Exploring Interactivity and Augmented Reality in Theater: A Case Study of Half Real

Michael R. Marner* University of South Australia
Sam Haren† The Border Project
Matthew Gardiner‡ Ars Electronica Futurelab
Bruce H. Thomas§ University of South Australia

Abstract

This paper presents a case study of Half Real, a live action, interactive theater show employing spatial augmented reality. Half Real is based on a murder investigation, where the live audience votes on how the investigation proceeds. Half Real immerses live actors into a virtual world projected onto the set. Using spatial augmented reality technology in theater brings new possibilities that would not be possible with simple video projection. Half Real models the set as a 3D virtual environment. Actors are tracked as they move about on stage, with the projected content responding to their movements. A real time AR system allows content to be generated procedurally. Half Real makes use of this by projecting vote options and results directly into the virtual environment. This paper describes Half Real, with a focus on the technology used to make the production possible. We describe the benefits of the techniques used and the challenges faced during production.

Keywords: Half Real, Spatial Augmented Reality, Interactive Theater, The Border Project.

Index Terms: J.5 [Computer Applications]: Arts and Humanities—Performing Arts; I.3.8 [Computing Methodologies]: Computer Graphics—Applications

1 Introduction

Who killed Violet Vario? That is the question posed to the audience of Half Real, an interactive theater show where the audience votes on how a murder investigation proceeds. Half Real utilizes Spatial Augmented Reality (SAR) [1] technology to immerse the actors in a projected virtual world (Figure 1). This use of SAR represents a departure from the simple video projection that has been used in performance art in the past. Half Real's projection system represents the set as a 3D environment. Actors are tracked in real time, allowing the projected content to react to actors' movements, and for virtual projected items to appear attached to actors. Half Real also takes advantage of this technology to support interactivity. The show puts the audience in control, voting on the path the story will take using wireless controllers.

This paper presents a case study of the development of Half Real. We provide a description of the technology used and how this technology made the production possible. We also discuss the lessons learned in developing a SAR system for performance art. The remainder of the paper is as follows: first, a brief discussion of previous work relating to audience interactivity, real-time procedural visual content for performance art, and spatial augmented reality is presented. Section 3 describes the SAR system developed for the show, how the stage is mapped as a 3D environment, and how projected content is created. Next, we discuss the interactive aspects of Half Real, and how the technology was used to support this. In Section 5 we describe how actors were tracked on stage throughout the show, and how this tracking data was used to drive the projected content. Finally, we provided an overview of the implementation details, discuss lessons learned from the production, and conclude with a look to the future of using SAR in performance art.

*e-mail: michael.marner@unisa.edu.au
†e-mail:sam@theborderproject.com.au
‡e-mail:Matthew.Gardiner@aec.at
§e-mail:bruce.thomas@unisa.edu.au

Figure 1: Half Real merges live action with a virtual world.

1.1 Creative Goals and Intentions

Half Real is the third theater work created by The Border Project which employed the audience interacting via the ZigZag controller system[1]. The Zigzags allow an audience to collectively make decisions to affect how the performance unfolds. The previous works (Trouble on Planet Earth and Escape from Peligro Island) were both inspired by the choose-your-own-adventure genre of books. In these works, the audience was able to choose the decisions of the protagonist, allowing them to navigate a pre-rehearsed narrative

[1]http://matthewgardiner.net/art/ZigZag

tree of possibilities.

Half-Real is a departure from the audience controlling narrative. Instead, the performance is an experience where the audience "investigates" a narrative world with the goal of determining which of three suspects murdered character Violet Vario. The aesthetic and mode of the work referenced contemporary gaming within the thriller genre, such as Heavy Rain[2]. The audience had three levels to investigate. The first was to investigate three pieces of evidence and their associated events; the second was to investigate two of the three suspects further (eliminating one); and the final was to interpret one of the characters' recent nightmares.

Half Real had some clear objectives for its video system design:

1. To integrate a Graphical User Interface (GUI) into the represented place onstage and embedded within the dramatic action of the scene. In previous interactive works, the narrative world was paused, and a vote sequence would occur where the options were posed to the audience, a screen would display the GUI options for the audience's choice and the vote would occur. For Half Real, the goal was to integrate the GUI within the visual world occupied by the performers, and for the GUI options to appear at the point an investigative pathway emerged dramatically in the scene, and remain present until the point of the vote.

2. As the narrative investigations continually jumped location, time, and sometimes into a characters dream or fantasy, a video system was required that could easily represent a vast multitudes of spaces with minimal changes to the physical space.

3. To assist in representing a diverse range of characters. As Half-Real had a cast of three performers (mirroring the three suspects to be investigated), the production sought a different theatrical language to represent the many minor characters from the key suspects.

4. To create a design that was sophisticated and intricate, but which was also highly tourable and easy to set up.

2 Background

Half Real was inspired by previous performances and art installations that employ projection. The work by Naimark [7] reprojects a captured physical environment onto a white painted version. Be Now Here used a similar technique to recreate an outdoor public plaza. Dance productions such as Glow [2] and Mortal Engine [3], use camera based tracking to generate projected animated content in real time based on the dancers' movements and the soundtrack. Half Real builds upon this concept by using 3D tracking and a 3D projection environment. Half Real also combines procedurally generated content with pre-made video to enhance the performance and support interactivity. Audience interactivity has previously been implemented using camera based techniques. Maynes-Aminzade et al. [6] demonstrate how computer vision can detect audience movements to control a virtual paddle in a game of pong, and how laser pointer tracking on a projection screen can be used for large scale interactive systems. Motion Swarms [8] uses image processing techniques to produce a particle system based on audience movements. This particle system is then used to control a virtual beach ball. One of the main limitations of these systems is that it is difficult to identify and respond to the actions of an individual audience member. Maynes-Aminzade et al. however note that this is not necessary for many types of interaction. The illusion of interaction is what is important from the audience's perspective. Rather than use computer vision techniques for broad audience participation, Half Real gives each audience member a physical controller. This allows much more fine grained control of the interactivity than other techniques.

Half Real uses a spatial augmented reality system based on Shader Lamps [10] to project perspectively correct virtual content onto the physical set. Previously Shader Lamps has been used to create physical cartoon dioramas [11] and life size immersive environments [5] which can be modified in real time by the user [9]. Superficially, the Half Real set is similar to the Cave Automatic Virtual Environment (CAVE) [4]. However, where a CAVE provides an immersive virtual environment to the users inside the CAVE, the Half Real set provides a virtual environment for the audience. The actors placed in the virtual environment are part of the illusion.

3 The Stage as a 3D Environment

Projectors are becoming more and more popular in performance art. In the past, projectors have been used for image projection and video playback. Whether they are used to project onto flat screens, or more complex objects on stage, 2D content has been created and used. This approach suffers from two main drawbacks:

1. The projected content is created for a specific projector. This makes it difficult to add additional projectors later in development, as new content needs to be created specifically for the new projector location.

2. Setup requires a tedious process of placing projectors in the correct locations, aligning them precisely, and performing keystone correction.

The approach taken in Half Real is fundamentally different. The entire set is modeled as a 3D scene (Figure 2(a)), and SAR technology is used to illuminate the set. Rather than creating unique content for each projector, the content is created for the scene. This approach requires a 3D model of the set to be created. However, 3D models are becoming more common to aid in tasks such as preproduction lighting design.

The 3D scene based approach taken for Half Real provides several advantages over 2D:

1. Projected content needs to only be created once for the scene, regardless of how many projectors are used.

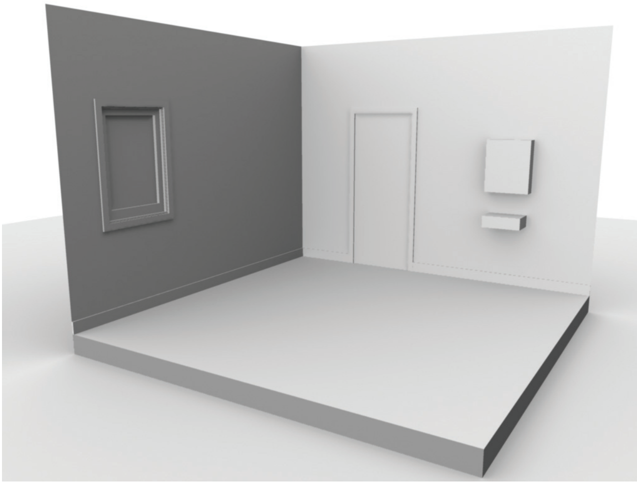
2. Projectors can easily be added as required.

3. Scenes can be previewed in 3D pre-production, before the physical sets have been built. This is similar to visualizations made in lighting design software.

4. Pre-show setup is much simpler. The projectors only need to be roughly placed, and the SAR calibration algorithm will calculate what content will be projected from each projector.

The use of a real-time spatial augmented reality system also brings new possibilities to the creative team. Tracking systems, cameras, and other sensors can be used to track actors' movements and other actions. Projected content generated procedurally, rather than pre-rendered video and animation, can be created to react to the data coming in from these sources. This enables much more dynamic projection effects to be created. In addition, because the projected content is controlled by a computer system, rather than using simple video playback, the performance does not need to follow a linear path. Half Real takes advantage of these possibilities by tracking actors, using procedurally generated projected content, and an interactive performance where the audience votes on the direction the story takes.

[2]http://www.heavyrainps3.com



(a)

(b)

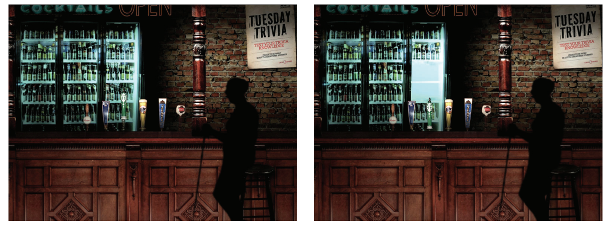
Figure 2: The 3D model of the Half Real set (a), and physical version with projected textures (b).

3.1 Creating Content

Once a 3D model has been created for the set, content can be authored. Texture maps for the 3D objects are used as guides, and the content to be projected is created in standard animation software.

Treating the entire set as a projected virtual world creates challenges for both content creators and live actors. As the walls are projected onto from the front, the actors inevitably cast shadows onto the walls. Ceiling height in venues meant high projector locations could not be used. Previous research [12] has shown that users are able to cope with shadows introduced in projected environments, although users prefer systems utilizing dual projectors to eliminate total occlusion. However, instead of trying to reduce shadows, Half Real embraces the interplay between the live action and the virtual world. Virtual characters are rendered as silhouettes rather than realistic people. Actors' movements on stage were developed to enhance the effect given by their shadows.

Projecting content onto the walls and floor caused problems with content creation. For example, a bold, bright color on the floor reflects onto the walls. Scenes that look good on a standard computer monitor could look dark and bland when projected, due to the limited color gamut and brightness of the projectors, and ambient light. The interplay between projected light and stage lighting also caused problems. Stage lighting is necessary for the live actors, but drowns out the virtual environment. Actor blocking, projected content, and lighting design were iterated simultaneously to achieve the best balance. In some instances the best approach was to modify the projected content so that it, rather than stage lighting, could illuminate the actors. This is shown in Figure 3.



(a) (b)

Figure 3: Stage lighting could drown out the projected images, particularly when actors needed to be close to the walls. In some cases, using the projectors to light the actors was necessary, and the projected content was modified to allow this. These images show the original content (a), and the modifications made to light the actor (b).

4 Interactivity in Half Real

Half Real is a live murder mystery theater production, with the investigation conducted by the audience. At key points during the show, the audience votes on how the investigation proceeds. This results in each performance being unique, with the audience seeing only a part of the story. The whole story contains approximately 4x the material as is seen in any single performance.

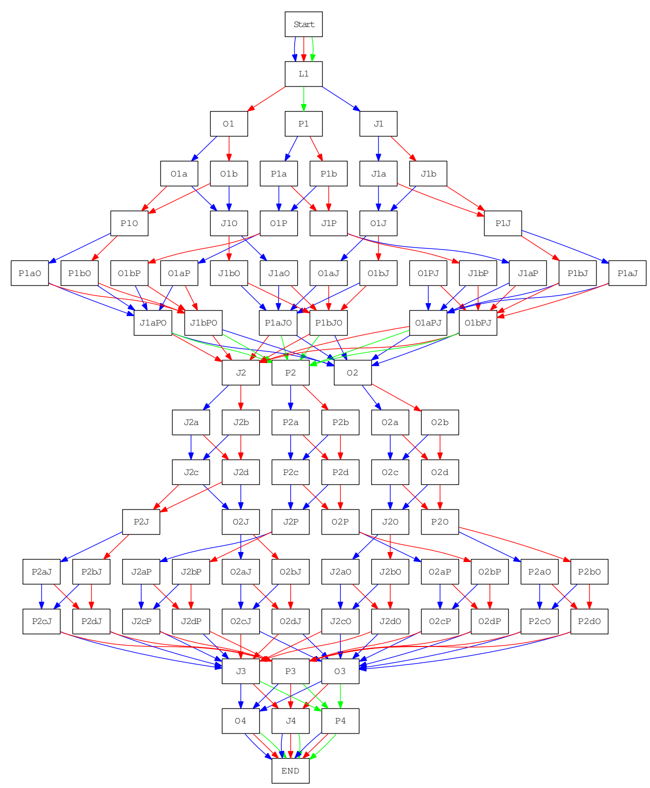


Figure 4: The voting graph of Half Real. Vertices represent scenes, and edges represent audience votes.

Half Real's interactivity created a unique context for the audience, where they shifted from being spectator to participant. The inclusion of the dynamic GUI within the dramatic action, coupled with the ZigZag voting sequences, created a more informal mode than traditionally experienced during a performance. The audience was extremely aware of its collective decision making during the work. At the point of voting, the audience were given a permission to "interact", be it through conversing with nearby members, observing what other people choose on their ZigZag, or vocally respond when their preferred choice was lost by one percentage point.

Choices that were extremely close or extremely divergent were often the most fascinating, as the audience often became vocal in the manner they might be at a sporting match or playing a boardgame with friends. This participatory aspect of the performance altered the audience behaviour traditionally associated with modern theatre where an audience quietly receives the creative authorship of playwright or director. Half-Real allowed the audience to playfully engage with the meta-theatrical context of a group democratically deciding what they'd like to see happen next, and also activated the audience to define what kind of theatrical experience they would have through their collective decision-making.

The interactivity also provided the audience with a "portrait" of their assumptions (i.e. who the killer was), with each node was a "litmus test" of the audiences suspicions, and each subsequent choice sought to build upon the audiences previous selection.

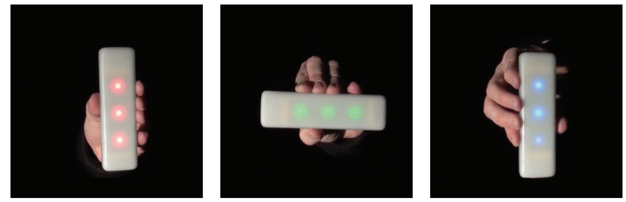
4.1 Show Structure

To accommodate this interactivity and variation, the computer system represents the show as a Directed, Acyclic Graph (DAG), shown in Figure 4. Each vertex is a scene, with audience vote results represented as edges. From the audience and actors' perspective, some scenes in the graph are duplicates. This means that some scenes can be seen in different orders. However, as the vote sequences into and out of particular instances of scenes are different, this looping is "unrolled" into a DAG.

The show is roughly structured into four acts. During each act, the story branches out as the audience decides which paths to take, before converging back to the start of the next act. The first act allows the audience to discover the origin of three pieces of evidence. Here, the suspects are introduced. In act two, the audience is able to investigate two of the three suspects in more detail. In act three a single character is investigated. Act four requires the audience to vote on who they believe was the murderer, based on the evidence uncovered in their investigation. During each scene, key pieces of evidence are uncovered. When this occurs, a vote option appears in the environment, attached to either the evidence or the character associated with the evidence. At the end of each scene, a voting sequence occurs to decide how the investigation proceeds.

4.2 Voting

Each audience member votes using a ZigZag. The ZigZag is a remote control style device without buttons. Instead, the device contains a three axis accelerometer, three red/green/blue LEDs, batteries, and an Xbee[3] wireless module that handles the wireless communication to and from the computer system. The computer system coordinates the tallying of votes, and initiation of voting sequences. At the beginning of a voting sequence, the LEDs in the ZigZag begin flashing. The audience then has ten seconds to cast a vote. As the ZigZag is rotated, the color of the LEDs changes to match one of the vote options, as shown in Figure 5. After ten seconds have passed, the current color is locked in and the vote is cast. A random seed minimizes communication timing collisions between the 200 ZigZags. The vote results are then tallied and shown to the audience through the projection system. At the start of the show, a short tutorial on the devices was presented, before a lesson vote, that did not affect the path through the story.



(a) (b) (c)

Figure 5: Audience members vote by orienting the ZigZag to select a color, either red (a), green (b), or blue). Colors correspond to vote options that change color at the beginning of a vote sequence.

5 Actor Tracking

A major advantage of representing the set as a 3D scene, rather than simply treating the walls and floor as 2D projection surfaces, is that virtual objects can be placed relative to any 3D location. This ability was used in Half Real to pin virtual information, such as vote options, to actors during the show (Figure 6).

Figure 6: A Vote Option attached to an actor.

Actor tracking was implemented using a Microsoft Kinect[4]. The Kinect was chosen for the following reasons:

[3]http://www.digi.com/xbee/

[4]http://www.xbox.com/en-US/Kinect

1. The Kinect operates in the infrared spectrum. Therefore, tracking is not affected by changes in the projected images or stage lighting, which would affect visible light cameras.

2. Actors were not required to wear or carry sensors or fiducial markers.

3. The Kinect is quite a robust tracking system, making it particularly suitable to theater.

4. The set of Half Real was small enough for the Kinect to be able to track almost the entire area.

5. The Kinect is inexpensive and readily available, so the hardware could be replaced if necessary while on tour.

The main downside to the Kinect is that it could not be used to track props on the set. It was not possible to project explicitly onto objects the actors carried or moved during the show. In addition, we did not perform skeletal tracking, as this requires a calibration step that would interrupt the performance.

5.1 Attaching Information to Actors

The Kinect and the 3D scene representation provided accurate position information for actors as they moved about the space. However, like any SAR environment, virtual information had to be projected onto the available physical surfaces, such as the walls of the set. The 3D actor positions need to be translated to a suitable location on the walls behind them. This position needed to give the illusion of the virtual information being attached to the actors.

Half Real uses a ray-casting algorithm to calculate a suitable location for attached content. A ray is cast from the audience location, through the actor's location, and onto the wall behind them. This position was then adjusted to accommodate smooth transitions between the left and right walls. In addition, the location information from the Kinect was averaged over the most recent 30 frames. This gives the moving virtual information a visually appealing slow-in, slow-out animation effect. It also greatly reduced jitter when actors moved close to the range of the Kinects tracking distance.

Like any projection system, view dependent rendering effects only appear correct from a single viewpoint. For Half Real, an ideal audience viewport was chosen at center-stage, approximately 1/3 back from the front row. This location was chosen to give a good visual effect for the majority of the audience.

5.2 Identifying Actors

While the Kinect is quite good at tracking people, it is not able to reliably identify them. Actors enter and exit the set many times throughout each performance of Half Real. Having actors tracked without interrupting the performance was an important goal when developing the projection system. Therefore, an explicit identification process could not be used. Instead, through the course of blocking during rehearsals, catch areas were identified in each scene. These catch areas are regions on the set that an actor would always walk through during the scene. Once the tracking system registered an actor passing through a catch area, that actor was associated with the correct virtual information.

In addition to catch areas, dead areas that never associated tracked objects in the system were needed. The set of Half Real was not simply a static scene, it contained a door and window that actors could use and walk or climb through, and a chair that was moved about on stage. The detection algorithm in the tracker would sometimes incorrectly register these objects as actors. By marking these areas as "dead areas", the system would ignore these objects when associating virtual information. As with catch areas, the dead areas were specifically defined for each scene, as in some scenes actors moved into the range of the window or door when the virtual information needed to appear.

6 Implementation Details

The show control system consisted of two subsystems. The Voting System, which managed the ZigZag devices and voting sequences, and the Projection System, which managed the virtual environment and actor tracking. This section describes the implementation details of these systems.

6.1 Projection System

The projection system ran on a single computer containing an Intel Core i7-2600 CPU with 8GB of RAM. Two Nvidia Geforce GTX560 graphics cards drove the projectors and the operator screen. The projection software was implemented using our SAR software framework, written in C++ using OpenGL. The software ran on Ubuntu 11.04 Linux.

6.1.1 Resource Management

Half Real's projected content consisted of images, video files, and procedurally generated content, such as the vote options. In all, 38GB of assets, mostly video, were used during the show. A "level loading" approach was used to manage these assets. Each scene was described in an XML file, which listed the assets required for the scene. A content manager was responsible for freeing data no longer required and loading new assets if required. Assets that were needed in consecutive scenes were reused, rather than reloaded. This approach was chosen due to the interactive nature of the show. The next scene, and therefore the assets required, would not be known until the audience voted. This meant that loading and unloading assets at predefined times was not possible.

6.2 Pre-show Calibration

A pre-show calibration process is required for both the projectors and the Kinect. The projectors are calibrated using the algorithm described by Raskar et al. [10], which finds the intrinsic and extrinsic parameters of the projectors. This process involves finding landmarks on the set with a projected crosshair (Figure 7). This process takes approximately 2-3 minutes per projector, and only needs to be performed when the projector locations change, such as moving to a new venue.

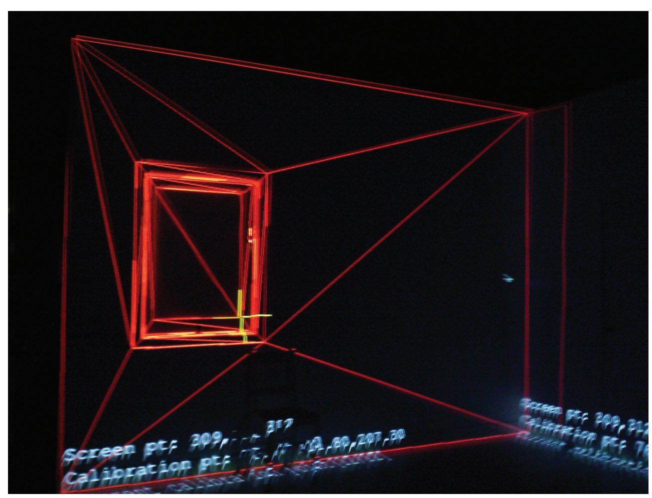


Figure 7: The operator performs projector calibration by marking points on the set with a projected crosshair.

Once the projectors are calibrated, the Kinect must also be calibrated. By default, the Kinect returns 3D locations as a distance from the Kinect. A coordinate space transform must therefore be calculated to bring the Kinect's tracking data into the world coordinate space of the tracking system. This calibration process involves finding crosshairs projected on the floor with the Kinect's visible light camera. 3D coordinates are obtained for these points from the Kinect and the transform can be calculated. Again, this process must only be performed if the projectors or Kinect are moved.

6.3 Decoupling The Kinect

During rehearsals, a software bug inside the OpenNI framework was discovered that caused the projection software to crash intermittently. As the scenes in the show are made entirely of projected artwork, the set would go dark if the software crashed. This was unacceptable for a live performance. The Kinect interface was rewritten to decouple the Kinect from the projection software.

The solution implemented involved two processes running simultaneously. The first process, tracker, handled interacting with the Kinect hardware via OpenNI. This process was run in a shell script that restarted the process whenever the software bug was encountered and the process crashed.

The second process was the main projection system. It accepted data from the Kinect via a TCP connection. This connection was automatically re-established as required. This solution resulted in a robust projection system. If the tracker process crashed, tracking would stop for approximately ten seconds, with the performance continuing regardless. Virtual information would be reattached to actors once tracking restarted based on their last known location.

7 Conclusion, Lessons Learned, and Future Outlook

This paper has provided a case study in using spatial augmented reality in live theater, enabling more dynamic projected content, and audience interactivity. Half Real successfully completed a tour of regional South Australia, before playing a three week, sold out season as part of the Melbourne Festival in 2011. We consider that achievement as proof the technology and software developed was a success. However, as with any production there are lessons learned and room for improvement.

One of the major issues that had to be overcame was reliability and robustness. In Half Real, if the projection software crashed, the stage went dark. We had to develop a system that would function correctly day after day, for extended periods of time. Decoupling subsystems was one of the most important factors in making the system robust. For example, it was important that the projection system kept running if the tracking system stopped responding. Another issue was sequencing content to be projected in each scene. The projection system used XML files for each scene. This effectively meant there was one scene description for the projection, and another for lighting and sound. In the future, we would like to improve the flexibility of our system by making it interoperable with existing stage management software, such as QLab[5], which would reduce the duplication, and make modifying the sequences of projected content much easier.

While Half Real has made an important step in using SAR for interactive performance art, we believe there are many more possibilities to be explored. For example, using projectors to simulate physical light sources, such as follow spot lights that automatically track the actor. Or, using the projectors to project directly onto the actors in order to change their appearance. We would like to extend the technology developed for Half Real to create much more sophisticated dynamic projection environments for performance and interactive art into the future.

Acknowledgements

The authors wish to thank the rest of the Half Real production team: Cameron Goodall, Duncan Graham, Geoff Cobham, Chris More, Chris Petridis, Tyson Hopprich, Ray Gardiner, Aaron Herczeg, Katherine Fyffe, David Heinrich, Alirio Zavarce, Amber McMahon, Caitlin Byrne, Steve Tilling, and Monica Hart. We would also like to thank Fee Plumley and the Australia Council for the Arts.

[5]http://figure53.com/qlab/

References

[1] O. Bimber and R. Raskar. *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A K Peters, Wellesley, 2005.

[2] Chunky Move. Glow. http://chunkymove.com.au/Our-Works/Current-Productions/Glow.aspx, 2006.

[3] Chunky Move. Mortal engine. http://chunkymove.com.au/Our-Works/Current-Productions/Mortal-Engine.aspx, 2008.

[4] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 135–142, Anaheim, CA, 1993. ACM.

[5] K. Low, G. Welch, A. Lastra, and H. Fuchs. Life-sized projector-based dioramas. In *VRST '01: Proceedings of the ACM symposium on Virtual reality software and technology*, page 93101, New York, NY, USA, 2001. ACM. inproceedings.

[6] D. Maynes-Aminzade, R. Pausch, and S. Seitz. Techniques for interactive audience participation. In *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces*, ICMI '02, page 15, Washington, DC, USA, 2002. IEEE Computer Society.

[7] M. Naimark. Two unusual projection spaces. *Presence: Teleoperators and Virtual Environments*, 14(5):597–605, 2005.

[8] Q. Nguyen, S. Novakowski, J. E. Boyd, C. Jacob, and G. Hushlak. Motion swarms: video interaction for art in complex environments. In *Proceedings of the 14th annual ACM international conference on Multimedia*, MULTIMEDIA '06, page 461469, New York, NY, USA, 2006. ACM.

[9] R. Raskar and K. Low. Interacting with spatially augmented reality. In *Proceedings of the 1st international conference on Computer graphics, virtual reality and visualisation*, pages 101–108, Camps Bay, Cape Town, South Africa, 2001. ACM.

[10] R. Raskar, G. Welch, K. Low, and D. Bandyopadhyay. Shader lamps: Animating real objects with Image-Based illumination. In *Rendering Techniques 2001: Proceedings of the Eurographics*, pages 89–102, 2001.

[11] R. Raskar, R. Ziegler, and T. Willwacher. Cartoon dioramas in motion. In *NPAR '02: Proceedings of the 2nd international symposium on Non-photorealistic animation and rendering*, page 7ff, New York, NY, USA, 2002. ACM. inproceedings.

[12] J. Summet, G. D. Abowd, G. M. Corso, and J. M. Rehg. Virtual rear projection: do shadows matter? In *CHI '05 extended abstracts on Human factors in computing systems*, pages 1997–2000, Portland, OR, USA, 2005. ACM.