Mixed Reality Immersive Design:  
A Study in Interactive Dance

João Beira  
University of Texas  
1 University Station A0800  
Austin, Texas, USA  
(+1) 5123516940  
joaobeira@gmail.com

Rodrigo Carvalho  
Universidade do Porto - FEUP  
Rua Dr. Roberto Frias, s/n  
200-465 Porto, Portugal  
rodriguedescarvalho@gmail.com

Sebastian Kox  
Oneseconds  
Brigittenstraat 18 3512 KM  
Utrecht, The Netherlands  
Sebastian@oneseconds.com

ABSTRACT

Dance performance interconnects the notion of space and movement, providing therefore the ideal framework to design and research immersive relations between the virtual and physical world: mixed reality performance. We start by discussing conceptual approaches to immersive and mixed reality environments, and reporting some of the current state of the art solutions.

We developed two proof-of-concept systems and frameworks to implement the use of spatial augmented reality applied to dance performance. Our systems establish a real time dialogue and interaction between 3D geometry and perspective grids calculations with body movement. We also present our framework design and implementation of our system, outlining relevant problems and solutions we find in the process.

Finally, the frameworks and design processes were validated through a field experiment - the performance - after which we gathered and analyzed the feedback from the performers, the audience and the choreographer.

Categories and Subject Descriptors
H.5.1. Multimedia Information Systems: Artificial, augmented, and virtual realities

General Terms  

Keywords  
Mixed Performance, Immersive Spaces, Spatial Augmented Reality, 3D Gesture Recognition, Projection Mapping, Interactive Dance

1. INTRODUCTION

As digital technologies transform the potential of live performance, digital media artists design new modalities of interactivity and create new relationships between physical and virtual spaces. Research on spatial immersive design in mixed reality is still scarce as the study of mediated digital live performance is an emerging topic within the field of human computer interaction. Augmented and virtual realities allow for the creation of immersive audiovisual spaces, places that work as alternative, parallel worlds: places that don't need to fulfill the same roles and functions as the real world, and where physics, logic and perception can be manipulated and augmented. These spaces work as metaphoric places of transition between real and virtual worlds in which virtual and real elements are interwoven.

The intention is not to simulate the real world, but instead to augment it, give new perceptions “and define arbitrary, abstract and otherwise impossible relationships between action and result.” [1].

Alice passing through the looking-glass is a recurrent metaphor for these immersive alternative spaces. In his article “The Ultimate Display,” Ivan Sutherland described the possibility of accessing places with “concepts not realizable in the physical world,” places with “forces in non-uniform fields, non projective geometric transformations, and high-inertia low friction motion... It is a looking glass into a mathematical wonderland” [2]. Places defined by a three dimensional digital matrix where data is materialized in space, just like the cyberspace described by William Gibson in Neuromancer, can be created and explored [3]. Nowadays cyberspace exists, but is not exactly as it was envisioned in the past. By the end of the 1990s the “original wonder of cyberspace so present in the early cyberpunk fiction of the 1980s and still evident in the original manifestos of VRML evangelists of the early 1990s - was almost completely lost” [4]. It has colonized our everyday life. It is no longer “the other place.” We don't need to “jack in” to visit it, because it is everywhere, all the time. The “other place” now is when we lose the omnipresent connection to cyberspace [5].

The use of these virtual spaces in stage performances allows, temporally, the return to the exciting action arenas of the original cyberspace. It has a huge potential to build metaphors, construct new narratives and explore perception and states of consciousness. As Davies writes, “Such environments can provide a new kind of "place" through which our minds may float among three-dimensionally extended yet virtual forms” [6]. Immersive spaces with total synchrony of movement, sound and light engage the audiences’ perceptions into temporal fantasies of alternative virtual worlds. During the few last decades, the notions of virtual reality inspired by theories of the exploratory potential of cyberspace focused on the virtual and synthetic environment. However, currently, new design strategies are now being extended to physical spaces: spatial augmented reality. Through the development of recent video projection techniques, such as 3D Video Mapping, it became possible to merge the virtual world with three-dimensional physical space.
2. IMMERSIVE SPACES IN PERFORMANCE

The use of augmented reality strategies to augment the expressive control of the performers in a performance context has shown to be an effective strategy to engage the audience and performers [7]. The work of Myron Krueger in the 70s with computer vision, silhouette extraction, responsive environments and interactivity with digital realities defined the vocabulary of human interaction with virtual worlds. His metaphors and interactions have formed a reference point for the field. In “Videoplace,” a body’s silhouette is tracked and transferred to the virtual world. There she is able to interact with others and manipulated elements through gestures and movements [1].

It is also important to highlight two more pioneers: Erkki Kurenniemi and David Roekey. In the 70s Kurenniemi created “DIMI-O,” an electronic organ with a 32-step sequencer memory unit from which the player could play and manipulate the notes sequences in the memory. As input it could use the feed of a video camera, where the captured picture was converted into black and white and then used as signal to control the notes in the memory unit [8]. DIMI Ballet was a dance performance made to demonstrate the use of this instrument with video input in stage [9]. Since the early 80s, Roekey has been working in interactive systems to capture body motion and translate it into sound. In “Very Nervous System,” human gestures and movements are captured through a video camera and translated in real time into an improvised music system that reflects and reacts to the qualities of the movements. He writes, “I use video cameras, image processors, computers, synthesizers and a sound system to create a space in which the movements of one’s body create sound and/or music.” [10]. The work of these three pioneers with camera tracking, body mapping and audio visual interaction laid the foundations in technology and vocabulary for contemporary interactive new media performances.

In the last two decades, the increase of computers’ processing capabilities, the decrease of hardware prices and the emergence of frameworks dedicated to real time interaction have pushed the possibilities of creation to new and seemingly infinite boundaries. Exciting new visions, achievements, metaphors and interactions have been made and provide new perspectives to immersive spaces and mixed and augmented virtual realities.

In “Messa di Voce” [11], by Levin and Lieberman, the sounds made from the actors’ voices are augmented into a virtual space. After the sound is created its graphic representation becomes an inhabitant of the interactive world. Then the actor can interact and play with these inhabitants that he just created. The performance is composed of different scenes, each one with different interactions. As an example, in the scene “Bounce,” the performer emits a special sound that creates series of bubbles that flow in the air. The performer can then interact physically with the virtual bubbles that are in fact the virtualization of his voice. A frontal camera detects the performer’s silhouette and uses the information about the performer’s body to place the elements in the virtual world (projected in display in the back of the stage).

In Obermaier's version of “Le Sacre Du Printemps” [12], a projection representing the virtual worlds appears with prominence in the main stage while the performer is dancing on a secondary stage and being tracked by a computer vision system in real time. The dancer's body is transferred to a virtual three-dimensional space where her body interacts with a matrix world and new perceptions of the body come to life.

Obermaier again in “Apparition” [13] develops an interesting relationship between the dancers' bodies and the virtual spaces. He uses their body shape as a surface for mapped projections. A computer vision system tracks the dancer's silhouette through a frontal camera. This shape is used as a mask for the mapped projections, and a kinetic projection surface is achieved. Aside from being and actors in this mixed reality, the performers are also layers of the virtual space itself and their body dynamics and features are transferred to the architecture of the space, which can be expanded and contracted and made fluid or rigid depending on the performer’s movement.

Another good example of manipulating a virtual space through the dynamics of the body is “Body Navigation” [14]. A floor projection and a top camera transform the floor of the stage in an interactive arena. The performer’s body is tracked and used to define the features of the virtual reactive floor. Shadows, lines, and other graphic elements are attached to the dancer position and are manipulated by his movements.

“Seventh Sense” [15], was an immersive performance by Anarchy Dance Theatre and Ultra Combos. The performers were in the middle of a box with projection on three walls and the floor. Their position was being constantly tracked by a computer vision system on the top, and every motion had repercussions in a 3D mesh geometry.

Researchers at Yamaguchi Center for Arts and Media in Japan have recently released the “Reactor for Awareness in Motion” [16], a toolkit to create environments for dancers that provides a series of functionalities to access, recognize and process motion data and provide real-time feedback and interactions with the audiovisual environment. Based on the idea of dialogue between bodies and environment, this toolkit acts as a kind of motion data manager that receives tracking inputs from different sensors, motion capturing systems or 3D cameras such as Kinect, manages the data, and outputs it to projectors, speakers or other actuators.

The body of the performer is now connected with the space of the interaction. A computer vision system, a sensor-based or any other motion-activation system is linked to his body in a sensory feedback loop. A library of mappings decodes the body movements and manipulates the space and the audiovisual content according to the predefined logic. As Dixon writes, “The body of the performer and the piece of the music or video are distinct entities or objects but … for cybernetics they are intimately connected within a communication or control system”.[17].

3. SPACE AUGMENTED REALITY

During the last few decades, the notion of virtual reality has focused on the virtual and synthetic environment. However, currently, new design strategies are now being extended to physical spaces: spatial Augmented Reality (SAR). Through the development of recent video projection techniques, such as 3D video mapping, it has become possible to incorporate the virtual world into three-dimensional physical space.

“The virtual reality community has oriented themselves away from head-mounted displays and towards spatial displays”, says Bimber and Raskar [18]. The main reason for this seems to be the focus on mobile use of augmented reality applications, but the authors argue that for “non-mobile applications a rich pallet of different spatial display configurations can be as beneficial for augmented reality, as they have been for virtual reality” [18]. The design of immersive spaces with projection-based surround screen displays, semi-immersive wall-like, cylindrical or spherical spatial
displays appears to be a good solution to achieve the desired immersion due to the fact that “spatial displays detach the display technology from the user and integrate it into the environment.” [18]. Recent projects such as “IllumiRoom” [20] show how video projections and light design can actively augment the existing physical environment and complement the content displayed on the television screen. The visual display combines the focus-plus-context concept with real-time projection mapping.

Therefore, there are several different strategies to create an immersive environment. However, in the specific case of a stage performance, the SAR approach emerges as an excellent option. Due to the fact that it joins in the same space features from different types and levels of immersion with virtual elements, it allows the creation of a mixed space that is able to provide a good level of immersion both for the performance and for the audience.

SAR is a mixed reality environment where the virtual elements are displayed in the space and mapped with architecture of the space. Like in augmented reality, the user (or audience) is not attached to body or head displays, and the visualizations are integrated into the physical environment. Through the use of projectors, light and images are mapped with the physical space, taking advantage from its geometrical features and viewers’ perspectives to insert objects, create new spaces or manipulate the real physical one [19]. It is not completely clear where SAR should placed in the virtual reality continuum [21] as it gathers elements from different classes and the level of immersion can vary depending on the projected surface area. Like in virtual reality, a person is immersed in a synthetic world. The level of immersion will depend of different factors related to the scale of the projection (related to the person) and the number of planes mapped (just a frontal plane, all around the audience, a corner of a room…). But this world is not completely synthetic; it overlaps the real one, just like in augmented reality. At the same time, like in augmented virtuality, real elements or persons can be inserted inside of the synthetic world through the use of cameras, sensors or other devices.

### 4. DESIGNING A SPACE AUGMENTED SYSTEM FOR THE STAGE AND DANCE PERFORMANCE

The solution and framework presented in this paper has been successfully used in five live performances during 2013. The show and process was organized in two main sections of augmented dance and visualization based on the movements of the dancers. Both sections had different levels of interactivity, and the overall goal of the project was to give control to the dance performer(s) and remove the scripted behaviour of predefined visualisations. Through the use of advanced gesture tracking and recognition, we explore strategies and possibilities for spontaneous adaptations and changes of the visual space in live performances.

There were three main issues for creating live interactions combined with artistic choices: capturing the movement, defining the interaction language and expression, and designing the output modalities. The narrative and storytelling of both sections is based in the exploration of augmented interactions from the performers. Both physical and virtual spaces merge in one same visual output and reality, providing optical illusions to the audience. Approaching the magical concepts not achievable in the physical world explored by multiple authors in the previously analyzed literature, the performers are able to trigger interactions and control the virtual space based on their physical activity. This augmented strategy is the core of dance performance expression.

#### 4.1 Projection

In order to generate a strategy to perform real time immersive mapping we used the method known as “cube mapping” [23], which uses a six-sided cube as the map shape of environment mapping. Each side is rendered individually and grouped in the final shape of the stage design. The proposed method of simulating a 3D environment on flat surfaces was established using one video projector, which results in a loss in pixel count and focus due to the depth of the stage and different surfaces. Due to the angle of the position and the angle of the video projector position to the stage, shadows are not really an issue as they drop on the back of the performers.

In order to be able to present this immersive space in a performance context, it was necessary to create a stage design where the three projection surfaces had an open angle of 90º between them. This specific stage design allows the performer to be immersed in the space and also allows various members of the audience who are positioned at different angles to watch the show (See Figure 1).

![Figure 1. © João Beira. Cube Mapping. Immersive space and 90° projection.](image)

#### 4.2 Setup

The stage is composed of a custom made structure that represents a half-cube and has a left, right and floor screen where the mapped projections are displayed. (See Figure 2.) In front of the stage, a Kinect camera scans and captures a 3D image. This information is managed by the motion capture software “Ni Mate” [24], which tracks the dancer’s position in 3D space, specifically the skeleton data. This information is sent by OSC protocol to the system network where customized software, based on Quartz Composer that generates and controls the visuals and the sound.

![Figure 2. © João Beira. System architecture and physical setup](image)
4.3 Motion Tracking

In our system, the motion tracking was made through the use of one Kinect camera, located in front of the performer. The skeleton tracking and real-time motion capture was established using the software NI Mate [24]. The use of an external sensor in the space to track the body, described by Axel Mulder [25] as "outside-in", provides information on the performance body relative to the space and releases the performance from obstructive sensors or devices, allowing the performer to move freely and more expressively.

The generation of 3D geometry and augmented space from the dancer’s motion led us to develop a tracking system in which the dancer can generate and manipulate 3D surfaces mostly with his hands. After the first tests with skeleton tracking, we realized this was the most effective strategy for the dance performer to control and manipulate the system. Mapping the physical body movement, skeleton tracking, with virtual three-dimensional perspective transformation allows the performer to be immersed in the virtual space as an extension of the physical space. This creates a mixed reality environment: a space where physical and digital objects co-exist and interact in real time.

According to the mapping strategies that Wanderly and Rovan defined [26], the correspondence between the dancer movements and the control parameters of the 3D geometry is “divergent mapping,” or one-to-many. An individual parameter in the performer’s body is assigned to simultaneous parameters in the space.

In each of the two scenes of the performance, two different mapping methods were used to create augmented dance strategies. The first method establishes a direct correspondence between the body of the performer with the mapping of a 3D object corresponding to the design stage, and the second allows the performer to dynamically trigger events.

4.3.1 Method - 1

In the first method there is a continuous stream of full body tracking information allowing the performer to modulate the geometry of the physical space in real time.

In order to generate a texture map of the stage design geometry from a live 3D scene, we used the visual programming framework Quartz Composer [27].

The output is projected on the three planes (left, right, and bottom) by warping quads and using the mapping technique (See Figure 3). In order to composite the output of three planes from quartz composer and quickly adjust the perspective transformation and the warping of different planes, we used the software Madmapper [28]. The communication between Quartz Composer and MadMapper was established through Syphon [29].

In this method, we tried to emulate the cube map method in Quartz Composer. By rendering one 3D scene in three different views (front, left and bottom) and applying the same rotation transformation on each instance, it formed a complete image when projected together on the three different projection planes. The limitations of Quartz Composer, however, didn't allow the proper 3D environment settings needed to make adjustments in the field of view or perspective transformations that would create a proper seamless cube map. To create a perspective transformation in each different view and simulate a proper overlap when the images were placed together, we used a plug in called Perspective Matrix from 1024 [30]. This method worked with the simple cube environment, but generated problems when introducing other 3D objects into the environment.
4.3.2 **Method - 2**

In the second part of the show we created pre-rendered 3D animations that are triggered according to gesture recognition. Predefined poses activate media content, but the dance performer has no control over the media content. This is an opportunity to develop elaborate relationships between physical actions of the performer(s) and complex 3D animations. Even though this method has a higher level of interactivity than the first method, the audience is exposed to more elaborate space design. The choreography plays a very important role on the process, and dancers react to clips and animations based on pre-designed collective or individual movements. Simple actions, like extracting solids from the screen, provide a very strong sense of augmented reality to the audience. Our method is established through the use of Kinect, Quartz Composer, MadMapper and Syphon. These methods limit expressive control of the dance performers over the media and deny any chance for modulation of the media in real time.

4.3.3 **Method - 3**

Future work involves focusing on merging these two techniques, establishing a dialogue between complex 3D animations triggered by the performer(s), and allowing them to modulate the geometry of the space. This way, the pre-planning process of the choreographer and the improvisation skills of the dance performer can create a better balance and flow of the augmented show and dance.

5. **CONCLUSION AND FUTURE WORK**

Several interviews that were conducted after the shows, with audience and project team members, allowed us to see our work from a variety of perspectives. A really interesting response from the audience interviews concerns the ways their perception is affected by this augmented strategies and interactions. Some viewers acknowledged that they felt confused and visually over stimulated, not always in the best way. Due to the nature of dance viewers’ expectations, audience members primarily focused their attention on the performer(s) and their motion. With the visual projections immersing the performers’ bodies and surroundings, there is a significantly larger amount of information to process.

The design process in immersive spaces for dance performances should avoid this obstructions for the audiences, optimizing the visual responses of the system according to the performers body position, and focusing the focus length on the performer action area. This response from the audience was more evident in the first part of the show, where even subtle movements from the performer triggered changes in the geometry immersive mapping.

One of the viewers refers to this augmented process somehow related to Butoh [31] due to the extreme emotional response from the audience to the physical of the performer motion and activity and the way that affects the environment.

The choreographer was interviewed for the project report. When asked about the future of dance and technology, he held that it should be used to enhance and augment the performance and help to reach a point that would be impossible without it. In this way, in the design process, the creators must ask themselves first what is the objective of their use of the technology. How can it truly augment the performance and not just be used as an accessory?

After months of work with the performers and the choreographer, it is interesting to notice how the workflow and rehearsals changed according to the tracking technology applied. Due to the nature of the first method, the choreographer assigned a more creative role to the dance performer. The work process was mostly developed through interactions from the performer with the system. In this sense, the performer engaged with the system, and similar to learning process with an music instrument, the outcome of the movements in the live shows were mostly established through honing those skills in practice. In the second method, more conventional strategies were organized and designed by the choreographer, as specific events were triggered by the system. The improvisation and creative role of the performers were clearly limited, but the visual outcome of the content was more complex and visually appealing to the audience.

The primary performer reported that she felt in control of the system. She could sense a good level of feedback. Every movement she made would affect the geometry and project her movements into the grid. “I can control my body and movement in specific directions in order to shift the environment,” she said.

She also reported some bugs that happened when she place one hand over the other without realizing it. In these moments, sometimes the skeleton tracking would lose track for some moments. Even so, she felt in control most of the time. When asked to make comparison between behind tracking by a camera (outside-in) and other methods like wearing sensors on the body (inside-in or inside-out), she stated that despite the fact that body sensors could give more precise control over the parameters, with the camera tracking she felt more expressive because she was able to move her body more freely.

When asked about the future of dance and technology, he held that it should be used to enhance and augment the performance and help to reach a point that would be impossible without it. In this way, in the design process, the creators must ask themselves first what is the objective of their use of the technology. How can it truly augment the performance and not just be used as an accessory?

After months of work with the performers and the choreographer, it is interesting to notice how the workflow and rehearsals changed according to the tracking technology applied. Due to the nature of the first method, the choreographer assigned a more creative role to the dance performer. The work process was mostly developed through interactions from the performer with the system. In this sense, the performer engaged with the system, and similar to learning process with an music instrument, the outcome of the movements in the live shows were mostly established through honing those skills in practice. In the second method, more conventional strategies were organized and designed by the choreographer, as specific events were triggered by the system. The improvisation and creative role of the performers were clearly limited, but the visual outcome of the content was more complex and visually appealing to the audience.

The primary performer reported that she felt in control of the system. She could sense a good level of feedback. Every movement she made would affect the geometry and project her movements into the grid. “I can control my body and movement in specific directions in order to shift the environment,” she said.

She also reported some bugs that happened when she place one hand over the other without realizing it. In these moments, sometimes the skeleton tracking would lose track for some moments. Even so, she felt in control most of the time. When asked to make comparison between behind tracking by a camera (outside-in) and other methods like wearing sensors on the body (inside-in or inside-out), she stated that despite the fact that body sensors could give more precise control over the parameters, with the camera tracking she felt more expressive because she was able to move her body more freely.

Future work involves focusing on merging these two techniques, establishing a dialogue between complex 3D animations triggered by the performer(s), and allowing them to modulate the geometry of the space. This way, the pre-planning process of the choreographer and the improvisation skills of the dance performer can create a better balance and flow of the augmented show and dance.

Adopting a more complex render framework or a multi-platform game engine, such as Unity3d [32] or 3DVIA Studio Pro [33], will allow a more complex, effective and detailed use of 3D shapes manipulated in real time. In particular, the use of shadows is highly effective for visual immersion and 3D illusions, as we confirmed in the second part of the show. Being able to render them in real time by tracking the performers’ movement seems to be an important milestone to achieve, establishing a more effective strategy for expressive control in augmented dance.
6. ACKNOWLEDGMENTS

Research under the UTAustin/Portugal Co Lab Program. Supported by FCT Grant [SFRH / BD / 33964 / 2009] and [SFRH / BD / 51747 / 2011]. A special gratitude to the choreographer Yacov Sharir, sound designer Bruce Pennycook, and the dance performers Chelsea Pribble, Tawny Garcia, Lily Amelie Hayes, Billie Secular and Reema Bounajem.

7. REFERENCES


