Fear Follower

Erfun Ackley*
University of Michigan
Department of Performing Arts Technology
Ann Arbor, MI 48109
ackleye@umich.edu

SinYu Deng†
University of Michigan
Department of Performing Arts Technology
Ann Arbor, MI 48109
dsudeng@umich.edu

Adam Schmidt‡
University of Michigan
Department of Performing Arts Technology
Ann Arbor, MI 48109
schmia@umich.edu

ABSTRACT

The *Fear Follower* is an interactive experience that integrates visual motion capture with robotics, dynamic visual projections, and immersive sound design. A visual projection of a large eyeball maintains an unwavering gaze on the user, tracking their every move around the room. Simultaneously, a robotic hand wheels itself towards the user, creating a dynamic and responsive connection between the installation and the individual. The culmination of these elements generates a strange ambiance that evokes a palpable sense of tension for the user. This synthesis of visual and auditory components is designed to immerse participants in a multi-sensory journey, pushing the boundaries of conventional interactive installations and redefining the relationship between users and their interactive environments.

Author Keywords

Motion Capture, Interactive Arts, Robotics, Audiovisuals

1. INTRODUCTION

Visual marker based motion tracking systems have long been used to track human motion for applications in entertainment and academic research, but these systems are increasingly being used in creative research to influence media such as audio or visuals. These systems can also be found in robotics research to study how robotic systems can be controlled, either in isolation or when sharing a space with humans to enable safe interaction. The Fear Follower explores the intersection of visual motion capture, interaction design, robotics, and audiovisual experiences. Users wear specialized glasses and gloves to facilitate interaction with the system. A projection of an eyeball gazes at the user, following their movements. A robotic agent points towards and follows the user to facilitate a dynamic interaction that creates a visual dialogue between the user and the robotic entity. An auditory aspect of the project adds an immersive layer, where the proximity of the robotic hand to the user results in a nuanced sound design, enhancing the sense of pressure and engagement.

The overarching goal of Fear Follower is to induce in the user a sense of being observed and followed. By employing visual motion capture, interactive projections, simple robotics, and dynamic sound design, Fear Follower aims to push the boundaries of conventional interactive installations, offering users a unique and compelling experience that blurs the lines between observer and observed.

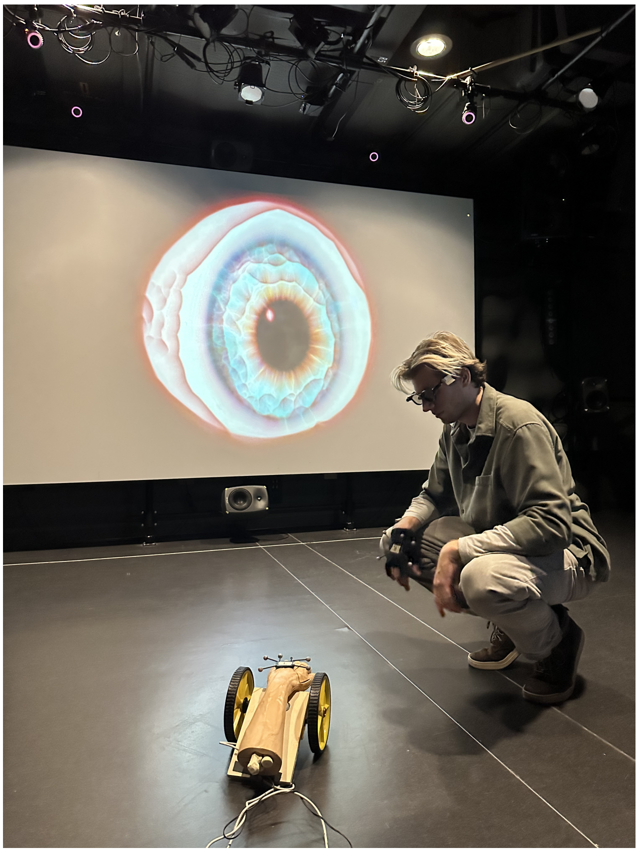

Figure 1: Fear Follower

*The responsibility for sound design was assumed by this author.

†This author undertook the responsibility of creating the animations and visualizations.

‡This author was responsible for both the construction and operation of the robot.

2. RELATED WORK

2.1 Follow-Me Robotics

There are numerous examples of systems designed to enable a robotic agent follow a user or subject, aptly named ‘follow-me’ systems [2, 1, 3]. Many systems use live Global Positioning System (GPS) data, computer vision, or a fusion of both to accomplish this task. There is a particularly marketable use case in which flying drones are used to follow individuals for recording video in an automated and hands-free way for extreme sports such as surfing, skiing, or speedboarding. Tuan et al’s research explores how to fuse GPS and video feed data to create a robust follow-me system that combines the long-term accuracy of GPS tracking and the short-term accuracy of visual tracking [1].

Visual-GPS sensor fusion works well for outdoor, large scale, body-centric systems, but visual marker-based tracking systems can output absolute position data in real-time at a higher resolution. Researchers have used these motion capture systems to control and stabilize robots in the past, providing robust closed loop control of a robotic agent rather than open-loop state estimation [6, 5]. Since the position of many points on these robots are being monitored at a high frequency, the robotic control algorithm can take into account the current position of the robot to enable precise, accurate, and stable movement control in real time.

2.2 Immersive Audiovisual Experiences

Granular Synthesis has emerged as a pivotal tool in one of the author’s endeavors to craft intricate sound textures and captivating soundscapes. The capacity to govern the synthesizer using diverse datasets, sourced from a spectrum including microphone-derived acoustic tracking, gyrometers, accelerometers, and visual motion capture systems, bestows upon the user a profound sense of engagement, thereby enhancing utility and immersive experiences.

The pursuit of an immersive audio environment has been standing as a primary objective, acknowledging the capacity of sound to serve as a potent stimulant within installations and human-computer interactions. As elucidated in Darin E. Hughes’ paper, ”Integrating and Delivering Sound Using Motion Capture,” the author notes, ”In fully immersive environments, sound provides a truly 360-degree, 3D sense. Humans can detect sounds all around them, in front or behind, as well as above and below head level. Eyes cannot see through walls or around corners, but ears can hear sounds that are occluded by such obstructions, thus providing a unique source of information that would otherwise be unavailable” [4].

3. USER INTERACTION

The user’s position is tracked when wearing a pair of glasses with attached motion capture markers. The projected eyeball follows the user’s position, constantly gazing at them. At the same time, a robot hand follows the user. As the robot hand gets closer, the soundscape becomes more tense. When wearing motion-tracked gloves, the amount of hand movement will modulate the intensity of the eyeball while also affecting the speed of the robot hand, increasing the overall sense of pressure in the room. All of the motion-tracked components of the system can be seen in Figure 1.

The utilization of 6 degree-of-freedom rigid bodies, coupled with our avoidance of creating full body AIM models, facilitates the seamless interchangeability of glasses and gloves among users. This adaptability allows different individuals to experience the interactive system.

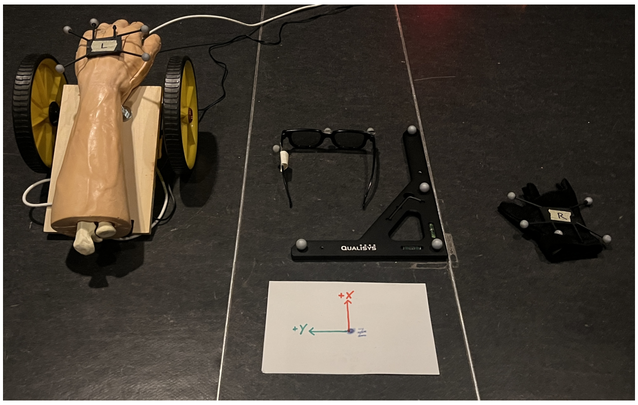

Figure 2: The physical components of Fear Follower

4. SYSTEM DESIGN

4.1 Signal Flow

A Qualysis visual motion capture system captures the position of carefully placed reflective markers. Numerous markers are placed on a pair of glasses that can be worn on a user’s head, a glove that can be work on user’s hand, and a small two wheel robot. Several markers are placed on each of these three subjects in unique arrangements so that the Qualysis Track Manager (QTM) Software can determine their Cartesian coordinates (X, Y, and Z) within the defined space as well as their angles of rotation (roll, pitch, and yaw) relative to the room.

The QTM Software streams Open Sound Control (OSC) Data over a User Datagram Protocol (UDP) port to a laptop on the same network and is received in Max/MSP. This first computer performs the sound generation and some initial signal processing before relaying data to two other computers via OSC. The second computer generates the visuals of the experience using TouchDesigner, and the third computer controls the robot over via Max/MSP and a serial port.

4.2 Sound Design

For sound design, recorded sounds of objects with metallic and bright qualities were incorporated, facilitating the application of a Low-Pass Filter to regulate sound brightness. This strategic choice stems from the comparative ease of subtracting various characteristics and frequencies rather than adding them.

Sound manipulation involved a diverse array of techniques, including nuanced adjustments to playback rates and the application of granular synthesis. These processes were executed using SuperCollider, with subsequent recording and exportation of the modified sounds in .WAV audio format. Following this, Audacity was employed to further refine the sounds and introduce reverb. In the final phase, the .WAV files were imported into the Max patch, featuring two primary audio files: 1. A background sound reminiscent of subterranean water flow. 2. The processed sound is designed to control intensity.

Both sounds underwent low-pass filtering, where the cutoff frequency was modulated by the positional difference between the Robot Hand and the user. Proximity to the robot correlated with increased tension in the sound.

Drawing inspiration from Murray Schafer's 'The Tuning of the World,' the soundscape is conceptualized through three key elements: figure, background, and field. The figure represents the listener's focal point, surrounded by background elements, while the field encompasses the entire auditory environment, establishing a dynamic figure-background relationship [7]. The positional variance between the user and the robot, coupled with the application of low-pass filters, serves as the field. Notably, the two filters associated with each audio file exhibit a negative correlation. This implies that an increase in brightness in the figure corresponds to a darkening of the background, and vice versa. This discriminative action within the sound space holds the potential to elicit a stimulating effect on the user's auditory system [8].

4.3 Visual Design

In the visual design component of the project, TouchDesigner was utilized to provide visual feedback (Figure 3). The design aimed at achieving an artistic and mysterious visual effect, incorporating spherical and eyeball-like materials with a feedback mechanism. Through the use of OSC linkage, values derived from Max/MSP, such as the angle between the projection screen and the user or the user's position on the z-axis, were mapped onto the transformation of the spherical geography. Additionally, the magnitude of the user's hand velocity was mapped onto the color attribute of the sphere's edges, transitioning to a red hue.

The projection serves to amplify the sensation of isolation. These enlarged visual elements play a significant role in the overall design, evoking a haunting and solitary atmosphere within the interactive space. Their magnified presence aims to immerse users in an environment that provokes feelings of seriousness, loneliness, and weirdness, enhancing the overall emotional impact of the experience.

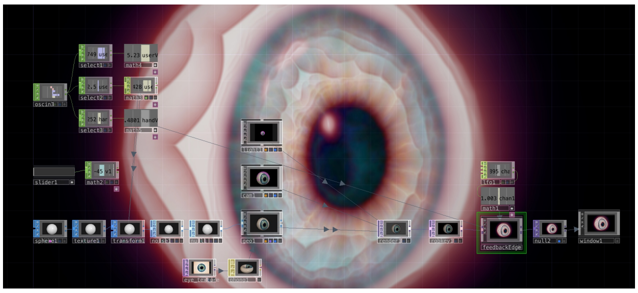

Figure 3: Visuals generated by TouchDesigner

4.4 Robotics

The robot hand is a two-wheeled robot that uses a continuous hobby servo motor (MG996) for each wheel. These digital servos are controlled via Pulse Width Modulation supplied by an Arduino Uno. Signal processing to create motor commands based on motion capture data was performed in Max/MSP and sent to the Arduino Uno over a USB port.

Motion capture markers were placed on both the user and the robotic agent. Rigid bodies were built to capture position and angle data from both the robot and the user. Using the absolute angle of the robot and the relative angle between the position of the robot and the user, a rudimentary control algorithm was created to minimize the error between these angles. This made the hand appear to constantly re-orient its heading towards the user. The robot has a very slow base speed but speeds up as the magnitude of the velocity of the user's right hand increases.

5. DISCUSSION

The authors had a specific aesthetic goal that subtly deviated from the users' interpretations of the project. All materials crafted or employed in the project were strategically chosen to enhance the likelihood of the project being perceived as a psychological horror experience. Despite these intentions, users noted the charm and character of the robot hand and how it resembled Thing from The Addams Family or the mutant toys from the Disney/Pixar animated movie Toy Story (Figure 4), introducing a comical element to the project. In general, users conveyed experiencing a compelling interaction during the demonstration, and their feedback challenge the authors to consider how the system could be improved if it were to be worked on further.

Figure 4: Mutant Toys from Toy Story

6. FUTURE IMPROVEMENTS

6.1 Visualization

Certain areas within the spatial setup pose challenges for accurate motion capture. The positioning of the user in specific corners can lead to errors in converting the folder angle accurately. Consequently, this issue affects the intended effect of the visual eyeballs accurately maintaining their focus on the user.

6.2 Robotics

The robot control algorithm used was rudimentary but functional. Only proportional feedback was used and although the robotic movement seemed stable, results could have been improved by adding integral or derivative feedback as found in many well-tuned feedback controllers for robots. However, we found the imperfect and less predictable motion endearing and attributed this to much of the robot's charm. To improve the system, we would have to choose whether our goal is to intimidate the user or lean into the endearing, pet-like quality it currently has. We imagine improving the control algorithm and increasing the robot's speed could yield more uncanny and creepy results like we originally intended, whereas adding randomness and less direct control in the algorithm could make the robot appear to have more exploratory agency and character.

6.3 Audio

In the realm of interactive simulations, the challenge faced by sound designers is distinct from their counterparts in film. Unlike a film sound designer who can meticulously refine each sound in the final product, a sound designer for interactive simulations confronts a perpetual absence of a fixed outcome. With every simulation run, a diverse array of outcomes and interactions reshapes the overall auditory presentation. To equip sound designers for interactive simulations with tools conducive to successful outcomes and effective final products, it becomes imperative to pioneer new capabilities. These innovations should enable designers to operate within an environment that mirrors the conditions experienced by the audience. In essence, the objective is to align the design process with the listeners' environment [4].

In the case of this project, users have the ability to move at different speeds and possess varying heights. In the computations conducted in the Max patch, all three spatial axes—X, Y, and Z—have been employed to measure the real-time difference between the user and the robot. Consequently, taller users experience more tense sounds and a wider range of dynamics. Omitting the Z axis from the equation could result in users of different heights encountering similar soundscapes, potentially aligning more closely with the sound designer's initial conception of the sound and its behavior.

Additionally, an audio interactive system capable of manipulating sound based on the data received from Motion Capture in real-time could be much more efficient if all audio manipulation and synthesis algorithms are implemented in a singular programming language and software. Unless all programs, such as Max/MSP and SuperCollider, can seamlessly communicate and exchange data without causing delays in the process. For the improvement of audio interactivity, all steps conducted in SuperCollider, Audacity, and Max/MSP will be solely executed in either SuperCollider or Max/MSP.

7. REFERENCES

[1] T. T. Do and H. Ahn. Visual-gps combined 'follow-me' tracking for selfie drones. volume 32, pages 1047–1060, 2018.

[2] W. Hedgecock, M. Maroti, J. Sallai, P. Völgyesi, and A. Ledeczi. High-accuracy differential tracking of low-cost gps receivers. 06 2013.

[3] S. S. Honig, T. Oron-Gilad, H. Zaichyk, V. Sarne-Fleischmann, S. Olatunji, and Y. Edan. Toward socially aware person-following robots. *IEEE Transactions on Cognitive and Developmental Systems*, 10(4):936–954, 2018.

[4] D. E. Hughes. Integrating and delivering sound using motion capture and multi-tiered speaker placement. In R. Shumaker, editor, *Virtual and Mixed Reality*, pages 179–185, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[5] X. Liang, M. Hirano, and Y. Yamakawa. Real-time marker-based tracking and pose estimation for a rotating object using high-speed vision. *Journal of Robotics and Mechatronics*, 34(5):1063–1072, 2022.

[6] M. Popescu, D. Mronga, I. Bergonzani, S. Kumar, and F. Kirchner. Experimental investigations into using motion capture state feedback for real-time control of a humanoid robot. *Sensors*, 22(24):9853, Dec. 2022.

[7] R. Schafer. *The Tuning of the World*. Borzoi book. Knopf, 1977.

[8] R. D. Shilling and B. Shinn-Cunningham. Virtual auditory displays. *Handbook of virtual environment technology*, pages 65–92, 2002.