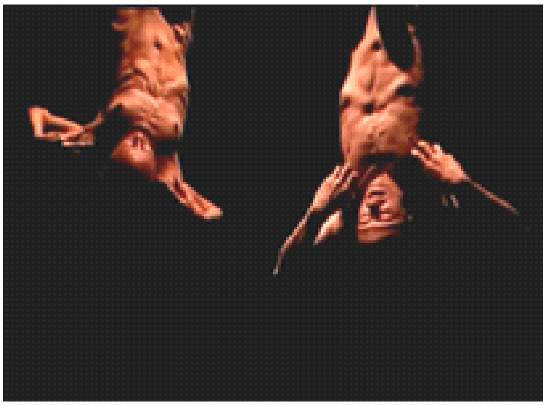
poster



coexistence

Rebecca Allen and Eitan Mendelowitz
Department of Design | Media Arts , University of California Los Angeles
rallen@arts.ucla.edu

Abstract

Recent advances in mobile, wearable technology expand our ability to communicate remotely and connect to networked environments while still participating in the physical world around us. As technology allows us to more easily participate in simultaneous realities, aesthetic issues take on a central role. There is an opportunity to create new artistic experiences that blend both virtual and physical worlds and to create unique interfaces that more fully involve the senses.

This paper briefly describes an interactive art installation, titled Coexistence, that blurs the boundaries between physical reality and virtual reality, between biological life and artificial life, between performer and audience.

Coexistence questions our fundamental understanding of reality while exploring new forms of interface and experience design using networked wearable technology. This work includes innovative approaches to interaction, performance and immersion utilizing see-through head mounted displays and panoramic video. An important component involves a unique sensory interface that integrates breathing sensors and haptic feedback.

Keywords: Mixed Reality; wearable computing; machine vision; head mounted display; experimental performance

Year the Work was created: 2001

Project Partners: Emergence Team: Eitan Mendelowitz/Damon Seeley; USC Integrated Media Systems Center; Performers: HannahSim/MarkSteger/Osseus Labyrint.
This work is partially funded by Intel Corporation and Interaction Design Institute Ivrea.

1. Context

Mobile devices such as cell phones and PDAs provide ways to access information or communicate remotely while still engaging in the physical reality around us. As interfaces improve, it will seem natural to move in and out of cyberspace as we go about our daily lives. Portable digital devices allow us to interact in the physical world while coexisting in other places and virtual environments.

A visual counterpart to this involves new developments in head mounted displays. Head mounted display technology is becoming more compact and less expensive. In the near future we will be able to wear glasses that allow us to see more than the world in front of us. These glasses will serve as screens with variable transparency, allowing us to become immersed in a virtual environment, to interact with video, or to view any combination of the physical world mixed with a virtual world.

As mobile technology becomes more pervasive, interface and interaction design must be re-examined. We are moving from keyboard and mouse to sensing technologies that more fully engage the body.

In addition, people are becoming familiar with the notion of a distributed presence - a distributed self that exists in many worlds and plays many roles at the same time. Emerging digital technology allows us to distribute our presence and interact with others over multiple channels. It also presents us with increasingly sophisticated re-presentations of artificial life in which objects, both virtual and physical, can behave in ways that suggest “aliveness”.

As technology allows us to more easily participate in simultaneous and mixed realities, aesthetic issues take on a central role. There is an opportunity to create new artistic experiences that blend both virtual and physical worlds and to create unique interfaces that more fully involve the senses.

Exploration into these areas led to the creation of an interactive art installation titled Coexistence.

2. Coexistence

Coexistence is a new work by Rebecca Allen in collaboration with her Emergence team at University of California Los Angeles, the Integrated Media Systems Center at University of Southern California and the performance group Osseus Labyrint.

This work blurs the boundaries between physical reality and virtual reality, between biological life and artificial life and between performer and viewer. It questions our fundamental understanding of reality through new forms of creative experience.

Using see-through head mounted displays and panoramic video, we explore innovative approaches to interaction, performance and immersion. An important component involves a unique sensory interface that integrates breathing sensors and haptic feedback.

To experience this work a person wears a head mounted display (HMD) that includes a position tracker to track head orientation, audio headphones and an attached camera to see the world in front of you. A computer vision system has been developed to track physical objects. This allows virtual and physical objects to appear to be in the same physical space.

The interface involves breath and tactile feedback. Using breath sensors, performers and artificial life forms come to life and respond to the participants' input. A hand-held controller provides haptic feedback as a form of communication between multiple participants or with virtual characters.

3. Mixed Reality

Fuelled by the rapid growth of networked digital technology and the parallel growth in mobile computing, there has been an increasing interest in techniques that combine real and virtual environments to create mixed realities.

Between the extremes of physical reality and virtual reality lies the spectrum of mixed reality. Unlike virtual reality that replaces the physical world, mixed reality research aims to develop technologies that mix computer generated virtual objects and characters with the real world.

In Coexistence a person experiences multiple realities that encourage interaction. Wearing a head mounted display the participants will at times see through the attached camera and view the space and physical objects around them. With the breath sensor one can bring computer generated images to life. These virtual objects appear in the physical space and seem to be connected to physical objects. For instance, a virtual object grows out of a physical container and eventually surrounds it.

The breath is used as a metaphor for breathing life into an object. By blowing into the breath sensor a person affects his or her environment. When a person blows into the sensor a computer generated image will come alive and metamorphose. Whenever one participant blows into the sensor the other participant will "feel" their breath through a haptic force-feedback controller.

The breath is also used as an interface to emphasize the connection of our physical body to mixed reality. Breathing is the ultimate expression of our physicality. In Coexistence we not only maintain our simultaneous awareness of physical and virtual space but of our own physicality intermingled with virtual life.

4. A New Platform for Performance Art and Interactive Experiences

During the Coexistence experience one will be transported to a strange, engaging, immersive performance experienced from a unique viewpoint. Instead of sitting with an audience watching a performance, you feel as if you are on stage, standing in the center of a unique performance. One is able to look around in any direction as the performers approach you and perform all around you.

This unique interactive performance utilizes a 360-degree panoramic video system developed by Professor Ulrich Neuman and his research team from the Integrated Media Systems Center at University of Southern California. The panoramic video can be viewed through a HMD. During playback the HMD head-tracking device allows participants to turn their heads freely to observe the desired portions of the panoramic scene.

We are using this panoramic video system to explore new forms of art and performance and to reconsider relationships between performer and audience. When viewing the performance with a HMD it seems as if you are in the center of a performance that takes place all around you. The performers are directing their actions to you, an audience of one.

During the performance one can use the breath sensor to activate and modify the actions of the performers. This presents a new type of immersive and interactive experience, connecting an audience member to a performance in a way that has not previously been explored.

The performance is choreographed and performed by Hannah Sim and Mark Steger, founders of the performance group Osseus Labyrint. They were invited to collaborate because of their extreme focus on the physical appearance and movements of the human body, an appropriate counterpart to a virtual experience.

Osseus Labyrint performances are otherworldly, very physical but bizarrely alien. As described by a reviewer, "...the death defying duo explores the outer limits of what the human body can achieve... Their joint image, twin-like and androgynous, is a science-fiction nightmare. As man and woman they create the illusion of a spare, post-human race."

5. Technical Description

This artistic work utilizes a proprietary software system called Emergence [2]. The Emergence Engine was designed to be modular and highly flexible. The engine's flexibility is achieved through the use of an innovative scene graph formulation. The Emergence scene graphs have more components than standard scene graphs such as those employed by Java3D [4], and VRML97 [1]. The Emergence scene graph does not merely store hierarchies of geometric primitives and their attributes, the scene graph contains complex data structures called topologies, assets, elements, and controllers. Descriptive information can be associated with each of these structures in the form of properties. The Emergence scene graph allows for the easy integration of new technologies, as they are required. The following paper gives an overview of the Emergence scene graph structure and explains how it has been used for Coexistence to add mixed reality technologies including head mounted displays, real-time video mapping, "breath" sensing, and machine vision systems.

6. Topologies

The topology is the basic building block of the Emergence scene graph. There are different types of topologies. Topology types include such types as "Immovable Object," "Moveable Object," and "Creature." These type definitions help to define characteristics of the topology. Each topology has a name and properties that are used to describe the topology. Many of a topology's properties describe physical characteristics such as location, mass, and elasticity. Topologies may have an arbitrary number of children. A topology's location in the virtual world is relative to that of its parent. As a result, any geometric transformation applied to a parent is inherited by its children. The emergence scene graph must be a tree (i.e. except for the root, every topology must have one and only one parent). There is one special topology called "World" which serves as the root node for the emergence scene graph.

Topologies, by themselves, are static and invisible - they have no geometry, no motion, and no behavior. Topologies may be associated with assets, elements, and controllers. It is these structures that provide the scene graph with substance.

7. Assets

Assets are structures that store data for topologies. Assets include such things as texture maps used to texture geometry, wave files used as sounds, and network sockets for TCP/IP network messages. For our new mixed-reality installation, titled Coexistence, we created a new asset that contains a real-time video stream that can be acquired from any Video For Windows or DirectShow compatible camera; currently we are using a FireWire web-cam. Like all assets, any member of the emergence scene graph can use the video stream asset.

8. Elements

In addition to assets, topologies can have Elements with which they are associated. Elements are responsible for the Emergence Engine's input and output. They define how topologies look, sound, and what they can sense from the physical world. Example output elements include geometry, particle systems, sound, and interfaces for haptic controllers. Each element of every topology is called once per frame at which point the element is responsible for its own output (usually performed though OpenGL rendering calls). Because every element is responsible for its own render, arbitrary new elements can be created and added to the scene graph without having to modify the Emergence Engine as a whole. Elements can have properties that effect how the element is rendered. Properties can be set either interactively though the Emergence Engine's GUI or through the emergence scripting. In addition elements can be disabled or enabled in the same manner.

New output elements added to the emergence engine for Coexistence include a real time video stream, and a panoramic video display system. The real time video steam acquires video frames from the video stream asset and texture maps them onto a plane. The location and orientation of the frame in the virtual world can be set through the element's properties.

The panoramic video element was created using DLLs provided by the Integrated Media Systems Center at USC based on their panoramic video system [3]. The panoramic video is captured by five separate camera feeds that are stitched together to form one ultra-wide image representing a 360-degree field of view. When active, the panoramic video element determines the participant's point of view (using information from the head mounted display input element) and maps an appropriate segment of the stitched image onto the inside of a cylinder. The result is an immersive panoramic video experience where the user can look around 360 degrees and up and down. Properties control such things as qualities as the current frame being rendered and the frame rate. Like all properties these values can be set through the Emergence GUI or the Emergence scripting language.

Like output elements, input elements are called once per frame. Input elements contain properties that represent the state of input devices. These properties can be displayed by the emergence GUI, read by a script, or accessed by other members of the scene graph. Examples of input elements include keyboards, joysticks, and game-pads. For Coexistence we added input elements for a head mounted display (HMD) and a "breath" sensor. For each frame the HMD input element sets its pitch, yaw, and roll properties by querying the motion trackers that are built into the HMD.

In order to explore the interface modality of breath the Emergence team created a "breath" sensor. The breath sensor is a standard microphone coupled with signal processing performed by the breath sensor input element. Each frame the breath sensor analyzes the waveform captured by the microphone. When a participant blows (even gently) on a microphone this can be distinguished from a loud sound by enumerating the samples above a set threshold. While load sounds are characterized by a high average amplitude they still have a wave structure and therefore tend to have a lower number of samples above the given threshold than one's breath. The breath sensor's properties include threshold, amplitude, and the number of samples above the threshold.

The addition of these complex elements point to a strength of Emergence's scene graph formulation. The real time video stream, panoramic video, HMD, and breath sensor were all added without any modification to the rest of the emergence system. In addition the use of properties by the scene graphs allows for control of these new elements without any extensions to the emergence scripting language or the emergence GUI.

9. Controllers

Controllers are responsible for creating change in Emergence worlds. A controller's sole task is modifying the properties of other components in the scene graph. There are many different types of controllers. Physics controllers implement the virtual world's physics including gravity, friction, and collisions. Other examples include animation controllers that control the key frame animation of models, scripting controllers that run Emergence scripts, and behavior controllers. Behavior controllers implement a behavior based artificial intelligence system that allow for the creation of autonomous situated agents in Emergence worlds [2].

For Coexistence we added two new controllers. The first controller modifies the orientation of the (virtual) camera in response to changes in pitch, yaw, and roll of the HMD as detected by the HMD element. This allows the Emergence participant to move his/her head and look around the virtual world (or a panoramic video sequence).

10. Machine Vision Controller

The second controller enables mixed (or augmented) reality. The machine vision controller examines the video stream from the video stream asset, identifies pre-defined targets placed in the room, and, from the location of the targets as observed in the video images, determines the physical camera's location in the room. The controller then uses the location of the physical camera to set the location of the virtual camera in the virtual world. Moving the camera in the real world corresponds to the same movement of the virtual camera in the virtual world. Using this controller we are able to allow objects in the virtual world to augment the real time video image of the physical world. For example, we can place a virtual flower in a physical pot.

The Emergence Engine was designed with artists in mind. It is part of the Emergence Lab's mission to make tools accessible to artists by making sure they run on standard consumer PCs. In addition, for the Emergence Engine to be immersive it must run at high frame rates (at least 20 frames per second). To these ends, the machine vision controller must run highly efficiently.

In order to minimize the vision controllers computational load constraints are placed on the targets. All targets must be of solid predefined colors. These particular colors cannot appear elsewhere in the installation space.

In addition, targets are required to be within a certain range of the camera. Too far away and the target may be ignored, too close and the targets position may not be accurately determined. By imposing these constraints, the machine vision system no longer has to look at each and every pixel of the video image (a time consuming task). Instead the vision controller is able to sample a fraction of the image and quickly find any targets that may be present.

Once a target is located, the vision system performs a more refined sample in the region of the target to determine the targets size in the image (and hence its distance from the camera). The location of the targets in the image and distance of targets from the camera are used to determine the location of the physical camera through basic geometry. This location is then used to set the location of the virtual camera.

Summary

Coexistence is an artistic experience that questions our fundamental understanding of reality by blurring the boundaries between physical reality and virtual reality, between biological life and artificial life and between performer and audience.

This work includes innovative approaches to interaction, performance and immersion utilizing see-through head mounted displays and panoramic video. An important component involves a unique sensory interface that integrates breathing sensors and haptic feedback.

In Coexistence we envision a future in which our distributed presence exists in multiple realities creating, perhaps, a richer life experience.

References

[1] ISO/IEC 14772-1:1997. Information technology -- Computer graphics and image processing -- The Virtual Reality Modeling Language (VRML) -- Part 1: Functional specification and UTF-8 encoding.

[2] Mendelowitz, E. 2000. The Emergence Engine: A Behavior Based Agent Development Environment for Artists. Proceeding of the 12th Innovative Applications of Artificial Intelligence Conference.

[3] Neumann, U., Pintaric, T., Rizzo, A. 2000. Immersive Panoramic Video. Multimedia: Proceedings of the 8th ACM International Conference on Multimedia, 493-494.

[4] Sun Microsystems, 2000. Java 3D API Specifications. http://java.sun.com/products/java media/3D/forDevelopers/J3D_1_2_API/j3dguide/index.html