

Three Angry Men: An Augmented-Reality Experiment in Point-of-View Drama

Blair MacIntyre¹, Jay David Bolter², Jeannie Vaughn², Brendan Hannigan¹,
Maribeth Gandy³, Emanuel Moreno², Markus Haas¹, Sin-Hwa Kang²,
David Krum¹ and Stephen Voids¹

¹ College of Computing

² Wesley Center for Media Center, School of Literature, Communication and Culture

³ Interactive Media Technology Center

Georgia Institute of Technology, Atlanta, Georgia, USA 30332

Abstract. “Three Angry Men” is an augmented reality experience, in which the user becomes a character in a radically abridged version of the play “Twelve Angry Men.” The other characters in this experience appear to the user as texture-mapped video mixed with the physical surroundings. At any time, the user may change seats to occupy a different character’s point of view and experience a different interpretation of the play. “Three Angry Men” suggests that augmented reality can be used in a new class of entertainment and informal education applications.

1 Background

Augmented Reality (AR) is a relatively young technology. Over the past decade, AR has been used or proposed for the delivery of information at the appropriate time and place in the physical environment. AR has been tested to support workers assembling wiring bundles on a Boeing 747 [1]. Various prototypes have shown how a tourist might walk around a city wearing some lightweight and relatively unobtrusive headset and receive just-in-time information about sights, restaurants, or hotels as she passes (e.g., [2]). In addition to these “practical” applications, our group is exploring AR as a medium for delivering dramatic experiences. In these experiences, the user sees and hears virtual characters who act out a drama in the physical world, the room or possibly the outdoor environment in which the user is located. The experience may be an educational one, or it may be purely entertaining or aesthetic.

The characteristics of AR give it special advantages for creating such experiences in comparison both to older media (such as film and drama) and to new media (such as as Virtual Reality or desktop DVD). Unlike almost all other media, AR can enhance the immediate physical environment with virtual characters, objects, or sensory data. Our group is exploring the combination of the physical and the virtual in a series of applications [3] [4] [5]. One such experiment is entitled “Three Angry Men,” a dramatic experience in multiple points of view.

2 Three Angry Men

The “Three Angry Men” AR experience dramatizes the way individuals perceive “facts” and other people differently based on their own background and systems of beliefs. As a basis for our application, we chose the classic television play (and subsequent film and stage play) “Twelve Angry Men,” written by Reginald Rose in the 1950s [6]. In Rose’s play, twelve men of various backgrounds are locked in a jury room to determine the guilt or innocence of a young Hispanic, accused of killing his father. Although the original vote is 11 to 1 for conviction, the lone holdout manages to convince all the others to put aside their prejudices and look at the case with both reason and compassion. The jurors work through the evidence and eventually all vote to acquit.

For technical reasons, “Three Angry Men” (TAM) is a radical abridgement of the original. Three jurors (instead of 12) debate the fate of the boy in a 6-minute excerpt. Each of the three jurors is a clearly drawn character who contrasts with the other two: Juror A is the reasonable and compassionate man who originally holds out for a discussion of the boy’s case; Juror B is an African-American, who originally votes guilty, but in fact feels sympathy for the defendant (who in our version is also imagined to be black); Juror C is a troubled man, who expresses violent prejudice against the defendant. In our excerpt, C clashes with A and B, but it eventually emerges that C has his own personal reasons for refusing to see that the case against the defendant is collapsing.

In TAM, the user wears a see-through head-worn display (see Section ??) and enters into a jury room, a physical space with a table and three chairs. When the user sits down in a chair, the drama begins. The user can now see and hear two other virtual characters occupying the other two chairs. These virtual characters are represented using texture-mapped video, mixed with the physical surround (the jury room). By sitting in one of the chairs, the user assumes the identity of one particular juror (A, B, or C) and experiences the scene from that juror’s point of view. The user can see and hear the other two characters as they deliver their dialogue; the user also hears the dialogue delivered by the juror whose chair he occupies. In addition, the user hears inner thoughts of that particular juror, representing that juror’s reflections on what the others are saying. At any time during the scene, the user can stand and change seats. When he or she gets up, the scene stops. When he or she sits down in another juror’s chair, the scene resumes, but the user now experiences the action from that juror’s point of view.

There are three principal design elements in TAM that warrant further explanation: changing point of view, the inner monologues, and a simple, implicit interactivity based on the user’s movement.

2.1 Changing Point of View

The single most important design element in TAM is the changing perception of the story based on point of view (POV). Users always see the drama from the viewpoint of one of the three jurors, and their perception of the scene reflects their expectations, beliefs, and prejudices. To represent the three viewpoints,

we shot and edited three different six-minute versions of the TAM excerpt. The actors in our excerpt interpreted their roles differently for each version, and they were encouraged to exaggerate the differences of each interpretation. For example, juror A (the compassionate juror) sees B (the African-American juror) as a potential ally, but he quickly recognizes C as prejudiced, as shown in Figure 1. So in the version from juror A's POV, B is well spoken, while C is loud, rude, and unreasonable. A himself is not visible from POV A, but he speaks his role with calm authority.



Fig. 1. Jurors B and C from A's point of view.

From Juror C's point of view, however, the roles are reversed. C naturally regards himself as reasonable, although frustrated and impatient with the other two jurors. Although he speaks the same dialogue (with minor variations) as in the other versions, his tone of voice is far calmer. Furthermore, he sees the other two with visible prejudice, as shown in Figure 2. The black juror looks to him like a "typical" inhabitant of the ghetto (in this version the juror's appearance and his dialect are stereotypical of economically deprived, inner-city blacks). Juror A looks effeminate and indecisive, a caricature of a liberal as a conservative might view him.

TAM is not a drama of subtle characterization. In fact, "Twelve Angry Men" was chosen as our model in part because the characterizations are clear-cut and extreme. The actors in TAM were asked to exaggerate their interpretations in order to be certain that the user would appreciate the effect of changing POV. Because of limitations of resolution in the video display technology and the impossibility of allowing close-ups, it was difficult for users to identify subtle facial expressions in our actors. Therefore, the actors had to achieve characterization through broad gestures and exaggerated tones of voice and accent.

2.2 Inner monologues

In addition to the external dialogue among the three characters, the user can hear inner monologues of the juror whose seat (and viewpoint) they occupy.



Fig. 2. Jurors A and B from C's point of view.

These monologues, delivered at various points during the experience, represent the thoughts or reflections of the juror about the ongoing discussion. If the user is occupying the chair (and therefore assuming the personality) of Juror A, he will hear Juror A's reflections; as Juror B or C, he will hear that juror's reflections. The external dialogue was adapted from the screenplay of "Twelve Angry Men," but the inner monologues were written by one of the project team (Jeannie Vaughn). Their purpose is to heighten the user's sense of having stepped into the particular character of the juror by hearing the juror's private thoughts.

When a monologue occurs, the external dialogue and action freezes. Dramatic time stops, as the user enters into the space of the juror's thinking. To indicate this arrested time and to differentiate the inner monologue from the ordinary dialogue, the other two characters become translucent and ghostlike, while the juror's thoughts are delivered to the user with a slight echo effect. Figure 3 depicts part of prejudiced Juror C's inner monologue, in response to Juror A's arguments.

The inner monologues are designed to reinforce the characterization of each juror and to emphasize their differing responses and biases toward each other. For example, sometimes a juror will contradict or criticize another juror more directly in his thoughts than he does in the external dialogue.

2.3 Simple, location-based interactivity

TAM is not a hypermedia experience: the user cannot make choices that redirect the action or change the outcome of the drama. The same result occurs (Juror C is discredited) no matter which seat the user occupies. Although users cannot change the events, they can alter the interpretation of those events by switching seats and, therefore, point of view. For example, a user might begin by viewing part of the drama from A's POV and then finish the drama from C's. This order of viewing would give the user a different and more negative experience than reversing the order.

Using position to determine the delivery of information is a technique commonly used in AR and VR. Researchers in VR have argued for years that position



Fig. 3. Juror C’s point of view, accompanied by C’s inner monologue: “I’ve seen a thousand kids like that and they’re all the same. Yeah, I could vote on his face. He’s got guilt written all over him. This guy knows it but he’s afraid to admit it; being politically correct, I guess. He’d be singing a different song if he ran into that kid in a dark alley.”

tracking is part of a “natural interface”—in contrast to the artificial interface of the GUI—because human beings are accustomed to finding their way around in a three-dimensional world. In AR, as in VR, it is often difficult to use symbolic or operational interface elements found in conventional GUIs. Menu and icon selection and text-typing are not good tools for AR interaction, due to the difficulty of using typical interaction devices (e.g. a mouse or a keyboard) while mobile, or simulating such devices by tracking and registering hand and finger movements with precision. More sophisticated input methods, such as voice input, have been proposed or tried for AR experiences. It is possible to imagine a version of TAM in which the user could actually speak his or her own words. Those words would elicit different responses from the other characters, who might have a repertoire of different behaviors. Thus, when participating in the experience as Juror A, the user might be able to convince Juror C to be less prejudiced. It is often assumed that the ideal work of digital narrative or drama is one in which there is a maximum range of interaction between the human user and virtual characters who are fully endowed with artificial intelligence.

Although research into virtual characters constitutes an important area in digital storytelling, we do not believe that virtual characters are the only way to achieve compelling forms of digital drama or narrative. We suggest that there is a spectrum of possible digital forms. At one end of the spectrum are highly linear or non-interactive forms (such as 3D computer graphic animation). At the other end, as people have imagined, but never succeeded in implementing, are fully interactive forms in which virtual characters interact seamlessly with human users or players (e.g., the Star Trek “Holodeck”). TAM occupies a point along the spectrum between these extremes. It is a linear drama that the user can interrupt and restart, as he or she explores three distinct points of view. We believe that any point on the spectrum from linearity to full interactivity could

produce effective experiences: there is no one single right technique. A variety of levels of interaction deserve to be explored.

3 Technical issues and constraints

TAM is currently implemented using mostly off-the-shelf technology. Our video-mixed head-worn display combines a low-end opaque, monocular display (US\$350 NTSC monocular i-Glasses) with a firewire webcam (US\$200 Videre Design DCAM). The experience runs on a laptop carried by the user in a backpack (a Dell Inspiron 8000, with a GeForce2GO graphics accelerator and 1Ghz PIII processor). The images in this paper were created in our lab, where we are using a fairly expensive (and precise) Intersence IS-600 Mark2 tracking system; we are currently in the process of building a wide area, vision-based tracker that employs a camera on the user's head to track passive fiducial markers in the environment. This tracker will allow us to deploy and test TAM in a more realistic setting (i.e. one that looks more like a jury room).

A key feature of TAM is the use of an off-the-shelf software product, Macro-media Director, to author the experience. We have integrated the tracking information and background video into Director, allowing the designers to use Director's existing facilities to control the presentation of the video texture-mapped polygons in a Shockwave3D environment. Director 8.5 provides a relatively simple 3D programming model that was adequate for this experience. The use of Director as an AR prototyping environment is important to our group's long term goal of making AR available to creative people who are not accomplished programmers.

4 AR experience design

In creating TAM and in other proposed dramatic experiences, we base our work in part on a media theory called "remediation" [7]. We regard AR as a new medium, similar in function to earlier media such as film and television. In creating experiences for this new medium, designers must depend on the similarities to other media (particularly film, television, and stage production). This dependence seems to be a general feature of the development of new media, for at least two related reasons. First, designers working in a new medium understand this new medium in terms of more familiar ones. They have generally been trained to work in some other medium, and therefore naturally incorporate practices and assumptions from their training into their work in the new medium. (For example, in the mid-1990s, when the World Wide Web was new, Web designers had often been trained as graphic designers and saw the Web as a "remediation" of the magazine page or the advertising poster.) In addition, viewers or users of a new medium will also understand the new medium explicitly or implicitly by comparison with earlier, more familiar media forms.

An AR application such as TAM remediates film (or television) and stage production. It borrows qualities from these media, while at same time providing

a novel experience that adds to what other media can provide. What this AR application adds is the intimate combination of the physical and the virtual. The user is seated in a physical space (one that ideally resembles a jury room) and so feels a sense of immersion in the drama. In designing AR applications like TAM, our strategy therefore is to look consciously for techniques in film and stage production as well as other media that we can borrow from and alter to create new experiences [3]. We have identified three major design elements in TAM (POV, inner monologues, and simple location-driven interaction); each of these design elements draws its significance for the user from its relationship to earlier media forms.

Our first key design element, the first-person point-of-view presentation, has been used for decades to tell stories in film and television. In the standard (sometimes called “Hollywood”) style, the camera occupies a variety of different viewing positions. Often the camera is used to represent what one character could see from his or her viewpoint: this is the so-called “subjective camera.” One effect of the subjective camera is that viewer identifies with the character. However, in Hollywood film this technique is generally used only briefly, for a few seconds at a time, in conjugation with third-person, objective shots. In TAM, the entire experience is viewed subjectively from the viewpoint of one of the three characters, which should intensify the user’s feeling of identification with the character. In the case of Juror C, this identification may arouse a particular conflict in the user, because of C’s blatant prejudice.

The second design element (inner monologues) is borrowed from both film and drama. Each inner monologue in TAM functions like a soliloquy in drama, where the action stops and one character expresses his or her thoughts or feelings arising from the action. The voice-over narrative in film performs a similar function, although usually in that case the visual action of the film continues while the character speaks. In this sense the voice-over can be a more “realistic” representation of the inner thoughts of a character. Users familiar with drama and film should immediately understand the technique of inner monologue that is used in TAM.

The third design element, simple interactivity through change of place, bears little relationship to the techniques used in film, but does draw upon some elements of the stage tradition. The very fact that the user can control his or her viewpoint is a radical departure from film. In film, of course, the shooting and editing determine the viewpoint, placing the viewpoint under the control of the director or editor, not the viewer. When considering user-controlled viewpoint, AR (and VR) draw more on the role that audience members play in watching a stage production. Just as members of the audience are free to pay attention to the action on any part of the stage during a theatrical production, TAM users are free to look anywhere in the room during the AR experience. As a result, each user receives a slightly different experience. The user’s control of the viewpoint also requires that the virtual characters in TAM behave more like stage actors than film characters; they must be continuously “alive” and contribute to a forward-moving narrative through their reactions to one another. They should

also provide cues, such as gestures and gaze direction, to guide the user's viewpoint to the most relevant action throughout the experience. In TAM, we take user control of the viewpoint one step further; aside from simply controlling their physical viewpoint in the environment, when the user gets up and moves to another chair (POV), they change the whole interpretation of the scene. The technique of interactivity through change of location is related not so much to older media, but to newer media such as video games and VR, in which the user's movement through the virtual space often controls the behavior of objects or the display of information. VR is probably better known to most users than AR, because VR has gotten far more attention and publicity for the last decade. As a result, users may approach the interactivity paradigm of TAM based on their knowledge of or assumptions about how VR works.

5 Conclusions and future work

We are currently reshooting and reediting TAM using new actors, and introducing a fourth character to add further dynamism and variety to the experience. But even in its current excerpted form, TAM suggests that AR can create engaging experiences by refashioning the treatment of filmic point of view. The combination of digitized "video actors" and location-based interactivity constitutes a new media form that may be effective in a variety of educational as well as entertainment applications. AR entertainment could include the repurposing of scripts originally written for the stage or the screen (as is the case with TAM), and the creation of entirely new dramatic works. Museums and exhibits could be ideal settings for educational AR experiences. For example, the Oakland Cemetery in Atlanta, Georgia was the burial site for common people and leaders of the city from the Civil War to the early 20th century. In an AR cemetery experience, former mayors might rise from their graves to acquaint a visitor with their moment in the history of the city. African Americans buried in one part of the cemetery might rise to confront the Confederate war veterans in another part.⁴

We believe that this technology could appeal to a broad spectrum of visual artists as well as museum and exhibit designers. In order to realize their design ideas, however, artists and designers must have a software environment that is familiar and relatively easy to use. We have adopted Macromedia Director as the authoring environment because it is probably the most commonly used platform for multimedia development. Because we continue to make usability and availability a main goal of our AR work, we are now considering other tools that might aid artists and designers in working in AR. Work on our first AR experiences have shown us the importance of giving designers effective tools for sketching out and visualizing their ideas prior to implementation. Designers of AR experiences need the equivalent of story-boarding, which has been used for

⁴ In the spring of 2003, we will explore some of these possibilities in a project-based AR Design class in which students will design and (at least partly) implement AR experiences for the Oakland Cemetery.

decades by directors and production people to visualize film. In the near future we hope to develop analogous tools and techniques for AR story-boarding and animatics. We believe that such tools, together with an off-the-shelf authoring environment, can lower the costs and barriers to the use of AR technology for dramatic experiences.

Finally, we realize the need for user testing of this system. One approach that seems promising is to use the concept of “presence” to evaluate TAM (and other dramatic AR experiences). Presence has been studied extensively in VR, but has seldom been studied in AR environments. Building on VR presence studies, we hope to learn whether dramatic AR experiences such as TAM can be sufficiently engaging that the user becomes present in the combined physical/virtual world of the experience, despite their continual awareness of the physical world.

6 Acknowledgements

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, by a grant from the College of Computing and by equipment donations from Hewlett-Packard.

References

1. Mizell, D., Curtis, W.D., Gruenbaum, P., Janin, A.: Several devils in the details: Making an ar app work in the airplane factory. In Behringer, K.M., ed.: International Workshop on Augmented Reality (IWAR 98), IEEE (1998) 47–60
2. Feiner, S., MacIntyre, B., Webster, T.H.A.: A touring machine: Prototyping 3d mobile augmented reality systems for exploring the urban environment. *Personal Technology* **1** (1997) 208–217
3. MacIntyre, B., Bolter, J.D., Moreno, E., Hannigan, B.: Augmented reality as a new media experience. In: In International Symposium on Augmented Reality (ISAR 2001). (2001) 197–206
4. Moreno, E., MacIntyre, B., Bolter, J.D.: Alices adventures in new media: An exploration of interactive narratives in augmented reality. In: In CAST’01 (also appears as a special issue of the web journal <http://netzspannung.org/journal/special/>), Bonn, Germany (2001) 149–152
5. MacIntyre, B., Lohse, M., Bolter, J.D., Moreno, E.: Integrating 2d video actors into 3d augmented-reality systems. *Presence* **11** (2002) 189–202
6. Rose, R.: *Twelve Angry Men: A Play in Three Acts*. Dramatic Publishing Company, Woodstock, IL (1983)
7. Bolter, J.D., Grusin, R.: *Remediation: Understanding New Media*. MIT Press, Cambridge, MA (1999)