The Pond: Interactive Multimedia Installation

Sean Follmer
HCI Group, Department of Computer Science
Stanford University
Stanford, California, USA
sfollmer@stanford.edu

Chris Warren
CCRMA, Department of Music
Stanford University
Stanford, California, USA
cdwarren@stanford.edu

Adnan Marquez-Borbon
CCRMA, Department of Music
Stanford University
Stanford, California, USA
adnanm@stanford.edu

ABSTRACT
In this paper, we describe an interactive multimedia installation, The Pond. It is a human-scaled, proximity-based non-linear sequencer. A ceiling-mounted projector fills the floor with light. As each player enters the pool of light, a circle surrounds them and follows them as they move. Each player carries a ROCK, an accelerometer-based wireless controller. A quick shake of the ROCK triggers the player's identifying sound and, like throwing a rock into an actual pond, causes their circle to expand. As the circle spreads to touch another player's circle, it triggers a convolution of both their sounds resulting in a complex melodic sequence.

Keywords
Interactive, installation, multimedia, wireless, convolution, Max/MSP, Open Sound Control, Eyepatch, VBAP, Flash, computer vision, UPIC.

1. INTRODUCTION
The metaphor behind the project is simple: when a rock is dropped into a pond, the waves created by the disruption of the surface move concentrically and collide with other waves. Trying to avoid typical new instrument design paradigms, we conceived of an extremely simple interface that allows the user to quickly learn how to operate it and that brings an element of entertainment to the installation. By shaking the ROCK the user sends a signal to the host computer that produces a predesignated sound and creates the concentric waves in the visual projection. Computer vision allows the users to be tracked and the waves to emanate from them. When the waves collide with the other players they in turn automatically trigger their own sound which will be convoluted with the initial sound. The result is a constantly changing convolutional 'melody' with each shake as the first player's sound passes through all the other players' sounds. Vector-based amplitude panning is used for proper spatial placement of the collisions. Because the installation is based on a human scale and its interaction informed by proximity, it creates a sort of game that effectively engages the users participation within the installation and with the rest of the users.

2. SYSTEM
2.1 Overview
The ROCK is a controller that houses a small circuit board comprised of a vibration sensor, a microprocessor chip, and a wireless radio transmitter. At the other end is a radio receiver that is connected via serial-to-USB port to a computer that feeds the information to Max/MSP and Flash for the audio and video processing. Computer vision software tracks the users and allows for a spatially based visual and musical experience. These software applications are connected using Open Sound Control.

2.2 Rock Hardware
The ROCK itself is not an instrument, but rather a control interface for the software. It is made of 16 laser cut plywood layers of various shapes and sizes that are glued together to produce the final form which houses the electronic components. Inside, a small circuit board consists of a piezo vibration sensor, an ATMega32 chip [1], and an Xbee-PRO [2] wireless transmitter. The vibration sensor converts the motion of the ROCK into voltages that are converted into digital signals via the ATMega32. Afterwards, they are forwarded to the wireless transmitter and finally sent to the computer.

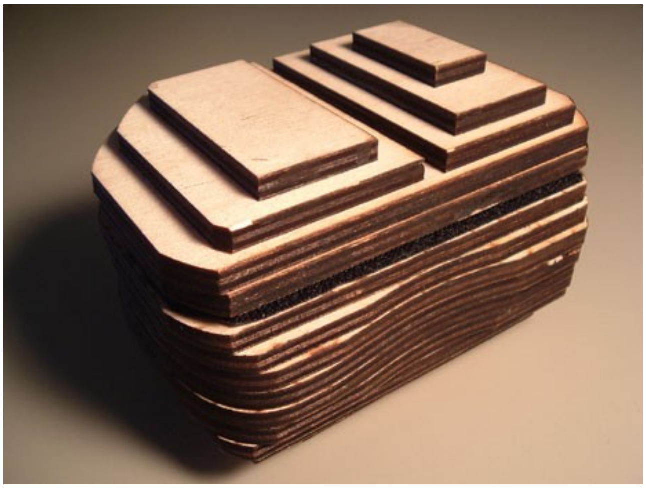

Figure 1. The ROCK

2.3 Transmitter Electronics
The front line of the circuit consists of a piezo film accelerometer and the ATMega32 chip in which the A/D conversion takes place. This conversion, thresholding and data transmission was coded in C. Once the signal has been digitized it is then sent to the Xbee for transmission to the computer. The Xbee is a low cost cost device for wireless use, using the Zigbee wireless standard. It is a powerful and simple device ideal for this sort of applications. The programming of the Xbee was done with the manufacturers software. Each ROCK can send one of four intensity values depending on the intensity with which it is shaken. In addition to the ROCK's particular identification number in order to sort out each message.

2.4 The Receiver
The receiver is another Xbee that has been configured as a network receptor. Since the Xbees only handle serial data, a serial-to-USB converter board is needed. Our choice was the Sparkfun FT232R breakout board.

Figure 2. The Transmitter Circuit

2.5 Computer Vision

To track the location of the users we use an web camera with the infrared (IR) filter replaced by a visible light filter. Two large flood lamps bathe the stage in IR. This allows for projection in the visible light spectrum and without interfering with the tracking in the IR spectrum. Eyepatch [3], a computer vision tool, tracks the users based on light intensity (humans absorb IR creating dark spots). Eyepatch does blob detection and sends out the location of the blobs over TCPIP in XML.

Figure 3. The Installation and Visuals

2.6 Visuals

The visuals were generated in Flash and then projected onto the floor from above on top of the users. Flash receives location data from Eyepatch and draws a circle around the person to let them know they are being tracked. Whenever a user shakes their ROCK, a signal is sent to Flash and an expanding circular wave is generated. Collision detection is done and if there is an intersection between a wave and a different person a signal is sent to Max/MSP. Users can see the waves projected on the floor as seen in figure 3 and can dodge them or run into them to create interesting timing.

2.7 Sound

The sounds themselves were chosen to be "water-like" in nature in order to be consistent with our metaphor. However, they are not samples of water-related sounds but rather were synthesized using the UPIC computer system. The intensity level triggered by each ROCK is used as a dynamic control for the volume of the triggered sound. After receiving a message from the transmitter, the receiver sends a message via OSC to Max/MSP to trigger the player's sound. When a wave collides with a person, Flash sends a message with the location and identities so Max/MSP can convolve the two sounds and VBAP [4] can properly localize them.

3. RELATED WORK

Much work has been done in projected multimedia experiences. Scott Snibbe's Boundary Functions [5], an installation that projects divisions between people on the floor inspired our work.

4. FUTURE WORK

We would like to add a small pager motor to the ROCK for an additional layer of haptic feedback in order to augment the interactivity. Additionally, we would like to port the visual tracking into other visual environments in order to increase the visual processing speed. Video of the work in progress is available at http://www.alloyelectric.com/thepond.html

5. ACKNOWLEDGMENTS

We would like to thank Michael Gurevich, Colin Gilboy and Sasha Leitman for their help and support.

6. REFERENCES

[1] ATMega32 Microcontroller, http://www.atmel.com.

[2] Xbee Wireless Transmiter, http://www.maxstream.com.

[3] Maynes-Aminzade, D., T. Winograd, and T. Igarashi. Eyepatch: Prototyping Camera-based Interaction through Examples. UIST: ACMSymposium on User Interface Software and Technology, 2007.

[4] Pulkki, V. Spatial Sound Generation and Perception by Amplitude Panning Techniques. Dissertation for the degree of Doctor of Science in Technology, Department of Electrical and Communications Engineering, Helsinki University of Technology. 2001.

[5] Snibbe, S. *Boundary Functions.* 1998.