

Augmented Reality as a New Media Experience

Blair MacIntyre¹, Jay David Bolter², Emmanuel Moreno², Brendan Hannigan¹

GVU Center

Georgia Institute of Technology

Atlanta, GA, USA

¹*College of Computing*
{blair, brendan}@cc.gatech.edu

²*School of Literature, Communication and Culture*
jay.bolter@lcc.gatech.edu
motiv-8@mindspring.com

Abstract

In this paper we discuss our work on applying media theory to the creation of narrative augmented reality (AR) experiences. We summarize the concepts of remediation and media forms as they relate to our work, argue for their importance to the development of a new medium such as AR, and present two example AR experiences we have designed using these conceptual tools. In particular, we focus on leveraging the interaction between the physical and virtual world, remediating existing media (film, stage and interactive CD-ROM), and building on the cultural expectations of our users.

Keywords: augmented reality, narrative, media forms, interaction design.

1. Introduction

In this paper, we present the initial results of a collaborative exploration, between researchers and students in New Media Design and Computing, investigating the creation of engaging, dramatic augmented reality (AR) experiences. For this paper, we are concerned with AR techniques that utilize *personal displays* (such as head-worn displays) that directly augment a person's perception of their surroundings. Our research goal is to simultaneously develop the narrative theory to understand how to design these experiences, and to build the authoring and runtime tools necessary to implement our designs. The purpose of this paper is to present the design method we have used to develop AR narratives, as well as specific examples of the application of this method.

The inspiration for this work is a belief that AR techniques will not move beyond specialized application areas, such as equipment maintenance, into areas of broader interest, such as entertainment and education, until we develop these theories and the corresponding tools, *and* that the best way to develop these theories and tools is through tight collaboration between technologists and New Media theorists.

Our method is to approach AR as a new *medium*, and examine it using media theory as one would any other medium (such as virtual reality, film or stage). In particular, we focus on *remediating* [2] established media in an attempt to understand and develop media forms for AR. (Media forms can be thought of as sets of conventions and design elements that can be used by authors and developers to create meaningful experiences for their target users.) Media studies teaches us that remediation is a critical tool because we never design in a vacuum, even when designing for a new medium. A user's expectations are (implicitly and explicitly) based on their experience with, and understanding of, all media forms; a lifetime of experiencing film, stage, tv and so on creates a starting point for their interpretation and understanding of any new experience. Understanding, and leveraging, these shared cultural expectations of the intended audience will allow us to create richer, more engaging and more understandable AR experiences.

Developing and understanding media forms is also fundamental to the eventual success of a new medium such as AR at a more basic level; by understanding common elements and conventions, we can begin to develop tools that explicitly support them, significantly easing AR content creation. A recent example of this process is the rapid development of HTML tagging and other web technologies after 1993, when graphic designers began to work on web sites. The introduction of the inline image tag, for example, made the web page a new media form that could be modeled on graphic design for print. Graphic designers then began to push on this new media and lead those responsible for standards to develop new versions of html and new markup languages (such as Cascading Style Sheets) to provide them with the tools they needed. This acceleration of tool development is a normal part of the development of any new medium, and happens as a matter of course once the low-level technology is practical and relatively widely available.

By involving designers while the low-level AR technology is still in its infancy, we hope to accelerate the

development of higher-level user-interface technology and media forms, as well as application and runtime architectures. In addition, since recent research has demonstrated that the common technical problems encountered when building AR systems, such as tracking and registration, are much more tractable in the context of a specific application or task domain [8, 32], we hope that by focusing on application-specific solutions to these problems we can create more robust systems and influence the development of more general solutions to these critical technical problems.

The remainder of this paper will be organized as follows. First, in Section 1.1 we give a brief overview of the importance of AR as a new medium. In Section 2, we explain how the tools of media studies can help us understand how to approach the creation of narrative AR experiences. Based on this understanding, in Section 3 we discuss the strategies we used to create two AR experiences. In Section 4 we give a brief overview of the tools we have created, and are in the process of creating, to facilitate these collaborative projects. Finally, in Sections 5 through 7 we discuss related work, our future plans and the conclusions we have drawn from our work.

1.1. A Unique New Medium

The importance and uniqueness of personal AR as a medium is the result of three features that combine to distinguish it from earlier media: blending the virtual and physical worlds, continuous and implicit user control of the point of view, and interactivity. While no one of these features is unique (except, to some extent, the blending of the virtual and physical worlds), the combination is.

These features derive from the personal nature of the human-computer interface. First, since personal AR displays (e.g., see-through head-worn displays) directly enhance a viewer's perception of the world around them, AR techniques can be used to display synthetic information *anywhere* in the user's environment, at any location or on any object. This fluid blend of the physical and the virtual, and the inevitable tension between them, offers rich dramatic possibilities that are impossible in any other medium. Second, since the user wears the display, they naturally control the point of view of the experience as they look and move around. And third, because the user is situated in the physical space being augmented, personal AR systems are inherently interactive; even if the virtual content is non-interactive, the user implicitly interacts with the physical space.

Although we will not delve into multi-user experiences in this paper, another important feature of personal AR is that the displays are only perceived by a single user, raising the possibility of creating rich, personal experiences for all occupants of a shared space. The dramatic and narrative

possibilities of AR experiences in traditionally static installations, such as theme parks, museums and historical sites, goes far beyond what is currently possible with non-AR technology (such as projection screens or the sequential or location-aware audio augmentations common in museums today).

2. AR and Media Theory

As described above, the goal of our research is to develop an understanding of narrative AR applications by considering AR as a new medium, and using the discipline of Media Studies to guide our exploration of this new medium.

While AR is unique in many ways, a medium is never just the technology itself. While the pure technology provides a set of features that can be exploited (what Norman would call "affordances" [33]), the features of the new technology will develop into one or more particular forms within a particular historical and cultural setting. Each of these media forms is the deployment of a technology according to certain conventions and practices. So, for example, the film camera, projection, and distribution system (invented and perfected from about 1890 to about 1930) constituted a new technology of representation. This technology was then developed into a large number of different film forms: the Hollywood narrative film (and its many subgenres), the documentary (and its subgenres), the propaganda film, the educational film, and so on. There is a whole literature that studies the genres and forms of various twentieth century media. (For example, for film forms see [4]).

This study of the history of media is important to the development of AR applications because AR is one of a series of new digital media (along with desktop multimedia, the WWW, VR, and so on) for which conventions, practices and user expectations are currently evolving, and which, taken together, are arguably of similar importance to print technology or to analog photography [31]. These new media can lead to new forms and conventions for both experiential and symbolic information [28]. These forms and conventions are related to the conventions of earlier media forms, in particular from film, stage, radio, and television, as well as print [2].

As discussed in the introduction, users of new digital media rely on their understanding of the conventions of the relevant earlier media forms; when the conventions diverge, users have to be "tutored" in how to interpret the new media. We seek to define appropriate conventions for AR that will allow developers to create meaningful information experiences for their users while avoiding the need for "tutoring" users. We believe these conventions must grow out of earlier media forms. Because Media

Studies provides explanations of these earlier conventions and their significance for viewers/users, we draw on this discipline to understand how to adapt these conventions to AR.

In the case of AR, we believe that film and stage production will be an important source of relevant earlier media forms, just as they have been for interactive narratives and virtual reality. For example, filmmakers have a century of experience in telling stories through characters and camerawork, and stage directors have an even longer tradition of controlling a story and the viewers attention without the luxury of camerawork. Some of their insights can be borrowed and taken over into AR for conveying narrative information. Techniques of characterization can be adapted to create characters who will appear before the AR user as if standing in the physical room (for example, rendered as “video actors”, as discussed in Section 3.1.1). Users will have expectations about how such characters should behave based on their experience of movies or stage plays (for example, the ghost movies form is based on expectations of how ghosts appear and behave in film).

Users may also have expectations about camera work based on film, but this is an area where adaptation is not as simple. In AR, unlike film, the user is in control of the “camera”, rendering many film techniques (panning, zooming, etc.) unusable. In this case, it is useful to look instead at stage to see how the narrative is controlled and advanced. For example, as the use of lighting (e.g., to dim the stage lights and cast spotlights on one or more characters) and actor staging (e.g., where an actor may walk across part of the stage while talking to draw the viewer’s attention) seem directly applicable, but because the AR viewer is “in the middle of things” and cannot see all of the “stage”, these common stage techniques will require careful adaptation.

3. Creating AR Experiences

In this section, we will discuss how Media Studies influenced the design of two example AR systems. The design issues can be broken down into interactions with the physical world, remediation, and cultural expectations.

3.1. Example AR Projects

Before discussing the specific design issues with these two projects, we will briefly summarize each of them.

3.1.1. The Ghosts of Sweet Auburn

In the *Ghosts of Sweet Auburn* project, we are exploring the creation of tours of the Sweet Auburn Historic District, part of the Martin Luther King, Jr., National Historic Site and Preservation District in Atlanta. Sweet Auburn was the cultural and economic center of the Black Community in



(a)



(b)

Figure 1. Two frames from an early mock-up of the “Ghosts of Sweet Auburn” AR experience. In (a), a ghostly narrator is sitting in a physical chair, speaking to the viewer. In (b), the user is holding a book the ghost directed them toward, and as they look at different pages, the blank pages have dynamic content mapped onto them (similar to [1]).

Atlanta before and during the Civil Rights era, and was home to leaders such as Martin Luther King, Jr., and John Wesley Dobbs. We chose this topic because it is culturally relevant to Atlanta (where we are located), because it would allow us to confront difficult social and cultural issues, and because of the ready availability of rich content.

While we have prototyped, or are in the process of prototyping, a variety of tours centered around Sweet Auburn, in this paper we will focus on examples related to a media form for historical tours that we have come to call *Ghost Movies*. Images from a student mock-up of a ghostly tour of John Wesley Dobbs home is shown in Figure 1. In this media form, the experience centers around ghostly figures from the past, who inhabit the space and interact with the user.

There are two ways we could incorporate ghostly figures (or any humanoid content) into a 3D environment: create a 3D model of characters and animate them, or create 2D video of an actor playing the character and then texture map the video onto a rectangle in the virtual world. While animated 3D models are potentially much more flexible, they are very difficult to create and the resulting animated

character rarely possesses the same qualities as a character acted out by a reasonably good actor (i.e. in terms of body motion, facial expressiveness, and overall appearance). Therefore, we have used texture-mapped video of actors, which we refer to as *video actors*, for all of our AR experiences. We discuss an early implementation of our video actors, including the advantages and disadvantages of using video, in [27].

3.1.2. A Mad Tea Party

The *Mad Tea-Party* AR experience is based on a chapter of the same name from Lewis Carroll's *Alice's Adventures in Wonderland* [16]. The user assumes the role of Alice and sits at the tea party with three interactive characters: the Mad Hatter, Dormouse, and March Hare, as shown in Figure 2 (for a more complete discussion of this experience, see [29] and [30]). The user's objective is to get directions to the garden, located somewhere in Wonderland. The characters interact with the user and with each other. Each has a set of primitive actions that they can perform, including serving tea, receiving tea, sipping tea, asking riddles, and various reactions to events that may occur in the story environment. If properly provoked, a character may splash the user (or another character) with tea. The user can address the other characters, and has a range of gestures for virtually serving, receiving, sipping and throwing tea¹.

The characters have procedural behaviors that govern how each character acts or reacts to the user or other characters in the scene. Each action represents a primitive story element—the progression of these elements builds the overall narrative experience. For this experiment, only a select set of procedural behaviors has been implemented. The tea-party setting allows the user to be seated, simplifying position tracking. The teacups and teapot are part of the story and provide physical objects for interaction. The characters in the scene are simple, yet each provides opportunities for dramatic gestures.

3.2. Leveraging The Physical World

One of the key problems with an AR narrative, shared by Virtual and Mixed Reality (VR and MR) narratives [7, 22], is creating a story-world that feels real and unconstrained, while at the same time serving as the setting for a narrative experience for the user. These two requirements are contradictory: if the world feels real and is unconstrained, it is difficult to ensure the user is at the right place at the

1. Initially, the gesture recognition and audio level sensing were simulated using a "Wizard of Oz" interface (where an operator pressed keys on the keyboard based on user actions). We have recently implemented simple gesture recognition using a Polhemus magnetic tracker attached to the props, and now use audio-level sensing to capture when the user addresses a character.

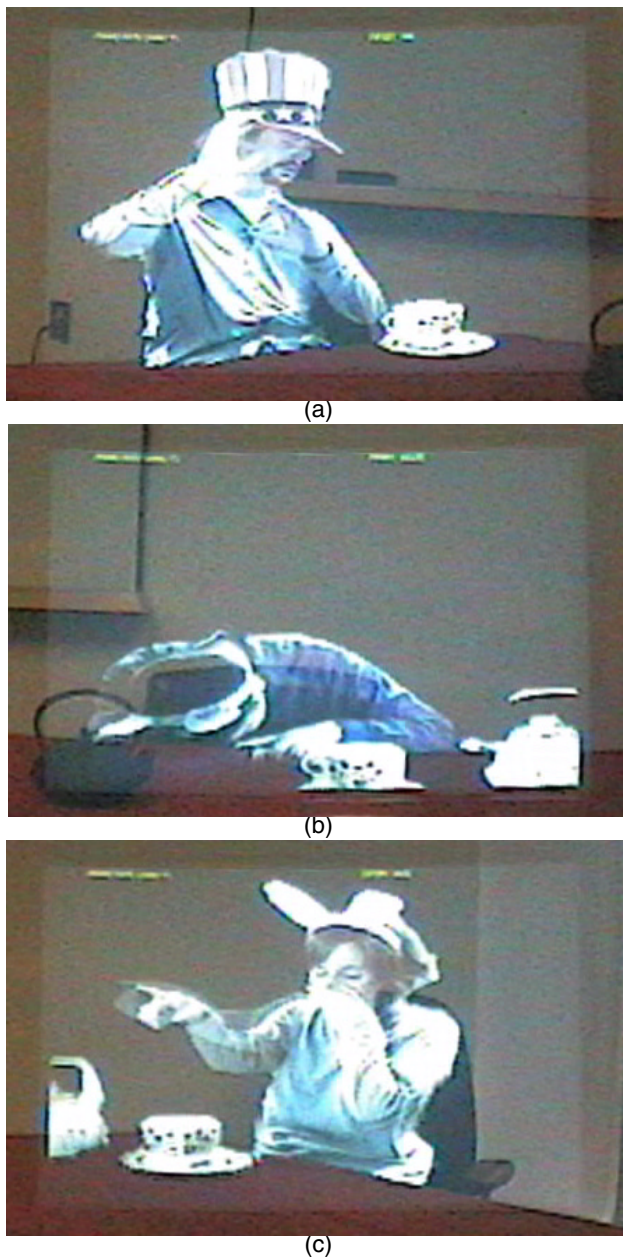


Figure 2. A Mad Tea Party. The user looks across the table at (a) the Mad Hatter to their left, (b) the Dormouse across from them, and (c) the March Hare to their right. The Mad Hatter has just been splashed with tea by the user, causing the March Hare to laugh at the Hatter. The Dormouse is asleep, but will soon wake up from the noise.

right time to experience the narrative, and that they understand how they can interact with the story.

In VR worlds, systems have resorted to crude techniques such as teleporting users to the correct location, and giving them virtual palettes of tools to support their interaction with the world. These techniques are both possible and necessary because of the disconnect between the user's

physical and virtual environments; neither are possible in AR. These location problems arise partially because of the novelty of the experience for some users (who may ignore the story to explore the space), but also because of the difficulty of using VR input devices. For example, in [22], the system operators would occasionally re-orient characters (without their knowledge) who were obviously disoriented or stuck.

AR has the advantage that users inherently know how to move about the physical world, but we must still help them understand what they can do to affect the world, and account for unexpected user behavior. For example, in the Ghosts of Sweet Auburn project, users may want to explore the historic setting of the experience without completely abandoning the narrative. We do not deal with this latter problem in our current prototypes, but discuss some possibilities in Section 6.

To help the user understand how they can affect the narrative, we rely on two things: the affordances of the carefully constructed physical space and its relationship to the virtual world, and symmetrical activities of the characters. In *A Mad Tea Party* (or, simply, *Alice*), the teacup and teapot are seen physically on the table and virtually within the story-world. This coexistence is a cue to the user that the objects have significance. The character actions provide further cues to the user, since they perform all of the activities available to the user (and very few others). As the user witnesses other characters sipping, serving and splashing tea, the significance of the objects is suggested. As the characters wake up the Dormouse, or address and answer each other, the user is subtly encouraged to participate in the same manner. Finally, the actions of the procedural characters discourage the user when acting inappropriately. If the Hatter is asking the user a riddle, the Hare's attention is turned towards the Hatter. The Hare's orientation towards the Hatter is a subtle cue that if the user tries to address the Hare during the Hatter's riddle, the Hare will ignore the user.

3.3. Remediating Film, Stage and Interactive CD-ROMs

A new media form often borrows simultaneously from two or more earlier forms, and this is the case with our AR ghost movies. The new form borrows from the conventions of film and drama (for techniques in defining and directing the user's attention). The borrowing from the stage convention is described in 3.3.1. AR ghost movies also borrow from the more recent forms of interactive games and fictions on CD-ROM. These games and fictions create story worlds with which the user can interact, although the interactions are typically defining and delimited by a basic linear plotline (spine). In 3.3.2 we describe how techniques for plot structuring can be modified to fit the AR

environment. Finally, in 3.3.3 we describe the use of procedural characters. Procedural characters are a technique for interactive fiction described by Murray [31]. In our case, we have combined Murray's ideas with the notion of objective-driven characterization that is common in traditional American film [9].

3.3.1. Co-opting Stage Direction

One problem with creating AR narrative story-worlds that feel real and unconstrained is that, even if the user is at the right place at the right time, there is no way to guarantee they are looking in the right direction to see what they need to see.

As mentioned in Section 2, we believe the two obvious candidates for remediation are film and stage. Film has an immediacy that is similar to AR, with the camera often at the viewer's location. Unfortunately, most of the narrative techniques in film center around camerawork (e.g., pans, cuts, and zooms), which makes them difficult to adapt to an environment, like AR, where the camera is under viewer control.

Stage direction, on the other hand, provides many useful techniques for AR environments. In a stage production, for example, if the user should be looking at some character or part of the stage, the other characters may look or move in that direction; the user naturally follows the gaze or motion of whatever character they are looking at. We use this technique in both productions. In *Alice*, when something interesting is happening with a character—such as when they react to being splashed with tea—the other characters look, laugh and point at them, making it clear where the interesting action is taking place. In the ghostly tour of the Dobbs home, the ghosts would move toward items of interest: if the ghost of Dobbs was talking about his treasured journal, he would move to stand beside a chair that the journal was sitting in front of, drawing the user's attention to the chair and (hopefully) the journal.

These techniques rely on, and are designed for, cooperative users; if a user does not want to take part in the narrative, there is little we can do to force them to do so. However, there are a significant set of users who want to experience the narrative, but occasionally get distracted or lost (by, for example, exploring the objects in the physical world). Well designed cues can help these users re-orient themselves.

3.3.2. Linear Spines, Interactive Cul-De-Sacs, and Procedural Nodes

The *Mad Tea Party* bases its narrative structure on a remediation of the common structure of desktop interactive narratives (i.e., CDROM games). Samsel and Wimberly argue that all coherent interactive stories are essentially linear narratives, with user actions simply modifying how

the story is presented while the author maintains control of the narrative [37]. They promote a technique they refer to as the *interactive cul-de-sac* (user-choice points, or nodes, located on a linear narrative spine, that contain different scenes and are chosen based on user action). This convention works well for many CDROM-based applications that are displayed on the computer screen, in part because of the separation between the physical world and the storyworld. However, the use of a linear spine and cul-de-sacs poses problems for AR, since the user would be required to be passive much of the time, watching while their chosen path unfolds as a scene variation in the story world.

When users assume roles in AR, typically their own physical bodies become the “avatars” for users-as-characters in the story-world. The author has no direct control of the users’ bodies; the users must actively perform (control their avatar) to move the story forward. The author must have a strategy for scripting the user on a level more basic than choosing occasional optional pathways or varied scenes. The individual actions of the user must be scripted (predetermined and encouraged), evaluated, and used in a manner that a conventional cul-de-sac cannot provide.

In Section 3.2 we discussed some of our strategies for constraining users to follow the story and discover what actions they can perform. Assuming these techniques work, users will find themselves in an environment where they can perform actions that affect other characters. The question is how to integrate these actions into a traditional narrative. While fully procedural authorship is conceptually attractive, it is untenable for complex narrative. Therefore, we maintain the concept of a linear spine, but replace the cul-de-sacs with *procedural nodes*. For the procedural node, we adopt Murray’s model of a procedurally authored story, where the author creates basic building blocks, or “primitives,” that can be arranged differently to construct a coherent story [31]. The primitives are the basic actions or gestures of the user as structured by the author. The computer as story-presenter responds to user’s gestures—first by capturing and analyzing the gestures, then by applying procedural rules to determine the appropriate story element (as designed by the author) to present.

Therefore, rather than producing several varied linear scenes for a cul-de-sac, we can focus on creating primitive story elements that are attached to the basic scripted actions of the user and other interactive characters present in a scene. The actions and the corresponding story primitives fit within the framework of the linear narrative spine; as certain combinations of actions and events occur, the story is advanced from node to node, all the while appearing to be a reactive, procedural system. Variation

occurs with the user’s actions, which result in varying arrangements of story primitives. The tools and mental framework used to organize a story along a linear spine are preserved, albeit in a slightly different form (e.g., the screenplay), while the static nature of the cul-de-sac is eliminated [29].

3.3.3. Objective-based Characters

To heighten the illusion that the story world is independent of the user, the story world must appear to continue whether or not the user takes action. If the behaviors of the interactive characters are solely reactive to the user, the story will pause when the user is inactive. Therefore, Alice borrows from film conventions to create objective-based characters that behave not only in reaction to the user, but also in pursuit of their objectives.

Decker’s model of the character structure and objective drive in the American film is used as a formulaic basis for describing procedural character behavior [9]. Each character is given an objective, a set of possible actions, and procedural rules (behaviors) that dictate how the character uses possible actions in pursuit of an objective, or in response to other characters and story events. While this model is similar to an agent architecture with a blackboard-based communication scheme [5], the intent is to create the illusion of independent character action, not to create truly intelligent agents.

We structure the agents using Decker’s terminology, dividing them into a main character (the user), opposition character, and window character(s). The essential element of the opposition character is that his or her objective pursuit ultimately hinders the main characters objective pursuit. Events that occur in the story world also provide obstacles in the objective pursuit. Clearly defining the objective obstacles is the essential step towards creating drives.

Character drive includes the decisions that the main character makes. Story drive includes the events that occur to the main character. Objective drive is the path that the user takes from the beginning of the movie to the objective at the end. The path consists of choices the main character makes (character drive) and/or choices forced upon the main character by the world (story drive). These formulaic conceptual tools are used for creating content in mainstream American films, and can be translated into programming terms useful in the creation of procedural characters, helping to bridge the gap between technologist and content designer.

In the creation of Alice, story drive is translated into a sub-set of environmental properties, procedures, and triggers beyond the control of the user or any other character. Character drive is translated into a sub-set of properties, rules, and possible actions for procedural

characters, combined with the possible actions and history of decisions the user has made during the experience. Finally, objective drive is translated into a set of procedures that evaluate the story and character properties and determine future story events and/or the actions of procedural characters.

For example, the Dormouse's objective is to sleep, but he cannot go to sleep with a full cup of tea. If his glass is full, he will empty it by splashing another character before going to sleep. The March Hare's objective is to serve tea—whenever someone's cup is empty, the March Hare will attempt to fill it. The Mad Hatter's objective is to tell riddles, which he cannot do unless another character is willing to listen to him. None of the other procedural characters can listen to him if tea is being splashed. It is likely that he will attempt to wake the Dormouse to tell him a riddle, which leads to the Dormouse attempting to sleep, splashing tea, and so forth.

Any one of the characters can provide Alice (the user) with directions to the garden, but certain conditions must be true. The character cannot be upset with Alice (from being splashed or ignored by Alice). The character's personal objective must already be met. The user must discover how and when to attend to a character in order to obtain the directions.

The characters have looping idle states that occur when no procedural rules apply. Neutral postures are utilized to minimize the appearance of "ticks" and "jumps" that are common in looping video segments. The technique is time consuming during video production, and does not completely eliminate the looping artifacts.

3.4. Cultural Expectations

As we noted above, when a media form is new, viewers or users will not yet have established expectations about how the form is supposed to operate. By appealing to these expectations, we can encourage the viewers to suspend their disbelief and enjoy a more emotional or a subtler relationship to the AR piece.

Thus, in the Ghosts of Sweet Auburn project, we build on our cultural familiarity with ghosts (from books, plays and films), and create a narrative by having ghostly characters from the past react to the user as he or she interacts with physical objects in the environment. The design idea was inspired by a discussion of the characteristics and limitations of the available technology: graphics displayed on optical see-through head-worn displays have a translucent, ghostly appearance, and the limitations of tracking technology can cause objects to jiggle and swim in the air as a ghost might do. Furthermore, the near impossibility of simulating realistic human-human conversation is avoided because people

"know" that ghosts often appear, deliver their message and leave without complex interactions.

In the case of the Mad Tea Party, we also rely on our audience's general cultural familiarity with the nature and content of this most famous children's story. Most users will have at least a general idea of Lewis Carroll's stories, and their background knowledge will help to put them in an appropriately playful mood, which facilitates interaction with the fanciful characters.

As mentioned in Section 3.1.1, video actors are an appealing addition to a 3D AR narrative because the expressiveness of a real actor is extremely hard to mimic with a 3D animation. In the context of narratives designed to evoke emotional reactions, such as tours of historical sites, actors in video form can help bridge the gap between "computer generated" and the physical environment. An actor dressed as a wounded soldier at a war memorial is far more likely to emotionally involve the viewer than an animated version of that soldier.

4. Design Tools

One of our initial goals was to use our collaboration to drive the creation of new AR tools and technologies. Early in our work, we had cross-disciplinary design teams working together, with the intent of having the technologists build the tools needed by the designers to implement their ideas. From these projects, we developed our initial version of the Video Actors environment in Java (using Swing and Java3D), discussed in [27].

However, this approach did not fit the way the New Media designers were used to working in more established media, where they have access to a host of tools with which they have remarkable proficiency. Over time, the designers began to co-opt familiar tools to prototype their ideas; both examples discussed in this paper were built using tools familiar to the designers. The tour of Dobbs' home was initially prototyped in a combination of video tools including Final Cut Pro, Adobe Premiere and Adobe After Effects. The Mad Tea Party was built using these same tools (to create the video actors) and Macromedia Director (to create an interactive experience); during the course of building the prototype, the designer implemented most of the functionality of our Java-based Video Actor environment in Director.

In the end, we recognized that the technologists need to learn to use the designers' tools (e.g., Director) and extend them to support AR (e.g., to import tracking information). In retrospect, this is the only reasonable solution; it is unlikely we will build an entire prototyping environment to rival the capabilities of tools like Director, and if we did, it is unlikely the designers could be as effective with them. Therefore, we are currently building native plug-ins

(“Xtras”) for Director to turn it into a full-fledged AR prototyping environment for optical and video-mixed AR, including support for tracking, spatialized sound, gesture recognition and so on. We are also extending our video actor tools to replace repetitive looping videos of idle characters with *video textures* (the video texture algorithms can generate arbitrarily long, natural looking video sequences of idle characters) [38].

One of the research contributions of the Alice VR System [6] illustrates the impact we hope to have by enabling designers to build AR systems using familiar tools. The seemingly straightforward definition of a few simple keywords (e.g., *asSeenBy*), that allow designers to specify 3D graphical relationships in ways that map directly to the relationships between the camera and objects in the VR world, greatly lowered the threshold of entry to VR programming. Similarly, Pausch has shown that tools designed by technologists for animators often do not match the tasks they want to do, such as 3D graphics lighting systems that automatically add highlights to handpainted textures (something the designers often want full control over) [35]. In both of these cases, working with real designers to create tools that matched their needs provided insights that were otherwise unavailable.

5. Related Work

AR has been put forward as a powerful tool for a variety of applications, but most research has focused on task-specific applications (e.g., in architecture and construction [15, 36], medicine [42], industrial applications [14, 32], situated tours [13, 18]), as well as for education and training in these same application domains. Some research has also focused on the use of AR as application-agnostic computing environment, allowing wearable computer users to be continuously enveloped in a 3D augmented world as they go about their daily life [23, 41].

Unfortunately, very little work has been done to develop non-trivial content for AR applications: notable exceptions include Höllerer and Feiner’s work on historical campus tours at Columbia [18] and Billingham and his colleagues’ work at the HIT Lab (e.g., [1]). This scarcity of complex applications is understandable, both because a significant set of AR applications are task-specific and can get by with simple content, and because many of the hard problems that must be dealt with when creating an AR system (e.g., tracking, registration, and interaction) are still active areas of research (e.g., [17, 19]).

In many ways, our current work and research plan has more in common with multi-disciplinary media design projects, such as those at the Computer Related Design program at The Royal College of Art (RCA) in London (e.g., [11]), and the Interactive Cinema group at the MIT

Media Lab (e.g. [40]). Each of these groups have been exploring the possibilities of various new media. Glorianna Davenport’s Interactive Cinema group has been looking at the possibilities of narrative or dramatic experiences in digital environments, including AR environments. Like us, they have recognized the importance of drawing on and enhancing earlier media forms, such as cinema and dance.

However, we believe that there is an important difference between our work and that of these other groups. Aside from the fact that we are focused on AR as a unique medium, the purpose of our collaboration is to advance the state of the art in user-interface technology for AR through an exploration of the media design issues of this new medium, while these other groups are primarily focused on exploring and understanding the possibilities of these new media with less of a focus on the underlying software architectures. This is not meant to denigrate the significance of their work, but rather to point out a difference of intent.

There is also support in the HCI and education literature for our belief in involving designers in the development of new media technology. Oren asserts that technologists cannot design media, it must be done by those trained in content creation. He makes the point that while Edison invented motion pictures, he did not invent film [34]. Others point out that what we need to understand when developing support for new media is the act of composing content, since the most critical aspect of a medium will be about composition, from experts down to everyday people [10, 39]. Since most of the people that end up composing content for any medium are neither technical nor expert designers, it is vital that we understand the common media forms so that tools can be built that make it easy for non-experts to create content that follows these forms.

The motivation for this work is the strong belief that our previous work, applying an understanding of earlier media forms to new media forms [2], will yield significant results in understanding the design space of AR. Part of this belief stems from the success of those mentioned above, and others, who have advanced the state of UI research by applying knowledge of media such as film (e.g., [20, 26]), comics (e.g., [24]), and traditional animation (e.g., [25]) to the design in user interface software and technology.

6. Discussion and Future Work

Our goal has been to empower designers with backgrounds in film and multimedia to create compelling AR narrative experiences. Therefore, we are extending a popular multimedia design application, Director, for scripting interactive experiences. But work can be done to give the designers a greater command of the tracking and display technologies needed for AR.

In addition, we need to continue to adapt the conventions of AR ghost movies to the existing technology, even as we work to improve that technology. For example, the problems of limited field of view and of latency in tracking have sometimes made it difficult to exploit stage conventions. Some users reported that it was hard for them to turn their heads quickly enough to see action outside of their field of view. Until tracking is improved, we will have to develop methods for “telegraphing” actions, perhaps by exploiting conventions from cartoon animation [25].

Our early design experiments have also raised a number of research questions around the general issue of adapting earlier media techniques to AR. These questions include:

1. How can we convey implicit information, such as foreshadowing or conversational flow? In film, such information is normally conveyed through editing (such as the *shot-reverse shot* technique). But editing in this sense is not available to us in AR.
2. Which genres of film or stage production can we carry over into the AR environment? Is the documentary genre effective for conveying information in AR? Should we use a disembodied narrator or a visually present narrator?
3. How should audio be handled? How important is 3D spatialized sound for conveying information? Is simple stereo sound adequate, and if so, for which situations?

One of the most intriguing areas in which to develop new conventions may be the control of the temporal dimension of the presentation. We can draw on film and on stage production for techniques of accelerating, slowing, or even stopping time as appropriate. For example, we are considering how to create an AR version of the soliloquy, in which one character delivers an extended monologue. The solution may be to stop the action, essentially freezing all the other characters, so that the one character can speak. In other situations, we might allow the user to stop the action, look or even walk around the frozen scene, and then resume the action. These pauses are just one example of the ways in which time might be manipulated in an AR experience. Another example would be the use flashback or flashforward, familiar from film and stage.

7. Conclusions

In this paper, we have presented the design method, based on conceptual frameworks from media studies, that we have used to develop AR narratives, together with specific examples of the application of this method, taken from our *Ghosts of Sweet Auburn* and *A Mad Tea Party* prototype experiences.

Our attempts to define conventions of viewing and participating in AR narratives are only first steps. We are at a moment in this new medium similar to the earliest

moments in narrative film at the turn of the previous century. We cannot yet know which conventions will seem compelling to users; only when creative artists have authored a significant number of AR experiences, will we be able to tell which of these media forms is viable.

Our early efforts are encouraging, however. Our prototypes already illustrate the value of a close working relationship between designers and technology developers. The needs of designers to remediate earlier forms place demands on the technology, and the development of technological solutions makes possible new design strategies. The prototypes also show how media theory from the humanities can be combined with a leading-edge technology to create AR media forms with potential applications in entertainment and education.

Acknowledgments

The authors would like to acknowledge the members of the Augmented Environments Lab, and the students in our AR Design classes and research projects at Georgia Tech, for their influence on this work. This work was supported by a GVU Research Seed Grant, Siemens via a GVU Industrial Affiliate Grant, ONR under Grant N000140010361, and equipment and software donations from Sun Microsystems and Microsoft.

References

- [1] Billinghurst, M., Kato, H., and Poupyrev, I. (2001) “The MagicBook—Moving seamlessly between reality and virtuality.” *IEEE Computer Graphics and Applications*, 21(3): pp. 6–8.
- [2] Bolter, J.D., and Grusin, R. (1999) *Remediation: Understanding New Media*. Cambridge MA, MIT Press.
- [3] Bolter, J.D., MacIntyre, B., Philip., K., and Harpold, T. (2000) “Cultural Narrative in Augmented Reality.” Presented at *Digital Arts and Culture (DAC) 2000*, Bergen, Norway. August 2–4, 2000.
- [4] Bordwell, D. and Thompson, K. (1997) *Film Art: An Introduction*. University of Wisconsin, Madison.
- [5] Cohen, P.R., Cheyer, A., Wang, M., and Baeg, S.C. (1994) “An open agent architecture.” In *AAAI Spring Symposium*, pp. 1–8, Mar. 1994.
- [6] Conway, M., Audia, S., Burnette, T., Cosgrove, D., Christiansen, K., Deline, R., Durbin, J., Gossweiler, R., Koga, S., Long, C., Mallory, B., Miale, S., Monkaitis, K., Patten, J., Pierce, J., Shochet, J., Staack, D., Stearns, B., Stoakley, R., Sturgill, C., Viega, J., White, J., Williams, G., and Pausch, R., (2000) “Alice: Lessons Learned from Building a 3D System For Novices.” In *Proc. CHI 2000*, pp. 486–493.
- [7] Craven, Mike, Taylor, I., Drozd, A., Purbrick, J., Greenhalgh, C., Benford, S., Fraser, M., Bowers, J., Jää-Aro, K., Lintermann, B., and Hoch, M. (2001) “Exploiting Interactivity, Influence, Space and Time to Explore non-Linear Drama in Virtual Worlds.” In *Proc. ACM CHI 2001*, Seattle, WA, March 31–April 5.

- [8] Curtis, D., Mizell, D., Gruenbaum, P., and Janin, A. (1998) "Several Devils in the Details: Making AR App Work in the Airplane Factory." In *Proc. IWAR '98*, pp.47–60.
- [9] Decker, D. (1998) *Anatomy of a Screenplay: Writing the American Screenplay from Character Structure to Convergence*. Chicago: The Screenwriters Group.
- [10] diSessa, A. (1985). "A principled design for an integrated computational environment." *Human-Computer Interaction*, 1(1): 1–47.
- [11] Djajadiningrat, J., Gaver, W., and Fres, J. (2000) "Interaction relabelling and extreme characters: methods for exploring aesthetic interactions." In *Conference proceedings on Designing interactive systems (DIS 2000)*, pp. 66–71.
- [12] Dubberly, H., and Mitch, D. (1987) "The Knowledge Navigator." Apple Computer, Inc., 1987, video.
- [13] Feiner, S., MacIntyre, B., Höllerer, T., and Webster, A. (1997) "A Touring Machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment." *Personal Technologies*, 1(4):208–217.
- [14] Feiner, S., MacIntyre, B., and Seligmann, D. (1993) "Knowledge-based augmented reality." *Communications of the ACM*, 36(7):52–63.
- [15] Feiner, S., Webster, A., Krueger, T., MacIntyre, B., and Keller, E. "Architectural anatomy." In *Presence*, 4(3), Summer 1995, 318–325.
- [16] Gardner, Martin (editor) (2000). *The Annotated Alice: The Definitive Edition*. New York: W. W. Norton and Company.
- [17] Hoff, B. and Azuma, R. (2000) "Autocalibration of an Electronic Compass in an Outdoor Augmented Reality System." In *Proceedings of ISAR '2000*, Munich, Germany.
- [18] Höllerer, T., Feiner, S., Terauchi, T., Rashid, G., and Hallaway, D. (1999) "Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality." *Computers & Graphics*: 23(6), pp. 779–785.
- [19] Jacobs, M., Livingston, M. A., and State, A. (1997) "Managing Latency in Complex Augmented Reality Systems." In *Proceedings of 1997 Symposium on Interactive 3D Graphics*, Providence, RI, April 27–30, 1997, pp. 49–54.
- [20] Karp, P. and Feiner, S. (1990) "Issues in the Automated Generation of Animated Presentations." In *Proc. Graphics Interface '90*, Halifax, Canada, May 14–18, 1990, pp. 39–48.
- [21] Kidd, C. K., Starner, T., Gandy, M., and Quay, A. (2000) "The Beware Home: A Contextually Aware Haunted House" *Georgia Tech GVU Tech Report 00-29*.
- [22] Koleva, B, Taylor, I., Benford, S., Fraser, M., Schnädelbach, H., vom Lehn, D., Heath, C., Row-Farr, J., and Adams, M. (2001) "Orchestrating a Mixed Reality Performance." In *Proc. ACM CHI 2001*, Seattle, WA, March 31–April 5.
- [23] Kooper, R., and MacIntyre, B. (2001) "An Interface for a Continuously Available, General Purpose, Spatialized Information Space." In *HCI International*, August, New Orleans, LA.
- [24] Kurlander, D., Skelly, T. and Salesin, D. (1996) "Comic Chat," In *Proceedings of the 23rd Annual Conference on Computer Graphics*, pp. 225–236.
- [25] Lasseter, J. (1987) "Principles of traditional animation applied to 3D computer animation." In *Proceedings of SIGGRAPH 87*, pp. 35–44.
- [26] Laurel, B. (1991) *Computers as Theatre*. Addison-Wesley Publishing Company, Reading, MA.
- [27] MacIntyre, B., Lohse, M., Bolter, J.D., and Moreno, E. (2001) "Ghosts in the Machine: Integrating 2D Video Actors into a 3D AR System." In *International Symposium on Mixed Reality 2001*, Yokohama, Japan, March 14–15.
- [28] Manovich, Lev. (2001) *The Language of New Media*. Cambridge MA: MIT Press, to be published.
- [29] Moreno, E. (2001) "Alice's Adventures in New Media: Towards a Collaborative Language for Augmented Reality." MA Thesis in Information Design and Technology, School of Literature, Communication and Culture, Georgia Tech.
- [30] Moreno, E., MacIntyre, B. and Bolter, J.D. (2001) "Alice's Adventure's in New Media: An Exploration of Interactive Narratives in Augmented Reality." To appear in *CAST'01*, Bonn, Germany, October 21–22.
- [31] Murray, Janet. (1996) *Hamlet on the Holodeck*. Cambridge MA: MIT Press.
- [32] Navab, N., Bascle, B., Appel, M., and Cubillo, E. (1999) "Scene augmentation via the fusion of industrial drawings and uncalibrated images with a view to marker-less calibration", In *Proc. IWAR '99*, pp.125–133.
- [33] Norman, D. (1990) *The design of everyday things*. Currency/Doubleday.
- [34] Oren, T. (1990). "Designing a new medium." In *B. Laurel (Ed.), The Art of Human-Computer Interface Design* (pp. 467–479). Reading, Mass.: Addison-Wesley.
- [35] Pausch, R., Snoddy, J., Hazeltine, E., Taylor, R., and Watson, S. "Disney's Aladdin: First Steps Toward Storytelling in Virtual Reality." In *Proceedings of SIGGRAPH 96*, pp. 193–204.
- [36] Reiners, D. Stricker, D., Klinker, G., and Mueller, S. (1998) "Augmented Reality for Construction Tasks: Doorlock Assembly." In *Proc. IWAR'98*, pp. 31–46.
- [37] Samsel, J. and Wimberley, D. (1998) *Writing for Interactive Media: The Complete Guide*. New York: Allworth.
- [38] Schödl, A., Szeliski, R, Salesin, D., and Essa, I. "Video textures." In *Proceedings of SIGGRAPH 2000*, pp. 489–498.
- [39] Soloway, E., Guzdial, M., and Hay, K. E. (1993). "Reading and writing in the 21st century." In *EDUCOM Review*, 28(1), 26–28.
- [40] Sparacino, F., Pentland, A., Davenport, G. (1997). "Wearable Performance." In *ISWC'97*, Cambridge, Massachusetts, October 13–14.
- [41] Spohrer, J.C. (1999) "Information in places." In *IBM Systems Journal*, vol. 38, pp. 602–628.
- [42] State, A, Livingston, M.A., Hirota, G., Garrett, W.F., Whitton, M.C., and Fuchs, H. "Technologies for Augmented-Reality Systems: Realizing Ultrasound-Guided Needle Biopsies." In *Proceedings of SIGGRAPH 96*, pp. 439–446.