrared signal transmission is not an option (e.g. because of interference with a chosen tracking or interaction technology), wired sync solutions can be employed. When the sync signal was present, the stereo effect was quite perceptible and comparable with passive stereo methods, but the passive stereo solutions were also more cost-effective.

In general, users were able to see stereo imagery on the FogScreen reliably, but some types of imagery required more effort on the user's part than others. In particular, with random dot stereograms it generally took users some time for the 3D scene to become visible, while regular 3D geometry was easy and instantaneous to perceive. While the FogScreen can be viewed from a wide field-of-view for 2D imagery, effective stereo correspondence was experimentally determined to be limited to a 15-20º viewing angle.

We also compared the stereo performance of the FogScreen with a traditional silvered screen and found that the FogScreen creates a more pronounced sense of depth than a traditional screen, whereas traditional screens naturally reproduce higher resolution images due to their more precise nature. We measured the pronounced depth effect by having users estimate the extent along the viewing direction of the same geometry on the FogScreen and the regular screen, and found that users consistently estimated the same objects as roughly 50% longer when displayed on the FogScreen. Our theory is that this effect is rooted in the lack of a reference plane, as objects on the FogScreen appear to float in mid-air, whereas objects on the regular screen are perceived as being anchored in front of the screen plane. Also, the projecting "cones" emanating from the projectors, visualized by participating media beyond the transparent screen (clearly visible in Figures 6 and 7) may have contributed to this exaggerated sense of depth.

Interaction technologies

The immaterial nature of the display is important for enabling new interaction possibilities. While two traditional displays mounted back-to-back could present a similar display to multiple users, the opaque displays would prohibit users from effectively collaborating across the two displays by obscuring users' views of each other, distorting speech, and making it difficult to pass physical objects across the display. This permeability of the FogScreen is important not only for collaboration but also for the potential for direct interaction with the 3D scene. To enable interaction with our immaterial display, we investigated a number of tracking systems and input devices.

The ability for a user to walk completely through the FogScreen makes it appropriate as a portal in virtual or mixed reality environments. A CAVE™ with a FogScreen as a wall, for instance, would allow the user to easily enter the virtual environment from the outside. The sense of real immersion would still be achieved since the user would be completely surrounded by displays. Furthermore, before entering the environment, the FogScreen portal could present an interface to the outside world that would allow a user to set parameters about the environment he or she was about to enter.

Collaboration among multiple people within a single application can be greatly enhanced by a dual-sided display. Multiple users that wish to cooperatively use an application with traditional display technologies must all stand in front of the same display, limiting the number of people who can effectively participate. Immersive workbenches and tabletop systems [30][32] allow more users to share a workspace, but they do not all view the data from the same side – while a user on one side of the table may see text correctly oriented, a user on the other side will be unable to read it as it will be upside-down. With our dual-sided display, the same interface and layout can be presented on both sides of the screen, but text and images can be properly oriented so all users can actively participate. Collaboration is also encouraged through its inherent support for face-to-face interaction and eye contact. Four people interacting on a shared scene, for example, can more effectively interact in groups of two on either side of the screen, seeing and reaching through to each other, than in one group of four in front of a one-sided display or tabletop system.

Interaction can also be an important component of perception – as the goal of the pseudo-3D FogScreen is to create the sensation that virtual objects exist in the physical space around the user, the ability to seamlessly interact with those objects would reinforce their integration. 3D perception is aided by interaction through directly mapping the user's 3D gestures to virtual objects, creating a proprioceptive depth cue.

6.1. 2D tracking

The FogScreen appears to intrigue people as a passive, immaterial walk-through screen, and the natural inclination upon first seeing it is often to attempt to play with the virtual objects on the screen by touching them. By turning it into an interactive 2D computer touch screen, the application possibilities for the screen significantly broaden.

6.1.1. Ultrasound tracking

One of our tracking solutions consists of a low-cost, off-the-shelf whiteboard tracker (eBeam Interactive). The device tracks an ultrasound emitter in a 2D plane using ultrasound sensors that are attached in one corner of the screen (see Figure 3). This allows one ultrasound emitting pointer to be tracked as long as line-of-sight is maintained. We had to make some minor modifications to the emitting pointer to remove the need to push the wand against a solid screen. The device has trouble capturing ultrasound across larger surfaces, such as our 100" diagonal display, since the sensors are located on one side only. The previous eBeam System 1 model worked reliably, as the sensors are on both sides of the screen and reasonable tracking accuracy is provided (typically ±2cm on a 2 m wide screen).

The fog flow and ultrasound emissions of the device itself have no noticeable effect on the tracking. The accuracy is adequate for most entertainment and business applications, save very detailed, high-precision work. In addition to the spatial inaccuracy, the ultrasound tracking introduces a temporal delay of about 100 ms. In typical ‘push-button’ interaction this is almost unnoticeable, but it might present a problem in a fast paced application such as a game. While eBeam is easy to use and install, its sensitivity to ambient noise can be a problem.

6.1.2. Laser range finder

We added the support for a laser range finder to enable more intuitive, unencumbered interaction. We used an eye-safe Sick™ LMS-200 laser scanner mounted on the FogScreen device (see Figure 8). It scans the environment by firing a series of short infrared laser pulses in a fan shape, and measuring the time-of-flight from the firing to the return of an optical echo. It provides an accuracy of 10 mm and a statistical error of just 5 mm, which is adequate for our purpose.

Fig. 8. Unobtrusive tracking. *left to right*: (a) The Sick laser range finder 2D tracking system, highlighted in the upper left corner, allows users to intuitively interact with the FogScreen using their bare hands. (b) In the depicted demo application, the system leaves slowly fading sparklers or fiery traces wherever the user is touching the screen.

The Sick scanner also requires line-of-sight, but an emitter is no longer needed and users are thus able to interact directly with their bare hands. The tracker is triggered by objects intersecting the plane and the largest detected object is chosen to represent the touch screen input. As long as an object is present, the left mouse button is emulated to be pressed.

The laser range finder may be triggered by the fog and thus needs to be mounted 10-20 cm from the screen. The non-intrusive tracking allows natural interaction, while currently limiting the system to one user and one preferred side per scanner device.

The scanner transmits its information via a serial RS-422 link at up to 500 kbps. This is both too high a data rate for ordinary PC serial ports to receive and too high a sampling rate to be useful for our purposes. We constructed a simple AVR microcontroller based interface card to receive the data from the scanner, compress it, and transmit the compressed data at a rate of 115,200 bps to the host. At the host, the data can either be directly used by custom software or by our emulator of a generic desktop mouse to control arbitrary legacy applications.

The laser scanner occasionally produces too high measurement values, mainly due to a laser beam only partially hitting a target, so filtering is applied to the data stream. Spurious single sample errors are discarded by a 3-tap 2D median filter. For the mouse emulator, the onset of the mouse button click is delayed to allow the hand to fully enter the beam and a similar delay is introduced for the mouse button release when the hand is removed. As the scanning rate is 75 Hz, the resulting delay is not noticeable to the user.

6.2. Vision-based 3D tracking

The two abovementioned tracking technologies provide compact 2D tracking solutions that are straightforward to install and calibrate. The single-user constraint and limitation to 2D interaction made us look into other tracking technologies, such as the WorldViz PPT, to enable multi-user 3D applications. The PPT is a wireless 3DOF vision-based system that uses from two to eight cameras to track small near-infrared (IR) LEDs in the environment. Our setup uses four cameras, two on each side of the display (see Figure 3). The FogScreen is practically translucent in the IR spectrum so it does not interfere with the visibility of our IR LEDs.

Each LED marker needs to be visible by at least two cameras at all times, and up to eight can be tracked simultaneously. We use one LED to track the user's head position, and one LED for the hand position. Multiple LEDs can also be combined to provide more degrees of freedom, which obviously increases the size of the marker due to the required distance between LEDs. The use of two LEDs would allow a user to specify a vector, which can be useful for orienting objects on the screen – three LEDs allow 6DOF tracking.

There is no robust means of uniquely identifying a particular LED. Thus, LEDs that are close to each other or LEDs that move too quickly can confuse the tracker, causing IDs to be swapped or the appearance of more or fewer LEDs than are actually present. These artifacts are particularly problematic for applications that need to maintain knowledge between frames of tracked object identity. To address this issue, we developed a proxy VRPN tracker server [33] that filters the PPT output into more reliable data. It analyzes the position and velocity of tracked objects to predict future positions, reducing swapping, and removes the duplicate report artifact by eliminating tracking results that have very similar position and velocity.

The use of IR markers also imposes the requirement of controlled lighting, since many regular light sources have IR radiation that will generate noise in the near-IR camera image. This issue is evident especially in environments with daylight or bright incandescent spotlights, whereas standard fluorescent lighting does not have this problem. Specific spotlights (i.e. with minimal IR radiation) or IR filters could make it possible to use incandescent light sources in the environment.

6.3. Input devices

Gesture-based interaction without the use of input devices is natural and direct, but is often limited to simple pointing, such as in the case of our laser range finder. Systems that recognize multiple and more complex pointing gestures still tend to be limited to a single user, and while manipulation gestures can be intuitive, system commands, such as mode changes, might not be as easy to represent.

Fig. 9. Wireless joysticks with integrated 3DOF PPT marker and Bluetooth. *left to right:* (a) Symmetrical joystick. (b) Right-hand joystick.

We developed various wireless controllers to facilitate interaction for applications where discrete input in a comfortable form factor is desired. Our controllers take the form of 3-button joystick handles with an integrated or detachable custom-made miniature version of the PPT marker for 3D position tracking (see Figure 9). The use of Bluetooth in combination with PPT markers enables multiple wireless 3DOF controllers to be simultaneously active in the system.

In addition to a trigger button, our controllers have two horizontally or vertically placed buttons that are conveniently accessible with the thumb. The controller with horizontal buttons is symmetrical and works for both left- and right-hand users (See Figure 9a). It is ideal for transformations, such as horizontal translation or rotation around the up-axis, as well as in applications that want to mimic left- and right-button mouse clicks. The other controller type is for right-hand users and has two buttons on the left side (See Figure 9b). The buttons are closer to each other and thus require less thumb movement, but the smaller separation makes it more likely for novice users to confuse them.

Additionally, our 2×3 cm miniature version of the PPT marker (see Figure 4b) is sufficiently small and light to be held between two fingers or to be attached to body parts as a lightweight 3D marker. Its small size allows us to simulate unencumbered 3D hand tracking for experiments with multi-user hand-tracked-style 3D interaction.

Interaction Experiments

We developed and tested several types of interfaces using our interactive FogScreen system. Most of these interfaces were part of demo applications that were presented at ACM SIGGRAPH 2005. All except the fiery characters demo are based on the vision-based 3D tracking described in Section 6. The goal with each interface was to explore a different interaction mechanism using the unique capabilities of the FogScreen and our input devices. Each test was examined to see how it affected users' perception of, and reaction to, the 3D content.

7.1. 2D manipulation

The first action most users take upon seeing the FogScreen is to insert their hands into the display to "play" with objects on the screen. We took advantage of this natural tendency by developing a set of interfaces that involve the user directly touching the screen to interact with content.

7.1.1. Fiery characters

Figure 8 shows the fiery characters demo, which allows users to play with slowly fading sparklers and lines of fire on the translucent screen. The density of fog is kept low, so only the bright fiery spots are visible and everything else is invisible. The demo uses 2D touch screen and mouse emulation by means of the eBeam tracker or Sick scanner – when a user touches the screen, a mouse button press is emulated, allowing users to draw in the air with their hands on the virtual screen.

7.1.2. Rigid body simulator

To create the sensation of manipulating real physical objects, we developed an application that lets users intuitively interact with realistically behaving virtual objects (see Figure 10). Using a straightforward implementation of standard mechanical dynamics [34], our application simulates a number of virtual rigid bodies that bounce around as if in a low-gravity environment.

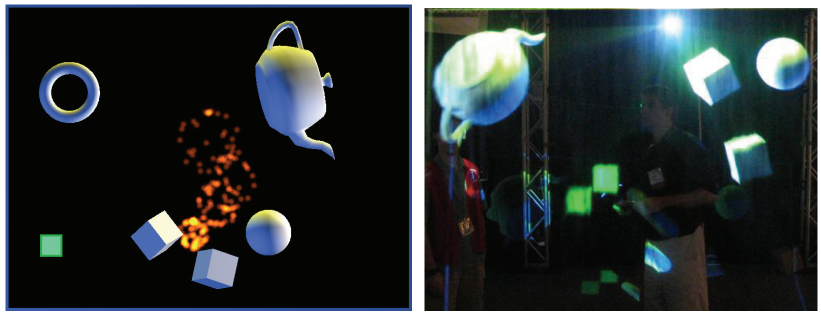

Fig. 10. The rigid body simulator has a physically intuitive interface which is easy even for small children to understand and play with. A 2DOF controller lets users move a green paddle which collides realistically with other virtual objects in the scene, including a teapot and a torus.

The interface is simple and straightforward, in an attempt to closely mimic manipulation of real objects – each user moves a single LED tracked by PPT to control a paddle. This paddle behaves as another rigid body in the simulation and allows the user to collide with the objects in the scene, directing their motion. The simulation and interaction are limited to 2D (all z-axis values are set to 0), as we found that without additional depth cues, perception of z-axis placement of objects was difficult for some users. However, the paddles can be controlled from any point in our 3D interaction space. Distance to the screen does not matter in this application. The 2D interface is so simple that even small children were able to immediately start playing with the simulation without any instruction. Because of the direct manipulation style of the interface, there is a sensation that the user is playing with real objects, rather than virtual ones. Unfortunately, this effect is somewhat reduced by the ability of the user to stand and interact with the screen at any depth because of the clamping of z-axis values.

7.1.3. Consigalo

The dual-side and reach-through capabilities of the FogScreen are used in Consigalo, an engaging multiplayer game (see Figure 10). Users control colored spherical cursors with the handheld PPT markers in the 3D space around the screen. The user touches the screen to grab objects at a location, or moves away from the screen to release objects. This allows players to pick up falling animals and sort them into the colored side goal areas, by dragging them across the display surface. The 3D tracking emulates a touch screen here: when moving the PPT marker (see Figure 9a) close to or through the screen, an event is triggered. This furthers the perception of virtual objects occupying the physical space of the screen by allowing users to 'grab' them when touching the screen. Consigalo also enables collaborative face-to-face 3D interaction by taking advantage of the screen transparency and dual-sidedness to allow play on either side of the screen – the players can even switch sides by moving through the screen while interacting with the animals.

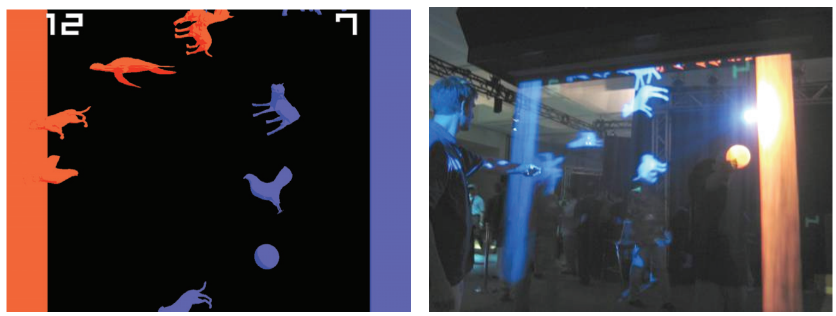

Fig. 11. The touchscreen-style interaction of Consigalo makes it very easy for users to intuitively understand the grabbing action necessary to catch animals and score points. The dual-sided display is instrumental in allowing many users to participate and engage each other from across the screen.

7.2. Navigation

To explore the concept of our display as a portal to a virtual environment, we created the virtual forest tour (see Figure 12). The interface is a first-person point of view situated inside the environment, with one of our 3DOF wireless controllers (see Figure 9) for interaction. The user can move the controller while holding the first button to change the velocity as the camera moves through the environment. The second button changes the viewing direction, and the third button controls the position of the light source in the scene. Altogether, the effect is a very natural game-style interface that makes it easy to navigate the environment.

While the perception of moving through a virtual environment is clear, it still requires some willful suspension of disbelief, as navigation does not use any natural locomotive methods. The presence of the virtual environment could be improved in a few ways. Using head-tracking to provide parallax in the environment would be a major step forward, but we haven't evaluated such hybrid setups yet. Some sort of natural locomotion interface such as the moving floor tiles [35], would also improve the sensation of navigating a real environment.

Fig. 12. The virtual forest tour acts as a portal from the real world into a virtual environment full of thousands of realistically rendered trees. People can explore the environment with a first-person game-style interface using a wireless 3DOF controller. We are currently exploring navigational interfaces taking into account the walk-through capability.

7.3. 3D manipulation

As we present 3D content on the pseudo-3D FogScreen, we are also interested in interfaces for 3D manipulation of scenes. We developed two applications that use different techniques to provide depth cues for a 3D cursor used to select regions of a curved 3D surface displayed by the screen.

7.3.1. Elastic surface deformer

Our first test was a single-user modeling application we developed to explore the combination of real 3D interaction and pseudo-3D visualization in our system. The elastic surface deformer uses a 3-button controller (see Figure 9b) to stretch and sculpt, as well as to move and rotate, an elastic 3D model of a human head (see Figure 13). The front and back views of the 3D model are projected on opposite sides of the screen, such that the user can walk through the screen and see what the object looks like from the other side in a pseudo-3D fashion.

We chose to use full 3D interaction in this application – users had to position the 3D cursor on the head surface to grasp and drag it. However, it proved slightly difficult to manipulate the 3D model, since the only available depth cue was cursor occlusion (perspective cues were not available as the head was orthographically projected to facilitate concurrent dual-sided view). While pure 2D interaction is too limited, a traditional solution to this problem would be to do selection in 2D and use relative 3D motion for dragging.


Fig. 13. The elastic face deformer's 3D interface gives users complete control over how they distort a virtual head, although selection was difficult due to insufficient depth cues.

7.3.2. Sound putty

The Sound Putty project extended the 3D manipulation of the elastic head deformer to a more abstract interactive art exhibit (see Figure 14). Multiple users are able to simultaneously influence the behavior of a virtual putty-like fluid by moving attractors and repellers around it in 3D. To provide additional depth cues, in the single-user case head-tracked rendering was used to provide motion parallax. Small motions of the user's head provide slight parallax which shows very clearly the depth of the surface, making correct 3D positioning possible. During interaction, the fluid will often move completely in front of the screen, no longer actually in the plane of the screen, but users are still able to effectively find and manipulate it.

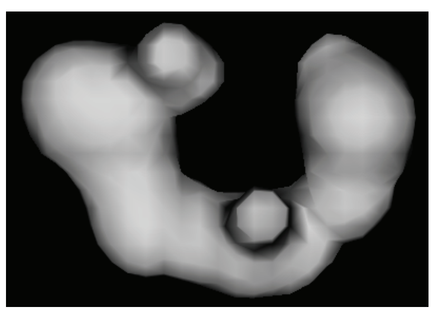

Fig. 14. Sound Putty presents the user with an abstract putty-like fluid that can be controlled to create interesting shapes and motions. The 3D interface was greatly enhanced by the addition of head-tracked rendering, which provides motion parallax depth cues to allow for 3D perception of the shape of the surface.

However, there was more of a learning curve associated with the head-tracked rendering, as users are not accustomed to that type of display. Proper calibration is critical to make a believable experience – when the calibration was slightly off, it distorted users' perception, making input more difficult than a regular 2D display, as they were grappling with figuring out what the image meant instead of focusing on the interaction. With proper calibration and after a short learning curve however, users had little difficulty interacting with this interface in 3D.

Future work

Our work with the FogScreen continues in many different areas. We are working towards improving the image quality of the display by further reducing turbulence in the fog flow. This will improve image fidelity and increase the effective field of view, as well as allow the realization of even larger screens. The quality of stereoscopic imagery, especially active stereo, should also improve with less turbulence.

We are also interested in improving the sense of presence of virtual objects in the physical world by integrating haptic feedback into the interface. A haptic device such as the SPIDAR [36] would work well with minimal discomfort – a normal display would interfere with the necessary wires, but they can go through the FogScreen, and a SPIDAR is capable of the large range of motion necessary. It would also provide 3DOF tracking input for at least a single user, removing the need for an additional tracking solution.

Finally, in the interest of developing a fully volumetric 3D display, we are currently investigating the use of multiple FogScreens in various configurations to allow images to occupy a tangible physical volume surrounding the user. This would overcome a fundamental limitation of existing volumetric displays, allowing the user to become fully immersed in the 3D visualization.

Conclusions

We have described a novel mechanism to create a pseudo-3D walk-through screen with interactive capabilities. The implemented system enables one to view and manipulate 3D objects in mid-air and observe them from different angles in a natural manner. Using it as an immaterial, head-tracked dual-sided display has led to an enhanced visualization experience. It creates a strong visual effect of 3D objects floating in air, even when the image is not stereoscopic. This is a first step in the direction of a truly volumetric walk-through display.

The addition of 2D and 3D interaction significantly expands the possibilities for applications of the FogScreen. It provides advantages over other displays by allowing unhindered multi-user collaboration, providing new interface potential, and subtly reinforcing the presence of virtual objects in the physical environment.

Unlike many volumetric displays, the pseudo-3D FogScreen can be very large and does not restrict the user from "touching" the objects, leading to a more immersive experience. Engaging interaction with immaterial 3D objects can be supported in a variety of ways, as we demonstrated at SIGGRAPH 2005 Emerging Technologies [37].

The FogScreen has proven itself as a captivating display technology that immediately generates interest and excitement in the audience. The feedback from our SIGGRAPH 2005 demonstration was unanimously enthusiastic about the dual-sided, interactive experience. Since then, our demos of head tracking and stereoscopy have been met with similar enthusiasm about the further improved perception of 3D imagery.

Acknowledgments

We wish to thank Karri Palovuori at Tampere University of Technology and Marc Breisinger, Nicola Candussi, Cha Lee, Rogerio Feris, Jason Wither and Alberto Candussi at the Four Eyes Lab, University of California at Santa Barbara. Special thanks to Andy Beall and Matthias Pusch from Worldviz LLC.

This research was in part funded by grants from NSF IGERT in Interactive Digital Multimedia, the Sweden-America Foundation, the Academy of Finland, the EC under contract FP6-511568 3DTV, TTY tukisäätiö, Emil Aaltonen Foundation, Tekes, Alfred Kordelin Foundation, Jenny and Antti Wihuri Foundation, and the Finnish Cultural Foundation and its Pirkanmaa Branch. Additional funding came from a research contract with the Korea Institute of Science and Technology (KIST) through the Tangible Space Initiative Project.

References

1. DiVerdi S, Rakkolainen I, Höllerer T, Olwal A (2006). A Novel Walk-through 3D Display. *Proc. SPIE Electronic Imaging, Stereoscopic Displays and Applications XVII*, San Jose, CA, USA, January 15–18, 2006, SPIE Vol. 6055, 605519, pp. 1–10.
2. Palovuori K, Rakkolainen I (2004), FogScreen. U.S. patent 6,819,487 B2. November 16, 2004.
3. Rakkolainen I, Palovuori K (2002), A Walk-thru Screen. IS&T / SPIE Electronic Imaging 2002, Proc. of Conference on Projection Displays VIII, San Jose, CA, USA, January 23-24, 2002, pp. 17-22.
4. FogScreen Inc. (2007), http://www.fogscreen.com. January 2007.

5. Pastoor S, Wopking M (1997). 3-D displays: A review of current technologies. Displays, Volume 17, Number 2, 1 April 1997, pp. 100-110.
6. Halle M (1997). Autostereoscopic displays and computer graphics. Computer Graphics, ACM SIGGRAPH, 31(2), May 1997, pp. 58-62.
7. Azuma R, Baillot Y, Behringer R, Feiner S, Julier S, MacIntyre B (2001). Recent Advances in Augmented Reality. IEEE Computer Graphics and Applications, 25(6):24-35, Nov-Dec 2001.
8. Sutherland I (1965). The Ultimate Display. Proc. of IFIP Congress 1965, Vol. 2, pp. 506-508.
9. Rekimoto J, Matsushita N (1997), Perceptual Surfaces: Towards a Human and Object Sensitive Interactive Display. In Workshop on Perceptual User Interfaces (PUI'97), October 1997, pp. 30-32.
10. Ullmer B, Ishii H (1997), The metaDESK: Models and Prototypes for Tangible User Interfaces. Proc. of the ACM UIST'97 Symposium, pp. 223-232.
11. Leibe B, Starner T, Ribarsky W, Wartell Z, Krum D, Singletary B, Hodges L (2000), The Perceptive Workbench: Towards Spontaneous and Natural Interaction in Semi-Immersive Virtual Environments, Proc. of IEEE Virtual Reality 2000, March 2000, New Brunswick, NJ, USA, pp. 13-20.
12. dnp Holo Screen (2007). DNP, http://www.dnp.dk/. January 2007.
13. HoloClear (2007). HoloDisplays, http://www.holodisplays.com/. January 2007.
14. Wilson A (2004). TouchLight: An Imaging Touch Screen and Display for Gesture-Based Interaction, Proc. of ICMI '04, 2004, pp. 69-76.
15. Hirakawa M, Koike S (2004). A Collaborative Augmented Reality System using Transparent Display. Proc. of ISMSE'04, 2004, pp. 410-416.
16. Olwal A, Lindfors C, Gustafsson J, Kjellberg T, Mattson L (2005). ASTOR: An Autostereoscopic Optical See-through Augmented Reality System. Proc. of IEEE and ACM ISMAR 2005, pp. 24-27.
17. Just PC (1899), Ornamental fountain. U.S. patent 620,592. March 7, 1899.
18. Sugihara Y, Tachi S (2000). Water Dome - An Augmented Environment. Proc. of the Information Visualization Conference, London, July 2000, pp. 548-553.
19. Aquatique (2007). Aquatique Show International, http://www.aquatic-show.com/. January 2007.
20. Fantasmic show. (2007) Disney, http://disneyworld.disney.go.com/wdw/entertainment/entertainmentDetail?id =FantasmicEntertainmentPage. January 2007.
21. Desert Rain project (2007), http://www.crg.cs.nott.ac.uk/events/rain/. January 2007.
22. Mee Fog Inc (2007)., http://www.meefog.com/. January 2007.
23. IO2 Technology LLC (2007). Heliodisplay, http://www.io2technology.com/. January 2007.
24. Rakkolainen I, Erdem T, Erdem Ç, Özkan M, Laitinen M (2006). Interactive "Immaterial" Screen for Performing Arts. ACM Multimedia 2006, Interactive Arts Program, Santa Barbara, CA, USA, October 23-27, 2006.

25. NEC (2007), WT-610 short throw projector. http://www.nec.co.uk/NEW_MultiSync_WT610.aspx. January 2007.
26. Infitec GmbH (2007), Infitec interference filters, http://www.infitec.net/. January 2007.
27. Bailey M, Clark D (1998). Using ChromaDepth to obtain inexpensive single-image stereovision for scientific visualization. Journal of Graphics Tools, 1998, 3(3), pp. 1-9.
28. DepthQ 3D projector (2007), Infocus, http://depthq.com/. January 2007.
29. WorldViz PPT 3D optical tracker (2007). http://www.worldviz.com/ppt/. January 2007.
30. Krueger W, Froehlich B (1994). The Responsive Workbench. IEEE Computer Graphics and Applications, 1994, Vol. 14, No. 3, pp. 12-15.
31. The Stanford 3D Scanning Repository (2007). http://www-graphics.stanford.edu/data/3Dscanrep/. January 2007.
32. Rekimoto J, Saitoh M (1999), Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments. Proc. of ACM CHI, 1999.
33. Taylor R (1998). VRPN: Virtual Reality Peripheral Network, http://www.cs.unc.edu/Research/vrpn/. 1998.
34. Baraff D (1989), Analytical methods for dynamic simulation of non-penetrating rigid bodies, Proc. ACM SIGGRAPH, 1989, pp 223-232.
35. Iwata H, Yano H, Fukushima H, Noma H (2005). CirculaFloor: A Locomotion Interface Using Circulation of Movable Tiles. Proc. IEEE Virtual Reality 2005, March 12-16, Bonn, Germany, pp. 223-230.
36. Kim S, Ishii M, Koike Y, Sato M (2000). Development of Tension Based Haptic Interface and Possibility of its Application to Virtual Reality. Proc. of ACM VRST, 2000, pp. 199-205.
37. Rakkolainen I, Laitinen M, Piirto M, Landkammer J, Palovuori K (2005), The Interactive FogScreen. A demonstration and associated abstract at ACM SIGGRAPH 2005 Program: Emerging Technologies, Los Angeles, CA, USA, July 31-August 4, 2005. See also http://ilab.cs.ucsb.edu/projects/ismo/fogscreen.html.