Initial Lessons From ARFaçade, An Interactive Augmented Reality Drama

Steven Dow, Manish Mehta, Annie Lausier, Blair MacIntyre, Micheal Mateas
College of Computing, GVU Center, Georgia Institute of Technology
Atlanta, GA 30332-0760, USA
{steven, mehtama1, gth654r, blair, michaelm}@cc.gatech.edu

ABSTRACT
In this paper, we describe an augmented reality version of the acclaimed desktop-based interactive drama, Façade [17]. Few entertainment experiences combine interactive virtual characters, non-linear narrative, and unconstrained embodied interaction. In ARFaçade players move through a physical apartment and use gestures and speech to interact with two autonomous characters, Trip and Grace. Our experience converting a desktop based game to augmented reality sheds light on the design challenges of developing mixed physical/virtual AI-based drama. We share our initial observations of players from a live demonstration and talk about our work moving forward.

Categories and Subject Descriptors
H.5.1 [Information Interfaces and Presentation (e.g., HCI)]: Multimedia Information Systems – Artificial, augmented, and virtual realities.

Keywords
Interactive drama, augmented reality, artificial intelligence, natural language, embodied interaction, experience design, entertainment, games

1. INTRODUCTION
In this paper, we reflect on our experience creating ARFaçade, an interactive augmented-reality (AR) experience based on the desktop-based interactive drama, Façade. Façade was created in 2005 by Mateas and Stern [17] and is notable for the advanced AI engine that enables the main characters, Trip and Grace, to react to a player's unconstrained movements and utterances while loosely following one of many possible dramatic arcs. On the surface, Façade is a natural candidate for conversion to AR: it takes place in a small world (two rooms in a small apartment), requires no fast-paced interaction that would be challenging with current hardware, and supports rich interaction between the player, the characters and the world. However, the conversion of Façade to ARFaçade sheds light on a range of architectural, interaction and content issues that will be faced by any experience that immerses a player in a mixed physical/virtual space.

The long-term goal of this project is to study the impact of embodiment on such rich interactive experiences. Many people have dreamed and written about embodied immersive experiences of the sort epitomized by the Star Trek "Holodeck" (e.g, [20]). Unfortunately, experiences with truly interactive characters situated in a rich non-linear story are rare because of the difficulty of creating the necessary artificial intelligence (AI) engines to drive them; none have been created (to our knowledge) where the player can move freely around a physical space and interact naturally with virtual characters that share the space with them.

Façade is one of the first examples of a non-trivial step toward the rich Holodeck-like experiences envisioned by authors such as Laurel [12] and Murray [20]. Like contemporary games, Façade is set in a simulated world with real-time 3D animation and sound, and offers the player a first-person, continuous, direct-interaction interface, with unconstrained navigation and ability to pick up and use objects. But like drama, particularly theatrical drama about personal relationships such as *Who's Afraid of Virginia Woolf?*, Façade uses unconstrained natural language and emotional gesture as a primary mode of expression for all characters, including the player. Rather than being about saving the world, fighting monsters or rescuing princesses, the story is about the emotional entanglements of human relationships, specifically about the dissolution of a marriage. There is unity of time and space – all action takes place in an apartment – and the overall event structure is modulated to align to a well-formed Aristotelian tension arc, i.e. inciting incident, rising tension, crisis, climax, and denouement, independent of the details of exactly what events occur in any one run-through of the experience.

ARFaçade moves the player from sitting in front of the computer display, where they interact with small virtual characters via a keyboard and mouse, into a shared physical room, where they interact with life-sized virtual characters via speech, gesture and physical movement (Figure 1). By leveraging the sophisticated technology underlying Façade, we have the opportunity to come closer to the dream of embodied, interactive experiences than has ever been achieved before.

Our work on ARFaçade is ongoing; an initial version of the experience is up and running, but there are still problems to be resolved before a complete evaluation can be performed. The goal of this paper is to convey what we have learned over the past year about the design and implementation of sophisticated embodied interactive experiences, both through our design iterations and from player feedback during a number of demonstration sessions. While some of the issues and lessons learned are specific to ARFaçade, many are applicable to any interactive AR experience.

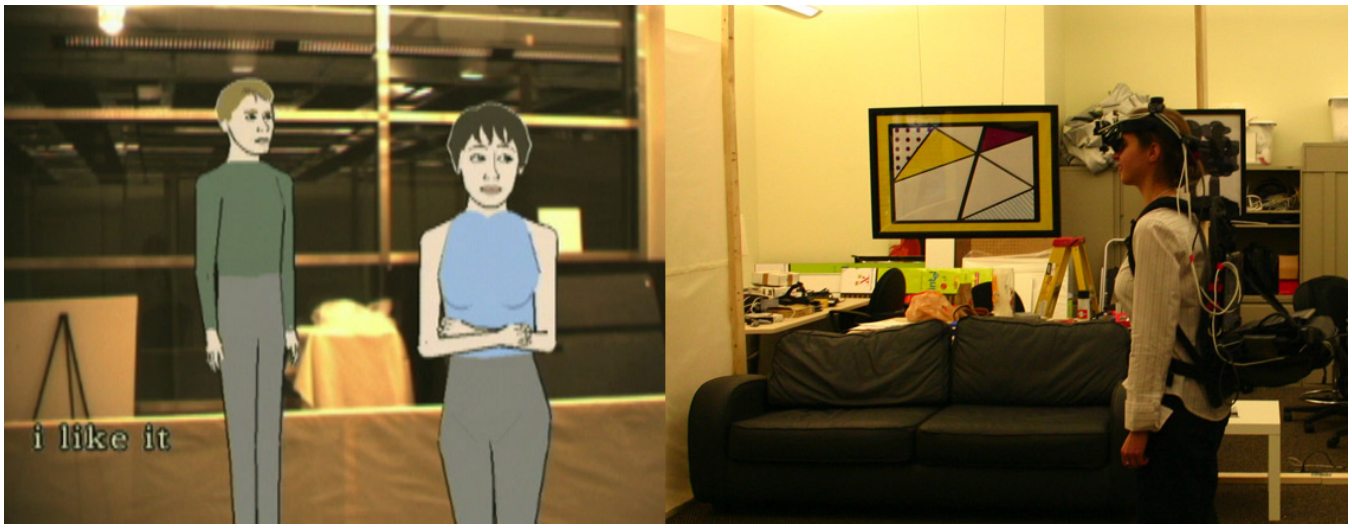

Figure 1: The ARFaçade experience—(left) the player's view of Trip and Grace through the head-mounted display with the text display showing their spoken statements, (right) a third person view of the player moving in the apartment.

First, we describe the details of Façade and the basic design for ARFaçade. We then discuss the challenges faced when building ARFaçade and relate them to larger questions of embodiment in interactive drama. We share ARFaçade's implementation details and our initial observations from informal public demonstrations. Finally we talk about our plans moving forward, including possible solutions to unresolved problems, the ongoing technical implementation and formal evaluation.

2. RELATED WORK

Although media theorists have talked about the possibilities for embodied, interactive narratives [12, 20], the gap between theory and practice is still quite wide. No other research project has come as close as ARFaçade to achieving the combination of three essential elements: interactive virtual characters, non-linear narrative, and unconstrained natural interaction (e.g., conversation, gesture, unconstrained physical movement throughout a large space and interaction with physical props). Due to the complexity of creating each component, and then integrating them into a system where the player feels agency (i.e., feels as if their actions truly affect the characters and the narrative), most projects focus on only one or two elements.

A number of research projects have also attempted to create rich embodied experiences, but none approach the complexity of ARFaçade. Cavazza, Charles and Mead have done a series of projects exploring interactive storytelling (e.g, [3, 4]). For example, in the Bond experience [3], the player has their image inserted into a virtual world where they interact with James Bond in a vignette from a longer story. They employ relatively sophisticated interaction recognition and narrative AI, but the story does not have the richness of Façade nor the first-person immersion of ARFaçade. A series of projects at USC ICT (e.g., [22]) and the Media Convergence Lab at UCF (e.g., [8]) have been aimed at creating synthetic experiences that integrate natural interaction and AI simulation, but they focus primarily on training contexts and have relatively simple narrative structures and interaction possibilities.

Conversational agents are now used in a wide variety of application areas, from personal assistants to e-learning. In the area of dramatic entertainment, a growing research community (e.g., [2, 6, 9]) is working on believable conversation agents, but generally their behavior is not integrated into a rich dramatic performance. Most work on interactive storytelling systems has focused exclusively on desktop interaction [14, 21]. ARFaçade integrates complex conversational agents within a complete dramatic experience, allowing the player to interact naturally with the characters in a physical space.

There are many examples of more constrained experiences situated in physical spaces. In KidsRoom, researchers created a physical children's playspace equipped with cameras to sense the movements of participants [1]. The researchers designed a simple, linear narrative that sufficiently constrained the space of possible interactions so that participants could enter the experience unencumbered by sensors. Similarly, the AR experience Four Angry Men and its predecessor Three Angry Men [13] situates a multiple point-of-view linear narrative in a mock jury room. Interaction is limited to head movements and seat changes around a table. Other well-known augmented reality experiences such as ARQuake [23] and ARPacman [5] have converted traditional action video games into embodied experiences in outdoor spaces, but these particular games do not have compelling narrative or conversational characters.

3. MAKING FAÇADE "EMBODIED"

In this section, we introduce the reader to Façade, and highlight the system features that helped and hindered with the conversion to augmented reality interaction. Physically, the desktop Façade experience occurs in a relatively small, fixed setting (Trip and Grace's apartment), so building a matching physical environment was feasible. The design of the apartment is purposefully minimal to focus the player's attention on the characters rather than the apartment. The room's objects, the post-modern décor, the character's appearance and utterances are all meant to create a certain mood and social backdrop in Façade.

All the objects have symbolic connections to the backstory; manipulating an object evokes conversational references to the associated backstory topics. However, the engine only monitors

Figure 2: Desktop Façade, the original game designed by Mateas and Stern in 2005.

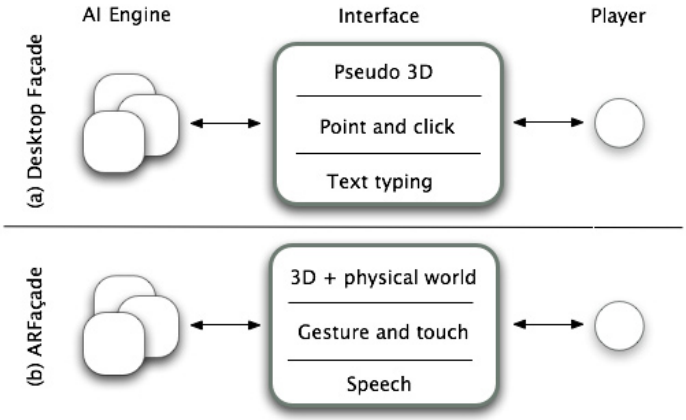

Figure 3: (a) The interface for desktop Façade (b) The interface for ARFaçade.

whether an object is being manipulated at all (e.g., picked up, looked at, pointed at), not details of how it moves through space. Since ARFaçade does not need to track the myriad physical objects precisely, simple approaches can be used to monitor them. These constraints, originally designed to focus the experience on character and story interaction and aid the AI storytelling engine, work to our advantage in conceiving the embodied version.

3.1 A SHORT PRIMER ON FAÇADE

As a friend invited over for drinks at a make-or-break moment in the collapsing marriage of the protagonists Grace and Trip, the player unwittingly becomes an antagonist of sorts, forced by Grace and Trip into playing psychological "head games" with them. The player, playing with her own name and gender, may react to the experience with hilarity or anger, or play a number of roles from councilor to devil's advocate. The experience is different each time the player plays it, and unlike most games, the players do not have a clear goal; the player invents goals for herself as the interaction with the characters unfolds. Although there are occasional breakdowns [9], the experience maintains a fluid interaction because the characters constantly respond to the player's unconstrained statements and movements with AI-generated speech and expressions.

Additionally, the story-level choices in Façade are intended to not feel like obvious branch points. We believe that when a player is faced with obvious choice points consisting of a small number of choices (for example, being given a menu of three different things to say to choose from), it detracts from the sense of agency; the player feels railroaded into doing what the designer has dictated. Instead, in Façade, the story progression changes in response to many small actions performed by the player throughout the experience.

Game players move through a 3D space with the arrow keys, interact with virtual items (to pick up glasses, statues, etc., or to hug/kiss/comfort the virtual characters) using the mouse, and speak to Trip and Grace by typing statements on a keyboard (Figure 2). The interface between the player and the AI engine consists of a graphics engine, keyboard text input, and mouse interaction with objects and characters in the space (Figure 3a).

The AI engine consists of three major components:

- the autonomous characters, implemented in the custom reactive planning language ABL (A Behavior Language); ABL supports the dynamic mixing of multiple, simultaneous behaviors, joint intentionality for multiple, cooperating agents, and meta-behaviors that can modify the runtime state of other running behaviors,
- the drama manager, which dynamically sequences dramatic beats as a function of the player's interaction history; the selected beat modulates the goals and behaviors of the autonomous characters,
- the natural language process system, consisting of a semantic parser that parses surface text typed by the player into the underlying discourse acts recognized by the system, and the discourse manager that keeps track of the current conversational context(s) and decides conversational responses to recognized discourse acts as a function of the active contexts.

3.2 Goals for ARFaçade

Our goal for embodied Façade is achieved using video see-through augmented reality. In ARFaçade, a modified graphics engine, physical interaction with objects, and speech handling (Figure 3b) replace the interface between the AI engine and the player. Several things about Façade's architecture made it very easy to adapt to augmented reality. The 3D environment navigated by arrow keys is nearly analogous to players freely walking around within a physical space. The space itself is a two-room apartment, not a fanciful or particularly large world. The AI engine in Façade only responds to the reference to objects manipulated by the player (objects pointed at, picked up, or looked at), rather than detailed object motion, thus simplifying the handling of physical interactions in ARFaçade; only the detailed position of the player must be tracked.

However, even though Façade's design facilitated it's conversion into an AR experience, there were a number of important design and technology challenges we had to overcome. In the next section, we discuss these challenges, our solutions to a number of them, and some remaining open questions.

4. DESIGN CHALLENGES AND SOLUTIONS

In this section we discuss the primary challenges faced when designing an embodied version of Façade. Many of these design questions encountered while creating ARFaçade will be common issues in designing any augmented reality drama: choosing how to render content, handling dialogue, interactions between physical and virtual objects, and facilitating movement in the space.

While discussing these issues, we also highlight important aesthetic and technical qualities of desktop Façade and how they impacted the embodied version. We propose solutions where a solution is possible and tie these issues back to larger questions of embodied interaction.

4.1 Character Rendering

Although desktop Façade does allow a player to navigate 3D space, a pseudo-perspective rendering system is used rather than standard 3D perspective projection. The Façade 2D animation engine generates the cartoonish image layers that make up Trip and Grace's expressive characters on the fly. As a player moves around the characters and as the autonomous characters move their bodies and change their facial expressions, the orthographically projected images are updated to provide the illusion of correct perspective. While dynamically generated 2D cartoon characters have expressive advantages over clunky 3D models or disjointed video content, it is unclear if they are appropriate for use in an AR experience. In particular, when the characters are pulled out of an environment with a similar cartoon appearance and overlaid on a "real" scene, will they still be believable?

In our initial experiments, players find the characters compelling when integrated with a video backdrop, creating an effect similar to *Who Framed Roger Rabbit?* While it might be interesting to explore the potential of realistic 3D models or video based characters in an interactive AR experience, much of the control and generative expressiveness would be lost. In ARFaçade, the complex mental and emotional state maintained by the autonomous character AI can actually be expressed in the visual realization of the characters through the procedural animation system. More importantly, it is unclear if we even want more realistic content–the theory of the Uncanny Valley [19] from robotics implies that more realistic artificial characters in augmented reality may end up being less believable. Whether they are is an open question.

4.2 Conversation with Characters

In desktop Façade the player starts typing a statement and letters appear on the screen. When the player is comfortable with the words on the screen, she hits the enter key to "say" the statement. A natural language processor [15] and AI engine [16] processes the utterance and cause the characters to react appropriately. Unfortunately, there are typically a few seconds of delay between the player hitting enter and seeing any effect, occasionally detracting from the experience.

For ARFaçade, speech recognition software would clearly introduce additional latency and (because of the large possible vocabulary and the context of use) errors. While it might seem that Wizard of Oz techniques (WOz) could be used in lieu of speech recognition, this turns out to be a challenge, as discussed by Maulsby in their emulation of an intelligent agent [18]. WOz can also potentially introduce latency and ambiguity into the interaction; it is unclear how additional delay will affect the flow of conversation.

Moreover, in desktop Façade, players adapt to the slight delay by strategically using the text buffer. Statements are often typed out and later retracted (the player backspaces over the text before hitting return), especially if Trip and Grace start talking about a new topic. Players quickly learn the limitations of the text buffer size (35 chars) because they see letters fill the width of the screen and they hear a beep when the text buffer is full. For the AI engine, this effectively constrains the amount of text and sentence structure that must be processed; for the player it provides a temporary buffer and a chance to reflect on the appropriateness of typed statements.

The problems with both speech recognition and WOz raise some hard questions for these kinds of experiences. Beyond just rapid, high-accuracy input, how do we provide equivalent affordances for players in ARFaçade and those available to players in Façade? Do we give players a chance to reflect on and revise verbalized statements? What feedback do we provide the player about the system's interpretation of their statements? How is the player made aware of system limitations, such as the maximum buffer size? Will an increase in latency diminish the conversational nature of the experience? Although there may not be a good solution for these issues, our current solution is to use the WOz method to type user statements as quickly as possible. Part of our ongoing work includes improving the feedback for the player and possibly providing the player an input method for retracting or entering their statements.

4.3 Physical/Virtual Interaction

In desktop Façade, the player can interact with the characters and objects, such as hugging, kissing and comforting Grace and Trip, or picking up drinks, trinkets, statues, phone, etc. The AI engine and characters adapt and react to such actions, sometimes apprehensively depending on the context, but it usually creates an engaging situation for the player. Façade loosely indexes conversation about objects so that any interaction (explicit touches, staring at an object, standing near an object, etc.) could cause the characters to converse or act on these objects. For example, if a player simply stares at the Italy photo on the wall, Trip will likely start talking about their recent holiday. Touching the characters and indexing objects is an enjoyable part of the desktop game that we do not want to leave out in ARFaçade.

For ARFaçade, conversation takes place around physical replicas of Façade objects (Figure 4). This interplay between physical and virtual items can make or break the experience. We discuss which virtual items are easy to deal with and why, and conversely which items present a significant challenge and how we hope to overcome them.

4.3.1 Easy Conversion to Physical World

Some aspects of desktop Façade were particularly well-suited for AR. The physical objects in the space are only referenced generally. While this may seem overly limiting considering the potential for fine-grained interaction, it actually allows for much smoother and contextualized conversation, and greatly simplified the infrastructure required for ARFaçade. A wizard operator can choose to index items if the player touches, looks at, or even comes close to a physical object of interest (e.g. art on the walls, statues, wine glasses, and the telephone). Some interactions are automatically handled by ARFaçade based on player movement (e.g., standing near, staring at), and a simple WOz interface with buttons to signal object references is adequate for covering the remaining interactions. (Additional WOz operators can be added if emulating speech and monitoring the interactions are too daunting.)

To increase the engagement of the experience, the audio for the apartment's answering machine and phone are played through physical props connected to additional computers hidden in the space. We expect that the experience will seem more real if we

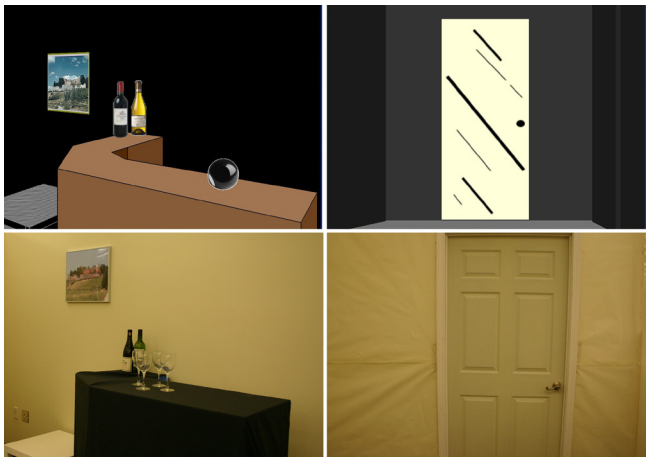

Figure 4: The top shows screenshots of the bar and door from Façade, the bottom the corresponding props in ARFaçade.

localize the sound of the phone ringing and the voice leaving a message. Similarly, if a player were to pick up the phone the voice would speak through the phone. The choice of audio is handled by the AI engine, based on the existence (or absense) of player interaction with the phone. In ARFaçade, we simply forward the commands to play particular audio files across the network to the remote machines.

We are particularly excited to investigate the larger questions about physical/virtual interaction in the desktop vs. the embodied experience: will players be conscious that the characters react to their position in the space, and how will this change the player's behavior? Will players feel comfortable actually manipulating objects in these spaces and how will this affect the experience?

4.3.2 Difficult to Convert to ARFaçade

Despite the control afforded by the WOz method, some interactions, such as touching the characters and touching objects that the characters can also touch, are very difficult to emulate in ARFaçade. It is possible to imagine workable, although inelegant, solutions allowing the player to physically interact with Trip and Grace. Whether detected with sensors or communicated through a WOz operator, an interaction protocol or gesture language could be designed for players to hug, kiss and comfort the characters. However, it would be awkward for players to perform these actions, and recognition of them might hard for both sensor-based automation or a wizard. Furthermore, without haptic feedback, it may seem unnatural to players to hug the air.

A number of items in Façade, such as the mixed drinks, magic 8-ball, and front door present a challenge because both players and characters can manipulate them. In ARFaçade, the virtual characters cannot pick up physical objects, nor can players directly manipulate virtual objects. While not ideal, our solution to these problems is to modify the story. For example, sometimes Trip makes a drink and either hands it to the player or places it on the bar. One possible way to change the interaction is to have Trip always place the drink on the bar and never touch it again, at which point a physical drink is revealed in the predetermined position. All such interactions will require experimentation and will likely be handled on a case-by-case basis.

4.4 Player Movement

Several challenges must be addressed with respect to player movement. In the desktop Façade, the player controls their position and a single orientation representing both their head and body (which are assumed coupled). In ARFaçade, we needed to de-couple head and body movement as is done in many first person games (such as in Quake where players can move independent of where they look/aim). Because Trip and Grace pay attention to their location relative to the player (e.g., they try to stay in front of the player when they want to talk), the player's body orientation should be used by the character engine rather than the player head orientation (which is used for rendering). The AI engine also watches the frequency of player movement to decide if they are acting "nervous," and has Trip ask them to leave if it decides they are moving too much. The stable movements of the body, not the rapid movements of the player's head, should be used in both cases.

More challenging are the times in desktop Façade where the AI engine moves the player viewpoint. Most automatic player movement (adjustments when sitting on the sofa, when too close to the walls, etc.) can simply be disabled in ARFaçade. However in one possible ending, Trip throws the player out of the apartment. Since Trip cannot manhandle a physical player, how do we recreate this ending in ARFaçade? One idea is to have the entire scene fade into a virtual space around the player, just long enough for the player to see herself get thrown out. If this is disorienting for the player we will try other techniques, such as simply fading to black.

In desktop Façade, we have observed a common trend of players to rapidly explore the virtual apartment before settling into interacting with Trip and Grace. However, since player movement provides interaction cues for Trip and Grace, this game-like exploration is contrary to the social setting of the experience. In ARFaçade, we are interested if the physical nature of the space will encourage more "appropriate" behavior, or if players will continue this inappropriate initial exploration of the space. Furthermore, would the AR gear (head-worn display, backpack with computer and sensors) reduce movement and exploration? Would the video mediated view of the world make it difficult to interact with physical objects, perhaps due to parallax between the display and the periphery?

5. IMPLEMENTATION

In this section we describe the technical implementation of ARFaçade. We refer readers to other papers to learn about the implementation of Façade [15, 16].

ARFaçade runs on a Windows XP laptop mounted on a lightweight external hiking frame with a small pouch to hold wires and a battery for the camera. The player wears an eMagin Z800 3DVisor head-mounted display, mounted on a medical headband. The display has two bright, high-contrast OLED displays and a 40 degree diagonal field of view, and is integrated with an extended-head Point Grey DragonFly camera (pointing forward) and an Intersense IS-1200 Vistracker (pointing upwards) (see Figure 5). We constructed the physical space to match Trip and Grace's apartment as close as possible, given the constraints existing physical room (e.g., Figure 4). The walls are made of wood and off-white, slightly stiff, linen fabric. The paintings are either original Façade images or recreations, printed on large poster printer paper and framed. We gathered a temporary collection of furniture and knick-knacks to

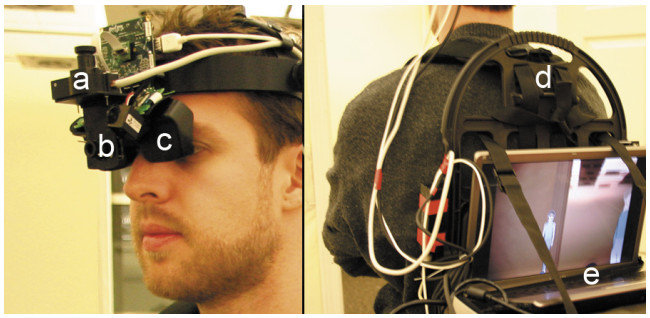

Figure 5: (a) IS-1200 VisTracker location tracker (b) Point Grey DragonFly camera (c) Head-mounted display (d) Inertiacube3 orientation sensor (e) Laptop computer

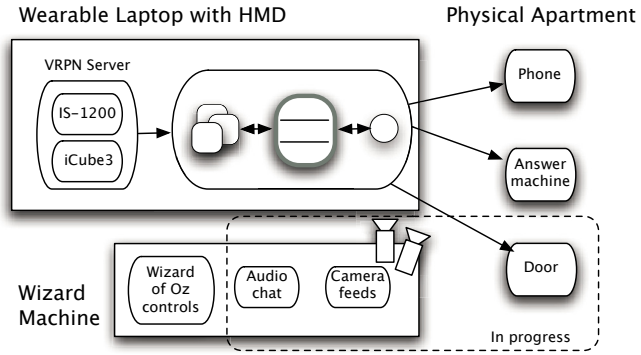

Figure 6: Architecture of ARFaçade

stand in for the rest of the physical items until more exact replicas can be found that match the stylized post-modern décor in Façade.

Several changes had to be made to the existing code base. The player's screen consists of video from the physical environment in his field of view overlaid with the virtual characters drawn at their correct location. The video feed is rendered into a background texture and most of the existing virtual objects are not drawn, but rendered into the Z-buffer. They exist as physical objects in the system and can occlude the characters–for example, Trip appears to go behind the physical bar when he is making drinks. Furthermore, the existing virtual map had to be changed to match the physical setup. All of the hardcoded locations for objects like the couch, tables, bar, etc. had to be modified in the graphics code. Likewise, the AI code contains hard coded values for character staging and path planning, and responses to player locations–these values had to be modified accordingly.

We used two trackers to decouple the body and head: an overhead hybrid inertial-vision tracker (IS-1200 Vistracker) to track the player's 6DOF head position and rotation, and an inertial orientation sensor (iCube3) to get the relative rotation of the body. The head tracker is accurate to within a few millimeters–good enough to be used for the user's viewpoint in the graphics engine.

A Wizard of Oz interface runs on a second computer at a desk outside the apartment, and lets a wizard handle speech input and references to objects in the space. The WOz interface has series of reference buttons, used by the wizard to signify a player's verbal or gestural reference to physical things and a text field to type the player's statements. The WOz Interface communicates with the wearable machine via TCP/IP. We plan to add cameras and a microphone to the space to give the wizard more visibility on what the player does.

6. INITIAL OBSERVATIONS

Although we did not conduct a formal evaluation, we demonstrated the experience for about 80 people during a game symposium in February. Many of the attendees are experts in the game industry and provided us high-level feedback similar in flavor to a group brainstorming session.

Over the course of several hours, several individuals ran through part of the experience wearing the backpack and head-mounted display. The audio was played through the backpack computer's speakers rather than through headphones, and the player's view was visible on the backpack computer's display in addition to the head-worn display. This allowed waiting visitors to enjoy the experience and discuss it with the research team. The wizard operator also listened to the context of conversation to help decrease errors in typing player statements–especially important in the noisy environment of the demo space.

Our impressions from preliminary demonstrations are extremely positive, as most players have stated a preference for the embodied version over the desktop version of Façade (all were familiar with Façade, and most had played it). Participants enjoyed the physical embodiment and ability to interact directly with the props, noting such things as how playing the audio from the phone and answering machine increased the experience's engagement. These initial demonstrations also illustrate how players modify their behavior and attitudes, as we discuss below.

6.1 Side Commentary

Most players spoke during the experience, but not always to Trip and Grace. Many players seemed to engage in a kind of "side commentary" with the crowd, the research team, and even themselves (e.g, "Did you see Trip's reaction?"). While this commentary often made it difficult for the wizard to decipher statements intended for Trip and Grace versus the general audience, it provided an valuable outlet for players to expose inner thoughts. One game designer even commented that "everyone talks to themselves out loud when playing these sorts of games."

If exposing inner thoughts is important to players, especially in a public showing, should we design for this explicitly? The WOz operator must have some way of knowing which statements should be passed on to the AI engine and which ones are just talking points. How does this change the conversation with the WOz operator? If we did provide a means to support side commentary, could we take advantage of this modality for other difficult interactions? For example, the player could say "I want to kiss Grace" and the wizard could press the "Kiss Grace" button. Such statements may be ambiguous and could be misinterpreted as a player's statement by the wizard. Players could also be given a way to back out of statements typed by the wizard, such as saying "no, no... erase that. I meant...".

6.2 Physical/Virtual Interaction

The initial evaluation illustrated how players want to be able to interact physically in the space. Several players touched glasses on the bar and sat on the sofa. People seemed more likely to move near the Italy picture when Trip invites the player over in ARFaçade versus the desktop version. People were less likely to rampantly

explore the space as players do in desktop Façade. In general, player movement in ARFaçade seems more calculated, closer to real-life patterns of movement. Clearly, virtually moving with keystrokes is much less daunting than physical movement. This was partly due to the context; it was a demo session with lots of onlookers and a bulky HMD and backpack.

More interestingly, several participants cited social reasons for not exploring the kitchen, saying that it seemed inappropriate to go into their friends' private space. The fact that the experience took place in a physical space, rather than a fantasy space, seemed to have an effect on some participants' sense of social boundaries. Going into the back room seems perfectly ok in a virtual game, but is suddenly taboo on a physical set.

Players in both versions tend to back away from the characters when they are too close. Backing away from friends is not normal social behavior, and likely a function of wanting a wider field of view since both desktop and AR versions have a much smaller field of view (~30-40 degrees) than our natural human vision (closer to 180 degrees). In subsequent evaluations we intend to record the position and head rotation of players to look for quantitative patterns and distinctions from desktop Façade. We are also considering the purchase of a much wider field of view head-worn display.

6.3 Group Experience

Having second hand visitors stand behind the player changed the experience in interesting ways (see Figure 7). Many players were aware of the crowd and would treat the experience more as a performance, saying and doing things to get funny reactions from Trip and Grace. For example, one player took an empty glass from the bar and pretended to drink a martini poured by Trip.

This sort of appropriation of the experience is natural and occurs with the desktop version. For example, one blog community has started trading "screen plays" generated from desktop Façade game play (e.g., one player pretended to be a zombie, saying nothing but "brains", and posted the resulting script). We are interested in understanding these group experiences and how appropriation plays out in an actual setting, in addition to conducting controlled, experimental comparisons with desktop Façade.

Another important question to consider is what people do while waiting to play. First person AR experiences are generally individual experiences where visitors have to wait in a queue for their turn in the space. Other mixed reality experiences, such as Desert Rain [11], explicitly talk about setting the back-story and handling the pre- and post-experience. One possibility would be to show waiting visitors the first-person viewpoint of the current player. During our demo session, visitors simply trailed behind the current player to see the laptop monitor on his or her back. This created a sort of party atmosphere and clearly affected the player's experience.

Another way to consider the group experience is to create alternative versions of ARFaçade itself. One powerful idea, that we tried during our demonstrations, was to have a second player control the wizard interface, driving the conversation to create a sort of improvisational karaoke for the in-situ player. While this dramatically changes the nature of the experience by taking away the player's control of what they say, it's creates new opportunities in the group setting. For example, a partner can type things the player would never want to say in front of a crowd, giving them the license to act out in ways they would normally not do. Interestingly, restructuring the experience this way significantly reduces the conversation lag time of Façade because the AI engine processes the WOz statements at the same time the player reads and speaks them aloud. Trip and Grace seemingly respond immediately to spoken words, rather than waiting several seconds for WOz typing and AI processing.

Figure 7: Group experience during the demostration day for ARFaçade. Visitors could see and hear the experience through the laptop.

We are not suggesting that the group ARFaçade experience was better than individual play, but it does raise interesting possibilities for design. In contrast to other styles of experiences (e.g., "Can you see me now?" [7]) which explore different interactions between players, actors, and audiences, ARFaçade points to a particular flavor of AR experiences: a single AR user with local, interactive audience participation.

7. MOVING FORWARD

While we look forward to the dream of the Holodeck and reflect on our experience with our preliminary design of ARFaçade, there is still work to be done before ARFaçade is fullly functional. Technically, we need to rewrite the 3D engine so that Trip or Grace can be accurately integrated in the environment, including adding shadows to visually attach the virtual characters to the physical floor. We would also like to get wider field of view cameras and displays, to test if our hypothesis of why people are backing away from the characters is correct.

We intend to provide more feedback to the wizard, including multiple external camera feeds and an open-mic audio connection from the player. We also need to improve the WOz interface and the feedback provided to the player. For example, the player should know when the wizard is typing, how much of the typing buffer they have used, and when they have exceeded it. We also plan to experiment with giving the player "enter/reject" control over what they have been saying. Players would not be able to edit their statements character by character, but it would allow them to preview and choose when to enter statements.

We need to refine interactions with the mixture of physical and virtual objects in the space. For example, we plan to mechanically open and close the apartment door when Trip normally opens the virtual door. To facilitate a way for players to hug, kiss, and comfort Grace and Trip we plan to teach players three simple gestures, likely

similar to real life motions and then use a WOz operator to detect the interaction.

We are currently planning evaluations from several angles. The larger question is: how do we evaluate something like Façade? We are approaching this question with novel methods and developing an evaluation procedure for understanding the emotional effectiveness of desktop Façade; we hope to apply the same metrics to ARFaçade. By comparing desktop interaction to augmented reality (and possibly a virtual reality version) with the same underlying content and AI engine, we hope to learn interesting things about the nature of embodied interaction. If embodied interaction does turn out to be important and engaging, then how do we optimize embodiment?

We also plan to continue exploring the group experiences possible through this type of augmented reality experience. This could prove to be an equally interesting evaluation because of what it can tell us about how people adopt and enjoy these kinds of experiences. For all of these evaluations we plan to record players' navigation through space, players' gestures and speech, WOz interaction with buttons and text input, and AI processing logs.

8. CONCLUSIONS

ARFaçade is the first AR experience that combines unconstrained movement, natural interaction and sophisticated, AI-based content. Our initial observations indicate players are excited about the embodied version of the experience, but we need to conduct a formal evaluation before drawing strong conclusions. The players' behavior during our demonstration sessions points to the importance of side-commentary during play, the challenge of mixed physical/ virtual interaction, and the mechanics of group experiences. We hope that our discussion of the design challenges we faced during the creation of ARFaçade provides insight to other researchers developing embodied, AI based games and experiences.

9. ACKNOWLEDGMENTS

The Graphics, Visualization and Usability Center for their support of this work through a Seed Grant program, and to Andrew Stern, Tom Amundsen, and Tobias Lang.

10. REFERENCES

1. Bobick, A.F., Intille, S.S., Davis, J.W., Baird, F., Campbell,, L.W., Ivanov, Y., Pinhanez, C.S., Schütte, A., and Wilson, A. The KidsRoom: A Perceptually-Based Interactive and Immersive Story Environment. *PRESENCE: Teleoperators and Virtual Environments 8, 4* (Aug. 1999), 367–391.
2. Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjalmsson, H., and Yan, H. Embodiment in conversational interfaces: Rea. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'99)* (Pittsburgh, PA, USA). ACM Press, New York, NY, 1999.
3. Cavazza, M., Charles, F. and Mead, S.J. Interacting with Virtual Characters in Interactive Storytelling. *In ACM Joint Conference on Autonomous Agents and Multi-Agent Systems,* (Bologna, Italy). 2002, 318-325.
4. Charles, F., Cavazza, M., Mead, S. J., Martin, O., Nandi, A., and Marichal, X. Compelling Experiences in Mixed Reality Interactive Storytelling. *In Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE'04)*, (Singapore). ACM Press, New York, NY, 2004.
5. Cheok, A. D., Goh, K. H., Liu, W., Farbiz, V, Fong, S. W., Teo, S. L., Li, Y., and Yang, X. Human Pacman: A Mobile, Wide-area Entertainment System Based on Physical, Social, and Ubiquitous Computing. *Personal Ubiquitous Computing, vol. 8,* (2004), 71-81.
6. Corradini, A., Mehta, M., Bernsen, N. O. and Charfuelán, M. Animating an Interactive Conversational Character for an Educational Game System. *In Proceedings of the International Conference on Intelligent User Interfaces (IUI'05)*, (San Diego, CA, USA). ACM Press, New York, NY, 2005, 183-190.
7. Flintham, M., Benford, S., Anastasi, R., Hemmings, V., Crabtree, A., Greenhalgh, C., Tandavanitj, N., Adams, M., and Row-Farr, J. Where On-line Meets on the Streets: Experiences with Mobile Mixed Reality Games. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'03),* (Ft. Lauderdale, FL, USA). ACM Press, New York, NY, 2003.
8. Hughes, C.E, Stapleton, C.B., Hughes, D.E., Smith, E. Mixed Reality in Education, Entertainment and Training: An Interdisciplinary Approach. *IEEE Computer Graphics and Applications, 26(6)* (2005), 24-30.
9. Johnson, W.L., Rickel, J.W. Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments. *International Journal of Artificial Intelligence in Education, 11,* (2000), 47-78.
10. Knickmeyer, R. L. and Mateas, M. Preliminary Evaluation of the Interactive Drama Façade. *In Extended Abstracts of the SIGGCHI Conference on Human Factors in Computing Systems (CHI'05)*, (Portland, OR, USA). ACM Press, New York, NY, 2005.
11. Koleva, B., Taylor, I., Benford, S., Fraser, M., Greenhalgh, C., Schndelbach, H., Lehn, D. v., Heath, V, Row-Farr, J., and Adams, M. Orchestrating a mixed reality performance. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'01)*, (Seattle, WA, USA). ACM Press, New York, NY, 2001.
12. Laurel, B., Computers as Theatre. Addison-Wesley, 1991.
13. MacIntyre, B., and Bolter, J.D. Single-Narrative, Multiple Point-of-View Dramatic Experiences in Augmented Reality. *In the Journal of Virtual Reality, 7(1) (*December 2003), 10-16.
14. Magerko, B. and Laird, J.E. Building an Interactive Drama Architecture. *In Proceedings of First International Conference on Technologies for Interactive Digital Storytelling and Entertainment*, (Darmstadt, Germany). 2003, 226-237.
15. Mateas, M. and Stern, A. Natural Language Processing in Façade: Surface-text processing. *In Technologies for Interactive Digital Storytelling and Entertainment (TIDSE)*, (Darmstadt, Germany). June 2004.
16. Mateas, M. and Stern, A. *A Behavior Language: Joint action and Behavioral Idioms.* In H. Prendinger and M. Ishizuka (Eds), Life-like Characters. Tools, Affective Functions and Applications, Springer, 2004.
17. Mateas, M. and Stern, A. Façade: An Experiment in Building a Fully-Realized Interactive Drama. *In Game Developer's Conference: Game Design Track,* (San Jose, CA). 2003.
18. Maulsby, D., Greenberg, S., and Mander, R. Prototyping an intelligent agent through Wizard of Oz. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'93),* (Amsterdam, The Netherlands). ACM Press, New York, NY, 1993.
19. Mori, M. On the Uncanny Valley. *In Proceedings of the Humanoids-2005 workshop: Views of the Uncanny Valley,* (Tsukuba, Japan). 5 December 2005.
20. Murray, J., Hamlet on the Holodeck: The Future of Narrative in Cyberspace. New York, NY: The Free Press, 1997.
21. Riedl, M., Saretto, C.J., and Young, R.M. Managing interaction between users and agents in a multiagent storytelling environment. *In Proceedings of the Second International Conference on Autonomous Agents and Multi-Agent Systems,* June 2003.
22. Swartout, W., Hill, R., Gratch, J., Johnson, W., Kyriakakis, C., LaBore, C., Lindheim, R., Marsella, S., Miraglia, D., Moore, B., Morie, J., Rickel, J., Thiebaux, M., Tuch, L., Whitney, R., and Douglas, J. Toward the holodeck: Integrating graphics, sound, character and story. *In Proceedings of Autonomous Agents*, 2001.
23. Thomas, B., Close, B., Donoghue, J., Squires, J., Bondi, P. D., and Piekarski, W. First Person Indoor/Outdoor Augmented Reality Application: ARQuake. *Personal Ubiquitous Computing, 6*, (2002), 75-86.