Bounce, Duck, Pogo, Dive

Pranesh Joshi[1], Eunyee Koh[1], YoungJoo Park[1], Randy Sappington[2], Andrew Webb[1]
Department of Computer Science[1] + Department of Psychology[2]
Texas A&M University
College Station, TX 77843 USA
pranesh@tamu.edu, {eunyee, awebb}@cs.tamu.edu, sappism@hotmail.com, yjoo9317@csdl.tamu.edu

ABSTRACT
A person's presence in an environment is often determined by physical action. Emotional presence can be delivered by physical action as well. A person's movement can convey a specific meaning to other people without the use of overt linguistic expression. Gesture is one of various types of interactions that people use in a typical social setting. A growing volume of work seen in art installations employs the use of gestures to express meaning and presence. The work presented here is a sample of how movement can be used to manipulate people's physical and emotional expression in an interactive art installation.

1. MOTIVATION
A person's presence in an environment is often determined by physical actions. If a disruptive child is jumping up and down in a classroom where everyone else is quietly seated, that child's presence in the room compared to other children in the room is more distinct. In turn, a person at a basketball game sitting calmly in a seat while all the surrounding persons are standing up and cheering also has a more distinct presence in comparison.

In the previous examples, physical actions were used to determine the physical presence of a person, but physical actions convey emotional presence as well. The disruptive child jumping up and down could be extremely excited about the successful completion of a task. This emotional presence of excitement is passed along to others in the classroom by the physical act of repetitive jumping along with other factors such as facial expression and vocal sounds.

Our primary motivation for this artwork is this idea that physical actions affect the various forms of person's presence such as physical or emotional presence.

2. RELATED WORKS
Most of related work involves capturing a hand gesture to trigger the systems [1, 3, 4, 5]. "Puccini Set Designer" [1] is an installation utilizing state-of-the-art technology to enhance the viewer's experience of the work of Puccini. This installation takes the viewer's participation in the process of experiencing the art into consideration. A user can step into a space and point to different parts of surrogate pictures to trigger a particular set in the Puccini's opera. Garver *et al.* [3] developed a hand gesture-based interface system, which uses the Tlib image processing library, to construct a simple and efficient interface environment which allows users to control a mouse through pointing gestures and execute commands with simple arm movements. Lenman et al. [4] presented a project that studied the use of hand gestures to remote control of electronic appliances in a home entertainment system (e.g. TV sets and DVD players). Ometa *et al.* [5] also proposes a hand gesture-based direct manipulation interface that can be used for data transfer among information artifacts such as computers, projectors or printers. In another related work, named "living-room", Galantay *et al.* [2] developed an installation for exploring interactive, space-related aspects of augmented-reality (AR) in art and design fields like architecture, interior design, or scenographic design. Using ARToolKit and MAX/MSP/Jitter, the installation captures the position of the user and the orientation of the user's gaze. The user's actions in the physical environment are used as input to modify the virtual scene.

3. DESCRIPTION
The project is an interactive art piece that captures the physical actions of jumping and crouching in a responsive environment and provides a visualization of an artistic image display, which reflects the interaction between the person's presence and the virtual environment.

The artistic image display is 3x3 square grid of images in which each square of the grid is continuously cross fading between different images. The squares cross fade at various moments such that all the squares do not change at the same time. The next image for a square is chosen at random.

The 3x3 grid of the display is mapped in the physical world by a 3x3 grid on the floor of the physical environment. The display is projected on a wall such that if a person is standing in the 3x3 grid on the floor is looking at the display, the each square on the floor maps geographically to a square in the display. In other words, a person standing in top-left corner of the grid is affecting the square in the top-left corner of the display.

As a person walks from square to square along the grid on the floor, a yellow border highlights the corresponding square in the display. Since a person can only be in one square at once, only one square in the display is ever highlighted. The display also renders two icons inside a highlighted square. The top-left icon is an image of a lock that represents whether a square is latched or not (locked or unlocked). If a square is latched, that square will not cross fade until the person returns to the square and unlatches it. The bottom-right icon is an image of the standard fast forward button (two triangles pointing to the right). The fast forward button allows the person to select a specific image for a square by transitioning through the images quickly.

The latch and fast forward actions are activated by jumping and crouching, respectively. When a person enters a square, the person can either jump to toggle the latch state of the square or crouch to fast forward through the images. Fast forwarding is deactivated when a person stands up after crouching. The reason for this mapping is primarily due to physical constraints (i.e. gravity). Since fast-forwarding can occur for an indefinite amount of time, jumping did not seem a suitable action to activate the fast-forward option since a person can only stay airborne as long as gravity allows.

The images used in the display were selected based on aesthetic quality and a relation to either jumping or crouching or both. Some of these relations were more metaphorical than others.

4. TECHNICAL DETAILS

The hardware used for this system includes three cameras, one projector, one Apple Power Mac G5 computer with dual 2 GHz G5 processors running Mac OS 10.3, and a white backdrop cloth (see Figure1). The user will employ the 3×3 grid in the physical environment to provide input to the system. The projection on the wall is where the response of the system is shown. Of the cameras on the right, the one on the floor is used to sense the jumping action and the other, on top of the stand, is used to sense the crouching action. The third camera, which is mounted on the ceiling, is used to sense the motion of the user walking within the bounds of the grid. The white curtain is used to maintain a noiseless background.

MAX/MSP/Jitter is used to process the signals from the cameras. The video signal from the camera in the ceiling is divided into a 3×3 grid by Cyclops. Each cell in the grid is assigned a single color. A tracking point is assigned to each cell, in which the brightness (value) of that cell is monitored. The MAX patch will send out the cell number that a person is in according to the drop in the brightness of a cell. The signal from the camera on the right top is divided into a 10×10 grid by Cyclops. This camera is used to sense the crouching action. The same principles of tracking the brightness applies here, however, the system will only consider the data from three adjacent rows that captures the upper body of a person in the space. At any point in time, brightness of at least one point should be below the threshold or else, the system will consider that the person is crouching provided that the person is in the view of the ceiling mounted camera. The last camera on the right bottom functions similar to the camera on the right top except that it captures the feet of a person in the space. The system will consider a person jumping if all the tracking point values are above the threshold provided that there is input from the ceiling mounted camera indicating the presence of a person in the space.

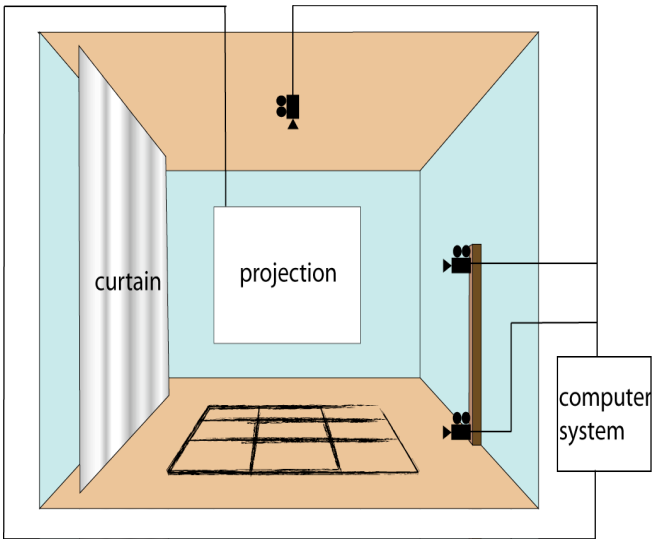

Figure 1: Flow Diagram

Java is used to implement a module that uses three values (x, y, z) from the three cameras mentioned above. The x is the cell that the person is currently in. The variables y and z will tell the java module about the action (i.e. jumping and crouching). This java module functions as a state machine, which will change the state of the visualization depending on the values of the three variables.

5. REFERENCES

[1] Flavia Sparacino, "Scenographies of the Past and Museums of the Future: From the Wunderkammer to Body-Driven Interactive Narrative Spaces, Proc MM '04, pg.72-79.

[2] Galantay, R., Torpus, J., Engeli, M., "living-room" Interactive, Space-Oriented Augmented Reality, *Proc MM 04*, pg. 64-71.

[3] Garver, R., Vision Based Gesture Recognition, College of Creative Studies, University of California, Santa Barbara. http://www.cs.ucsb.edu/~rgarver/gesture_current.pdf

[4] Lenman, S., Bretzner, L., Thuresson, B., Using Marking Menus to Develop Command Sets for Computer Vision Based Hand Gesture Interfaces, Proceedings of the second Nordic conference on Human-computer interaction, 2002, pg. 239-242.

[5] Omata, M., Go, K., Imamiya, A., A Gesture-Based Interface for Seamless Communication between Real and Virtual Worlds, 6th ERCIM Workshop on "User Interfaces for All", 2000