# Effect of Dynamic Camera Control on Spatial Reasoning in 3D Spaces

Ogechi Nnadi *
CoC
Georgia Institute of Technology

Ute Fischer †, Michael Boyce ‡, Michael Nitsche #
LCC
Georgia Institute of Technology

## ABSTRACT

Game worlds have to be presented to players via some form of visualization to be accessible. Consequently, there is a direct dependency between the virtual game world and this form of presentation. But how does the camera work affect our understanding of the game space? We implemented a dynamic camera system that procedurally switches camera styles depending on the type of region the player is in. We then tested to what degree this camera behavior affects the players' understanding of the game world and its zoning in comparison with a control group playing the same zoned environment with a default camera. In both cases, recognition of the zones was lower than expected but our results show that after an initial learning phase the recognition was significantly faster when the dynamic camera system was active. Players also appeared to be less "lost" in the game world. The results validate the role of the camera in virtual spaces and suggest a stronger role for visualization strategies in 3D game worlds.

## Keywords

cinematography, interactive camera, procedural space, video game

## 1. BACKGROUND

Cameras in video games are often used to convey a certain mood or atmosphere. Intuitively, we believe this to be true based on our experience with other moving image media where cinematographers use the camera to elicit, for example, tension of a situation, intimacy between characters, or a character's personal feelings such as vertigo. In video games, this task of the camera as an emotional agent has to be carefully balanced with their functional value to present a playable game level. We believe that in interactive 3D games the camera provides the game designer with an extra channel for conveying spatial information to a player. We argue that a dynamic camera behavior can add an additional layer to the gameplay as it affects the perception of the 3D game space and allows to re-position the player in relation to it. The goal of the present study was to investigate this influence of a dynamic camera on the player's perception of a 3D game world.

Current research into the use of camera work in video games tends to focus on the camera as an Artificial Intelligence-driven system that computes different perspectives based on given parameters that usually copy existing cinematic traditions such as shot-reverse-shot patterns and live TV broadcasting styles. Instead of focusing on the value (or lack thereof) of cinematic conventions, our hypothesis is more basic. We argue that dynamic camera behavior can affect a player's understanding of a virtual space. Our claim is that a dynamic camera behavior facilitates the way we understand a given virtual space. More precisely, we argue that dynamic visualization can assist players to distinguish definite areas in the game space and thus increases their ability to recognize zones within a larger virtual environment. In our experiments, we examined players' detection of different spatial zones to track how their spatial comprehension of the 3D environments was influenced by the camera's behavior.

Since many modern video games take place in elaborate 3D environments that require constant re-assessment of the current situation this set up is relevant for numerous game titles. Virtual environments demand the gradual assembly of visual components into rather complex cognitive maps. If one can prove the importance of the camera in our understanding of game spaces, the results would also have implications for research in the area of wayfinding and navigation in virtual environments.

## 2. RELATED WORK

A number of projects have been implemented to experiment with dynamic camera behavior. Although our approach does not directly suggest a new camera AI method but concentrates more on the role of the camera for the player's comprehension, these projects serve as references for our own implementation of the test environment.

### 2.1 Camera Control in Virtual Worlds

A first tier of research focuses on AI-driven camera behavior and how it might relate to existing cinematic traditions. Drucker implemented a methodology of camera encapsulation modules, CamDroid, in order to achieve task level interaction within the virtual space. Each module is assembled with four primary components: a local state vector, an initializer, controller, and a series of constraints which are drawn from traditional cinematography. The modules are incorporated into a camera framework which includes an interpreter allowing low level manipulation and creation of virtual prototypes, along with a renderer and object database [Drucker/ Zeltzer 1995]. The result is a gradual simulation of existing cinematic techniques into an AI system.

---

* email: ogechi.nnadi@gatech.edu † email: ute.fischer@gatech.edu
‡ email: mboyce@gatech.edu # email: michael.nitsche@lcc.gatech.edu

Copyright © 2008 by the Association for Computing Machinery, Inc.
Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions Dept, ACM Inc., fax +1 (212) 869-0481 or e-mail permissions@acm.org.
Sandbox Symposium 2008, Los Angeles, California, August 9–10, 2008.
© 2008 ACM 978-1-60558-173-6/08/0008 $5.00

Building on Drucker's concepts, Tomlinson designed an autonomous cinematography architecture for interactive environments. The architecture divides shots into elements, with an emphasis on cinematic characteristic i.e. camera focus or camera angle. It uses hierarchical cross-exclusion groupings to prioritize elements. The CameraCreature, an autonomous actor drives the visualization in the virtual world. It drives the camera behavior and other visual elements based on the environmental emotional state. As an example, the sky color is modified based on the mood of the character [Tomlinson/ Blumberg/ Nain 2000].

Bares focuses on constraint control within the virtual environment. The ConstraintCam, a visualization interface device for virtual environments, manages constraint decision making based on desired visual goal states. Techniques such as picture in picture permit simultaneous multi-shot visibility [Bares/ Lester 1999].

He suggests to automatically generate camera specifications through a series of finite state machines called Idioms, which correspond to a type of scene. Using cinematographic heuristics the system chooses the best shot for the scene. The overall framework called Virtual Cinematographer also provides occlusion detection and actor-centered camera transitioning [He/ Cohen/ Salesin 1996].

A second tier of research might be characterized as more interested in the narrative tasks of a dynamic camera system. For example, Steiner focuses on story coherence versus user control. By developing a platform for managing and presenting narratives in virtual environments, a sense of balance can be upheld. The platform is divided into sections: world management states, story management states, and player management states. Camera positioning serves as one factor in these states [Steiner/ Tomkins 2004].

Young's research creates an architecture for building intelligent narrative environments. The Mimesis system is a suite of control tools used with *Unreal Tournament*. Completing successful implementation of sentinel threads, it allows for enhanced temporal events such as simultaneously occurring events. In addition the system allows for derivations from the traditional storyline, aiming to deliver dynamic experiences to the individual user [Young/ Riedl 2003].

Finally, a third research tier is interested in the intersection of spatial positioning and representation in dynamic camera systems. Nitsche combines play elements and visual representation through interactive montage. Placing specific emphasis on spatial reinforcement, he explains how adherence or derivation from traditional cinematic techniques can affect the reinforcement of the entire environment. Player-character positioning serves as one of the most dominating references in each individual shot [Nitsche 2005].

Calderon et al. investigate a dynamic system for camera positioning optimized for architectural representations of virtual spaces. Their goal is to designate specific camera work to show certain architectural structures in a form of spatially mediated cinematography [Calderon/ Worley/ Nyman 2006]. This presentation, then, should improve the perception of the given architectural design.

Our own system is less focused on a dramatization or narrative impact (as described in the second tier). Instead, as will be outlined below, it combines the spatial approach with a reference to existing cinematic traditions.

## 2.2 Game Cameras

There is a notable body of work analyzing camera control in video games that is far too rich to be covered here. In contrast to the more general principles mentioned above, many approaches in Game Studies present analytical matrixes for the predominant camera style used in a specific video game title. However, some researchers have suggested wider approaches. As our experiment focuses on spatial recognition in combination with camera work we will refer only to a selected few publications that deal with the question of space and camera use.

[Jensen 2001] and [Meadows 2002], while concerned with virtual environments in general, provide some discussion of how games create the illusion of a virtual space through the use of perspective drawing of the image. The construction of a spatial effect from a 2D image is traced back to visual expressions such as perspective rendering, colors, and shapes. These expressions, furthermore, can provide the basis for a kind of "visual story" [Block 2001] operating on the level of the image.

Wolf applies film specific methods in his analysis of how game spaces are created in commercial video games [Wolf 2001]. His approach is insofar of relevance for our project as he looks at the way in which video games use cinematic elements such as on- and off-screen space and moving cameras to generate the experience of a specific virtual space. Although 3D navigable space is only one category in his analysis, his approach provides valuable insight into the use of virtual cameras for the visual creation of game environments.

Numerous publications address the use of the virtual camera in specific games and detail how it contributes to a shaping of the virtual environment at hand. For example, Krzywinska offers interesting perspectives on a useful limitation of spatial access [Krzywinska 2002]. In her work she combines spatial analysis with the visual aesthetics of horror movies to show how they jointly provide for a mix of choice and determinism.

The necessary balance between game functionality and cinematic presentation of the events is still being debated. Exactly how much attention should we spend on the cinematic elements versus the functional gameplay mechanics? Aarseth has identified a "genre trouble" in this discussion [Aarseth 2004] – that Atkins tried to address [Atkins 2006] by pointing toward a future-oriented gaze of players. According to Atkins, players understand a video game image in terms of its potential for future interactions. This orientation of players towards possible interactions afforded by the depicted space is unlike the way film audiences read movie imagery. While such an interaction-based orientation in game design differs from a film approach, all means of cinematic imagery can still be applied. Images function as pointers to interactive options in the game space ahead.

Our prototype did not include spatial changes in the gameplay as identified by Krzywinska or alluded by Atkins. While we acknowledge that such a combination would likely enhance the perception of structure in a game space, we deliberately excluded most interactive options except for navigation to concentrate on players' spatial reasoning based on pure visual (camera) behavior.

# 3. APPROACH

Our research focuses on the issue of how players read a game space in dependency to the active virtual camera work. If a game's camera work affects the way a player perceives and reads the game world, then we should be able to trace this effect in the perception of the 3D game space.

## 3.1 *Charbitat*

To explore our research question, our project builds on the technology developed for an earlier experimental game project, *Charbitat*, which dealt with multiple aspects of procedurally generated game worlds [Nitsche et al. 2006]. The background story in *Charbitat* is that the game's hero, a Chinese princess has been poisoned and has fallen into a deep sleep. In her dreams, she explores the dream world of *Charbitat*. This dream world is generated as the player moves through it and thus provides a potentially infinite procedurally-generated 3D-environment. The underlying design for this game world encapsulates the heroine's poisoned condition as it mirrors the turmoil of her elemental balance. *Charbitat* uses a reference to the five Taoist elements of Wood, Fire, Earth, Metal, and Water that are believed to be defining forces not only for physical but also for psychological states. All objects, regions, and Non-Player-Characters (NPC) are defined by their elemental base values. That means that certain regions will always differ in terrain shape and object population from others. The same five elements are used as the necessary seed values to drive further space generation. Whenever a player defeats a "poisoned" elemental NPC, the values adjust, the underlying seed values change, and the next world generation will react to that. In that way, players drive the world generation through their in-world interactions and thus co-define the development of the overall virtual playground. In order to succeed in the game, one has to actively shape the game environment to reach specialized areas and "cure" the poisoned elements. To structure the player's progress, *Charbitat* includes procedurally generated key-and-lock puzzles, which provide further challenges to the game.

**Figure 1: *Charbitat* world seen from the default camera position**

3D game environment, objects and NPCs within this world, and conditional quest settings are all generated procedurally. For our investigation we decided to use this generative basis and extend it to the use of cameras inside the game world.

## 3.2 *Charbitat* Camera

The approach we chose to examine our hypothesis was to design two separate camera environments. Both use a version of the *Charbitat* world but were set up very differently in terms of their visualization strategies. One uses a following camera (see fig. 1), which had been used in previous iterations of the game prototype and resembles a de facto game camera convention that is not tied to the virtual space but to the central character. The other setup included a new form of camera system, wherein the camera adjusts to different perspectives dependent on the particular element that dominates the region of the game world the player navigates in.

What cameras to assign to which elemental value was an almost arbitrary decision. Although each camera points back to a specific element, the artistic representation of this element can take on many different manifestations. The overall game world has an Asian theme: the heroine is a Chinese princess, the only lettering used in-game is Chinese, and most importantly, the basis for the world generation derives from the five elements of Taoism. For this reason we turned to Ang Lee's film, *Crouching Tiger, Hidden Dragon* (cinematography by Peter Pau) [Lee 2000] to select the elemental camera positions. Lee's film stands in the tradition of the Eastern action movie but it is also accessible and well known in the West. Its wide-spread appeal and success made it a useful model to follow in our use of cameras. The selection of camera perspective remained subjective but informed by existing cinematic references. Our project does not intend to promote a single camera technique; instead, its aim was to investigate how players understand the virtual game environment based on camera changes. While the actual camera shots themselves were not the focus, they had to provide for sufficient visual differences in order to generate sufficient change in the presentation.

For example in the *Crouching Tiger, Hidden Dragon* sequences that carried references to Earth, elements appear to use high elevation and far distance with the camera angled slightly downward. The earth is seen as a vast landscape, stretching into the distance, and open for exploration. We replicated this positioning for our own "earth" camera. In a similar vein, the other cameras had their distinct combinations of distance, angle, and height, corresponding to the element that was referenced. In the experimental setup, these cameras were activated whenever a player enters a region dominated by the specific elemental value. For instance, if a player moved from a Metal-dominated region to a Wood-dominated one, the camera position would change from a "metal" camera to a "wood" camera, as indicated in fig. 2.

# 4. TECHNICAL IMPLEMENTATION

Our prototype system was implemented as a modification to the *Unreal Tournament 2004* mod, *Charbitat*. The 3D world used during the testing was divided up into a 5x5 grid of square regions. Each region had a particular element associated with it. As outlined, there were 5 elements in the game (Water, Wood, Fire, Earth, Metal) that served as underlying seed values for tile generation and affected terrain structuring and object selection in the individual tile. As a result the mix of different flora, fauna and terrain differed in each of the tiles depending on the dominating element for this zone. The dominating elemental value of each tile defined the overall appearance of this zone inside the game world and differentiated it from the neighboring zones. A Fire region, for instance, included different objects than a Metal region. In practice, most tiles were mixed environments but the dominant element defined most of the visual cues. Our experimental modifications added to this design the dynamic camera position seen in fig. 2. Whenever the player stepped into a region, the camera gradually transitioned into the perspective for the

**Figure 2:*Charbitat* camera shots and the shots from *Crouching Tiger, Hidden Dragon* that inspired them**

corresponding element. These transitions took about 5 seconds to complete.

For our experiment, we installed the modification of *Charbitat* on two computers and recruited 27 student volunteers (undergraduate and graduate) to play the game. We had each student play *Charbitat* for about 10 minutes. They were told to explore the *Charbitat* world and to press a button whenever they felt they were in a new region, i.e. a zone with a different element. This instruction directed the player to concentrate on the depicted space and to notice differences in the various locations; rather than steering his or her attention to reading the camera's visualization which would have distorted the results. For this experiment we pre-generated a single world of *Charbitat* so that the regions north, east, west, or south of any region had different elements. This way we included enough variety in the tested game world and made sure that all participants experienced the same world structure.

Study participants were randomly assigned to the experimental or the control group. The experimental group played the game with the dynamic camera system at work. The control group played exactly the same game environment but instead of the elemental cameras they used a neutral and continuous following camera whose behavior did not change during gameplay. Thus, the control group had to distinguish between different regions solely based on the different objects and terrain shapes. Both groups of players were informed that there were different zones in the virtual game space and asked to identify the different regions inside the game environment. They had to press a button on the keyboard whenever they thought they had entered a new zone within the overall game space. If, as we hypothesized, the camera affected how players understood the game space, then changes of the element-camera, which were triggered whenever players entered a different elemental zone, should support their spatial identification.

Specifically, players in the experimental group were expected to be faster and more accurate in identifying the spatial zones of the game world than players in the control group. Our system registered players' button presses in addition to tracking and time-stamping the movements of their avatars inside the game word. We were thus able to detect when and where players crossed into another region, and more importantly, we could determine whether this moment coincided with players' realization of the new location; i.e., whether this moment was accompanied by players pressing the designated "change-zone" key. After the game each participant filled out a questionnaire regarding their play experience.

## 5. RESULTS

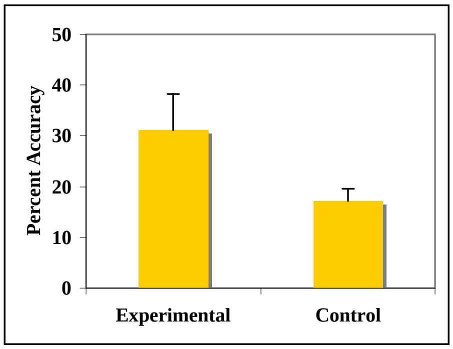
There were 13 study participants in the experimental group who played the game with cameras focusing on region elements, and 14 participants in the control group for whom camera perspective remained neutral throughout the game.

An alpha level of .05 was used for all statistical tests. Significance testing involved Analysis of Variance with game condition (control versus experimental) as independent variable, or when the assumption of normality was violated, the two-sample Kolmogorov-Smirnov (K-S) Test. Control and experimental subjects did not differ significantly in terms of the distance their avatar traveled in space ($M_{\text{Control}}$ = 162,620 units; $M_{\text{Experimental}}$ = 137,660 units; $F(1,25)$ = 2.13; $p$ = .157.); however, significant differences were observed with respect to the number of region crossings by their avatars. Avatars of experimental subjects had on average 36.39 zone crossings, 10 crossings fewer than avatars of control subjects ($M_{\text{Control}}$ = 46.64; $F(1,25)$ = 4.821, $p$ = .038).

### 5.1 Ease of Zone Recognition

To determine whether a dynamic camera system focusing on zone-specific elements faciliates players' orientation in a game world, we compared the zone recognition times for control and experimental participants. Zone recognition time measured the time difference between a player's avatar entering a new game region and the player recognizing that the avatar had crossed into a new zone. If elemental cameras support player orientation, then zone recognition times should be shorter for particiapants in the experimental group than for those in the control group.

**Figure 3: Average time participants in experimental and**



**control group took to indicate recognition of initial zone crossing versus subsequent crossings**

Data screening revealed that all study participants took significantly longer to recognize a new game region at the start of the game than later on (*K-S Z*(27) = 2.112; *p* = .000, 2-tailed)[1]. Specifically, the first time participants indicated they had crossed into a new zone occurred after a considerably longer delay compared to the average time it took them to identify subsequent zone crossings ($M_{\text{Recognition Time Initial}} = 2{,}228$; $M_{\text{Recognition Time Subsequent}} = 125.71$).

These finding suggests that participants' first response times may be conflated, reflecting both familiarization with the game environment and recognition of a change in game region. Subsequent responses, in contrast, seem to reflect participants' "true" recognition of a change in region. Moreover as shown in Fig. 3, while control and experimental subjects did not differ significantly in their initial recognition times (*K-S Z*(27) = .599; *p* = .865, 2-tailed), experimental subjects were considerably faster than control subjects to recognize subsequent zone changes (*F*(1,25) = 6.937, *p* = .014).

## 5.2 Accuracy of Zone Recognition

A participant's accuracy of zone recognition was conceptualized as the proportion of (actual) zone crossings that he or she correctly identified. We hypothesized that experimental subjects whose navigation was accompanied by elemental cameras should be more accurate in identifying zone transitions than control subjects for whom camera perspective remained neutral. Our findings support this hypothesis, albeit with marginal statistical significance (*F*(1,25) = 3.53, *p* = .072).

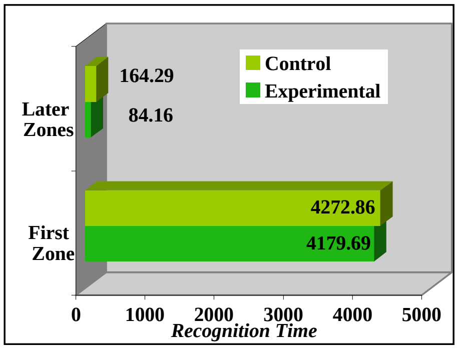


**Figure 4: Average percentage of zones participants in experimental and control group crossed and correctly recognized**

Fig. 4 shows that experimental subjects were almost twice as accurate as control subjects in recognizing a zone. Recall, however, that experimental subjects had fewer region crossings than control subjects; i.e., they had fewer opportunities to miss a crossing than controls.

# 6. DISCUSSION

Providing players with a camera perspective that dynamically adjusts to region-specific elements seems to support their spatial understanding. Once they were accustomed to the game environment, players in the experimental (elemental camera) condition were faster, than players in the control (neutral camera) condition to recognize that their avatar had moved into a new game region and they tended to be more accurate in these judgments. They also seemed less "lost" in the environment as indicated by their smaller number of factual region crossings. Players in the control group, it seems, wandered more (and longer) back and forth between different areas.

On the other hand, recognition accuracy was fairly poor, even for experimental subjects. Most region crossings went unnoticed by experimental and control subjects alike. One reason for this finding may be that the different regions were not as distinct as we had thought. Different regions shared a number of elements making it difficult for subjects, even for those in the experimental condition, to identify region-specific features. If the terrain and fauna in the different regions were not very differentiable to begin with, then players may have made the connection between the camera changes and the region changes.

However, poor recognition accuracy may also indicate that spatial perception aided by camera work is an unconscious process. Players might have been influenced by changes in camera perspective while they were concentrating on a completely different task (spatial identification). In that case, they might not have realized that it functioned as a regional marker but still reacted to it. Furthermore, it is possible that players in the experimental group confused the transiton from one camera position to the next (which lasted several seconds) with the normal perceptual motion of scenery that occurs when an avatar is running. This suggests that the transition between cameras, especially their timing, is more important than we anticipated during game design.

# 7. OUTLOOK

While our findings suggest that camera changes support a faster, more precise, and more efficient spatial reasoning, additional research is needed to address further the impact of region discriminability and camera transition time on players' spatial orientation. Our experiments have shown that cameras affect players' spatial understanding but also point toward issues of timing and visual differentiation that seem to have their impact on the recognition. Movement speed of re-adjusting cameras, the level of difference of the various visual properties between two cameras, and extreme angles might all affect the perception and reasoning of the depicted space.

Another important issue to consider is whether camera changes impact male and female players alike. Our pool of participants was not big enough to allow for a more detailed analysis of this question. However, gender differences have been noted in certain orientation and navigation tasks and it would be interesting to see whether similar differences occur when camera changes affect field of view and viewing angle.

# 8. REFERENCES

AARSETH, ESPEN. Genre Trouble. In N. Wardrip-Fruin & P. Harrigan (eds.), *First person: New media as story, performance and game*. MIT Press, Cambridge, MA, 2004.

ATKINS, BARRY What Are We Really Looking At? The Future-Orientation of Video Game Play, *Games and Culture* 1, no. 2 (2006) 127-40.

---

[1] A one-sample Kolmogorov-Smirnov Test was conduced with initial recognition time and subsequent (average) recognition time as repeated measures.

BLOCK, BRUCE *The Visual Story: Seeing the Structure of Film, Tv, and New Media*. Focal Press, Boston et al., 2001.

*Crouching Tiger, Hidden Dragon*. Dir. Ang Lee. Asia Union Film & Entertainment Ltd. 2000.

CALDERON, C., WORLEY, N. AND NYMAN, K. Spatial Cinematic Mediation in Real-Time Architectural Walkthroughs. ITcon, 11 *Architectural informatics* (special issue). (2006) 343-360.

DRUCKER, STEVEN M. AND ZELTZER, D. CamDroid: a System for Implementing Intelligent Camera Control, *Proceedings of the 1995 Symposium on Interactive 3D Graphics 1995*. ACM, New York. 139-144.

BARES, WILLIAM H. AND LESTER, JAMES C. Intelligent Multi-shot Visualization Interfaces for Dynamic 3D Worlds, *Proceedings of the 1999 International Conference on Intelligent User Interfaces*, 1999. ACM, New York.119-126

HE, LI-WEI, COHEN, MICHAEL F., SALESIN, DAVID H. The Virtual Cinematographer, *Proceedings of the 23rd annual conference on Computer Graphics and Interactive Techniques*, 1996, ACM, New York. 217-224.

KRZYWINSKA, TANYA."Hands-on Horror." in *Screenplay. Cinema/ Videogames/ Interfaces*, ed. by Geoff King and Tanja Krzywinska, Wallflower Press, London, 2002. 206-225.

MEADOWS, MARK S. *Pause & Effect. The Art of Interactive Narrative*. New Riders, Indianapolis, 2002.

NITSCHE, M. Games, Montage, and the First Person Point of View *DiGRA 2005 Conference Proceedings: Changing Views - Worlds of Play*. (Vancouver, CAN June 16-20, 2005) DiGRA,Vancouver. 29-36.

NITSCHE, M. / ASHMORE, C. / HANKINSON, W. / FITZPATRICK, R./ KELLY, J. AND MARGENAU, K. 'Designing Procedural Game Spaces: A Case Study' in: *Proceedings of FuturePlay* 2006 (London, Ontario October 10-12, 2006) (digital proceedings)

STEINER, K. E. AND TOMKINS, J. 2004. Narrative Event Adaptation in Virtual Environments. *Proceedings of the 9th international Conference on Intelligent User Interfaces* (Funchal, Madeira, Portugal, January 13 - 16, 2004). IUI '04. ACM, New York. 46-53.

TOMLINSON, B., BLUMBERG, B., AND NAIN, D. Expressive Autonomous Cinematography for Interactive Virtual Environments. *Proceedings of the Fourth International Conference on Autonomous Agents* (Barcelona, ESP June 2000), ACM, New York. 317-324.

YOUNG, R. M. AND RIEDL, M. 2003. Towards an architecture for intelligent control of narrative in interactive virtual worlds. In Proceedings of the 8th international Conference on intelligent User interfaces (Miami, Florida, USA, January 12 - 15, 2003). IUI '03. ACM, New York, NY, 310-312.