

# Consigalo: Multi-user Face-to-face Interaction on Immaterial Displays

Alex Olwal

PDC, KTH  
Teknikringen 14  
100 44 Stockholm, Sweden

alx@csc.kth.se

Stephen DiVerdi

UC Santa Barbara  
Santa Barbara, CA  
93106, USA

sdiverdi@cs.ucsb.edu

Ismo Rakkolainen

Tampere University of  
Technology  
33101 Tampere, Finland

ismo.rakkolainen@tut.fi

Tobias Höllerer

UC Santa Barbara  
Santa Barbara, CA  
93106, USA

holl@cs.ucsb.edu



**Figure 1. The Consigalo game is a highly interactive game for multiple users on an immaterial display. The system allows face-to-face interaction with minimal restrictions to user communication and movement, since the computer graphics are projected on a seemingly transparent sheet of dry fog that the user can reach and move through.**

## ABSTRACT

In this paper, we describe and discuss interaction techniques and interfaces enabled by immaterial displays. Dual-sided projection allows casual face-to-face interaction between users, with computer-generated imagery in-between them. The immaterial display imposes minimal restrictions to the movements or communication of the users.

As an example of these novel possibilities, we provide a detailed description of our Consigalo gaming system, which creates an enhanced gaming experience featuring sporadic and unencumbered interaction. Consigalo utilizes a robust 3D tracking system, which supports multiple simultaneous users on either side of the projection surface. Users manipulate graphics that are floating in mid-air with natural gestures. We have also added a responsive and adaptive sound track to further immerse the users in the interactive experience. We describe the technology used in the system, the innovative aspects compared to previous large-screen gaming systems, the gameplay and our lessons learned from designing and implementing the interactions, visuals and the auditory feedback.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

The Second International Conference on Intelligent Technologies for Interactive Entertainment (ICST INTETAIN '08). January 8–10, 2008, Cancun, Mexico. Copyright 2008 ICST. ISBN 978-963-9799-13-4.

## Categories and Subject Descriptors

H.5.2 [User Interfaces]: Graphical user interfaces, Input devices and strategies, Interaction styles

## General Terms

Design, Experimentation, Human Factors

## Keywords

Multi-user, interaction, dual-sided, FogScreen, immaterial, adaptive audio.

## 1. INTRODUCTION

The first interactive projection screen was Myron Krueger's Videoplace [13]. Since then, numerous interactive large display systems have been presented, such as the CAVE [4], Powerwall [20], ImmersaDesk and Infinity Wall systems [5]. For many of them, the user is at a distance from the screen. With others, the projection screen works as a touch screen. There has also been a great interest in developing games that are played on large displays over the years, with several Virtual Reality ports of first person shooters, such as Doom, Quake and Unreal Tournament [1][2][3][12] and sports games [9][11][17][26]. Nevertheless, no previous systems have enabled the gamers to reach or walk straight through the screen.

In this paper we detail the development and characteristics of Consigalo, a multi-user game that uses an immaterial dual-sided wall-size display to allow face-to-face collaboration in an intriguing and highly interactive experience.

Our motivation for the work was the goal of creating a casual, yet highly engaging user experience where we wished to accommodate a large number of possible simultaneous users. The contribution of this work is as a case-study into the design of user interfaces and gameplay mechanics that are particularly tailored to an immaterial display system. The result is an effective example of a new type of engaging, interactive multimedia content, with significant potential for entertainment and collaboration applications.

Our game is the first known game for an immaterial, dual-sided projection screen. The game, which was first demonstrated at ACM SIGGRAPH Emerging Technologies 2005, is based on an interactive system that combines an immaterial display with dual-sided projection, a 3D tracking system and responsive audio. It also features face-to-face interaction between the players, with interactive computer graphics appearing in-between them.

The system presents 3D graphics intriguingly floating in mid-air on a projection screen that one can walk through. The graphics are viewable from either side of the screen, and are interesting to look at, as well as to interact with. In contrast to other games developed for immaterial displays [14][23], we especially wished to leverage the capabilities of dual-sided visualization, multi-user interaction and an adaptive soundtrack.

The objective of the game is to catch shapes that are floating in the air and to collect them in the players' bins on the side of the screen. Given that the screen is immaterial, the display itself poses no restriction of movement – on the contrary, it encourages and invites real direct-manipulation with the users' hands reaching into the projected graphics. From our observations at ACM SIGGRAPH 2005 and other demonstrations, we find our system to be captivating and resulting in an intriguing gaming experience. As it is also easy to use, our design guidelines might apply equally well for other public applications, for example in science museums, trade shows, as digital signage, for advertising

and for various types of special events.

In Section 2, we provide a brief overview of the enabling technologies for our system; more detail on the various components can be found in [6] and [7]. We describe Consigalo's game play in Section 3. Section 4 focuses on the impact of the visual quality, and Section 5 describes the unique multi-user interaction enabled by our immaterial system. Sections 6 and 7 discuss visual and auditory feedback, followed by Conclusions and Future Work in Section 8.

## 2. THE INTERACTIVE SYSTEM

The main components in our system are:

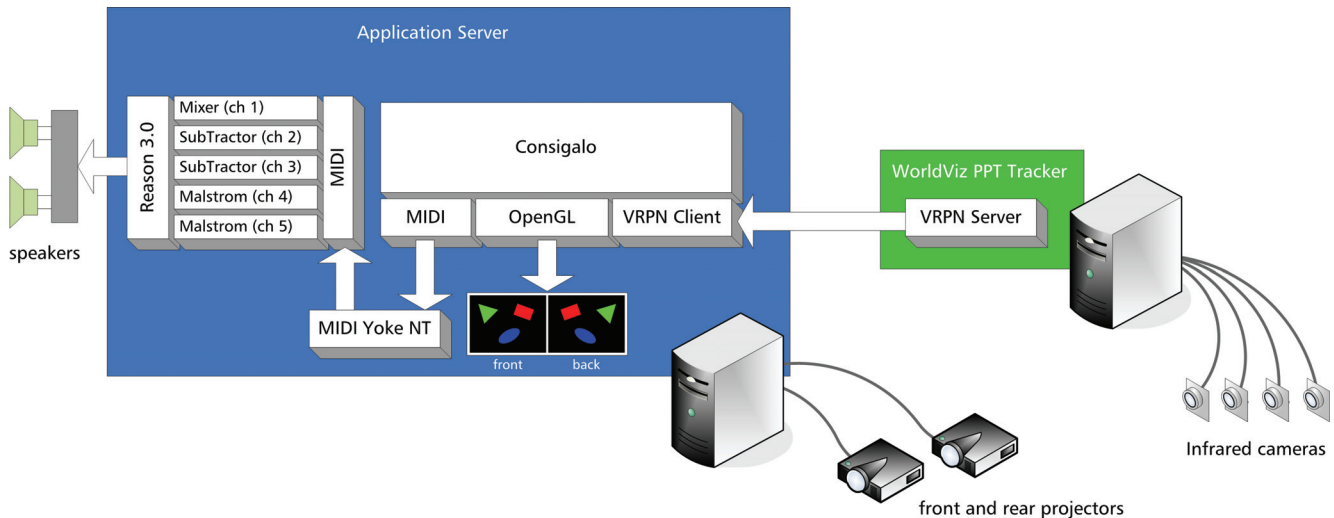
- The FogScreen display device
- Two projectors for dual-sided projection
- A 3D tracking system
- A software-controlled audio synthesizer

By combining these core technologies we obtain the unique characteristics of a highly interactive dual-sided immaterial display with engaging visual and auditory feedback. Figure 2 provides an overview of our system architecture.

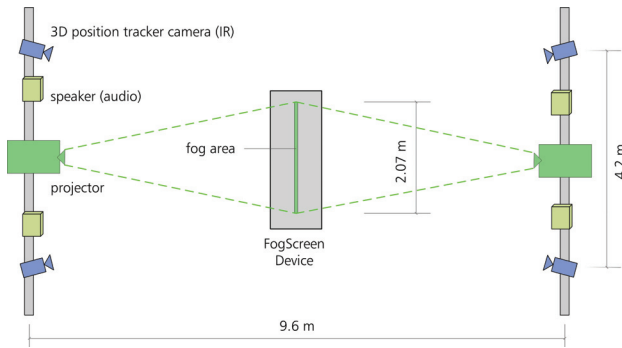
### 2.1 Dual-sided projection on the FogScreen

As the display component in our system we use the FogScreen [8], a large wall-sized display where a projection screen is created through controlled use of vaporized water.

It has primarily two unique capabilities that we make use of. First, its permeability allows users to casually walk and reach through the projection surface. Second, the fog diffuses light primarily in the forward-direction, making rear-projected imagery bright and front-projected imagery faint. This enables the combination of two projectors for individual, yet coordinated, imagery on both sides of the projection screen [6].



**Figure 2. System architecture. Consigalo communicates with the tracking server (which tracks the IR LEDs using four infrared cameras) using the VRPN protocol. It sends auditory feedback based on the user's interaction to a configured set of Reason 3.0 devices (music application we use to generate audio) using MIDI, via the MIDI Yoke NT patch driver. Computer Graphics are rendered from front and back perspectives in viewports that are being sent to two projectors, for dual-sided projection.**



**Figure 3. System setup, top view. Two projectors display visuals from each side of the FogScreen. Four cameras in the corners of the interaction space track the IR LEDs held by the users. Four speakers in the space provide the auditory feedback from the adaptive audio system.**

## 2.2 Rendering

The game is implemented as an OpenGL application using C++ on the Windows platform. The graphics are orthographically rendered in two 800×600 viewports (front and back perspective) in a 1600×600 hardware-accelerated fullscreen window. The two viewports are split by the two outputs of the graphics card, which displays one viewport on each projector.

## 2.3 Touch Screen Style Interaction

In order to support 3D interaction, we employ the WorldViz PPT Tracker [30], a tracking system that uses multiple cameras to track the 3D position of up to 8 infrared LEDs. We use four cameras that are placed in the corners of our interaction space, as shown in Figure 3.

Most of the 2D and 3D tracking systems that we tried with the FogScreen need modifications and even then can't reliably track the user's hand or a pointer near the FogScreen image plane – either due to visual reflections from the fog plane, or due to limited tracking range. The WorldViz PPT tracker, on the other hand, has no difficulties in giving reliable results near and on the image plane consisting of flowing fog.

The PPT tracking system is controlled by a separate PC that runs a VRPN [28] server. Our application implements a VRPN client and connects over TCP/IP to continuously receive the 3D position of the currently visible LEDs.

To minimize the footprint of the markers, we constructed a custom 2×3 cm miniature version that could be casually held between two fingers or worn as a ring, resulting in an experience that mimics bare-hands interaction.

Using the 3D position data, we detect when the player's hand is within a few centimeters of the plane of the display and use that as an indication of “touch”.

## 3. CONSIGALO

Consigalo translates to “Obtain it” from Spanish and is a minimalistic game, designed to take advantage of the FogScreen's dual-side and reach-through properties, and especially to explore its interactive capabilities when combined with a 3D tracking system.

From the top of the display, 3D shapes slowly fall down, and the objective of the game is to pick and collect shapes of a specific color.

### 3.1 Picking Falling Shapes

Each player controls a colored sphere, *the cursor*, which can be moved in the 3D space around the screen. As different colored shapes fall down from the top, rotating around their axis, the player reaches into the screen and picks shapes that match the cursor color. The interaction takes primarily place on the surface, where the 3D position of the cursor is mapped to a projected 2D location.

### 3.2 Dragging Shapes to Goal Areas

Shapes that are touched by the cursor will stick to it and this allows the user to drag shapes on the surface. The user can pick one shape at a time or collect multiple ones, before dragging them to one of the goal areas. The goal area is a colored vertical strip located on the left or right side of the screen that matches the player's color. After successfully having delivered shapes to the goal area, the player receives points for them and starts picking new shapes.

### 3.3 Ensuring Contact and Avoiding Collisions

While dragging the shapes, the player must maintain contact with the screen and avoid touching or being hit by falling shapes of another color. If the player stops touching the display, all accumulated shapes will start falling again, so the player needs to quickly pick them again. If the player accidentally touches a shape of another color, the players of that color will receive points, while all currently dragged shapes will be lost.

### 3.4 Interacting on Both Sides

Since the screen is made of dry, flowing fog in a thin plane, the user can easily move *through* the display to interact with screen content on both sides, even as the player is dragging shapes. Screen content is projected from two sides, such that one can go through the screen to see what the shapes look like from the other side, or play against or together with people on the other side of the screen, as well as with people on the same side.

### 3.5 Feedback

While interaction with the content happens in 2D on the display surface, we use the recovered 3D position of the cursor to provide visual and auditory feedback.

Besides visual cues, we also provide subtle changes in the music as feedback to the player, in addition to traditional sound effects. The volume of the player's individual beat is, for example, increased as the cursor approaches the plane of the screen, and individual audio tracks are intensified as the player successfully picks and drags shapes. We discuss the adaptive sound in more detail in Section 7.

## 4. VISUALS: RENDERING AND IMAGE QUALITY

The FogScreen strives to minimize the turbulence in its flow of vaporized water particles, but currently there is still a considerable difference between its image quality and that of a high definition projection system. With a focus on interactivity, this did not pose a problem in our scenarios, but we found it



**Figure 4. Multi-user, face-to-face interaction, and walk-through in Consigalo. Left and middle: Two users manipulate objects on opposite sides of the display surface. Right: A user moves through the screen to the other side.**

important to take this into account in order to provide a consistent and compelling experience.

To avoid confusion, we ensure that we render clear, high-contrast computer graphics in colors that are easy to distinguish and restricted ourselves to the use of white, black (perceived as transparent on the FogScreen), red, blue, green and yellow. We also make sure the interface does not rely on small text or fine details, to accommodate the lower image quality.

Ambient air currents affect the display as well, by causing the screen to deflect, making it not a smooth plane [6][7][19]. It would be reasonable to expect that users moving around the screen would exacerbate this problem, but in our experience, the strength of air currents from waving arms or walking is insignificant compared to other disturbances. The draft of air that follows a user walking quickly through the screen will cause a momentary deflection that lasts around half a second as new fog falls from the screen.

A potentially larger problem is that a limb penetrating the screen causes the fog below to become disturbed until the limb is removed. The design of Consigalo reduces the impact of this effect as objects fall from above, directing user focus to portions of the screen above where they are currently interacting.

The game supports any set of VRML models that are stored on disk. In the SIGGRAPH Emerging Technologies demonstration we used a set of animal VRML models. Each new object is assigned a random shape, color (one of the player colors), rotation speed, falling speed and size.

The player's cursor is rendered as a large sphere that is positioned on the projection surface based on the 2D projection of the marker held in the user's hand. As feedback to the player, the cursor color is highlighted if a shape is being dragged.

As mentioned in Section 2.2, the graphics are rendered from two perspectives, rendering from the back of the scene and from the front of the scene. Thus, a 3D model gets projected from opposite sides onto the projection surface, resulting in a "sandwiched" billboard that is floating in the air.

The dual-sided rendering is important, as it provides individual imagery, which can be coordinated to provide consistent visuals on both sides. Applications can take advantage of this to adapt the content for one of the two viewing directions. It can be used to ensure geometrical consistencies, such as rendering correctly faced text on both sides, or to provide different information for different users on opposite sides of the screen.

One of our tested scenarios is a wall-sized card game, where we intended to display shared open cards and the deck to all users whereas players on opposite sides of the screen should only see their own private cards. In general, this works well, but we identified some cases where information may "leak" to the other side. This depends primarily on the user's viewing angle, the projection angle, and the nature of the content (especially brightness and contrast). Under normal use, where the user stands in front of the screen and the projector is only slightly off-axis, it is practically impossible to see the content on the other side. However, if the user looks at the screen from a large viewing angle, it can be possible to see faint graphics from the other side. By coordinating the content on both sides, one could address this type of leakage by presenting high contrast information in the same location on the user's side.

## 5. MULTI-USER INTERACTION

One of the greatest advantages of the FogScreen is its ability to accommodate a large number of simultaneous users, with less impact on human-human communication than conventional wall-sized display systems. (See Figure 4.) Such systems typically employ a large *physical* display, forcing the users to face away from each-other, such that they all look at the screen when they collaborate on the shared display. Systems where the users can be on opposite sides of a transparent display have also been proposed, but while they provide the possibility of face-to-face collaboration, they impose physical restrictions that limit communication and movement in space.

### 5.1 Face-to-Face Interaction

A high level of immersion and interactivity can be achieved with engaging tracking systems. Two compelling examples can be found in recent video game systems. The Sony PlayStation 2 EyeToy [27] employs a camera to bring the player into the game. The players can see themselves on the TV with overlaid game elements that they can indirectly manipulate. The Nintendo Wii gaming system [18] on the other hand, uses motion-sensing controllers to stimulate a physical gaming experience. But a similar disembodiment is seen in both cases, since the player's motions typically control an avatar interacting with other players' avatars. As the number of players increase (e.g., four players playing tennis on the Nintendo Wii), the space in front of the display becomes limited and the horizontal alignment of the people make inter-human interaction complex.

With the FogScreen, two users can stand opposite each other with computer-generated imagery floating in-between them. They are able to talk freely, see each other through the display, and can have either common (like in Consigalo) or personal views of the

screen. We find this to be a very important aspect of the system, as it adheres to the social protocols of human-human communication. If the users are near the screen, in some cases they may block some light from the projectors. This can largely be avoided by placing the projectors at a sufficient height.

Face-to-face interaction has been used for a long time in, e.g., tele-presence, video-conferencing [10] and in sports over a distance [17], but has been harder to support when the users are co-located and collaborate in a shared interactive graphics environment. For a player, being able to directly look at and see the opponent manipulate imagery in front of them, as well as directly reacting to the opponent's motion, is of course a more natural experience, like in sports or board games.

With two persons on each side, four people can comfortably discuss and interact with the content on the screen and at this point, the computer graphics starts to blend in as a natural extension of the physical environment. The screen does not isolate players, but rather creates a collaborative, mixed reality visualization space. Although it is currently possible in many interfaces for four people to interact simultaneously, the ability for users to spread themselves around the interface, rather than crowding in front of it, is more comfortable and more closely emulates collaboration around a table or shared workspace. The same interface concessions as are made with traditional displays to enable 4 or more users could be used with the FogScreen to comfortably enable 8 or more.

In Consigalo, this feature allows players to be on opposite sides while they collect falling objects, as shown in Figure 5. This applies both to players on the same and opposite teams, such that the possibility of face-to-face communication can be fully exploited for collaborative or competitive purposes. During our public demonstrations of Consigalo, we found that users tended to interact on the same side of the screen by default, but once they saw that face-to-face interaction was possible, they found it to be a more natural experience. Indeed, standing across a playing field from one's opponent more closely mimics most competitive games and sports, and enables human-human interaction through body language that can be difficult to communicate when players are looking at a screen instead of at one another. Such a configuration makes it easier to demonstrate and instruct during game play, and also helps increase one's awareness of the other player's actions without demanding additional attention that would distract from the screen content.

## 5.2 Reach- and Walk-through

As previously discussed, the dual-sidedness of the system would have little practicality if not for the important immaterial properties of the projection surface. These properties effectively alleviate the communicative aspects of the interaction since a user can at any given time walk and reach through the projection surface, and look at or speak to a person on the other side of the surface.

By being able to reach through, more robustness is achieved and various types of props that actually may go through the projection screen can be supported. As a comparison one may consider the many reports on property damage caused by excited activity in the Nintendo Wii gaming system and a defective first version of the controller wrist strap. Many incidents were initially reported where, for example, the controller had flown out of the player's hand and into the TV screen [29], causing significant damage.

The physical limitations of traditional display technologies, as well as newer ones (such as most commercial volumetric displays) can thus be avoided as the screen itself is no longer a physical barrier, with a direct result that the user's work space is greatly expanded. While in our system, a flying controller may strike a user instead, the possibility is opened for interactions that take place in or behind the screen plane, which were not possible before. For example, rather than manipulating a 3D object via indirect manipulation in front of a screen, direct manipulation of the space the 3D object "occupies" is possible. This could allow the user to hit a ball with a physical tennis racket, baseball bat or golf club by swinging straight through the display. In Consigalo, this is taken advantage of by allowing users to grab the falling objects directly, increasing the intuitiveness of the interaction through accommodation.

## 5.3 Degrees of Freedom

For our purposes, it was critical that the system could support multiple users simultaneously and recover, at least, the 2D position of the user's hand, with an indication of whether the display was touched or not.

We also found it important that players could quickly join and leave the interaction, as we intended to support a high throughput of people. For these purposes, it might seem tempting to use a bare-hands interaction approach, but from our experiences with a laser range finder we found that it could not track hands very near



**Figure 5. Consigalo's immaterial, dual-sidedness and transparent qualities, allows it to accommodate more people than normally possible in traditional large display systems.**

to the projection surface, as the fog would often accidentally trigger a false positive. From an interaction perspective, it is also limiting to not be able to detect a hover state, as appropriate feedback cannot be shown to the users before they have already touched the surface. However, the biggest problem with the tested version of bare-hands interaction was, for our purposes, the problem of tracking multiple hands. In our experience with groups of many users, many of them will try to touch the display simultaneously and the situation can quickly become unmanageable.

A physical prop on the other hand, inherently provides delegated control and limits the number of simultaneous users in a controlled fashion – when a user is done, the prop is simply passed on to the next person.

To ensure the sporadic and intuitive nature of the interaction, a small and wireless form factor is desirable, which we were able to achieve using the tracking system described in Section 2.3 and a custom marker design. It supports 3DOF position tracking of a current maximum of 8 small LEDs. In our application there was no need for detecting the orientation of the device, although this could be achieved by combining multiple 3DOF sensors in a rigid structure or through the use of hybrid sensors.

It is always desirable to track the user’s interaction with as many degrees of freedom as possible and the technology that one chooses typically depends on factors such as intrusiveness, robustness, price and performance. One must however obviously be careful with how those degrees of freedom are mapped in the user interface.

In our case we had a tracking system that worked in a large area outside of the dual-sided 2D projection surface and it was initially tempting to provide the user with 3D control over the content on the screen, even though the visuals were monoscopic. Not surprisingly, our experiences with the “Elastic Surface Deformer” [6][7][19] confirmed that it was difficult to control a 2.5D orthographic visualization with full isomorphic 3D control. In that application, a user controls a spherical cursor to manipulate an elastic 3D mesh of a human head. Despite feedback in form of highlighting and visualizing, the intersection between the sphere and the mesh, it was somewhat difficult for users to assess the cursor’s distance from the rendered model. We did not wish to use 3D glasses for stereoscopic rendering, given our goal of an experience that was as unencumbering and sporadic as possible.

In Consigalo, we thus provided interactive control only when the user’s hand was very close to the screen. We found that the 3D position data could still be effectively used for actuation and feedback, as described below.

## 5.4 Actuation

In an interactive system, a clutching mechanism is a powerful and desirable property that allows for a wider range of control. In our system we were specifically interested in tracking the cursor at all times, while being able to have the user control when to grab or release a shape.

For the “Elastic Surface Deformer” and “The Virtual Forest Tour” [6][7][19] we developed a custom 3-button joystick with software mapped functionality. While the device was useful and flexible in its configurability, the meaning of the different buttons has to be learned. We were looking for something more straightforward and

less complex for our multi-user application. We were especially interested in something that did not require extensive explicit instruction.

Given our 3D position tracking, we decided to emulate a touch screen, with additional hovering capabilities. We triggered a “grab” when the user’s hand was very close to the screen. We found this mapping reasonable as it mimicked the actual touching of the surface, while it allowed us to provide feedback at all times.

From our informal observations at SIGGRAPH, the touch interface was easy to comprehend and, as expected, required no explicit instructions. Anyone who had watched another person interact with the system could quickly start playing the game.

## 6. VISUAL FEEDBACK

While commercial touch screens exist in many public places, they all tend to have very simple button-based user interfaces, where the user interacts with the system through discrete presses of large buttons on the screen. In such systems interactivity is low and so is the required feedback.

In a system with a much richer sensing and higher level of interactivity, it is important to provide adequate feedback in response to the user’s actions. This is especially relevant for actions with a subtle or less obvious meaning to the user.

We note that the spherical cursor could have been used in a richer way to help users assess the impact of their actions. The highlighting of the cursor upon grabbing an object might be a bit too subtle, and sometimes the users didn’t realize they lost contact with the screen while dragging.

Instead, it would have been interesting to investigate the use of various cursors that had a number of clear states which the system could smoothly interpolate between during the course of interaction. One example would be to use a hand cursor that had different states. For instance, the hand could be open as it was not touching the screen, have the palm facing up when it was touching the screen to indicate that falling shapes could land on it, and use a closed grabbing hand when shapes were held. As the user started moving into the screen, the hand would smoothly transition from open hand to palm up. When the hand was holding objects, the closed fist would smoothly open up as the user moved away from the screen, to indicate that the grabbed objects were about to be dropped. Feedback could of course also be provided through animated symbols and icons that briefly appear on the screen, but that could also be distracting and introduce unnecessary clutter, so careful design and consideration would be needed.

While we have reduced the direct use of tracking data to the surface of the screen, we still have full 3D position data of the user’s hand and this information can be successfully utilized for ambient effects, such as environmental changes or audio. This encourages exploration of the system and maintains the interest of a more experienced user, as system behavior that might seem random is discovered to be user-influenced. We describe how we developed a responsive audio sound track in the next section.

## 7. AUDITORY FEEDBACK

In order to add another dimension to the experience, we provide an adaptive soundtrack. As briefly mentioned in Section 3.5, various aspects of the soundscape change in response to the user’s actions. The goal of this feedback is to further improve the

accommodative aspect of Consigalo’s user interface and we find it particularly useful in a system where the passive haptic feedback provided by a physical interaction surface is missing. Since users are unfamiliar with the FogScreen at first, they are often unsure of how to approach the device and the types of interaction that are expected of them. The adaptive audio feedback encourages the user during their initial experimentation by creating more energetic or pleasing sounds when the user’s actions come closer to causing effects in the game state. This way through natural playing with the system, users are subtly guided to learn the correct way of using Consigalo on the FogScreen.

### 7.1 Software

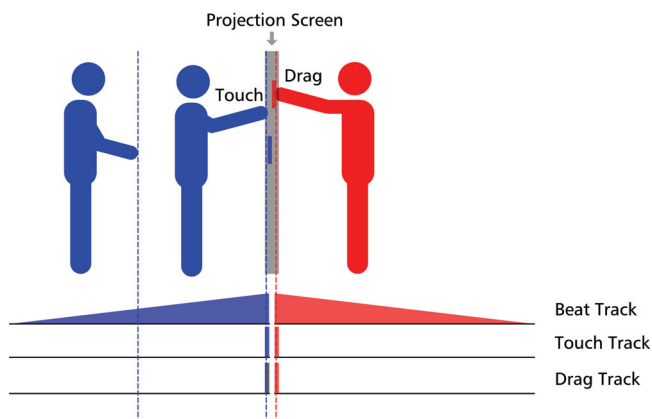
The soundtrack was designed using Reason 3.0 [21], an advanced music program that emulates a rack of hardware audio components, such as synthesizers, drum machines, signal processors, sequencers and mixers. Many aspects of the program can be automated through built-in functionality, or controlled from external sources through the MIDI protocol (Musical Instrument Digital Interface), a standard for communicating event messages between computers, electronic instruments and other equipment [15].

### 7.2 Audio Design

Each of the eleven stereo audio channels we used was created using a combination of the components in Reason.

For instance, the output of a synthesizer configured to sound like an organ, is routed through a delay unit, to a distortion unit. The output of each such combination is then routed to a virtual multi-track mixer where balance and volume can be controlled for each of the stereo channels used. The mixer finally outputs a combined stereo audio signal, which is transmitted through the PC sound card to the speakers on the left and right sides of the projection surface.

After the sound of each channel has been customized, one can either play that sound directly by sending a note from an external source, or preprogram a loop that plays continuously using the built-in Matrix step sequencer. In the latter case one can still



**Figure 6. Side view. As the user gets closer to the screen, the beat track intensifies. Upon touching the screen the touch track is activated. When the user grabs and starts dragging an object, the drag track is enabled as additional auditory feedback.**

control all parameters of the sound depending on how one chooses to map input to controls and parameters in the emulated audio devices.

### 7.3 Responsive Sound Tracks and Effects

Two of the stereo audio channels are always on and not modified, providing an ambient background sound through two sequenced loops. Each player has an individual drum machine track, a touch track, and a dragging track, which are played to the left and right speaker for the players respectively. The drum machine track gives an indication of the player’s proximity to the projection screen; the closer the player’s tracked hand is, the louder the beat gets. When the projection screen is being touched, the player’s touch track is activated. The dragging track is activated when the user is touching and dragging a shape, and it intensifies with the number of shapes that are currently being dragged. The drum machine track is a programmed beat, and the touch and drag tracks use a programmed pattern on the Matrix step sequencer to play a melody on the connected instruments. (See Figure 6.)

There are also different sound effects that trigger when the user picks up a shape, gets points for dragging shapes to the goal area, or accidentally collides with the other player’s shapes.

There are thus two background stereo tracks, three tracks for each user (beat, touch, drag), and three shared sound effects in the current version of the game, resulting in a total of eleven stereo audio channels.

### 7.4 Programmatic Audio Control

To control the various parameters in the Reason devices from our C++ application, we use the MIDI protocol. In Reason, MIDI is used to control the emulated devices using an external device. Data comes in through a hardware MIDI interface, for instance receiving values from knobs and sliders on a MIDI controller connected to the MIDI port on the PC sound card, or from a MIDI keyboard that connects over USB, for example. MIDI is less commonly used on the same computer to communicate between programs; in those cases other standards are typically used (such as ReWire [24], for example).

We avoid the use of external hardware by employing MIDI Yoke NT [16], a “software MIDI patch cable” that emulates a hardware MIDI interface and allows data sent to an input port to be automatically passed to one of its output ports.

After installation, MIDI Yoke appears in the system as a hardware MIDI interface and we configure Reason to listen to its output ports. Our application can now use the low level MIDI API in Windows to send MIDI messages from C++ to MIDI Yoke’s input ports, with the messages getting relayed to Reason, providing us with full control of our Reason devices.

The advantage of MIDI is its generic, well-established protocol for controlling parameters and functionality in the target application, independent of whether it is a software or hardware device, as long as it supports MIDI.

“Looped devices” that continuously play a sequence, such as the drum machine, the touch and drag tracks, are controlled through the Mixer device in Reason. By continuously sending MIDI messages that control the level for each of those channels, we can manipulate the intensity of the tracks. For example, in our C++ application, we map the distance between the screen and the

player's cursor to a 7-bit (0-127) MIDI level. The MIDI message is sent to the MIDI Yoke application, which in turn affects the volume slider in the Mixer device on the beat channel.

The sound effect devices are triggered by being directly sent a MIDI note message for on (127), followed by off (0) after a short 50 ms delay. This can be considered the equivalent of manually hitting a key on a MIDI keyboard.

As we only use two output audio channels in this version of the system, adaptive audio is only available for two players. It would however be straightforward to extend that as there is virtually no limitation in how many devices or mixers Reason can handle. The constraints are set by the audio hardware. Through the use of spatialized 3D sound, a virtually unlimited number of individual sound channels could be achieved, each of which is directed to the location of the respective player.

## 8. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented our Consigalo gaming system, which combines several novel technologies into an enhanced gaming experience. Consigalo specifically focuses on a sporadic, unencumbering experience, where the visual presentation and interaction with computer graphics floating in mid-air blends into the physical environment.

Highly interactive collaborative gaming is empowered by the dual-sided nature of our system. Users can stand opposite each other, with the graphics in-between them, and still freely engage in the tasks with minimal restrictions to their movement or communication. A 3D tracking system provides full 3D position data for a hand-held LED. We have considered and experimented with various approaches for the interaction and our experiences led us to the final design of a touch screen style interface, where feedback is provided in the full 3D space, but actuation happens on the projection surface only. The 3D position data also controls user-specific audio feedback that responds to the individual user's actions.

The design of Consigalo's interface was carefully tailored to take advantage of the strengths of the FogScreen display, and to limit the impact of its weaknesses. The engaging game play we wanted to enable called for direct manipulation in a physically intuitive manner, with accommodative feedback to overcome users' initial trepidation about an unfamiliar device. The game uses high-contrast visuals and avoids relying on fine visual details for play mechanics. Finally, face-to-face interaction is leveraged to improve the social aspect of user collaboration and competition. Altogether, Consigalo is an effective realization of the possibilities for interactive content using the FogScreen, enabling a new variety of multimedia entertainment.

It would be interesting to extend the system with spatialized audio, as we can use the 3D location of each player to create an even more engaging experience. It is also worth noting that the display allows sound to travel freely through the screen. There are no reflections, occlusions, or resonances, as is the case with normal projection screens.

The image quality on the FogScreen is reasonable for entertainment applications, especially with a careful design of the visual content, but we are working on improving it significantly. Our vision for the coming years is that of a high-definition, dual-sided, immaterial, interactive information display.

## 9. ACKNOWLEDGMENTS

This work was supported in part by the Sweden-America Foundation, and the Academy of Finland under grant 114006.

## 10. REFERENCES

- [1] CAVE Doom. <http://www.cavernus.org/DOOM/Doom.html>. November 2007.
- [2] CAVE Quake II. <http://www.cavernus.org/QUAKE/Quake.html>. November 2007.
- [3] CAVE Quake Arena. <http://www.visbox.com/cq3a/>. November 2007.
- [4] Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A. 1993. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. *Proc. SIGGRAPH '93*. ACM Press, New York, NY, pp. 135-142.
- [5] Czernuszenko, M., Pape, D., Sandin, D., DeFanti, T., Dawe, G. L., and Brown, M. D. 1997. The ImmersaDesk and Infinity Wall projection-based virtual reality displays. *SIGGRAPH Comput. Graph.* 31, 2 (May. 1997), pp. 46-49.
- [6] DiVerdi, S., Rakkolainen, I., Höllerer, T. and Olwal, A. 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.
- [7] DiVerdi, S., Olwal, A., Rakkolainen, I. and Höllerer, T. An Immaterial Pseudo-3D Display with 3D Interaction. *Three-Dimensional Television: Capture, Transmission, and Display*. Editors: Haldun M. Ozaktas and Levent Onural. Springer, Heidelberg, 2007. In press.
- [8] FogScreen Inc, <http://www.fogscreen.com/>. November 2007.
- [9] Hämäläinen, P., Ilmonen, T., Höysniemi, J., Lindholm, M., and Nykänen, A. 2005. Martial arts in artificial reality. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Portland, Oregon, USA, April 02 - 07, 2005)*. CHI '05. ACM Press, New York, NY, 781-790.
- [10] Ishii, H., Kobayashi, M., and Grudin, J. 1993. Integration of interpersonal space and shared workspace: ClearBoard design and experiments. *ACM Trans. Inf. Syst.* 11, 4 (Oct. 1993), 349-375.
- [11] Ishii, H., Wisneski, C., Orbanes, J., Chun, B., and Paradiso, J. 1999. PingPongPlus: design of an athletic-tangible interface for computer-supported cooperative play. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: the CHI Is the Limit (Pittsburgh, Pennsylvania, United States, May 15 - 20, 1999)*. CHI '99. ACM Press, New York, NY, 394-401.
- [12] Jacobson, J., Le Renard, M., Lugin, J., and Cavazza, M. 2005. The CaveUT system: immersive entertainment based on a game engine. In *Proceedings of the 2005 ACM SIGCHI international Conference on Advances in Computer Entertainment Technology (Valencia, Spain, June 15 - 17, 2005)*. ACE '05. ACM Press, New York, NY, 184-187.

- [13] Krueger, M., Gionfriddo, T. and Hinrichsen, K., VIDEOPLACE – An Artificial Reality. In Proc. of CHI'85. ACM 1985, pp. 35-40.
- [14] Lugmayr, A. and Piirto, M., Implementation of "Looney": designing an interactive game for children on FogScreens. ACM SIGGRAPH 2006 Research posters, Boston, MA, USA, 30 July - 3 August 2006. 47.
- [15] MIDI. <http://www.midi.org/>. November 2007.
- [16] MIDI Yoke NT. <http://www.midiox.com/>. November 2007.
- [17] Mueller, F., Agamanolis, S., and Picard, R. 2003. Exertion interfaces: sports over a distance for social bonding and fun. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Ft. Lauderdale, FL, USA, April 05 - 10, 2003). CHI '03. ACM Press, New York, NY, 561-568.
- [18] Nintendo Wii. <http://www.wii.com/>. November 2007.
- [19] Olwal, A., DiVerdi, S., Candussi, N., Rakkolainen, I., and Höllner, T. An Immaterial, Dual-sided Display System with 3D Interaction. Proceedings of VR 2006 (IEEE Virtual Reality Conference 2006), Alexandria, VA, Mar 25-29, 2006, pp. 279-280.
- [20] Powerwall. <http://www.lcse.umn.edu/research/>. November 2007.
- [21] Propellerheads Reason 3.0. <http://www.propellerheads.se/>. November 2007.
- [22] Rakkolainen, I., Laitinen, M., Piirto, M., Landkammer, J. and Palovuori, K. 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>.
- [23] Rakkolainen, I., Erdem, T., Utku, B., Erdem, Ç., and Özkan, M., Mid-air Display for Physical Exercise and Gaming. IEEE 3DTV Conference, Kos, Greece, May 7-9, 2007.
- [24] ReWire. <http://www.propellerheads.se/technologies/rewire/>. November 2007.
- [25] Rakkolainen, I. and Palovuori, K. 2005. Laser scanning for the interactive walk-through FogScreen. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (Monterey, CA, USA, November 07 - 09, 2005). VRST '05.
- [26] Sidharta, R. and Cruz-Neira, C. 2005. Cyclone UpperCut, a boxing game for an immersive environment. In Proceedings of the 2005 ACM SIGCHI international Conference on Advances in Computer Entertainment Technology (Valencia, Spain, June 15 - 17, 2005). ACE '05. ACM Press, New York, NY, 363-364.
- [27] Sony Playstation 2 EyeToy. <http://www.eyetoy.com/>. November 2007.
- [28] Taylor, R. VRPN: Virtual Reality Peripheral Network, <http://www.cs.unc.edu/Research/vrpn/>. 1998.
- [29] Wii Damage. <http://www.wiidamage.com/>. November 2007.
- [30] WorldViz PPT 3D optical tracker. <http://www.worldviz.com/ppt/>. November 2007.