

Expressive Autonomous Cinematography for Interactive Virtual Environments

Bill Tomlinson

Synthetic Characters Group
MIT Media Lab

Room E15-320G / 20 Ames St.
Cambridge, MA 02139
+1 617 253 5109

badger@media.mit.edu

Bruce Blumberg

Synthetic Characters Group
MIT Media Lab

Room E15-309 / 20 Ames St.
Cambridge, MA 02139
+1 617 253 7441

bruce@media.mit.edu

Delphine Nain

Synthetic Characters Group
MIT Media Lab

Room E15-320 / 20 Ames St.
Cambridge, MA 02139
+1 617 253 5109

delfin@media.mit.edu

ABSTRACT

We have created an automatic cinematography system for interactive virtual environments. This system controls a virtual camera and lights in a three-dimensional virtual world inhabited by a group of autonomous and user-controlled characters. By dynamically changing the camera and the lights, our system facilitates the interaction of human participants with this world and displays the emotional content of the digital scene.

Building on the tradition of cinema, modern video games, and autonomous behavior systems, we have constructed this cinematography system with an ethologically-inspired structure of sensors, emotions, motivations, and action-selection mechanisms. Our system breaks shots into *elements*, such as which actors the camera should focus on or the angle it should use to watch them. Hierarchically arranged cross-exclusion groups mediate between the various options, arriving at the best shot at each moment in time. Our cinematography system uses the same approach that we use for our virtual actors. This eases the cross-over of information between them, and ultimately leads to a richer and more unified installation.

As digital visualizations grow more complex, cinematography must keep pace with the new breeds of characters and scenarios. A behavior-based autonomous cinematography system is an effective tool in the creation of interesting virtual worlds. Our work takes first steps toward a future of interactive, emotional cinematography.

Keywords

Autonomous cinematography, behavior-based agents.

1. INTRODUCTION

We have implemented an autonomous cinematography system based on the autonomous character design work of the Synthetic Characters Group at the MIT Media Lab [2][14]. Our system controls a virtual camera and several virtual lights in our three-dimensional virtual world. The virtual camera chooses the perspective from which the world is displayed on a flat screen being watched by participants in our installations. By dynamically changing the camera and the lights,



our system makes it easier for participants to interact with our world and displays the emotional content of the digital scene.

Building a cinematography system for an interactive virtual world presents several challenges:

- First, how can a machine generate expressive cinematography for a bunch of unpredictable actors?
- Second, how can cinematography facilitate participants' interactions with the characters in a virtual world?
- Finally, can we solve the first two problems with the same cinematography system? Are interactivity and expressivity mutually exclusive in a virtual environment?

To answer these questions, we have developed a cinematography system from the ground up with interactive, emotional characters in mind. Using the same ethologically-inspired approach that we use to construct our characters, we created the CameraCreature, an autonomous character who lives behind the camera rather than in front of it. With a wide dynamic range of behavior, the CameraCreature controls the placement and attributes of the camera and lights in real time to display each scene in the most advantageous manner for both interactive and dramatic elements. Only such a system can effectively present events in an interactive three-dimensional world full of dynamic and unpredictable digital actors.

The CameraCreature exists in part to make it easier to interact with our installations. It chooses shots and positions lights in ways that make it easier for participants to explore our worlds

and interact with the inhabitants. The CameraCreature works closely with the interface and its gesture-recognition software to ensure that the means of controlling our characters is as intuitive as possible.

The CameraCreature also seeks to display the emotions of the characters in each scene through an assortment of expressive channels. Emotional effects influence the camera angle from which the scene is displayed. The CameraCreature's emotional state affects the motion characteristics of the camera and the transition styles between shots. Finally, the emotions of the CameraCreature influence a variety of parameters of the scene's lighting design. By layering a variety of emotional modifiers onto a basic shooting scheme designed to enable interactivity, the CameraCreature demonstrates that emotion and interactivity are not mutually exclusive.

There is a delicate balance that must be maintained between the level of interactive control in an installation and the means of conveying characters' emotions and relationships. The more direct the control, the more of a servant the cinematography must be to that control. Intentional control, by which a participant provides high-level input to a character and allows the character's autonomy to address lower-level action-selection, permits great flexibility of shot choice and therefore greater emotive possibilities.

Several cinematography systems built with the architecture described in this article have been shown to the public. Our first piece, "Swamped!", an interactive installation at SIGGRAPH '98, allows a participant to control a virtual chicken by means of a stuffed animal interface. [13] In "Swamped", the chicken seeks to torment and elude an autonomous raccoon. "(void*): A Cast of Characters", which appeared at SIGGRAPH 99, is a Charlie Chaplin-esque piece in which participants can make characters dance by manipulating two dinner rolls with forks stuck in them. Both of these projects and several smaller ones have been shown to visitors to the MIT Media Lab over the past two years. We have used these frequent opportunities for feedback as a valuable resource in the iteration and revision of the methodology described in this paper.

In this following sections, we explore the main theoretical underpinnings of interactive and emotional autonomous cinematography in several sections, provide a background of related works, describe and evaluate the cinematography system, and offer some directions for future work. Accompanying this paper is a video which depicts the "(void*)" installation.

As long as images appear on a screen, someone or something will need to choose which images to put there. With this cinematography system, we seek to create a system that is a hybrid of someone and something – an autonomous character controlling a set of digital tools that arrange the virtual camera and lights in a scene. By having an autonomous cinematographer as complex as our other characters, we hope to show off our current installations to participants, and also to make them think about the role of the autonomous cinematographer beyond the current limits imposed by time and technology.

2. RELATED WORK

In this section, I discuss the main disciplines that have inspired and influenced this work.

2.1 Behavior Systems

In constructing the behavior system for the CameraCreature, we built on the work of Bruce Blumberg [2]. His autonomous behavior systems provide an action selection mechanism inspired by animal behavior. His work seeks to create characters who behave intelligently and are capable of displaying emotions. Christopher Kline has furthered Blumberg's work, creating the underlying behavior structure that we have used in this cinematography system. [14] In this work, characters combine sensory information, motivations and emotions, all of which influence a hierarchical organization of cross-exclusion groups, that determine which actions the creature will take.

The Improv system [18] also offers inspiration in building interactive characters. By applying noise functions to the actions of their characters, the Improv system generates natural, organic-looking movement. Since this system was created for use with scripted scenarios, it is less useful in our worlds, where stories emerge from the unscripted interactions of our characters.

Maes [15] and Brooks [3] have also provided relevant resources for developing robust autonomous agents.

2.2 Autonomous Cinematography Research

The Virtual Cinematographer [12] uses the concept of the idiom, a sub-unit of cinematographic expertise, as a means of capturing the essence of a scene. The system developed a means of encoding techniques for conveying certain scenarios effectively. It is wonderful to see the example of film cinematography held up as an example of how to do autonomous cinematography. By creating an assortment of fairly rigid structures to shoot different kinds of scenes, the Virtual Cinematographer is limiting itself in two ways: first, it is only able to create effective shots for scenarios that it is familiar with, and second, each transition between two idioms will break the continuity of the scene. The *Virtual Cinematographer* also fails to address the topics of lighting design and interactivity.

Several others have also explored digital camera control. Steven Drucker [8] developed a system that helps the user perform visual tasks, such as exploring a virtual world. Using an assortment of camera primitives, he created a framework that gives the user higher-level control of a virtual camera, addressing the problem of shot selection by considering it to be a series of small constrained optimization problems. Tinsley Galyean [11] explored the area of interactivity as it is influenced by story. Of particular interest to our work, he examined the effect of the plot on the camera – how story line changes the presentation of the scene. Bares and Lester [1] addressed the problem of simultaneously taking actions in a virtual environment and controlling the camera. Their system creates models of the user's cinematographic preferences to create camera shots and sequences that show the user what they would have chosen themselves. Others have also considered ways of presenting interactive media. [5] [6]

2.3 FILM/TELEVISION

The silver screen gave cinematography its birth. For the last century, individuals and studios have made all kinds of films, from back-lot epics to back-street independents. The heritage of film provides much of the cultural and technical background

informing our research. Most movies adhere to some basic conventions about shot choice, sequence assembly, scene construction and lighting. These visual conventions help develop the themes that the director is emphasizing in each section of the film. Awareness of these means of directing (and misdirecting) an audience's attention can help the system reveal important elements of our virtual environments. Examples of these conventions include: looking over the shoulder of a character to see what it is seeing, placing a moving character in the frame such that it is moving toward the center of screen, and choosing a shot of a character's face to show that character's emotion. The huge difference between films and our medium is interactivity. In films, there is none. In an interactive environment, the experience is different every time.

When deciding where to put the camera, cinematographers consider the movements, relationships and emotions of the characters; the arrangement of the set; the ambient light and opportunities for adding or subtracting light. Cinematographers have a toolkit for their trade – camera, film, lenses, lights, gels. [16] Similarly, the Synthetic Characters Group has constructed a set of tools appropriate to the interactive, digital kind of cinematography. [20]

Documentary film making is much like narrative feature film making in technology, but quite different in technique. It is closer to a real-time virtual environment, in that the cinematographer is often trying to capture events as they happen “for real”, rather than having the luxury of a fully orchestrated film set. Although they often document real-time events, documentaries eventually have the luxury of the cutting room when crafting a final product; our system only gets one chance.

Televised live sporting events offer another source of inspiration. While sports do occur live, they are not completely random events occurring live. There is an element of constrained unpredictability to them. A running back is going to run toward the end zone, but he's not going to keep going out of the stadium and down the street. Our installations are similar to this, in that our characters may walk or swim or dance, but they're not going to climb a tree unless we've done an animation for it.

When shooting a soap opera, there is usually a three-camera set-up, with a director choosing which of the cameras to send to the recorder. There is a strong emphasis on emotion that pervades soap operas. However, they are scripted and rehearsed, even if they are ultimately shot in real-time. Camera moves, too, may be scripted, just as the dialogue is. And there is always the recourse to a re-shoot if an actor flubs a line, since it won't be seen by an audience until after the final product is complete.

2.4 Video Games

Video games are interactive in real-time. There have been great advances in playability and interactivity since the first days of *Pong* and *Pac-Man*. Modern games feature several basic paradigms for interactive camera systems.

The games *Zelda* and *SuperMario 64* allow a player to possess a character and explore a variety of scenarios in a virtual world. The player's character can navigate, collect things, look around, and perform a variety of other actions. Both games have exceedingly competent camera systems that choose shots to show off the actions of your character. Navigation is made amazingly

easy; travelling through the world is intuitive after only a few moments with the controls. The characters in these games are fairly simple, so it makes sense that neither one tries to convey the character's emotional state through the cinematic arts.

Tomb Raider lets the player control Lara Croft, a buxom, pistol-packing, female Indiana Jones. The camera follows her with a high degree of intelligence, making navigation passably easy. When asked to draw her guns, Ms. Croft automatically aims at whatever seems appropriate. This makes interacting with other characters in the world quite easy – she aims, you shoot. Once again, though, there is a pronounced lack of any emotional commitment required on the part of the player, except perhaps for the distaste registered at being required to shoot at tigers.

Ultima Online offers the top-down camera style featured by many adventure games. As the player's party of explorers wanders around the world, the camera watches them from high above. This makes navigation exceedingly easy, but creates a great feeling of detachment between the player and her characters.

Grim Fandango uses fixed camera angles, specially crafted for each scene. This creates a very cinematic feel to the game, but is rather rigid. With this cinematic style, there is no room for improvisation or interactions outside of those for which camera angles have been crafted. In that capacity, *Grim Fandango's* cinematography is fairly inflexible.

In *Doom*, the player is a gun-toting soldier in a multi-level dungeon. *Doom* is a “first-person shooter” type game; aside from a few statistic about your status, the entire screen shows a straight ahead view of what you are seeing. This first-person view allows the player to have complete control of the camera. Another game, *Thief*, is similar in format to *Doom*, but the story line changes the feel of the interaction strikingly. While *Doom* encouraged a guns-blaring assault, *Thief* forces the player to sneak around, since any frontal assault inevitably leads to your death. Although it creates only one emotion, fear, it is still a big step towards a full emotional repertoire in video games.

3. IMPLEMENTATION

In this section, we propose an implementation for a behavior-based autonomous cinematographer. In developing this system, we considered a variety of paradigms. We decided early on to implement a reactive system rather than one with planning. We felt that a planning system in which scripted camera motion attempted to cover our ever-changing scenes would probably be too brittle. We also wanted emotion and motivation to be central to the cinematography system. These are difficult to integrate into a planning approach. Finally, we wanted to employ the same architecture in the CameraCreature that we use in our characters, to make it possible for us to leverage all the work that has been done in our group for developing characters.

Our system takes the *shot* as the most important level of understanding cinematography. Shots are composed of *elements*, such as an actor or a motion style. Elements, in turn, are modified by a variety of settable *parameters*, such as the spring constant that controls the motion style or the height of the lead actor. The behavior system consists of four main elements – sensors, emotions, motivations, and actions. These work together much as they might in an animal – sensory information

is combined with motivations, modulated by emotions and fed into action-selection mechanisms.

3.1 Sensors

The CameraCreature is able to extract information about the states of the other creatures in the world. It chooses its actions by combining this data with its own internal state. The CameraCreature uses sensors to make the connections with the other creatures in the world, through which it can find out what emotions they are feeling, determine what actions they are taking, and gather some knowledge of their motivational state. The cinematography system is somewhat privileged in this regard – in order to present characters effectively on screen, it needs access to information about the internal state of the actors.

In order for there to be an exchange of information between characters and camera, some conventions must be imposed on that information. This is both a plus and a minus. Abstraction of relevant information allows the camera to see different creatures as ‘Actors’ from whom it can extract position, orientation, size, motivational data and emotional state. However, forcing a wide assortment of characters into the same structure also opens up some problems. For example, assessing the ‘Height’ of a virtual snake is a real challenge.

3.1.1 RequestShot

Actors can request shots by setting a feature that the camera then extracts with its sensors. For example an actor could set a feature `READY_FOR_MY_CLOSE_UP_MR_DEMILLE`. The value of this feature is weighted as an input to the actor element for that character, and to the angle `CLOSE_UP`, in order to help them win over competing elements. Allowing direct communication between actors and camera has tightened the link between characters’ actions and moods and their expression on the screen.

3.2 Emotions

Like our other characters, the CameraCreature has a simple model of emotion that affects which shots are chosen and sets parameters within those shots. By influencing elements such as motion style and camera angle, each emotional state causes a corresponding visual style. Each emotion is described by a function that takes into account the CameraCreature’s default temperament, the emotion’s rate of change over time (moodiness), and internal and external factors which affect the emotion’s level.

Since the emotions of all the characters in the scene factor prominently into the cinematography system’s emotion system, the shooting style reflects the current feelings of the characters. The amount that a character influences the mood of the CameraCreature is weighted to reflect how much that character has been on screen recently, how important that character is, or how powerfully the character is feeling that emotion.

Emotions have an impact throughout the cinematography system. For example, a happy CameraCreature might cut more frequently, spend more time in close-up shots, move with a bouncy, swooping motion, and brightly illuminate the scene. We’ve tried to allow emotion to percolate through the entire system.

Our emotion system has a flat arrangement with happy, sad, angry, surprised, fearful and disgusted as the six elements. [10] This structure is useful in that people are used to thinking of emotions in these terms, and it is therefore fairly easy to create effects at this level. We also considered another emotional structure that defines emotion-space into stance, valence and arousal [19], but found ourselves frequently trying to ‘map back’ to the previous flat arrangement. Finally, we returned to the original system which operates in terms that is more accessible for both designers and participants.

3.3 Motivations

Whereas emotions have broad-scale effects on the choosing of shots, motivations have a stronger, localized effect. Motivations have a similar formula to emotions, with base levels, inputs, gains and rates of change. However, they tend to be expressed in only one section of the action-selection mechanism, rather than across the board like an emotion.

This is the section where higher-level organization of shots occurs. While the action-selection section below is broken down by functional elements – actor, angle, motion style, etc. – the motivation section is arranged by conceptual effects – `DesireToEstablish`, `DesireForCloseUp`, `DesireForActionShot`, `DesireForTwoShot`.

Each motivation has a function that determine its value. This value is then used as the input for various shot elements in the action selection mechanism below.

- `DesireForTwoShot` is the default motivational state, so it has a fairly high constant value. This causes the cinematography system to make relationships between characters the primary object of its interest.
- `DesireToEstablish` starts even higher, so that the first shot of an interaction is a wide shot that lets participants get their bearings, but it is self-inhibiting, so that its value soon drops off sharply. It has a constant growth, though, so it will slowly rise again until it is satiated by being allowed to express itself.
- `DesireForCloseUp` takes inputs from requests made by the actors and is more likely when the camera system has a high value for its primary emotion.
- `DesireForActionShot` also takes inputs from requests made by the actors. These requests occur when a character does something they think deserves a shot to show it.

There are other lower-level motivations, that affect the action selection mechanism in various ways. For example, the cinematography system would look broken if it cut again in less than a second or so of another cut. We use a motivation to solve this problem. The value of `MaintainCurrentActorAndAngle` becomes very high when either the actor or the angle has changed recently. Since in the action selection mechanism a behavior must have a value of at least *twice* the value of the currently-winning behavior in order to take over, adding the very large constant value of `MaintainCurrentActorAndAngle` to *all* of the actors and angles makes it much less likely that another actor or angle will take over. This is a robust way of preventing cuts from happening too rapidly without hard-coding an actual minimum length.

3.4 Action Selection

The action selection mechanism is the means by which mutually exclusive behaviors can be organized into groups with cross-exclusion and mutual inhibition semantics and forced to ‘compete’ on the basis of their output values. A behavior is a routine that sends a message (i.e. “Earl is the lead actor.”) to the code that calculates where in world coordinates the camera should be positioned and oriented.[14] By arranging these in hierarchical groups, complex behavioral patterns can occur by recombining a variety of simple components.

At each clock tick, all behaviors in a mutual-exclusion group calculate their values to determine which of them will become active. It is a value based process, with each behavior’s value being a combination of motivations and emotions. The process is weighted in favor of currently active behaviors, so that two behaviors with closely matched values will not switch rapidly back and forth (behavioral aliasing).

The structure of sensors, motivations, emotions and action-selection described above is the functional means by which any of our autonomous characters choose their behaviors.

3.5 Camera Shot Elements

We break cinematography down into two main areas – camera and lighting. In this section, we discuss the elements that combine to make up each shot. The CameraCreature must decide where to put the camera, and which direction it should be facing, at every clock tick. There are two main parts of the camera’s decision process – shot choice and the motion style selection. Shot choice involves choosing an actor or actors to look at and an angle from which to look at them. The motion style section chooses the characteristic feel with which the camera moves through the world.

3.5.1 Actor

The most important decision that the camera makes is which actors to watch. Often, the camera watches the character who is controlled by the participant. People want to see their character, so the CameraCreature tends to skew its actor-picking section toward the participant-controlled character. However, it is necessary that the camera be able to cut to another character if that character is performing a really interesting action (such as when the raccoon eats the chicken’s eggs in the *Swamped!* environment).

The camera’s job is to reinforce the relationships that are created between the actors on screen. There are three kinds of relationships that we have found to be relevant.

- Character to Character: A two-shot between the two characters can help establish this. In a two-shot, the camera calculates the axis between the two most interesting actors, and chooses its angle as a differential from this axis, looking at a point somewhere along the line between the two actors. This assures that both actors are in the field of view.
- Character to World: For this, we enabled the camera to select pieces of the set as actors, and create two-shots between the actor and the set piece.

- Character to Participant: To show this, we use the single character shot, where the leading actor is the only relevant element in positioning the shot.

Through the rest of this section, we will refer to the actor (or pair of actors, or actor-and-set-piece) as the “target” of the camera. When the graphics system is called every tick, it is given a camera position and a target position, and from this it determines how to render the scene.

3.5.2 Angle

In addition to deciding which actors it is interested in, the camera also chooses an angle with respect to its target. This angle is calculated in the coordinate system of the lead actor (or the axis defined by two actors). There are a variety of angle types that we have developed, each of which serves a different main purpose.

Wide, establishing angles are useful for orienting a participant in the virtual world. Navigation angles position the camera to track the participant’s character, so that the participant can see where the character is going. Close-ups are very expressive and can be framed in a variety of ways to show off a specific emotion. For example, a low angle, looking up at the actor, makes it appear threatening, while a higher angle makes it appear fearful.

Each of these angles has a variety of parameters that let it adjust camera and target positions: the amount it should rotate around the target; the distance it should move away from the actor (a proportion of that character’s height); how far ahead of the character it should look; how high the camera should be (this is often emotionally determined, but again, it’s a percentage of the actor’s height); whether the camera should stay fixed for the duration of the shot or track with the character; and, if not tracking the target, how much the camera should drift from its initial position.

3.5.3 Motion

Since camera angle selection is often focused on enhancing interactivity, the motion characteristics of the camera are the main conduit of emotional expressivity. The section on motion decides on the parameters of a dynamical system of springs and dampers, that affects how the camera moves through the virtual world. By changing the settings on the spring dynamics system, the camera may be made to move with an expressive range of emotional effects. For example, an angry camera might move very abruptly, a sad camera might move in slow arcs, and a happy camera might have a bouncy, slightly oscillating feel to it.

3.5.4 Transition

Whenever the camera changes its current choice of actors or angle, a transition to a new shot occurs. This transition can have a variety of different styles. *Cut* causes the camera to go immediately to its new position. This is the transition most frequently used in movies. *Whip-pan* causes the camera to swoop through space rapidly to its new position, giving participants a strong feeling of motion. This helps keep participants oriented in the world.

3.5.5 Occlusion Detection

Once all the shot elements have been determined, the system casts a ray from camera to target, to see if the shot is occluded by anything. It is possible for other characters or set pieces to be in

the way, and thereby ruin any shot. The cinematography system checks to make sure that the target is the first thing that the ray encounters; this ensures that the camera's line of sight is clear. If the path is occluded, the camera must reposition itself. For example, the camera in "Swamped!" goes straight up until its forward line of sight is clear. Having the camera calculate occlusion as a two dimensional problem (it continues to look straight ahead, rather than down at the target, as it moves up) helps avoid the possibility that, if the lead actor happened to walk under an overhang, the camera would go up forever in an attempt to get that actor in an unoccluded shot.

3.6 Lighting Elements

Camera work is the most obvious element of cinematography. Lighting design is important in more subtle ways. Simply putting some lights on a scene will make events visible to the camera; carefully arranging those lights can have a myriad of emotional effects and provide subconscious cues to participants. [17]

We have split the lighting design of each scene into two parts – global lights and personal lights. The global lights are fixed in position in the world, while the personal lights travel with the characters and provide them with specially-tailored lighting.

There is interplay between the global lights, the personal lights, and the camera. When the camera moves in for a close-up emotion shot, that character's personal light increases its intensity, and the global lights dim a bit. This causes that character's lighting design to dominate the illumination of the scene. This provides for more extreme emotional effects when they are appropriate, and less extreme effects when normal illumination would work better (e.g. for navigation).

3.6.1 Global

The global lights have a default lighting scheme, with several lights providing the key sources of illumination. The global lighting scheme allows the world to maintain basic continuity, and helps orient the user by showing them where they are with respect to the lights. The global lights affect the overall coloration and illumination of the world.

Each global light has three parameters that can be controlled independently – color, intensity and position. Color sets the hue of that light, and also the baseline intensity. Intensity is used to modulate that light, based on camera position. It varies from 0 to 1, with full illumination coming when the camera is far from the lead actor. Global lights are positioned with the dominant light sources in the world – the sun, a camp-fire, etc.

3.6.2 Personal

The personal lights allow each character to be specially illuminated beyond the effect of the global lights. Each character has a light that changes its color, intensity and position based on the emotional state of that character. The positions of these lights are determined in their characters' coordinate systems, so that they appear to travel with their characters.

Whereas global lights are set in the behavior system of the camera, personal lights are distributed throughout the characters. This has two main benefits: it eases the communication between a character's emotions and lighting and it makes a character's

lighting design transportable (it's already a part of the character). Transportable personal lighting design is quite valuable, since it allows characters to keep their personal light if they are moved to a different scenario.

The CameraCreature also has control over the color of the sky. Sky color is a special kind of personal light – the personal light of the CameraCreature. This light lives in the CameraCreature's behavior system, and causes the sky color to change to reflect the CameraCreature's emotional state (and therefore some blending of all the actors' emotional states). Changing the sky color sounds quite drastic, but it occurs in cartoons all the time. Substantial changes in sky color are not shocking in a virtual world where the laws of physics do not always apply.

3.7 Camera and Sound

We have also taken the first steps to enable our cinematography system to cooperate with the sound design and composition. Currently, the score and sound effects change based on camera position. In addition, giving the camera the ability to make sounds made it seem more like an active participant in the scene. Simple whooshes make the camera seem much more alive, and made it less of an invisible observer. We soon hope to have the camera able to cut on the beat of the music.

4. ASSESSMENT

"It was great! I didn't notice it!"

-Steven Drucker, at SIGGRAPH 99

Having made several autonomous cinematography systems, we have tried to find techniques for effectively judging cinematography. We considered examining task-completion ease and other metrics for testing our system. However, cinematography remains a subjective discipline, in which success and failure rely upon how audiences feel about the scene. Ultimately, we found that listening to the subjective comments of our participants was the most effective way of judging our system.

We have shown our cinematography systems to more than a thousand people by now, during SIGGRAPH 98 and 99 and at demonstrations to our corporate sponsors at the Media Lab. Developing our autonomous cinematography system has been an extremely iterative process, with many demonstrations leading to revisions to the cinematography system (from small tweaks to complete re-writes). Rather than taking the form of an experiment, the life of the CameraCreature has been an evolution, with audience feedback defining the fitness function.

In deciding what revisions were most pressing, we watched for certain elements in people's interactions with our system. Seeing whether participants are able to navigate through the world points to cinematography and the control mechanism – both must be working well together for specific interactions like intentional steering to occur. It is also possible to see when an emotional impact has been achieved. For example, when a character is sad and the camera helps show this, people often laugh at the over-the-top performance being shown.

Most people appear to enjoy our interactions, which is a first level of accomplishment. For people to have fun, all elements of our system need to be sufficiently functional that they do not

simply annoy the participants. We found that people were able to navigate around our virtual space, and generally cause the on-screen chicken to do what they wanted. A few people asked, “Where did my chicken go?” if the camera cut away to show the raccoon eating her eggs, or doing something else that the camera deemed interesting. However, the camera almost invariably returned to a shot that satisfied them before the words were fully out of their mouths.

We have found that cinematography systems should be transparent. If the average person is paying attention to the cinematography rather than to the actors and the interaction, the system has failed. A successful cinematography system will never be a superstar. As Steven Drucker commented at SIGGRAPH 99, when we asked him what he thought of the cinematography on “(void*)”, he replied “It was great! I didn’t notice it!”

We found a quite striking distinction between *our* need for cinematography during the development process versus the final product we hoped to deliver. Character-builders prefer to have a button interface to the camera and lights so that they can see how the characters look from different angles and under different lighting conditions. People interacting with our system, though, preferred not to think about cinematography, focusing instead on the characters and interactions.

Is there one cinematography system that works equally well for everyone? Some people love to see the other things that are going on around our virtual worlds, and are less interested in having to specifically navigate to see it all. Others want the control that absolute ease-of-navigation brings. Some day, self-customizing autonomous cinematography systems will tailor themselves to the tastes of their audience. For now, a system that can blend interactive control with automated expressivity is a good answer for broadly applicable interactive cinematography.

5. FUTURE WORK

5.1 STORY

Creating adaptive stories is one of the great challenges facing the entertainment media. [7] We find a story to be a combination of a setting and a character or characters who undergo an emotional arc. By allowing participants to influence the arcs of our characters, we make those people integral to the stories that emerge. An autonomous cinematography system that is aware of the state of the characters can highlight their interactions by means of framing, pacing, montage and lighting.

5.2 Further Lighting

Although I have designed a simplistic lighting segment of this cinematography system, the field of interactive lighting design is sorely under-studied. This is, I imagine, primarily due to hardware limitations that prevented elaborate, dynamic lighting schemes in past interactive worlds. However, there now exists the hardware and software to support interactive lighting design; it is time for “camera system” to stop being synonymous with “cinematography system”. Lighting can be a powerful tool for developing characters and stories. [4] Until we have control over elements like shadow and depth-of-field, we will continue to envy our traditional filmmaking comrades.

5.3 Multi-User Interaction

Another realm that will become important as our installations evolve will be multi-user interactions. This will have grave impact on cinematography design, as interaction no longer means satisfying one person, but instead means having two or more primary participants. This can be worked around in many cases, by means of tactful interaction design. (A dragon with two heads, each controlled by a different person, would not be much more difficult to shoot than a single character.) But a world featuring two independent characters controlled by two participants both watching the same screen opens up a myriad of issues. What should the camera do if the two participants start running in opposite directions? It won’t be long before that camera is either so far away that it’s impossible to see the characters or else is cutting back and forth between the two characters, annoying both participants by having each person’s character on screen only half the time. Multi-user interactions will be an interesting challenge for interactive cinematography.

5.4 Montage

Including some techniques for influencing shot choice based on montage theory would increase the cinematic potential of this system. [9] There is already a part of the necessary underpinnings for this, built into the behavior-based structure from which the CameraCreature is made. Since every creature has an object of interest that is the main focus of its current action, it would be trivial to pass that pronome to the cinematography system, and have the camera prefer to switch to a shot of that object at the next cut.

5.5 Adaptable Screen Size

There is often a disparity between the format that is used to develop much of an interactive installation, and the ultimate display medium. If the CameraCreature had a way of detecting the technology that is being used to display its cinematography, it could change shooting style to match screen size. A tiny screen might favor emotion shots, while a wide screen would allow it to be more comfortable with establishing shots. This would be another step toward having a smart and interesting cinematography system.

5.6 Cheating

A final, somewhat longer-term project would be to allow the camera to cheat reality sometimes. Since nothing in a digital system is absolutely fixed, the camera could be given the power to halt a certain course of events if it just isn’t going to be able to show it yet, and resume those events once it will be able to cover them. The camera might also be able to move characters and set pieces slightly to get better-framed shots. Allowing the camera to alter “reality” in our virtual worlds might cause lots of problems, but it would be exciting to work with a cinematography system that was more than just a passive observer.

5.7 Learning

In a longer time-scale, it would be invaluable to have a camera that could learn. It could learn which participants it has worked with before, what shots work well, how different actors tend to behave, and a whole variety of other relevant information. A cinematography system that had a model of the participant, of all

the actors, of the music system, and even of itself could be quite powerful indeed.

6. CONCLUSIONS

In the past two years, we have implemented several major cinematography systems with the architecture described in this paper. During the course of these implementations, we have learned a few things about the needs of an autonomous cinematography system, a few more things about the process of creating them, and perhaps most importantly, a lot of things *not* to do when creating a system of this kind. In addition, we've found that a good cinematography system makes the rest of your interactive installation look a lot better, too. An actor is only as good as her cinematographer. We've tried to implement a CameraCreature that fades into the background to let our actors shine.

Cinematography helps establish relationships between characters. Now that the characters in interactive environments are starting to have enough personality to build simple relationships, it is exciting to be working in a field that is helping to highlight those relationships. While building this behavior-based autonomous cinematography system, we've found that a system that has relationships of its own is more capable of expressing the relationships between other virtual characters.

A cinematography system that is able to adapt to changing circumstances can cover a wider range of possible interactions. That is one of the challenges in dealing with unpredictable actors. Their behavior is constantly changing and making it very difficult for a camera to keep up. A planning system might be unable to cover all of the emergent phenomena that arise when dynamic characters interact. A robust cinematography system that is running around with the other characters can show them off more effectively. A system that can balance the actions occurring on screen with the emotions that the characters are feeling will create a *complete* experience.

In addition to showing off the personalities of our characters, the cinematography system needs to make the participant feel comfortable. To do this, participants need to be shown enough of the world not to lose their bearings, but have enough personal contact with each lead actor to get to know them. The participants need to be able to have their characters take the kinds of actions that they expect in each world, be it steering or waving or jumping. Therefore, the cinematography system needs to work closely with the interface connecting the participant to the interaction. In fact, the cinematography system *is* an interface, just as much a keyboard or a joystick. It's the only visual output device for most interactive installations.

A behavior-based autonomous cinematography system is an effective tool in the creation of interesting virtual worlds. Our work takes first steps toward a future of interactive, emotional cinematography.

7. ACKNOWLEDGMENTS

This implementation relies on two main bodies of work, developed by the Synthetic Characters Group over the last two years. The first is the behavior system architecture originated by

Bruce Blumberg [2] and continued by Christopher Kline, Song-Yee Yoon, and the rest of the group. [14] The second is the graphics system designed by Kenneth Russell, Michael Patrick Johnson, Bruce Blumberg and the rest of the Synthetic Characters Group. [20] Without these two substantial bodies of work, the research described in this paper would not have been possible.

In addition, all of the installations described herein have been very collaborative efforts, so many thanks to everyone who worked on them: Michael Patrick Johnson, Michal Hlavac, Chris Kline, Ken Russell, Andrew Wilson, Teresa Marrin, Song-Yee Yoon, Marc Downie, Ari Benbasat, Dan Stiehl, Matt Berlin, and Wally Holland.

8. REFERENCES

- [1] Bares, W. H. and J. C. Lester. Realtime Generation of Customized 3D Animated Explanations for Knowledge-Based Learning Environments. in Proceedings of the 14th AAAI National Conference on Artificial Intelligence (1997).
- [2] Blumberg, B. Old Tricks, New Dogs: Ethology and Interactive Creatures. Ph.D. Dissertation. MIT Media Laboratory (1996).
- [3] Brooks, R. Challenges for Complete Creature Architectures. In From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behavior (1990).
- [4] Callahan, S. Storytelling through Lighting. in SIGGRAPH 96 Course Notes, Pixel Cinematography (1996).
- [5] Christianson, D. B., S. E. Anderson, L. He, D. S. Weld, M. F. Cohen, D. H. Salesin. Declarative camera control for automatic cinematography in Proceedings of AAAI '96 (Portland, OR, 1996), 148-155.
- [6] Davenport, G., T. Aguiere Smith, N. Pincever. Cinematic Primitives for Multimedia. IEEE Computer Graphics and Animation special issue on multimedia, (July 1991), 67-74.
- [7] Davenport, G.. Seeking Dynamic, Adaptive Story Environments. Visions and Views. IEEE Multimedia, 1(3) (Fall 1994) 9-13.
- [8] Drucker, S. M.. Intelligent Camera Control for Graphical Environments. Ph.D. Dissertation, MIT Media Laboratory (1994).
- [9] Eisenstein, S. Film Form and Film Sense, edited and translated by Jay Leyda. Meridian Books. New York, 1960.
- [10] Ekman, P. Emotion in the human face. Cambridge University Press, Cambridge, UK, 1982.
- [11] Galyean, T. Narrative Guidance of Interactivity. Ph.D. Dissertation, MIT Media Laboratory (1995).
- [12] He, L., M. F. Cohen, D. H. Salesin. The virtual cinematographer: a paradigm for automatic real-time camera control and directing in Proceedings of SIGGRAPH 96 (August 1996), Computer Graphics Proceedings, Annual Conference Series, 217-224,.
- [13] Johnson, M. P., A. Wilson, C. Kline, B. Blumberg and A. Bobick. Sympathetic Interfaces: Using a Plush Toy to Direct

- Synthetic Characters in Proceedings of CHI '99 (1999). To appear.
- [14] Kline, C. and B. Blumberg. The Art and Science of Synthetic Character Design in Proceedings of the AISB 1999 Symposium on AI and Creativity in Entertainment and Visual Art, (Edinburgh, Scotland, 1999).
- [15] Maes, P. How to Do The Right Thing. Connection Science Journal 1(3) (1989).
- [16] Malkiewicz, K. Cinematography. Simon & Schuster, New York, 1989.
- [17] Malkiewicz, K. Film Lighting. Simon & Schuster, New York, 1986.
- [18] Perlin, K. and A. Goldberg. Improv: A System for Scripting Interactive Actors in Virtual Worlds. Computer Graphics, 29(3) (1996).
- [19] Russell, J. A circumplex model of affect. Journal of Personality and Social Psychology 39(6) (1980) 1161-1178
- [20] Russell, K and Blumberg, B. Behavior-Friendly Graphics. Computer Graphics 1999.