Towards Intelligent Authoring Tools for Machinima Creation

Brian O'Neill
GVU Center
School of Interactive Computing
Georgia Institute of Technology
Atlanta, GA, USA
boneill@cc.gatech.edu

Mark O. Riedl
GVU Center
School of Interactive Computing
Georgia Institute of Technology
Atlanta, GA, USA
riedl@cc.gatech.edu

Michael Nitsche
GVU Center
School of Literature,
Communication, and Culture
Georgia Institute of Technology
Atlanta, GA, USA
michael.nitsche@lcc.gatech.edu

CHI 2009, April 4 – 9, 2009, Boston, MA, USA
ACM 978-1-60558-247-7/09/04.

Abstract
As user-created content increasingly becomes an ever more prominent element of modern game design, tools have been developed to aide in the creative process for several forms of digital media, including machinima. Because creating content that will be valued by the community is a challenging process, tools are needed that will assist novices in both technical realization and optimization of content. We are exploring tools for machinima authoring that use a 3-pronged approach: authoring via metaphor, performance, and automation. Future work involves using AI to provide feedback to machinima authors, suggesting sensible attributes for scenes based on prior input by acting as a surrogate audience.

Keywords
Machinima, Story authoring, User-created content, Computational creativity support

ACM Classification Keywords
I.3.6. Artificial Intelligence: Methodology and Techniques—*Interaction techniques*; J.5 Computer Applications: Arts and Humanities

Introduction

A growing trend in the world is the desire to create media, rather than simply consume. Today, tools exist for creating nearly every media imaginable. However, the mere ability to create does not grant that a particular community will value the new work, or respect its quality. Because of this gap between creating and creating with value, tools are needed to support non-experts in their efforts at producing artistic artifacts. Tools that focus on the process of authoring content, rather than simply its feasibility, are valuable when that process is particularly challenging or costly.

One example of technologically-supported, user-generated content is *machinima*. Machinima refers to the use of video game technology to create animated cinema. Video game environments are used as virtual sound stages and provide resources that are otherwise costly or inaccessible to most filmmakers, such as locations, actors, effects, and a wide array of camera angles. However, though tools exist to make the process of machinima authoring technically feasible, it is still a difficult process to design and produce a cinematic visual narrative.

We seek to encourage machinima creation through a three-pronged approach to the process. We aim to support potential machinima authors who are specifically interested in *conventional cinematic storytelling*, as opposed to those who are interested in demonstrating an in-game accomplishment or experimentation with new media. In conventional cinematic storytelling there are common cinematic patterns called idioms. We give novice authors – those who are not trained, or are not experienced in conventional cinematic arts – the ability to create machinima through a simple, wizard-like interface (as opposed to working in the 3D graphical environment). This mode provides constraints on their decisions in order to help them avoid decisions that may result in artifacts that might be considered visually unappealing in accordance with conventional cinematic storytelling. We also provide expert authors – those with training and/or experience in cinematic arts – with a method of machinima creation by allowing one or more directors the opportunity to manually carry out and record the actions of the film characters. Finally, our system will provide a fully automated method of machinima authoring using a virtual cinematography tool. This automated mode can be used for one scene or for an entire film. With these three approaches, we will provide the capacity for machinima creation to a potentially greater population of users with varying degrees of expertise. Future work includes using artificial intelligence in a second way: to aide human creativity by acting as a surrogate audience, providing authors with feedback on the movie as it is being authored.

Related Work

Machinima creation started within video game environments. Authors typically manipulated the characters using standard control options, and recorded the movements of their avatars. These recordings can be edited together to build a scene, correct for errors, or splice together conversations. Sounds, including dialogue, can be dubbed over.

Some video game developers allow users access to necessary game modifications needed to create the virtual film scene. Thanks to control over all the assets, machinima authors can design their own worlds and

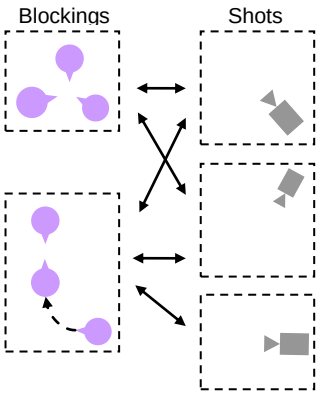

Figure 1. The relationship between, shots and blockings in the layering metaphor. Certain blockings can only be layered with certain stages or shots.

characters as well as control camera placement. While these tools allow for increased degrees of freedom for machinima authors, they do little to encourage conventional cinematic storytelling.

Machinima creation tools exist that take a computer-aided design (CAD) approach to the task. Like video game editing tools, these tools provide the capacity for designing worlds and characters, but also make use of cinematographic terminology and metaphor and focus on issues related to filmmaking, rather than trying to create a film in an environment originally intended for other purposes. However, while some tools have "quick-start" sets and characters available for first-time users, these systems still require authors to possess a degree of cinematic and directorial know-how to be effective in creating professional-looking works.

Work has been done on camera control in virtual worlds using techniques such as visual fly-throughs, finite-state encoding of cinematic idioms, constraint satisfaction, genetic algorithms, autonomous agents, and planning (See [2,5] for more information). With the exception of [2,5], these systems are not specifically designed for cinematic renderings. In particular, Cambot [2] models the complete process of filmmaking as a constraint satisfaction problem, endeavoring to create the film that meets an author's constraints as provided in the scripts. Cambot's algorithms are used for automating shooting of story content when the user desires. Previous work [9] envisioned a system in which authors were responsible for scripting a story with the help of an authoring assistant before sending the story to Cambot to be automatically rendered to machinima. In the work described here, the author edits story and cinematic

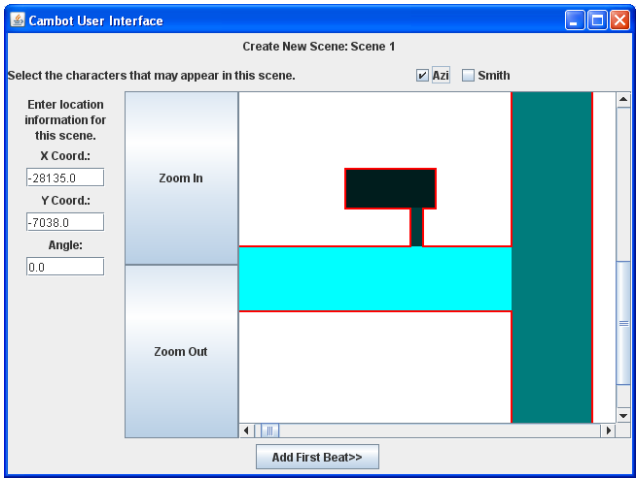

Figure 2. The user interface for our authoring tool. The region to the right provides authors with an overhead view of the movie set to help them choose a location.

elements simultaneously, with varying degrees of assistance.

A Machinima Authoring Tool

The approaches to machinima creation described in the previous section suffer from certain flaws. When creating machinima from within video games, users are frequently constrained by the rules and logic of the game. Additionally, many aspects of filmmaking such as camera placement are difficult. Standalone tools are frequently inaccessible to novice machinima authors, requiring a deep knowledge of the filmmaking process. In this section, we describe work toward a tool for machinima creation that is less constraining than a standard video game environment, but more accessible to novices than standalone tools.

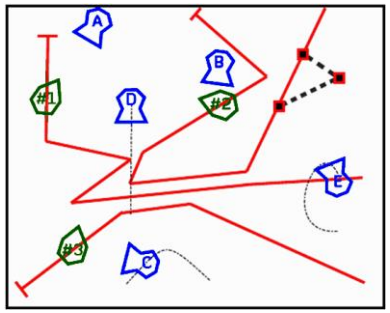

Figure 3. Performance Mode provides an overhead view of the movement of players. Here we see the movements of three players and the placement and movement of five cameras.

Authoring Metaphor

Our authoring tool provides a user interface for machinima creation. In order to make machinima creation more accessible to novices, our system uses a metaphor of layers to represent the structure of movies. There are three layers in our metaphor: movies, scenes, and beats. The movie is the final artifact, and the level at which characters are specified. A scene is a portion of a movie that takes place in a single location using a particular subset of characters. Scenes are made up of beats, the smallest divisible segment of a scene, typically encompassing one line of dialogue or a moment of action [7]. Thus, a movie is a sequence of short beats. Beats are authored by layering independent components: blockings and shots. Blockings are pre-specified configurations of character positions, orientations, and movements. For example, there are common ways of configuring actors on a stage for conversations involving 2, 3, or more people. Shots are pre-specified configurations of camera positions for capturing action. For example there are common (idiomatic) techniques for shooting conversations involving over-the-shoulder and establishing shots, tracking, panning, and zooming.

Throughout this process, a limited set of constraints are placed on the author. Certain shots can only be layered with certain blockings (Figure 1). Some moving shots require a minimum length of time for the overall beat. Additionally, we require that if a beat has two people in it, the blocking must fit exactly two people. Future work includes providing feedback and imposing constraints based on prior decisions, and is discussed later in this paper. We believe that this metaphor of layers makes machinima creation more approachable for novice authors.

Performance Mode

We wish to also support expert-style authoring, or at least to support authoring of scenes that cannot be easily captured with the layering metaphor. Performance mode allows people to interact directly with the game environment – in this case Unreal Tournament – to perform scenes through direct manipulation of avatars. Avatar movements and actions are logged for editing. From this log, we can produce an overhead view of the movements of the players (see Figure 3). From this view, the director can choose camera placements at specific times during the scene. With this information, the scene could be played back, taking one or more views from the perspectives specified by the user. Because of the difficulty of controlling the precise actions of a video game avatar, particularly if trying to simultaneously control the actions of several characters, we see performance mode as a tool for more experienced machinima authors.

Automated Mode

As a final option, we want to provide users with the ability to generate a scene automatically, which can then be edited to better match the intentions of the author using the metaphor mode. Automated scene generation uses Cambot [2].

As a standalone tool, Cambot takes a script as input, and generates a machinima film, choosing scene locations, character blocking, and camera shots from constraints provided by the author. These constraints specify the author's intentions by requiring or recommending certain conditions for the various traits of the movie, scene, or beat. Cambot uses heuristics to analyze the script and constraints against its library of

Figure 4. A series of stills from a movie produced using Cambot.

shots and blockings, finding the best location and sequence of shots that would make the best possible movie.

In our machinima authoring tool, authors input the characters, actions, and dialogue just as they would when creating the scene completely manually. However, when it comes time to select location, shots, and blockings, Cambot will make the decisions, using constraints provided by the user. Constraints can be provided through the authoring tool user interface. After the scene is completed, authors can go back and modify the product of the automation using the layered mode. This ability to switch from authoring mode to automated mode and to edit the final product separates this work from past work (c.f. [9]).

Integration of Modes

Our authoring tool provides machinima authoring capability to users of varying degrees of expertise. Novice authors – those without training or experience in the cinematic arts – are able to quickly produce machinima that conforms to idiomatic cinema conventions via the layering metaphor. Expert authors – those with training and experience in the cinematic arts – can act out their stories with more fine-grained control, which is more akin to how actual films are made. Novices and experts can use the automated mode to produce visuals that can be later edited. The user is allowed to manually select the style of authoring, and in the future will also be able to manually switch between authoring modes on a scene to scene basis. At the time of writing, the manual authoring mode and automated mode are complete, although not completely seamlessly integrated. Performance mode is implemented as a separate module to be integrated soon. The following section describes our future vision for intelligent feedback.

Intelligent Authoring Assistance

In the future, we propose providing users of our authoring tool with recommendations based on their decisions in previous beats and scenes. The system should be able to recognize elements of dramatic scenes and infer the attributes that should be applied to the subsequent beat or scene to match the author's theme. What is the most effective way to use artificial intelligence to support content creation? Lubart [6] enumerates four ways in which computer interfaces can support creativity:

- **Computer as nanny**: The computer provides organizational and classification services and performs routine operations on behalf of the user.
- **Computer as pen-pal**: The computer facilitates brainstorming with functionality that captures and transmits to collaborators the user's thoughts.
- **Computer as coach**: The computer is knowledgeable about the process and can offer suggestions and stimulate creativity
- **Computer as colleague**: The computer forms half of a human-computer team by contributing to the solution.

The computer-as-coach and computer-as-colleague metaphors are most applicable to intelligent systems for supporting human creativity. The computer-as-coach metaphor is used extensively in intelligent tutoring systems [10]. However, it is not our intention to teach a user how to produce machinima. The computer-as-colleague approach introduces automation via an expert system into the creative process. Cambot can be seen in this context as an expert system.

We are investigating a third alternative for using artificial intelligence to support human creativity: computer-as-audience [8]. Rather than serving as an expert on machinima creation, our goal is to act as a surrogate audience for the author. We consider this *computer-as-audience* role to fall between Lubart's coach and colleague positions. A computer-as-audience system could act as an advisor and a member of the target community. Additionally, the feedback provided by such a system could come continuously, sporadically, or only when specifically requested by the user. A computer-as-audience system must be able to provide feedback similar to that which would come from humans in the same role.

One of the significant challenges of modeling audience feedback in this instance is the difficulty of forming a model of a film in-progress. Audiences typically see final products; a computer-as-audience system must infer a model of the story by taking the perspective of a viewer *of an incomplete story*. To be useful, the system must be able to provide feedback on that incomplete artifact. Our investigation of computer-as-audience currently focuses on the development of computational equivalents to (a) cognitive models of narrative [3], (b) cognitive inferencing by readers of narratives [4], and (c) the discursive linkages between visual medium and narrative [1]. Reasoning about narrative from the perspective of the intended audience is hard; we believe that this capability for feedback from the viewer perspective will further strengthen the creative abilities of novice machinima authors. An early version of the cognitive model was implemented with limited inferencing [9] and is now being updated to reflect the full set of capabilities required for a computer-as-audience assistant.

References

[1] Branigan, E. *Narrative Comprehension and Film.* Routledge, 1992.

[2] Elson, D.K. and Riedl, M.O. A Lightweight Intelligent Virtual Cinematography System for Machinima Production. In *Proc. AIIDE 2007*, AAAI Press (2007).

[3] Graesser, A., Lang, K., and Roberts, R. Question Answering in the Context of Stories. *Journal of Experimental Psychology: General 120*, 3 (1991), 254-277.

[4] Graesser, A., Singer, M., and Trabasso, T. Constructing Inferences During Narrative Text Comprehension. *Psychological Review 101*, 3 (1994), 371-395.

[5] Jhala, A.H. and Young, R.M. A Discourse Planning Approach for Cinematic Camera Control for Narratives in Virtual Environments. In *Proc. AAAI 2005*, AAAI Press (2005).

[6] Lubart, T. How Can Computers be Partners in the Creative Process: Classification and Commentary on the Special Issue*. Int. Journal of Human-Computer Studies 63*, 4-5 (2005).

[7] McKee, R. *Story: Substance, Structure, Styles, and the Principles of Screenwriting.* HarperCollins, 1997.

[8] Riedl, M.O. and O'Neill, B. Computer as Audience: A Strategy for Artificial Intelligence Support of Human Creativity. In *Computational Creativity Support Workshop at CHI '09*, ACM Press (2009).

[9] Riedl, M.O., Rowe, J.P., and Elson, D.K. Toward intelligent support of authoring machinima media content: story and visualization. In *Proc. of the 2nd INTETAIN Conference, (*2008).

[10] Shute, V.J. and Psotka, J. Intelligent Tutoring Systems: Past, Present and Future. In D. Jonassen, ed., *Handbook of Research on Educational Communications and Technology.* Simon & Shuster MacMillan, 1996.