Interactivity and Emotion through Cinematography

by

William Michael Tomlinson, Jr.
M.F.A. Experimental Animation, California Institute of the Arts, 1996
A.B. Biology, Harvard College, 1994

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of
Master of Science
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 1999

Author____

William Michael Tomlinson, Jr.
Program in Media Arts and Sciences
May 7, 1999

Certified by____

Bruce M. Blumberg
Assistant Professor of Media Arts and Sciences
Asahi Broadcasting Corporation Career Development Professor
of Media Arts and Sciences

Accepted by____

Stephen A. Benton
Chairman
Department Committee on Graduate Students
Program in Media Arts and Sciences

Interactivity and Emotion through Cinematography

by

William Michael Tomlinson, Jr.

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on May 7, 1999
in partial fulfillment of the requirements for the degree of
Master of Science

ABSTRACT

I have created an automatic cinematography system for an interactive virtual environment. This system controls a virtual camera and several virtual lights in a three-dimensional "world" inhabited by a group of autonomous and user-controlled characters. The virtual camera chooses the perspective from which the world is displayed on a flat screen. The lights control how the three-dimensional digital objects in the world are illuminated. By dynamically changing the camera and the lights, my system facilitates the interaction of humans 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, I have constructed this cinematography system with the same approach that the Synthetic Characters Group uses when developing our virtual characters – an ethologically-inspired structure of sensors, emotions, motivations, and action-selection mechanisms. Using the same approach for all the different elements of our virtual worlds eases the cross-over of information between them, and ultimately leads to a richer and more unified installation.

By helping people interact with our characters, and by showing the emotions of virtual characters, the cinematography system described in this thesis attempts to enhance participants' enjoyment and understanding of our screen-based virtual environments. As digital visualizations grow more complex, cinematography must pace with the new breeds of characters and scenarios. This thesis takes first steps toward a future of interactive, emotional cinematography.

Thesis Advisor: Bruce M. Blumberg
Title: Asahi Broadcasting Corporation Career Development
Assistant Professor of Media Arts and Sciences

Thesis Committee

Advisor____

Bruce M. Blumberg
Assistant Professor of Media Arts and Sciences
Asahi Broadcasting Corporation Career Development Professor
of Media Arts and Sciences

Reader____

Glorianna Davenport
Principal Research Associate
Director of Interactive Cinema
MIT Media Laboratory

Reader____

Thomas Porter
Supervising Technical Director
Pixar Animation Studios

May 1999

TABLE OF CONTENTS:

1 INTRODUCTION

"Did you ever get the feeling that you were bein' watched?"

-Bugs Bunny

The *Swamped!* chicken

As movies and video games come together into a new medium, the art of cinematography needs to evolve. Movies play to the viewers' emotions, evoking humor, fear, love and excitement. Games are meant to be responsive and challenging, focusing on interactivity. In both media, events are displayed for the audience on a screen; in movies by a cinematographer, in games by an algorithm.

In the Synthetic Characters Group at the MIT Media Laboratory, we are creating a hybrid of these two media in which a three-dimensional virtual world is populated by a group of autonomous and human-controlled characters. This new medium needs cinematography that can satisfy both the emotional and the interactive aspects that arise from the interplay of human participants and virtual characters.

I have created an automatic cinematography system for our interactive virtual environments. My 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. The lights control how the three-dimensional digital objects in the world are illuminated. By dynamically changing the camera and the lights, my system facilitates the interaction of humans with this world and displays the emotional content of the digital scene.

Building a cinematography system for an interactive virtual world presents several challenges (described below with references to an interactive installation in which a human controls a chicken in her attempts to protect her eggs from a hungry raccoon):

- First, how do you set up effective cinematography for a bunch of unpredictable actors? If the chicken starts to get scared because the raccoon is after her eggs, how can the camera and lights adapt to show her fear?

- Second, how do you use cinematography to facilitate a participant's interaction with the characters in a virtual world? If the participant's chicken is running away from the raccoon, how should the camera and lights keep the user engrossed in the chase?

- Finally, how do you solve the first two problems with the same cinematography system? Are interactivity and emotion mutually exclusive in a virtual environment?

To answer these questions, I 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, I have 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 wily raccoon from *Swamped!*

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, changing the way it moves through the world to show how it's feeling. 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.

Two characters dancing together in *(void*)*

There is a delicate balance that must be maintained between the level of interactive control in an installation and the means of conveying emotion. The more direct the control, the more of a servant the cinematography must be to that control. Intentional control, though, 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.

In this thesis, I 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 as I have implemented it, and close with a discussion of future work and the relevance of interactive cinematography to modern media.

In the rest of this section, I define a few terms that will be used throughout this thesis, and present the installations that have featured versions of the CameraCreature. In “The Needs of Cinematography”, I discuss the elements that were considered in building the CameraCreature, focusing particularly on participant interactivity and emotional expressivity. In “Related Work”, I outline the strengths and weaknesses of disciplines that have inspired this effort – film and television, video games, and research in autonomous cameras and behavior systems. In the “Implementation” section, I detail how the CameraCreature works, breaking it down into its component parts. In the “Evaluation” section, I consider ways in which this effort has succeeded and failed. In the “Future Work” section, I offer a variety of directions that this research could be continued. In the “Conclusion”, I recap the work that I have done and discuss its relevance in the broader scope of the future of interactive media.

The cinematography system that I describe in this thesis has been custom-tailored to work with specific installations created by the Synthetic Characters Group. However, I have tried to create a system that extends to a broader spectrum of applicability than we currently use it for. Also, while designing it, I have thought a bit about the greater context of this work – how it fits in with other media that currently exist, and what its place might be among media yet to come. This thesis is meant to be a discussion that sheds light on these larger issues, using one specific implementation as a springboard for that discussion.

Cinematography will have a place in modern media well beyond our current horizons of foreseeable media. As long as images appear on a screen, someone or something will need to choose which images to put there. In this thesis, I 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. This CameraCreature can have personality of its own, and have an impact on the interaction based on its own style and mood. By having an autonomous cinematographer as complex as our other characters, I 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.

1.1 Terminology

In this section, I define several terms that will recur throughout this thesis.

A “creature” is an inhabitant of our virtual world who has sensors, drives, emotions, and actions. This word applies to any element of our world that uses the behavior-based approach to action selection originated by Bruce Blumberg [Blumberg, 96] and advanced by our group over the last several years [Kline, 99].

A “character” is a subset of “creature”, including any who have personalities that may be perceived by participants. This includes all the on-screen actors, the cinematography system, and the music system.

The “actors” are all those characters who can appear on screen. These are the main targets of the cinematography system.

“CameraCreature” is the proper name that I will use when referring to the cinematography system as a character, rather than simply as a collection of functionality.

A “designer” is a member of any group of computer scientists, animators, behavior-makers, engineers and interface designers who are working together to make an interactive virtual environment or similar creation.

A “participant” is a human who is interacting with our system. Our system is designed to accommodate several participants at once, some of whom may be controlling actors, others of whom may be interacting with the installation in other ways (influencing music, etc.) Designers are quite often participants, especially in the development stage.

The “audience” consists of any humans who are viewing the installation, whether or not they are participating. Often our installations have a few participants, but a substantially larger audience.

"Interactivity" defines the reciprocal action of a participant providing input to an installation by means of some interface, and in turn attending to the events that occur in the virtual world as a result of that action.

"Action" refers to anything that one or more actors do, that might be of interest to the CameraCreature.

1.2 Examples

We have designed our cinematography system to be useful in a wide range of interactive worlds. By experimenting with different scenarios, we hope to make a system that can be used to set up new systems rapidly in the future.

1.2.1 Swamped!

In the Synthetic Characters' first major project, we chose a cartoon-esque scenario that features a lively chicken and a hungry raccoon. While the chicken runs and flaps and squawks, the raccoon sneaks around the farmyard, exploring and searching for opportunities to eat the chicken's eggs. The only way for the chicken to protect her eggs is to annoy the raccoon enough that he chases her, instead of going after her eggs. This installation, *Swamped!*, premiered at SIGGRAPH 98.

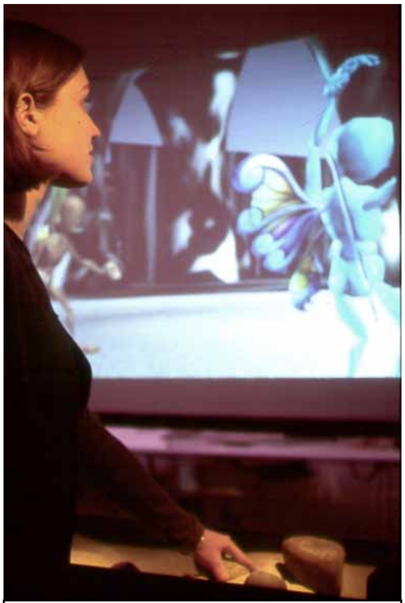
The chicken is semi-autonomous, choosing her actions by blending between her internal drives and external input from a human participant. To provide this input, the participant manipulates a sensored stuffed animal, which contains electronics that detect how the person is moving the toy's body, feet, wings and head. The raccoon is fully autonomous, doing what he must to satisfy his hunger, his curiosity, and his dislike of that chicken. Both characters have an expressive range of motion by which they can take action in the world and express their emotions.

***Swamped!*: A plush-toy tangible interface**

The cinematography system's job is to capture the wacky antics that ensue as the chicken and raccoon interact in the farmyard. Most of the actions of the two characters are broad of scope, modeled after the physical comedy of Buster Keaton, Charlie Chaplin, Wile E. Coyote and the Roadrunner. Therefore, the camera needs to be prepared for actions of this kind, and ready to show them on the screen.

1.2.2 (void): A Cast of Characters*

Our second major installation features a larger group of winged characters, each of whom represents a distinct behavioral archetype in a virtual community. The setting is tribal – a coming-of-age ritual where the human participant guides an undifferentiated character, called the Grub, in its efforts to become a full member of the group of characters. Depending on how the participant influences the Grub, it can develop into any of the archetypes. Each participant's character then remains in the world for several additional cycles of interaction (with different humans guiding new grubs). This installation's name is *(void*): A Cast of Characters*. It is currently under development for SIGGRAPH 99.



(void*): A Cast of Characters

The Grub is a simple grayish humanoid character, with stubby little wings. While it has a certain limited behavioral repertoire of his own, it relies on the participant for the majority of its actions during the ritual. The four adult archetypes (who exist in different numbers depending on what kinds were developed by the last several participants) represent Earth, Air, Fire and Water. Earth is kind and gentle, moving with a steady, constrained style and a rustling sound. Air is bold and a bit aloof, whooshing from place to place with quick fluid motion. Fire is erratic of motion and impulsive in its behaviors, crackling through the world. Water's moves are graceful and sure, flowing to the sound of a running stream.

This assortment of characters, with a wide dynamic range of motion and behavior, poses the greatest challenge yet for my cinematography system. Such different characters require the camera to be able to accommodate a range of expressivity, as participants guide the Grub in one of the four directions. The camera should help show the participants where they are going not only in the virtual space that the Grub inhabits, but also in the behavioral space of that character. By moving quickly or slowly, erratically or smoothly, the camera should help participants to understand which of the archetypes they are currently expressing for the impressionable Grub.

1.2.3 Others

In addition to these two main installations, our group has created a variety of small projects for which I have built cinematography systems. We've made several featuring the above characters in a range of scenarios, a few with a beaver as the lead character, and one with two surfers trying to catch waves. Each has had its own needs and concerns, and has helped me flesh out the important components for a more general solution to the issue that makes up this thesis – how to build a cinematography system that facilitates the interaction of humans with each digital world and displays the emotional content of each scene.

A campfire scenario, starring a beaver and a raccoon

2 THE NEEDS OF CINEMATOGRAPHY

2.1 The Medium

Whenever a moving visual work is created – a movie, cartoon, video game, or virtual environment – the authors need to consider how their work will be seen by the audience. By far the most common medium for this kind of work is the flat screen, in one of a few aspect ratios (1.33:1, 1.85:1, 2.21:1, etc.). Televisions, movies and video games have embraced these formats in a range of scales – from the three-inch screens of Nintendo Gameboys to movie screens fifty feet across. In our installations, the primary visual surface is a projected video image in the 1.33:1 aspect ratio of television and 16mm film.

We chose a single flat screen with continuous moving pictures on it for several main reasons. First, it permits us to exploit the experience of our viewers, who have lived for years in our media-rich culture. Using a screen as the primary visual output device allows this cinematography system to draw on the cultural heritage of film, television, computing, and video games. Second, it allows for a collective experience, where several people watch and interact simultaneously; head-mounted displays and other more immersive display technologies tend to isolate participants. Finally, a system that works on a pre-existing infrastructure is far more likely to be useful in the world. This thesis seeks to explore ideas that may be applied in a wide range of interactive pursuits; the single flat screen is currently the default medium for displaying moving images.

When we decided on the flat screen, I realized that it carried a lot of perceptual repercussions along with the cultural ones discussed above. First among these is the limitation of human perception. One person cannot see everything that happens in a real-time complex environment. Things happen behind us, and behind trees, and too far away. In the real world, we direct our gaze and focus our eyes on what we want to see. In a screen-based virtual environment, though, it is necessary for someone or something to decide what to put on the screen, so that viewers can comprehend the events occurring in that environment. This requires the elimination of an infinity of other possibilities. (On a movie set, for example, there are unlimited viable camera angles from which the director and cinematographer choose.)

Cinematography acts as a visual guide for the audience. By making decisions about camera and lights, we lead the audience to understand events in a certain way. The images that occur on the screen are the only visual element of that world that the audience sees. Therefore, there is the opportunity to synthesize a fictional reality by assembling images in a way that leads (or misleads) a viewer.

Camera work and lighting do more than simply offer up events as they occur. By altering how people perceive those events, they contribute to the meaning of the piece as well. Different cinematic styles can make the same series of events seem comical or sinister, romantic or bittersweet. The combination of shot choice, camera movement and lighting design can have a striking effect on the meaning of a scene.

Framing is of cardinal importance in the way characters are presented. Different compositions (placing characters in different positions and physical relationships to each other on screen) have different connotations in contemporary film grammar. Camera angles and lighting decisions can convey emotions, psychological relationships, and a variety of other important information about a scene. Similarly, camera motion is a carrier of content, with different styles conveying a range of meanings. Framing and camera motion are discussed in more depth in the Interactivity and Emotion subsections below, and in the Implementation section.

In virtual worlds, it is quite possible to have one continuous, smooth fly-through camera style. Because this is so easy, it is a pitfall that claims many victims in their first efforts at digital animation. The CameraCreature breaks up its continuous movement with a variety of transitions (cuts, fades, whip-pans). These transitions are another conduit through which the emotions of the characters can be expressed. There are certain conventions that have been passed down from cinema that constrain transitioning. For example, if a character is walking in a certain direction in one shot, it should be walking in the same direction in the next shot. Jump cuts, where the camera cuts from one shot to another that is very similar (nearly the same position and direction), are generally avoided in film. By remaining aware of cinematic conventions while creating cinematography systems for interactive virtual environments, I hope to profit from the experience of generations of film makers and draw on the cultural heritage that they have created.

An important role of cinematography is to make the lead characters stand out from the rest of the world. A strong contrast between important characters and the background is an excellent way of making them "pop" out of the scene. The two main possibilities for contrast are to have a generally light world with dark characters, or a dark world with light characters. One of the pitfalls we fell into with *Swamped!* was that the characters and the set had a similar intensity level, and were only separated by color, rather than by intensity. This makes it difficult (especially in still shots) to separate our characters from our background. While this is mitigated somewhat by motion, it is still a problem in the fully animated world. In *(void*)*, we're working to ensure that the characters stay brighter than the set and therefore stand out.

To make characters readily distinguishable from their setting, it is important that the cinematography and the visual design of the world work well together. For example, the very saturated colors of the *Swamped!* world made it quite difficult to do anything striking with lighting. A green floor that is (R=0, G=1, B=0) will only ever reflect green light, no matter how much red light is aimed at it. When the set is being designed, it is useful to have a knowledge of the range of lighting conditions it may ultimately be subjected to, and modified accordingly. Similarly, character design benefits from advance knowledge of how the creatures will be shot. For example, the human eye and mind are acutely effective at looking for eyes and faces; if a character will never get near the camera, then it might need more striking pupils and irises to allow its face to stand out from a distance.

A character who doesn't stand out from its background

All of these conventions passed down from film can help the CameraCreature to gain the trust of its audiences. By adhering to a consistent set of principles (especially a set that many audiences are already familiar with), the

CameraCreature makes participants feel at home watching events. Participants who feel comfortable with the way the scene is being presented on the screen are freed from the need to think about cinematography at all, and instead are able to concentrate on interacting and understanding our characters.

In a virtual world, someone or something must determine what portions of our world make it onto the screen. This could be a human with control over some tool set for simultaneously controlling the camera and lighting elements. It could be a script, where the camera takes proscribed actions in a certain succession. It could be a single wide shot that shows the entire world. Or it could be a "smart" computer program – a dynamic artificial system that makes these decisions by means of some higher level mechanism.

Such a mechanism needs to be as smart as the characters it is assigned to display. A world full of autonomous and user-directed characters demands a cinematography system that can keep up with their constantly-changing emotions, motivations and actions. This mechanism would have to address two main issues that lie at the heart of interactive virtual environments – interactivity and emotion. These are the core of the experiences that we construct. Our works are interactive – participants' actions shape the course of events in our virtual world. And they are emotional, as the moods of our autonomous and directable characters change in response to the actions that the participant takes. In the next two sections, I'll discuss ways in which interactivity and emotion can influence cinematography systems for interactive virtual environments.

2.2 Interactivity

Attempting to build a cinematography system that addresses interactivity opens up a large array of new possibilities and difficulties. When one or more participants are permitted to influence a scene, certain constraints must be placed on the cinematography system to ensure that those participants are *able* to influence that scene.

2.2.1 Interfaces

In order to understand the constraints that interactivity places on cinematography, it is necessary to understand the role that physical interfaces play in the interactive process. Physical interfaces, such as the stuffed chicken in

Swamped!, are the mechanism by which participants provide input to our system. These interfaces allow participants to affect some element of the virtual world. Interfaces like the chicken, which permit a participant to control a character, pose a variety of questions about how to affect that character. Is there a direct mapping from the anatomy of the interface to the anatomy of the virtual character? Is this interface a means by which to navigate? How do participants tend to hold the interface? Is the interface meant to be *the same as*, or *a symbol for*, the onscreen character?

The plush toy interface

When a person picks up an interface to control a virtual character, there exists a relationship between the three. This is a physical relationship, resulting from the configuration of the person, the interface, and the screen. And there is a psychological relationship, in which the person either *associates with*, *possesses*, or *is* the virtual character. This relationship has a bearing on the interactive nature of the installation, as participants have differing expectations of interactivity, depending on the relationship that they perceive.

In an interactive installation, participants are thrust into an active point of view. They are the focal point, and have to take action in order for events to occur. This is a strikingly different feeling from the passive point of view that is inherent to television and other non-interactive media. Tangible interfaces heighten this active sensation – holding a fuzzy bird is not like holding a remote control.

Tangible interfaces have bearing on cinematography, since participants must be able to make the correlation between what they are holding in their hands and what is occurring on screen. This can be a simple correlation ("When I flap the wings, the on-screen chicken flies.") or a more complex mapping ("When I turn the stuffed animal to the left, the on-screen chicken turns, too.") Navigation is an excellent example of a kind of interactivity that can be made thoroughly intuitive or nearly impossible by properly or improperly combining an interface with a camera system. In the next section, I discuss how interfaces, participants and virtual characters work together.

2.2.2 Coordinate Systems

One of the great problems facing cinematography in an interactive domain lies in coordinate systems. When a person is using a physical device to control an on-screen character, there are three different coordinate systems – that of the person, that of the device, and that of the on-screen character. For example, imagine that our interface is a stuffed animal. The participant is holding the stuffed animal facing toward the screen, and the corresponding character on the screen is facing away from the camera. The interface and the character are aligned, so naturally, turning the stuffed animal to the left should cause the on-screen character to turn left. However, imagine if the actor on the screen were facing toward the camera. All of a sudden, there's a problem with the control mapping; if the participant turns the stuffed animal to the left. should the actor turn to *its* left (which will cause it to appear to go the wrong direction) or should it turn to its right? Whose coordinate system is correct – the person's, the camera's or the interface's?

To address this issue, when a participant appears to be trying to navigate (e.g. making the chicken walk) the camera favors shots that are behind the participant's character. By repositioning the camera whenever the on-screen character turns more than a certain amount either way (about 60°), this navigation shot can ensure that the coordinate systems of the participant and the on-screen creature match. We expect that the participant will be holding the stuffed animal facing the screen, so that the three coordinate systems match. If the participant is holding it facing the opposite direction, the tangible interface will not correlate to the on-screen character. This is a sacrifice we make to ease navigation in the most common circumstance.

Cinematography cannot solve the interaction problem alone. To make it easy for participants to interact with an installation, it is important that the interface design, gesture recognition software and cinematography system all work together. [Johnson, 99] Participants expect to be able to interact easily, and it is the job of the designers to make that possible.

2.2.3 Expectations

When first-time participants approach one of our installations, they have certain expectations. These expectations come from a lifetime of interacting with natural, mechanical and digital systems. Within the first few years of life, everyone accumulates a substantial knowledge of *how things work*, that is carried into every new experience. Most participants are not aware of their own expectations until they are not satisfied. In the next several paragraphs, I will discuss two scales along which people have expectations of interactivity – time and space.

2.2.3.1 Temporal

There is a strong time element to interactivity. In exploring a new interaction, participants unconsciously test the system, to see what it is supposed to do, what it can do, and especially what it can't do. I'll address these in order from shortest-term to longest-term.

The first thing people inherently test for is *responsiveness*. If they pick up an interface and move it, they expect something to happen. If nothing happens, they look at the demo-giver quizzically, then back at the screen. If still nothing has happened, they say, "It's broken." People assess a digital system as *broken* with amazing enthusiasm. An installation has only a few seconds before it is branded as *broken* and dismissed.

Once they have verified to their satisfaction that a system is not broken, participants look for some *intelligible* response to their action. For example, just as each mouse click should make a laser fire, each squeeze of our stuffed animal interface should make the on-screen chicken squawk. If behavior is random in response to the participant's interaction, that participant will try to glean some pattern from the randomness. An installation should provide intelligible feedback within a few tens of seconds. If this does not occur, the interaction is *confusing*.

Once participants have a grasp of the basic means of interaction, they expect there to be "more to it than that." *Complexity* is my third level of participant expectation. Clicking the mouse fires a laser, but clicking-and-holding the mouse fires the super-duper-laser. Or flapping the wings in different contexts has different effects. Complexity can emerge during the first minutes of an interaction. If complexity does not emerge, the interaction is *boring*.

After participants feel confident in their understanding of the complexity of the interaction, they begin to expect *subtlety*. Human conversation is wonderfully subtle. I would argue that this is a stage that has not yet been reached in digital media.

With regard to cinematography (which is an interface as much as a keyboard is, though output technologies often get ignored as interfaces, in favor of input devices) participant expectations are quite stringent. If the camera holds on an image that is not a character for more than a few seconds, or stops suddenly in the middle of a smooth pan, or goes inside of a tree, it is broken. If the camera moves around so much that it is impossible for participants to get their bearings, it is confusing. And if it shows the same shots over and over and over again, it is boring. Hopefully, this thesis will help build a basis from which someone might someday craft a system capable of subtle cinematography.

2.2.3.2 Spatial

People also have expectations about how they will be able to interact in space. This applies just as readily to a virtual space as to a real one. If their characters have hands, participants will want to be able to pick things up. If there's a tree, they'll want to climb it. When people are placed in a novel environment, they often begin to act more like children, exploring everything. From watching so many people interact with our installations, I've come up with a rough breakdown of what people expect from interacting with a space.

An establishing shot, which helps with orientation

On the largest scale, participants desire *orientation*. They want to understand the gross orientation of the objects in the world and the basic scope of behavior that they might be able to undertake (global campaign vs. village community). They want to see where their characters are, what they are near, and what might be interesting for them to interact with.

Once participants have established their position in the world, and become familiar with the overall layout of their sphere-of-influence, they want to be able to navigate. If there's a building, they wants to go there. If there's a well, they wants to look down it. *Navigation* requires that they be able to get from Point A to Point B with a minimum of difficulty.

When their characters arrive at a location or approach another character, the scale of interaction decreases again. Now *communication* is the key. They expect their characters to be able to communicate with creatures they encounter.

The smallest scale of interaction that people expect is *manipulation*. If there are objects in the world, participants expect to be able to pick them up, use them, put them on, drink out of them, throw them, break them, spin them or do any number of other things to them.

An effective cinematography system should enhance these kinds of interactions. The camera and lights can help to show participants their surroundings, let them navigate, facilitate their ability to communicate with characters and enable them to manipulate objects. Wide establishing shots help people see their creature's surroundings. A strong light source (such as the bonfire in one of our worlds) can help people keep their bearings by watching the illumination of objects. Over-the-shoulder shots show off the orientation of two people talking to each other. Close-ups show faces and hands. Each case is appropriate for certain kinds of interactivity, but strikingly wrong for others.

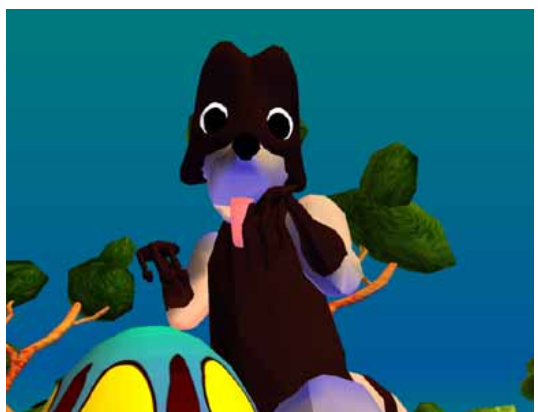
Some video games handle all these varied kinds of interactivity by breaking the interaction into a variety of modes (see Related Work Section). These modes allow a user to look around OR steer OR interact with autonomous characters. Modes are somewhat limiting to interaction. In any given mode, the participant can use only a small subset of their character's dynamic range. In addition, there must be periods of transition as users adjust to the new paradigm. In the CameraCreature, I've tried to build a system that addresses all of these types of interactivity in one continuous interaction.

It is possible to design installations that help in this effort. In order to limit the range of interactivity that a participant expects from an installation, we have learned to craft an interaction that concentrates on the sub-units of time and space that we want to address. If we're making a demo about dancing, there should not be any shots that are close-ups of hands. Such a shot would only lead the participant to expect more from the hands. If we are making an installation about origami folding, though, the camera should not zoom out to an extreme wide shot; it should stay in close on the hands.

Facilitating interactivity is a difficult problem on its own, with an array of competing elements calling for different camera shots and lighting configurations. In the next section, we add in the element of emotion, and discuss how a cinematography system, especially one that can focus on interactivity, can also attend to the emotions of the characters.

2.3 Emotion

Giving characters a range of expressive motion and behavior creates an emotional experience that can be helped or hindered by camera and lights. Camera work and lighting can create emotional effects, but they are far more effective at enhancing emotion already present in the characters.



A surprised raccoon

The Synthetic Characters Group seeks to create characters who can express their emotional and motivational states. In this thesis, I'm presenting a cinematography system that can enhance participants' perception of characters' emotions by presenting them in an appropriate fashion. The challenge is to show the characters' emotions without hindering the interactive essence of the installations. How can our installations have the emotional impact of classic cinema without the director having control over the course of events?

There are several main ways that film makers show emotion through cinematography. We convey how a character is feeling by the camera angles and lighting effects that we choose. We juxtapose shots for dramatic effect. We change the motion characteristics of the camera when it moves. By controlling various parameters of camera and lights, we embed emotion into cinematography.

2.3.1 Framing

The most obvious way to show a character's emotional state is framing – where the camera is placed in relationship to the character. Conventions have developed over the century that cinematography has existed to show a variety of emotional effects. For example, placing the camera below a character and in front of him makes him look powerful, imposing or threatening. Lighting augments this effect – a harsh red light can make a character look positively demonic. Camera framing also establishes relationships between characters, and in the interactive domain, between the participants and the characters. Seeing their character's face causes participants to identify with it.

2.3.2 Montage

Another way to convey the internal state of characters is by juxtaposing images. The famous Odessa Steps sequence from Sergei Eisenstein's 1925 film *Battleship Potemkin* is the classic example of drama being heightened by means of editing. An image that by itself might seem neutral can take on strikingly different meanings when placed next to different images. This is

A still from the Odessa Steps sequence of *Battleship Potemkin*, a 1925 film directed by Sergei Eisenstein

known as the Kuleshov effect, based on work by Russian director Lev Kuleshov. [Gianetti, 93]

Eisenstein's theory of montage suggests that the shots surrounding a certain shot give that shot a great deal of its meaning. [Eisenstein, 1960] While this cinematography system does not actively construct scenes by intentionally juxtaposing images, it nevertheless bears mentioning as one of the major techniques of film making. Creating a cinematography system that constructed sequences by considering montage would be a very interesting extension of the work of this thesis.

2.3.3 Camera Motion

A third way to show emotion is to change the motion characteristics of the camera. The guidelines for this area are similar to those for character motion – just as happy characters tend to walk with a bounce in their step, so, too, our camera follows them with a bouncy motion. A scared character might act skittish, and the camera showing them will zip from spot to spot, stopping abruptly. By layering an emotion-rich layer of motion onto a shot selection mechanism, I've tried to allow emotion and interactivity to complement each other rather than vying for the same elements.

2.3.4 Dynamic Range

In all of these techniques, it is important for the camera to have a wide dynamic range. In framing, the camera and lights should be able to move around with respect to the characters. In cutting between shots, there should be a variety of transition styles. For camera motion to be expressive, the camera must be able to move fast or slow, to change speed gradually or abruptly.

In our current system, the six primary emotions – happiness, anger, sadness, fear, surprise and disgust [Picard, 1998] – each have a corresponding visual style. Both camera and lights contribute to these emotions; for example, happiness has a bouncy, fluid camera style and bright lights, a fearful shot features a jittery camera, faster cutting, and strong underlighting. In each scenario, though, it is necessary to adjust these basic templates of camera and lights to match the rest of the piece.

2.3.5 Continuity

In the Interactivity Section, I discussed the notion of “modes”, which allow the camera to switch among several different models of cinematography. This works passably well in satisfying interactive requirements, but it fails in the emotion department. Building drama in a scene requires the “suspension of disbelief”; it is difficult for an audience to suspend its disbelief when the camera switches to a different style several times each minute. The system I’ve built uses the same camera technique at all times, to maintain continuity throughout the interactive experience.

All of the above techniques are designed to help the audience understand the emotions and motivations of the characters, and the interactions between characters. Understanding the feelings of the characters draws the participants into the scene, and makes it a more enjoyable and engrossing experience.

In order to have a cinematography system that is influenced by the characters’ emotions, there must be a crossover of information between it and the characters. In films, the cinematographer asks the director when he wants information about the characters. In our interactive installations, where there is no director, the cinematographer needs to communicate with the actors to find out what emotions they are feeling.

In our efforts at emotional cinematography in an interactive piece, we strive first to achieve cliché. Cliched camera techniques became cliched because they work well. They are often the simple, obvious ways of shooting. Painters start with years of life drawing in order to perfect their ability to control their medium; I feel that it is important to ensure that simplicity is possible before complexity is introduced.

By embedding emotional effects throughout the cinematography system, and layering characteristic camera and lighting styles over an interaction-based cinematography system, I hope to have created an autonomous cinematographer that effectively serves both emotion and interaction.

2.4 Balance

When creating an interactive installation, it is important to understand the balance between interactivity and emotion. For example, an interaction that features absolute, direct control over a character needs the camera to watch that character constantly. Otherwise, the character might run amok, since it has no "intelligence" except the input from the participant, whose control is incapacitated if the character is not on screen. This restriction on camera motion limits the range of emotional effects that the camera can exhibit. If the participant is controlling a character by less direct means (e.g. providing high-level direction to a semi-autonomous character), there is much more leeway for the emotional elements to explore. A world filled with autonomous and semi-autonomous characters allows for a much more expressive cinematography system. Characters who can look after themselves, at least temporarily, allow the camera and lights to be much more daring and interesting.

In this work, I want interactivity and emotion to be equal partners in influencing cinematography. However, there is a difference in how the two are perceived by an audience. While successfully showing emotion and enabling interactivity are both very beneficial to the experience, failing to enable interactivity at a certain point is far more problematic than failing to show an emotion. The cost of failure is much higher in interactivity.

Therefore, the broad positioning in shot choice often depends predominantly on interactive concerns. Once interactivity chooses the general domain that the camera will inhabit, emotional factors can contribute to the specifics of how it gets there and where it will be within the broad confines set out by interactivity.

Decisions made during the development of an installation – in character design, interface construction, and intended emotional impact – need to be made with an awareness of the others. A movie is an installation where all the focus has been placed on emotional impact, to the exclusion of interaction and interface. Many video games are examples of the exact opposite. There is a continuum between them, though; fertile ground where interactivity and emotional effects can coexist. Intentional control can contribute greatly to an installation that wants to balance interactivity and emotion, since it allows for complete control without requiring constant control. Once constant control is no longer necessary, emotional effects can have more room to make their magic.

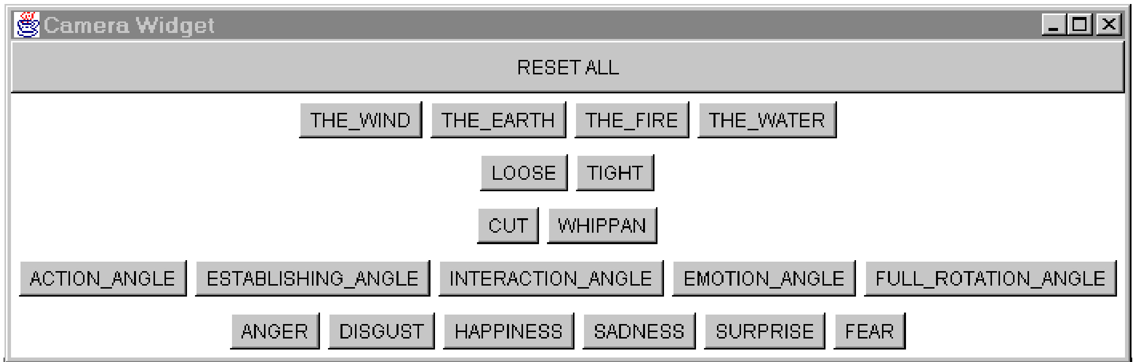
2.5 Development

Up to this point, I have discussed the cinematography system in its ultimate role – doing its best to capture interactivity and emotion by rendering a scene for an audience. However, the system must also serve during the development of the installation. To create a compelling, unified virtual world, it is essential to build the main components simultaneously, so that they can be revised over many iterations to work well together. Characters, camera, music, set – these especially need to be developed with an eye and an ear to each of the others.

A shot of an under-development raccoon by an under-development CameraCreature

When developing characters, it is helpful to have a cinematography system that assists in interactivity and displays emotion. Prior to the CameraCreature, we used a camera that could be moved around in space by means of a widget that let the developer zoom, scale, rotate and the like. Now, early stages of a CameraCreature takes input from a panel of buttons, so that developers can tell it which character to track, and whether to watch it from the front or back. Just this simple level of camera control greatly speeds up the early development of characters, since developers can immediately see what the characters are doing, rather than spending a large chunk of their time moving around the camera.

In later stages, interactive, emotional lighting design can be quite helpful as well. When debugging creatures' emotion systems, it is often difficult to tell how they are feeling, either because the animation is still being tweaked, or because the emotion system isn't working correctly yet. In these cases, having the appropriate character's



A button panel that can control the cinematography system. Includes rows for actors, motion style, transition style, angle, and emotion, as well as a "reset all" button at the top to return the camera to a fully autonomous state.

emotion system hook directly into the color of the sky, with different colors representing different emotions, can show what that character is feeling in an immediate and fairly intuitive way.

In the same way, when developing the cinematography system, it is useful to have the characters under simple button control, so that I can ensure that the camera and lights are doing what they are supposed to do. When a character is told via buttons to "walk" with the emotion "sad", it is easy to see whether or not the camera is helping convey that emotion and showing that action.

Even in the final, running installation, the camera still has a widget available, in case the demo-givers want to demonstrate some element of a specific creature, or somehow influence the shooting style. Having a "Reset" button is useful, too, so that the system can be returned to its fully autonomous state on the fly. The button-widget development tool is quite valuable in different ways, from conception to finalization of an installation.

Another technique that has been incredibly helpful in debugging the cinematography system has been the use of camera sounds. By using sound effects to convey its intentions, my system frees the designer from having to watch a visual representation of the cinematography system's state. Forcing the designer to watch text debug output or a graphical depiction of the system's decision process is far less effective than a voice that plays a sound clip – "Action Shot." Subtler camera sounds, perhaps different pitched whooshes, could even be allowed to stay in the final installation, to give the CameraCreature added personality. Creating the cinematography system as a character makes it as easy to give it sound effects as it is to do the same for any other character; the mechanism has already been implemented. Multimodal design tools are very helpful when working with autonomous cinematography.

A dynamic cinematography system that attends to both interactivity and emotion is a valuable building block in the development of virtual environments. Having it develop alongside the other elements of the world allows it to contribute to the creation process in a variety of ways.

2.6 Story

As our name suggests, the Synthetic Characters group focuses on building strong characters. Cinematography helps us show off our characters. When several of these characters are brought together, we hope that a story emerges from their interactions. In developing story, cinematography takes on added significance, helping to create dramatic arcs in the social equilibria between the characters. In addition to the emotional effects discussed above, we have several techniques for building stories through camera work.

An establishing shot of the *Swamped!* world

The cinematography system has the ability to fade up and down, and to start or end with a certain shot. These effects can provide an interaction with a beginning and an end, and therefore differentiate each participant's interaction into a kind of story, rather than a haphazard subset of a continuous process. Since most people interact with our system for only a short time, usually between three and ten minutes, we tailor our interactions to focus on short vignettes. If participants are interested in a longer interaction, they stay on for several vignettes, and find out more about the characters by watching them in the new situation.

This is the paradigm in many cartoons, where a few characters (Wile E. Coyote and the Roadrunner, Bugs Bunny and Elmer Fudd, Mickey and Minnie Mouse) have a series of short interactions that together contribute to their character exposition. Interestingly, though, these vignettes are just as valid when viewed in different orders. This means that there can be no character development from vignette to vignette, but rather only within each scene. Characters must return to their original state, so that they can interact again some other day. This often occurs by means of bending the laws of nature (Wile E. Coyote gets crushed by something at the end of most of his vignettes, but always returns to his haggard-but-enthusiastic self by the start of the next scene.) The vignette approach, while it offers a view into the characters, and allows for story within each vignette, does not allow for character growth. It is, instead, a character study, viewing an essentially static character from outside.

A camp-fire lighting design

In our current project, we are working on creating characters who change and grow in complexity over time. This will cause our interactions to feel more like movies in their treatment of emotion; a character will start off in one state, experience a series of events, and end in another state. Whereas the series of events is scripted in films, and therefore the character must always arrive in one certain state at the end of the movie, we give control over the series of events to the participant, and therefore allow the character to end up in a wide range of final states.

This change in focus for our construction of story will have the effect of requiring longer interactions to see and affect the development of the new character. In our current installation, we are building a system where a participant controls a character for a few minutes; at the end of that participant's time as active user of the interface, his character becomes fully autonomous, and exists as the archetypal character most similar to the way the participant directed it.

In both of these story paradigms – one with a series of vignettes displaying static characters, the other with characters who persist and change over a series of vignettes – the camera only needs to have knowledge of the characters themselves, rather than having an abstract notion of story. In our simple vignettes, the difficulty is to figure out when a vignette is over. Beginnings are easy. The first beginning is the when the system is turned on. After that, each beginning occurs right after the previous ending. Deciding when to end is the challenge.

We've given certain characters in our installations knowledge of what constitutes an ending. For example, in the *Swamped!* installation, the raccoon knows when a vignette is over because he gets hit with an anvil. The raccoon then informs the cinematography system that it should do an ending – zoom out, fade to black, or whatever the ending should be. This method abstracts to larger numbers and varieties of vignettes, as long as some character knows that an ending is occurring, and notifies the camera. Once the cinematography system has performed its ending, it can then inform all the other characters that an ending has occurred, and they can reset their positions, states and any other relevant information, in order to prepare for the beginning of another interaction.

This greatly simplifies the way that we deal with story. By allowing our stories to emerge from our complex characters, it becomes unnecessary to impose a story on a scenario from the top down. Enabling this communication between the actors and the cinematography system is one of the reasons that I chose to implement this system with the same behavior-based approach as the actors, as I describe below in the Behavior Section.

2.7 Behavior

When we build virtual environments in the Synthetic Characters Group, we keep hold of one goal above all others – creating a unified and consistent interactive environment. To this end, I've built the cinematography system using the same behavior-based approach that we use to build our characters. [Blumberg, 96] This enables the camera to choose actions effectively in real time, increases its possibilities for expressivity, and facilitates interaction between characters and camera.

Like animals in a natural environment, our characters need to figure out what actions to take at each moment of their lives, with limited time and limited information about the world. Similarly, the cinematographer has to choose shots in real time. [He, 96] Just as we use a behavior-based approach when creating our characters, I have found that the same structure of drives, motivations, emotions and actions provides a compelling means of doing expressive cinematography in real time.

An action shot of the raccoon chasing the chicken

To create a cinematography system that resembles our characters in its structure, I envision the camera as an organism who *desires* to record the events that take place in our virtual world. Rather than having hunger for food, it hungers for appropriate shots. The CameraCreature combines several elements when choosing its position – the actor it is currently interested in, the angle from which it wants to view that actor, the motion style that it will use to get to that position, and the kind of inter-shot transition that it prefers. The timing of action-selection in these four sections is affected by the camera's internal motivations, by its current emotional state, and by events that it perceives in the world.

In our on-screen characters, we've noticed what animators and actors have known for years – that *where* they move isn't nearly as expressive as *how* they move. When they choose what to do, our characters layer an emotional style onto basic actions (e.g. sad walk). Similarly, the CameraCreature layers characteristic styles of motion over broader movements that capture the actions and interactions of the characters. By affecting the movement of the camera on both broad and fine scales, the CameraCreature carries information about the events occurring on-screen *and* conveys the emotional content of the scene.

Making the cinematography system a behavior-based creature facilitates overlap and interplay between the characters and the camera. Once the camera is a creature, it is as easy for a character to make eye contact with it (and therefore with the audience on the other side of the screen!) as it is for that character to look at any other character. This realm of possibilities opens up a whole bundle of new options when building a camera – do characters avoid it? Could a character be a camera hog, always trying to be center stage? Can the camera get out of a character's way? Does the camera like one character more than another? Suddenly, the camera has personality and depth of character.

Using the behavior-based approach in building this system facilitates interactivity, because the ease of interplay between the characters and the camera let a participant-controlled character request shots as it sees fit. If that character is navigating, it can request shot that aids in navigation; if it notices a person steering erratically, it can call for an establishing shot that will help orient the participant. Ease of information transfer between the characters and the cinematographer makes it possible for participants' control to be assisted by shot choice.

Considering the cinematographer to be an autonomous character also helps during the development process of the world. A cinematographer who knows how the characters are feeling can show that in its style of shooting. This can help character designers understand the behavior systems of their characters more easily. From the earliest stages of development all the way through the finished product, a tied-in cinematographer smoothes the creative process.

A cinematography system built as a creature gives our world a consistent feel across the different elements. The emotion of the scene can be felt in the motion of the camera as well as in the actions of the characters. When the characters look out through the screen, they seem to come to life. By linking the camera and the characters to each other, a behavior-based cinematography system increases the expressivity of both.

3 RELATED WORK

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

3.1 Film/Television

3.1.1 Movies

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. Together, these films create a rich matrix for visual media yet to come. The heritage of film provides one of the main cultural and technical focal points for this thesis. All the tricks and habits and traditions of how to put images on a screen are crucial background for any effort into such a closely-related endeavor as building an autonomous cinematography system.

Most movies adhere to some basic conventions about shot choice, sequence assembly, and scene construction. These visual conventions help develop the themes that the director is emphasizing in each section of the film. Cunning camera work can help expose subtler themes and nuances, by directing the audience's attention. Awareness of these means of direction (and misdirection) 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.

A camera's most basic task is to aim at interesting things; the first purpose of lights is to illuminate those things so that the camera can see them. Like all good tools, lights allow a functional obligation (providing light) to become an art form (lighting design). Just as there are conventions in camera position, there is a long heritage of lighting in film. These conventions allow the lights to optimize reality – maintaining visual continuity while providing striking images that allow us to feel with the actors and immerse ourselves in the space.

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. [Malkiewicz, 89] Similarly, the Synthetic Characters Group has constructed a set of tools appropriate to the interactive, digital kind of cinematography [Russell, 99] These resemble the film cinematographer's tools in function. There is a virtual camera, sending a live feed to the screen rather than to unexposed film. There are virtual lights, which have values for the intensities of red, green and blue light, rather than a suite of gels, screens and gaffer's tape.

In addition to filling the job of the cinematographer, cinematography systems in real-time virtual environments are the *de facto* editor, too. Editors choose how to arrange the shots provided by the cinematographer in a way that tells the story. In this process, they have substantial creative discretion, since many moments are shot from several angles, to provide "coverage", lest the editor get stuck without a way of making the scene cut well.

In a real-time world, "right now" is the only chance to show the scene to the participant. Shots, transitions, and cuts happen, and if they look bad, it is already too late to change it. There is no independent editor – by choosing the next shot, the cinematography system is inherently the editor. As such it needs similar tools to a film editor: the ability to cut, a way to fade in and out, a way of knowing how long or short a shot can be. In my system, the editor is interspersed throughout the cinematography system, living predominantly in the "Transition" section, but also residing in the timing decisions that choose when to cut and when to hold a shot.

Films are often shot out of sequence, with all the scenes that occur in one location being shot at the same time, to cut down on production time and expense. While the CameraCreature must render scenes in order (real-time is like that), in digital systems it is quite possible to change any element of the world instantaneously. While the current version of the CameraCreature does not warp reality, future versions may be made to control many of the elements that are, in film, unchangeables. A cinematography system that could subtly alter character position for the duration of a single shot would be quite powerful, if it had a way of assessing the composition of the frame.

The huge difference between films and our medium is interactivity. In films, there is none. The director and her crew make a movie, the film lab makes prints of that film, and everyone in every theater sees the same movie (modulo scratches, etc.). In an interactive environment, the experience is different every time, depending on how the participants interact. Aside from tying the cinematographer's hands with the real-time elements described above, it also adds a whole suite of difficulties in order to allow the participant to easily and effectively interact. (Spoken like a good-old-fashioned film-guy, I suppose.)

3.1.2 Documentaries

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. While a fiction film might have a crew of hundreds, performing tasks from set construction to make-up to dolly-pushing, the typical documentary film crew is three – director, cinematographer, and sound. In our installations, there aren't even three people, just an assortment of smart tools.

The real-time shooting style of many documentaries gives the work a feeling of urgency, which helps to accentuate the subject matter. Narrative features occasionally exploit this urgency, using hand-held cameras to invoke the documentary feel. In early versions of the CameraCreature, I found that a similar feel was almost inevitable when dealing with interactivity; things change, and the camera lived in a constant state of *reacting*. Therefore, I chose to embrace the documentary as my primary stylistic inspiration.

In fiction film making, the cinematographer is usually meant to be transparent; in documentaries, there is often the feel of a live person behind the camera. Having the cinematographer of our virtual worlds be a biologically-inspired virtual character like our actors causes it to have a documentary-like organic feel. The live decision-making of a behavior-based creature has been invaluable in embracing the documentary as a stylistic inspiration.

Once again, though, the great difference between documentary films and our interactive installations is the real-time element. Without recourse to post-production, my CameraCreature must contend with events as they happen, and document them in real-time. Although they often document real-time events, documentaries eventually have the luxury of the cutting room when crafting a final product.

3.1.3 Televised live sporting events

The advent of video technology allowed there to be a real-time element to television. Whereas film must be developed in chemicals before it can be viewed, and will therefore never be a real-time format, television has embraced live shows from its very beginnings. Today, televised live sporting events are the most valuable advertising time, with a 30-second spot during the Super Bowl costing over $1.6 million. People clearly place value on real-time, live events.

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. There is a general domain of variability that the cinematography system must cover, but it doesn’t have to be prepared for completely novel actions.

To cover the range of possibilities, sporting events are usually documented by means of a number of human-operated cameras, the output from which appears on a row of screens being watched by a human director, who chooses from among them to decide which one to send to the viewing audience. The CameraCreature’s behavior system is similar to this, considering a variety of kinds of shots before choosing the most appropriate.

The great difference between sporting events and our subject matter is that emotions aren’t crucial in sports. While the director might cut to a shot of a disappointed athlete on the sidelines during a time-out, he’s not going to cut to that emotional shot while a wide receiver is sprinting down the field. In a movie about football, the director might choose to show a touchdown by watching the expression on the losing coach's face, rather than showing the player crossing the goal line. But in live sports, emotion very clearly takes a back seat to action. In the work of the Synthetic Characters Group, we want emotion to play a central role, standing alongside interactivity as equal partners. Balancing these two is the challenge set before my cinematography system.

3.1.4 Live-to-tape shows

Live-to-tape shows, such as soap operas, are also worth mentioning. 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.

3.2 Video Games

Video games are interactive in real-time. It would not be a video game if it did not change in response to a player's actions; it would be a video tape. Video games are a huge industry, with Nintendo, Sony Playstation, Sega and PC-based games firmly entrenched in the world's digital cultures. 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. In this section, I'll discuss a few of the archetypal games for these paradigms. Some games combine several paradigms, to varying degrees of success.

3.2.1 Zelda: The Ocarina of Time

In *Zelda*, you play an elf-like creature, exploring a variety of scenarios in a virtual world. Your character can navigate, talk to people, shoot at things, play an ocarina and look around. Each of these actions has an accompanying camera style; choosing your action determines which of several camera paradigms the game uses. For navigation, the camera follows your character around, doing smart things like staying out of walls and looking over cliffs when you get close, but generally staying behind your character so you can see where you're going. When you decide to talk to a person, it smoothly transitions into specified view of the two-character interaction. It then stays fixed in that position, so you aren't distracted while reading what the character is saying. When you take out your sling-shot, it switches into aiming mode, with a first-person shot through the triangle of your weapon. When you play the ocarina, it chooses a specific shot that stays fixed on your character so you can choose your notes without being distracted. Finally, there is a mode where you can get control over the camera, so that you can look around from the point of view of your character.

***Zelda*: a navigation-enabling shot**

The interaction in *Zelda* is fairly seamless, with a multi-skilled character who is entertaining and easy to control, without being simplistic. However, the use of different camera modes, despite having smooth transitions, breaks a sense of continuity as a whole. It is interesting, but it is not exciting or emotionally-charged. In order for events on a screen to transfix the viewer, there need to be long periods during which the audience can suspend its disbelief. When the camera changes style so frequently, it is difficult to lose yourself in the events that occur.

In addition, no effort was put into creating any emotion, even if there had been continuity. The player's character is a lifeless shell (which could be okay, if the player is meant to imbue him with personality) and all the other characters in the world are flat of character, with a few stock phrases typed in as text.

Zelda has an exceedingly competent camera system. Technically, it is rarely "broken" (going through walls or the like). It chooses shots with some intelligence, and has easy-to-use modes. But the modes are its downfall in the emotional realm.

3.2.2 Super Mario 64

Super Mario 64 allows the player to become Mario, and travel through a series of quests in a castle setting, in order to free Princess Toadstool. *Super Mario* is a wonder of navigation. Travelling through the world is intuitive after only a few moments at the controls. The camera system in *Super Mario* is a big part of this ease of navigation; the camera moves around quite intelligently to show Mario and his environment, without disorienting the player. The camera system is so smart, in fact, that at the beginning of game play the camera is shown being carried by a winged bee-like creature, and the cinematography system is introduced as a set of brothers working to document your travels.

Mario preparing for a dismount

It is possible for the user to take control of the camera and look around with it. It is necessary to do this in order to solve some quests. This looking-around mode is always available, but isn't at all necessary for navigation. Having navigation be so easy, and yet so interestingly shot is quite refreshing in a video game.

However, like most video games, it has the unfortunate shortcoming of being devoid of emotion. Mario has a fair amount of character, in his bustling little run and enthusiastic leaps, but he always behaves the same way. He is flat, the villains are flat, the camera system is flat. Until video games overcome this problem, all the interactivity in the world won't make them as enthralling as a good movie.

3.2.3 Tomb Raider

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 I registered at being required to shoot at tigers.


Lara Croft fires her pistols

3.2.4 Ultima Online

The top-down view of *Ultima Online*

Ultima Online offers the top-down camera style featured by many adventure games. As your group 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 you and your characters.

3.2.5 Grim Fandango

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.

Fixed camera angles in *Grim Fandango*

3.2.6 Doom

In *Doom*, the player is a gun-toting soldier in a multi-level dungeon. The camera system is first-person; while playing, you see the barrel of your weapon (there are many) in the lower portion of the screen so that you can aim.

Point-of-view camera in *Doom*

Aside from a few statistic about your status, the entire screen shows a straight ahead view of what you are seeing. Steering and shooting are done with a combination of mouse and keyboard.

Doom's first-person view allows the player to have complete control of the camera. Camera and character are exactly linked – your character *is* the camera. Once you get the hang of it, it is a fast interface, and in the game, reaction-time counts. However, it takes a while to learn the mapping between keyboard/mouse and what happens on-screen.

While *Doom* offers excitement galore and a rapidly-responding kind of interactivity, it is completely barren of any emotion. Characters run and shoot and die, and enemies kill you or you kill them. The camera system is not at all autonomous, since the player has absolute control over it.

3.2.7 Thief

In *Thief*, the player takes the role of a master burglar sneaking into an enemy fortress. It is similar in format to *Doom*, but the story line changes the mode of 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. By leveraging the simplistic continuity of the game play, *Thief* creates a remarkably compelling experience. Continuity breeds suspense; *Thief* is quite visceral in its capacity to make the player scared. Although it creates only one emotion, it is still a big step towards a full emotional repertoire in video games.

***Thief* builds suspense through continuity**

3.3 Research

The Virtual Cinematographer [He, 96] 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. However, this system has the same basic problem that so many video games have – different modes for different action. 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 topic of lighting design, traditionally the role of the cinematographer alongside camera positioning. Our system attempts to capture the same feel with a less constrained approach, where shots are constructed from multiple sub-parts, and a single system, the CameraCreature, is always in control of the camera and the lights.

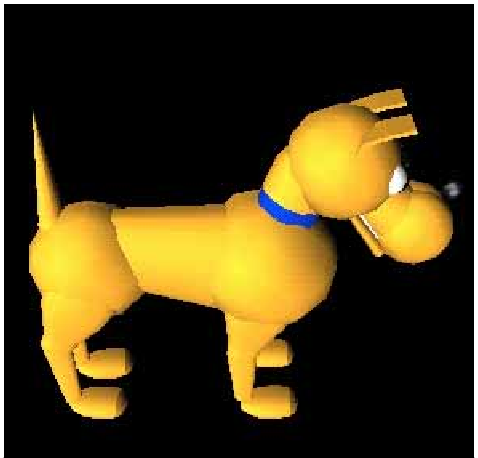
Also, the Virtual Cinematographer fails to address the topic of interactivity. The paper does not mention the problems and possibilities inherent with having live human participants. The Virtual Cinematographer may create a stunning sequence of shots, but I'm concerned about whether it would be possible to *interact* with those shots.

Steven Drucker [Drucker, 94] 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. Drucker addresses the problem of shot selection by considering it to be a series of small constrained optimization problems. He deals primarily with the problem of how an interactor might more effectively control a camera, and how an autonomous element of that camera system might find the best way of satisfying that interactor. In this thesis, I present a system that is fully autonomous and chooses shots continuously without any direct influence by the participant.

Tinsley Galyean [Galyean, 95] 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. In this thesis, I present a cinematography system that shoots a scene based primarily on its knowledge of character, rather than having a separate notion of story.

Bares and Lester [Bares, 97] 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.

3.4 Behavior Systems



Silas the Dog, Bruce Blumberg's Ph.D. Thesis

In constructing the behavior system for the CameraCreature, we built on the work of Bruce Blumberg [Blumberg, 96]. 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.

The Improv system [Perlin, 96] also offers inspiration in building interactive characters. This system seeks to create believable autonomous actors for use in scripted interactions. The two parts of Improv enable authors to readily craft rules by which their actors behave and develop procedures for automatically generating motions. By applying noise functions to actions of the 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.

Christopher Kline has furthered Blumberg's work, creating the underlying behavior structure that I have used in this cinematography system. [Kline, 99] 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. A more elaborate treatment of how I use this system follows in the "Implementation" section.

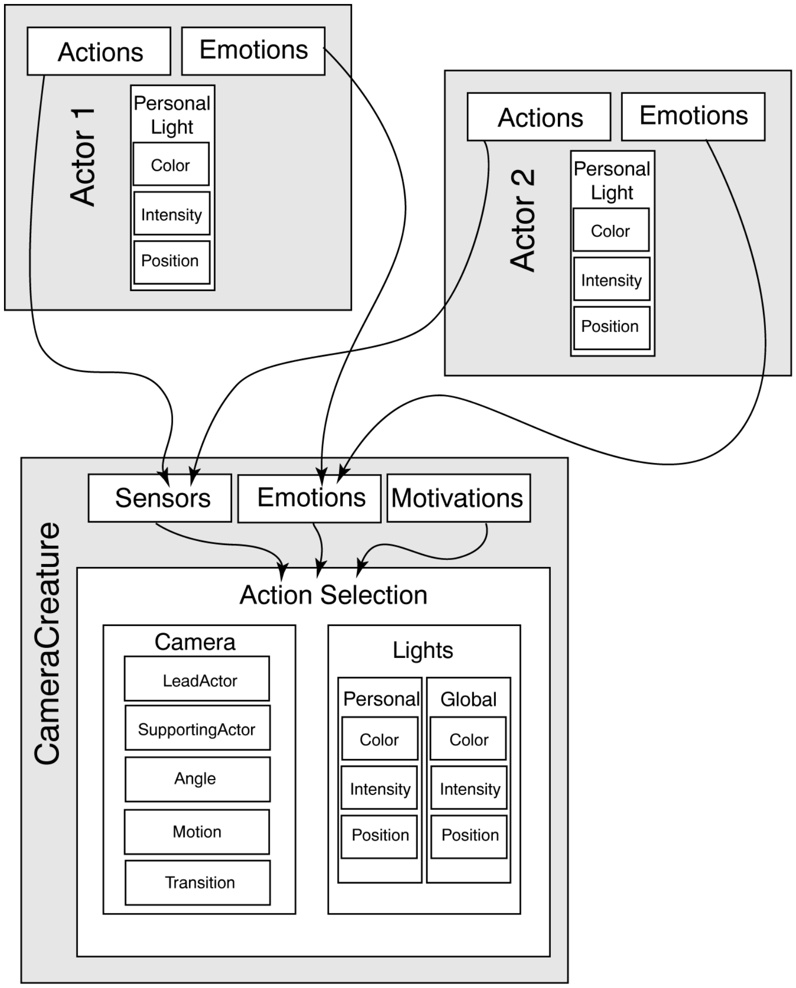
4 IMPLEMENTATION

In this section, I propose an implementation for the system described in the above sections. The implementation below is an abstract view of how the system works – a composite of several actual implementations (*Swamped!*, *(void*)*, and others). All share the same basic structure – an autonomous behavior system combining a variety of elements to choose its shots. I have consolidated them for the sake of clarity.

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 [Blumberg, 96] and continued by Christopher Kline, Song-Yee Yoon, and the rest of the group. [Kline, 99] The second is the graphics system designed by Kenneth Russell, Michael Patrick Johnson, Bruce Blumberg and the rest of the Synthetic Characters Group. [Russell, 99] The CameraCreature's personality has been designed using the behavior system architecture, and its functionality has been made possible by the graphics system. In the implementation details below, I will primarily consider higher-level concerns, since much of the graphical and behavioral underpinnings were coded by other members of the Synthetic Characters Group.

The CameraCreature is a combination of a behavior system and some Java code that executes the decisions of that behavior system. The behavior system was written using the VRML file format to represent a network of hierarchically-organized cross exclusion groups. The Java side takes the decisions of the various elements of the behavior system, performs the appropriate math, and delivers the resulting information about the camera and lights to the graphics system.

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 being combined with motivations, modulated by emotions and fed into action-selection mechanisms. Each action and group of actions has a set of inputs and gains for these inputs, by which the inputs are modulated.



A flow-chart diagram of two actors and the CameraCreature

4.1 Sensors

The sensory apparatus of the CameraCreature detects all of the relevant creatures in the world. Actors in our virtual worlds often have limits placed on their sensory inputs, just as animals are limited in what they can see, hear or smell; the CameraCreature, though, is free from those limits, so that it can have omniscient knowledge of the events occurring in the world. Being able to sense the actors allows the cinematography system to be influenced by their emotions. The sensor section establishes the link between the characters and the cinematographer.

4.2 Emotions

Like our other characters, our CameraCreature has an emotion system that affects which shots are chosen, and also affects parameters within those shots. We use six cardinal emotions in our system – anger, disgust, fear, happiness, sadness, and surprise. [Picard, 98] Each of these has a corresponding visual style, and affects the values that are fed into many levels of the action selection mechanism.

Each emotion has two main elements that constitute it – a button element and an autonomous element. The button allows a developer to dope the emotion system with artificial values. This makes it is easy to see the CameraCreature behave in its various emotional styles.

The autonomous component of each emotion is made up of several parts. There is a base value of the emotion, which is a function that describes that emotion in terms of growth and decay rates. This is the CameraCreature's default temperament.

To this is added emotional inputs from the actors in the scenes, perhaps weighted toward certain characters to make them lead actors. At each clock tick, every actor tells the CameraCreature its values for each emotion. By blending the emotional states of the various characters in the world, the cinematography system is influenced by the emotions of the actors – if an actor is very happy, it will provide a strong input to the CameraCreature's happiness, and the CameraCreature will be more likely to behave in a happy manner.

In addition to the base value for each emotion, there exists an amount by which that value is maintained or inhibited once it becomes active. For example, a shy character might become angry, but anger would self-inhibit, or possibly increase fear, so that anger would not remain the dominant emotion for long. By altering how different emotions maintain and inhibit each other, I have tried to craft CameraCreatures with distinct personalities.

Each of the CameraCreature's emotions has a corresponding visual style.

“Anger” has the following effects on the camera – abrupt cuts, tight tracking of moving characters, a motion style that tries to capture constrained violence. The sky turns red when the cinematography system is angry.

“Disgust” tightens the camera’s springs, so that it pans tightly, and whip-pan between shots. Disgust gives the sky a reddish-orange cast.

“Fear” causes the camera to become somewhat jittery, to simulate the shaky-cam (used by the Cohen Brothers in their film *Raising Arizona*) that denotes someone stalking someone else in modern film cinematography. The sky becomes a dark greenish-blue.

“Happiness” causes the camera to whip-pan fairly frequently with a bouncy motion. The sky becomes an upbeat gradient of cyan.

“Sadness” makes the camera track characters with a melancholy slowness. Whip-pans are transitions, but are quite slow due to the loose springs. The sky becomes dark blue.

“Surprise” causes the camera to cut frequently and move rapidly. The sky lights become brighter.

4.3 Motivations

Whereas emotions alter the way the system chooses actions at all levels, motivations affect only a subset of the action selection mechanism. Hunger is a fine example of a motivation – it greatly increases a creature’s probability of performing some kind of “eat” action, but is not a factor in other parts of that creature’s behavior system (except in that the “eat” section may win, and therefore disable all the other areas). Each motivation, like each emotion, has two sections – a button input and an autonomous component.

The buttons that control motivation constitute the main level at which artificial control is inserted. Each motivation that a developer might want to control has a corresponding button. When clicked, a button provides very strong inputs to that motivation, and therefore to its corresponding action. This makes the button widget an integral part of the creature, rather than an add-on that circumvents an existing system.

The autonomous section of each of the CameraCreature's motivations begins with a formula that represents the CameraCreature's natural tendency toward or away from that behavior. Actors are able to add values to motivations by making a request. This is a way in which motivations are different from emotions. Whereas each actor is expected to tell the CameraCreature its emotional state at every tick, actors only affect motivations by means of an active request (see below – RequestShot).

In addition to the base value and the actors' inputs, there is also a mechanism by which different motivations can reinforce or inhibit each other. This is the heart of the autonomy of the creature – the place where motivations affect each other and then feed into the action-selection mechanism.

4.4 Action Selection

The action-selection mechanism allows for two different kinds of hierarchical organization – BehaviorGroups and BlumbergHierarchies. A BehaviorGroup is a collection of behaviors and groups of behaviors, any of which may be active. Multiple parts of a BehaviorGroup (any that have a non-zero value) may be active simultaneously. A BlumbergHierarchy is a collection of behaviors and groups only one of which may be active at any given time. Within a BlumbergHierarchy, cross-exclusion weeds out the lower-valued actions.

The top level of the action-selection mechanism splits the process into two main parts – camera and lights. This is a logical split, since the two have relatively little cross-over. There are a few elements that cross the camera/lighting gap (for example, distance of the camera to the lead actor affects the intensity of that actor's light). And camera and lights are both affected strongly by the emotion system. But overall, splitting camera and lights is the obvious first level of division.

4.4.1 Camera

The CameraCreature must decide where to put the camera, and which direction it should be facing, at every clock tick. It does this by deciding where it would ideally be at each tick, comparing that to its current position, and then determining how far towards the ideal position it is appropriate to move. To choose its ideal position, it decides which actors it wants to look at and what angle it should use to look at them. Then it decides what motion and transition style will be properly expressive for that shot. Once it has determined all these elements, it sends the answers to the Java side of our system, which actually performs the calculations on the chosen elements, and tells the graphics system where to put the camera.

There are two main parts of the camera's decision process – the shot choice and the motion style selection. Shot choice is primarily the domain of interactivity, since camera position relative to the character determines the ease of interaction. Motion style is the vehicle for emotional content. There is overlap between the two, since small perturbations around a basic shot can carry emotional content without sacrificing the interactive intention, and since motion style affects how rapidly the camera achieves its desired position.

Timing of transitions between shots is another means of conveying the emotional mood of the scene. Fast cutting feels angrier and more exciting than slow cutting. Since the timing of transitions is determined by the selection of actors and angles (whereas the style of transition is determined by the "Transitions" section), the emotion of the CameraCreature should affect the rate at which actors and angles grow and decay. This would allow the rate of turnover among actors and angles to depend on the mood of the scene. Currently, this has not been implemented, but would be a valuable addition to the system.

4.4.1.1 LeadActor

The most important decision that the camera makes is which actor should be the lead. Most frequently, this is the character who is controlled by the participant. Inevitably, 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 air-creature is the lead actor

In addition to being able to choose any of the actors, the LeadActor section may also choose to look at the world. This enables the camera designer to craft specialized shots, where the camera is placed in a certain position and orientation in world coordinates. These specialty shots, when used sparingly, can give the impression of great intelligence to the camera, since it periodically frames shots very precisely to match a specific invariable event. An example of this is the egg-ejection shot in *Swamped!* In this shot, the Acme Truck spits out a big Easter egg, which lands in the foreground of the camera. In order to create this shot, it was necessary to specify the camera position in world coordinates (since we knew in advance where the egg was going to land.) Overuse of specialty shots, though, gives the interaction a wooden, scripted feel.

I will refer to the leading actor (or the world, if it has been chosen) as the "target" of the camera for much of the remaining discussion. 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.

4.4.1.2 SupportingActor

The SupportingActor section is much like the LeadActor section. Its inputs have essentially the same values (although they have different button inputs), but the actor who wins in the LeadActor section is inhibited here. Therefore, the SupportingActor is the actor who has the second highest value for becoming LeadActor. The supporting actor is used in our interaction shots, in which the camera is framed around two characters rather than just one. While this is less useful in terms of human interactivity, it is very useful for showing off the physical and emotional relationship between two characters.

4.4.1.3 Angle

In addition to deciding which actors it is interested in, the camera also chooses an angle with respect to its target. There are six main angle types that I've developed, each of which serves a different main purpose. Within each angle type, there may be several versions of that angle, corresponding to different emotions or some other element.

Each of these angles has a variety of parameters that let it adjust camera and target positions. Each Angle behavior must decide the following: whether it is interested in the LeadActor or in the axis between the LeadActor and the SupportingActor; how far to rotate around the LeadActor or that axis; the distance it should move from the actor (a proportion of that character's height); how far ahead of the character it should look (see the next paragraph on leading); how high the camera should be (this may be emotionally determined, but again, it's a percentage of LeadActor height); whether the camera should stay fixed for the duration of the shot or track with the character; and whether the target should stay fixed or track the character.

Allowing the camera to look at a point a bit in front of the actor, or "leading" that actor, is a useful technique for creating better shots. It is customary in cinematography to have a character positioned somewhat to the left on-screen if he is walking right, and somewhat to the right if he is walking left. Causing the camera to look at a point that is in front of the character (for example, half a body length along that actor's vector of forward motion) moves the character to the proper side of the screen. This has become common in films because it lets the audience see more of the screen in front of the character, and less of what is behind him. This is an effective aid to interactivity, too, since seeing what is in front of the character is helpful in navigation and other kinds of interaction.

Establishing angles are useful for orienting a participant in the virtual world. By moving the camera far overhead, it gives a bird's-eye view of the world, allowing participants to see where their characters are, where other characters are, and what structures (buildings, boulders, etc.) are nearby. Establishing angles tend to be single actor shots, with a fixed camera position for the length of the shot. Allowing the camera to move in order to keep its position fixed with respect to the character would be disastrous, since a small turn on the part of the actor would result in an enormous movement of the camera.

Navigation angles position the camera behind the participant's character, so that the participant can see where the character is going. This also ensures that there is consistency between the coordinate systems of the participant and the on-screen character. With navigation angles, it is important to reposition the camera whenever the character turns more than a certain amount toward the screen, so that participants are always able to see in front of the character. I prefer navigation angles with fixed cameras that track targets, since it limits the overall amount of motion in the participant's visual field. This is especially important with larger projection screens; a camera style that seems perfectly fine on a smaller screen that we use for development may be nauseating when put up on a 50-inch display.

Interaction angles allow two actors to influence the camera position simultaneously. By placing the camera at a certain position with respect to the axis between the two actors, it can establish a relationship between those two actors. This relationship is both physical (since it shows where the two actors are, whether they are looking at each other, and how far they are apart) and psychological (since a shot framed around both actors shows how they act toward each other).

Action angles show what one character is doing. They position the camera in front of the actor, looking at it. The camera should be far enough away that the actor's entire body can be seen. This shot is useful if a character is performing an interesting action, like opening a box, but makes it very difficult for a participant to navigate.

Emotion angles are close-ups of one character's head and shoulders. Close-ups are very expressive, when they work. Getting them to work, though, is quite difficult. Whether the camera and target are fixed or tracking, a rapidly-moving actor is very hard to frame in an emotion angle. Therefore it is useful to place some limit on when actors can get emotion angles, depending on the speed of the actions they are taking. Emotion shots 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.

Specialty angles are used in a few limited cases by my CameraCreatures. By choosing the world as their leading actor, these shots frame events from a non-character-based perspective. Specialty shots are used to frame a shot exactly in the global coordinate system, to ensure that a certain moment of action gets seen from an appropriate angle. Since these are the only shots the CameraCreature chooses that are not oriented around one or more actors, it is important that they have the same "feel", so that they do not stand out from other framing. To help these shots fit in, they have a slight camera motion built in, so that they are not completely static. Since the other shots are based around characters, they derive their motion from the movements of the actors they are watching.

In addition to the above angles, I also have a No-op in this section. This can be fired to cause a cut from a certain angle of an actor *to the same exact shot*. This is useful especially in the case of the navigation shot, since often it is necessary to transition to another navigation shot of the same actor. A No-op is fired if the LeadActor turns more than a certain amount (e.g. 60°) away from the forward direction of the camera. If the participant appears to be steering for an extended time, several navigation shots may occur back-to-back. Since the camera is fixed within a navigation shot (to reduce motion in the participant's visual field), it is necessary to break the shot several times, to realign the camera behind the actor.

Most of the camera angles in the CameraCreature are set to track the LeadActor. This ensures that characters do not walk off screen. In one version of the CameraCreature, the CameraCreature watches the LeadActor to see if it is more than a certain angle away from the forward direction of the camera. If it is, the CameraCreature fires a No-op, to realign the camera.

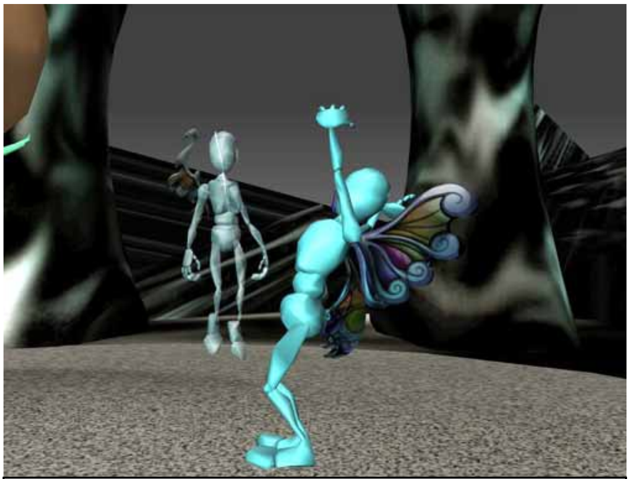
Each of these camera angles has a different bearing on interactivity. Navigation shots and establishing shots serve well for navigation. Action shots are useful for characters doing many non-navigation behaviors (wave, dance, jump). Emotion shots are too close to be of use to any interactivity unless an interface allows control of emotions (which none of our installations does). Interaction shots are useful in a variety of kinds of interactivity, since they show the relationship between two characters. If one is chasing the other, or if they are talking, an interaction shot may serve different purposes.

Depending on the kind of interactivity that a shot serves, coordinate systems may be dealt with differently. In navigation-enabling shots, it is useful to maintain a strong correlation among the coordinate systems of the on-screen character, the participant, and the interface. In action-enabling shots coordinate systems are less central to the interactivity, and it is not necessary to readjust the shot (via the No-op) to keep them registered.

4.4.1.4 Motion

While camera angles are primarily concerned with interactivity, the motion characteristics of the camera are the main conduit of emotional expressivity. I break the movements of the camera into two main areas – intra-shot motion and transitions between shots.

The section of the CameraCreature's action-selection mechanism that deals with intra-shot 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.



The water-creature dances at the air-creature

Because a real camera in the physical world has to deal with physics on its side (and therefore instant acceleration is impossible), it could never keep a position *exactly* with respect to a certain piece of its target's anatomy. Our camera can do this, though. And since our characters are able to move abruptly, those motions would simply be magnified by the camera.

Before it gets sent to the graphics system , our camera and target positions are piped through the spring dynamics system. Controlling the parameters of this system can create very distinct motion styles.

“None” allows the camera to track a character directly, without any smoothing of the motion. This results in a rather jittery look, that is useful for a fearful shot to make it seem as if the actor is being stalked by the camera.

“Loose” smoothes the motion with loose springs and strong damping, so that the camera feels lethargic and slowly comes to its new position. This motion style is useful for a sad camera system.

“Medium” smoothes the camera motion in a bouncy, happy-feeling way, with medium springs and loose dampers. As the camera comes to its new position, it often oscillates a bit before finally coming to a rest. This motion style also makes the camera bounce if its dynamical system tries to send the camera through the ground plane.

“Tight” smoothes the motion with tight springs and tight dampers. This results in a rather angry feel, since the camera moves abruptly and without much excess motion.

In addition to motion effects that are controlled by the spring dynamics system, this section is also home to specialty effects, where the camera displays a certain unique kind of movement. For example, “Shake” is a specialty case which becomes active when something heavy falls and hits the ground. In our implementation, inspired by the tradition of Warner Brothers’ cartoons, an anvil falls on the raccoon as a result of a prank by the chicken. When the anvil hits, the ground seems to shake as the camera shakes up and down. This ties the camera to the physical world in a way that is quite convincing.

4.4.1.5 Transition

Whenever the camera changes its LeadActor, SupportingActor or Angle, a transition to a new shot occurs. This transition can have a variety of different styles. We’ve implemented three in our various CameraCreatures.

“Cut” causes the camera to go immediately to its new position. This is the transition most frequently used in movies. Since a cut feels more abrupt than whip-pan, we use it for angry and surprised emotions.

"Whip-pan" sets the camera to a new position, but pipes it through the dynamical system described above (in Motion). This causes the camera to move through the world to its new position. This can occur rapidly if the springs are set tightly, or slowly if they are loose. Generally, whip-pans take between 1/4 and 1 second. The speed of the whip-pan depends on the motion style; a loose spring makes a whip-pan feel sad, while a bouncy spring feels happy.

"Fade Out" is a specialty case which occurs at the end of a demo or a particular vignette. For example, in the spirit of the Road Runner, after the anvil falls, the camera fades to black to signal the end of a segment. While the camera is black, the anvil has a chance to disappear and the raccoon to recover from his wound.

For each transition, there are a few parameters that should be set. Sometimes a transition should not happen right away, but at some definable time in the future. Also, certain effects should last for a specified amount of time. Therefore, we have parameters for "delay before transition" and "transition length".

Whenever a transition is fired, I've often found it useful to inhibit changes in the currently-active LeadActor, SupportingActor, and Angle. This keeps the CameraCreature from cutting again too soon. Having cuts follow too quickly after one another makes the camera feel like it's confused or broken.

4.4.1.6 RequestShot

In order for actors to inform the cinematographer of their desires for shots, I created a mechanism, called RequestShot, by which actors can communicate with the cinematographer. RequestShots are distributed among the actors – each actor can ask for shots whenever it feels the need. Actors may request each of the following elements with a specified value – LeadActor, SupportingActor, Angle, Motion, Transition. Since they are able to request any element that the cinematographer controls, the actors are very influential in the shot-selection process. During the development of each scenario, it is important to balance the input of the actors with the autonomous drives and motivations within the CameraCreature itself.

4.4.1.7 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 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, that the camera would go up forever in an attempt to get that actor in an unoccluded shot.

4.4.2 Lighting

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.

A skittish raccoon with a cyan sky

An angry raccoon with a lurid sky

I've 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.

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 (steering by the sun – phototaxis – has been with us since organisms first began to move). The global lights affect the overall coloration and illumination of the world.

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 by calculating them in their characters' coordinate systems, so that they appear to travel with their characters.

Although they are split, conceptually, 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).

4.4.2.1 Global

When creating the global lighting scheme for each virtual environment, I begin with a key light source. This might be the sun, a bonfire, or some other strong source of illumination. Then I fill out the world illumination with other lights, often of a different color from the key light. For example, if a fire is shedding reddish light, the fill lights would be predominantly blue hues.

The global lights are exceedingly important for setting the mood of the scene. Using a deep blue light as the key light, and filling with low-intensity white will make the scene feel like a moonlit night. Strong blue and pink lights from opposite sides give a theatrical, carnivalesque look. By tailoring the world lighting design to work with the set, a convincing theme for the scene can be constructed.

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 (a deep blue light might be (R=0, G=0, B=0.4)). 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. Position places the camera in the world's coordinate system. Global lights tend to be fixed in position.

4.4.2.2 Personal

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, considering the Synthetic Characters group's views on virtual environment design. Personality is seen not only through the behavior and emotions exhibited by the actor, but also through camera and lighting and music. Making every actor's personal lights a part of its own behavior system supports this premise, integrating expressive lighting design implicitly with other forms of expressivity for that actor.

Each personal light, like each global light, has a color, an intensity, and a position. In personal lights, color correlates to the emotion of that character. For example, red generally means angry, blue means sad, white means surprised. Other emotions are assigned to other colors depending on the lighting design of the world. Lighting intensity changes depending on how far the leading actor is from the camera, with full intensity when the camera is very close to that actor. A personal light's position is another channel for emotional expressivity – a light placed below the character's face tends to make that character look scary or angry, while one placed above the face makes a character look sad or kind.

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 themselves are up for grabs.

5 EVALUATION

While cinematography is meant to be essentially transparent to the end user, it is possible to determine how people are affected by it, and whether our system has succeeded. 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.

Emotional effects are a bit harder to see, but watching participants has allowed me to tell when emotional effects are working, versus when they are just distracting. For example, when a character is sad and the camera helps show this, people often laugh at the over-the-top performance being shown. (Why people laugh at sad characters is a topic for another thesis.)

We've had several main venues through which I've been able to watch participants with an eye to their interactions and reactions. When *Swamped!* premiered at SIGGRAPH 98, several hundred people from a range of backgrounds interacted with it. We had groups of animators, armies of programmers, an assortment of students, and a few small children. Our installation has been up-and-running at the Media Laboratory ever since; during that time, several hundred more people have had occasion to interact with it. This has been an even more mixed bunch than the SIGGRAPH crowd, with far more children and international participants. Finally, we'll have a chance to show our new project, *(void*)*, at SIGGRAPH 99. This will be another wonderful chance to evaluate this thesis work in a crowded setting.

Overall, this cinematography systems that I've created while researching this thesis have been well received. 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. Experiences like this, repeated many times during the several hundred times the CameraCreature has been demonstrated, are the basis for much of the interactivity section. From personal observation, I came to expect certain behavioral patterns from participants.

The development team found various parts of the cinematography system quite successful and helpful, and other parts more problematic. The button widget has been quite well received especially during the earliest stages of a project, when there was no automatic cinematography to rely on. Camera sounds were found to be fairly annoying for those people who were not using them to debug. Since most of the development team on a project work on small PC-based windows, perhaps three inches diagonally, there was sometimes a disparity between what the developers need and what would eventually be appropriate for the large installation screen. A system that looked good on the big screen made the characters feel like tiny ants to the developers. On the other hand, a system with close shots designed for tiny windows was terrifying and nauseating when used on the big screen. A future possibility would be to design a system that could sense what size display technology a person was using to view the system, and adapting the cinematography accordingly.

A hungry raccoon

There are several main ways that various stages of this cinematography system have broken during development. Simply reading this list might enable a future designer of cinematography systems to anticipate many of the problems that I have encountered in the last two years. Breaking is an important part of the development process; understanding how it broke made it possible to fix. So here's a list of the main ways that my system tended to break.

- The camera would choose a shot that had no character in it.
- The lights would be so bright that all subtlety in character design was lost.
- The camera would jitter as it tracked a character.
- The camera would oscillate as it came to rest.
- A character's personal light would shine on other pieces of geometry as that character walked by.
- The camera would be too far or too close.
- The camera would be inside a character or a piece of set.
- The camera would be blocked from seeing its target by a piece of set.
- The camera would try for a close-up, but the character would be gone by the time the camera got into position.
- The camera would jump-cut.
- The lights would change suddenly while the camera was still.
- The lights would oscillate in color-space.

While this list is not all-encompassing, it gives a feel for the kind of problems that came up during the creation of the CameraCreature.

In the process of designing this cinematography system, I found myself wishing for several bits of functionality, which would greatly increase the system's expressive range. First among these is shadows. Currently, every light illuminates any surface whose normals aim toward it. A graphics system that could render shadows would be a great boon for interactive lighting design. Also, being able to control what surfaces are illuminated by lights would be helpful. Having personal lights only shed light on that character would be a nice possibility. Currently, I had some difficulties keeping the throw of personal lights off other objects. Third, I'd have loved to have control over depth-of-field – the distance from the camera that is in sharp focus. Being able to focus on some things and let others go soft is a powerful way of directing an audience's attention. Despite these requests, I have found it to be wonderful to work with a team composed of artists and computer scientists who have provided such thorough supporting technologies that I have used in building this system.

5.1 Who are we designing for?

In general I've found that, among novice participants, age is inversely proportional to skill in interacting with our installation. While this is a broad generalization, it has become clear that a seven-year-old can steer our chicken within moments, while a middle-aged person takes far longer to master navigation. In the same vein, children are much quicker to assess when the camera is "broken".

This brings up an important point – who are we designing for? Should my cinematography system work well for adults, who have very strong preconceptions about what kinds of things appear on screens, and how it is socially appropriate to interact with them? (For many older people, the idea of interacting with a screen is thoroughly alien. Children, however, accept what they see much more readily.) Should I be designing for adults 20 years from now, who will be used to interacting with screens by then? Or should I design for children, who learn so rapidly that many more paradigms of interaction are viable. Finally, could it be possible to have an adaptive system that does both?

For example: it is possible to navigate by only seeing behind your character. It can be a sort of game – how can participants make forward progress when all they can see is the view behind their character. A child might enjoy such a game thoroughly. The process of navigating is the fun. But most adults, it seems, are more interested in having a quick and easy *process*, and want to "cut to the chase."

To satisfy a range of age groups, there must be something that is interesting to each person. The stuffed animal interface is a big winner among kids, so much so that the screen may become almost irrelevant. Many adults, though, seem to tolerate holding the chicken interface as if it were a punishment they have to endure on the way to understanding our research. (Thankfully, most adults at SIGGRAPH and the Media Lab are sufficiently enlightened that they feel comfortable around stuffed chickens!)

The issue being: is there one cinematography system that works equally well for everyone? Having worked on this thesis for two years, I feel the answer is no. 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 the best answer for broadly applicable interactive cinematography.

6 FUTURE WORK

This thesis has provoked several directions of continued research. These vary in scope and implementation time. I've split this section into two parts – short-term alterations and long-term ideas.

6.1 Short-Term Improvements

6.1.1 Work with Sound

It would greatly enhance an experience to have the cinematography cooperate with the sound design and composition. Therefore, I'd like to improve the CameraCreature to be able to cut to the beat of the music, and generally be aware of the scene's dynamic score. I implemented a few tests this way, and it is quite technically feasible. The same technology by which actors request shots could allow the music system (itself also a creature) to inhibit any changes until an appropriate beat occurs.

In addition, camera sounds, which I mentioned in the Development subsection of Cinematography, could have a place in the final installations. Simple whooshes make the camera seem much more alive, and contribute to its feeling more like a character in the interaction rather than just an invisible observer. This could be made to work with the music system as well, with the camera and the music working together to integrate camera sounds with the musical score.

6.1.2 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".

6.1.3 Multi-Character Composition

An important area of cinematography of which I've only scratched the surface is multi-character composition. Currently, most of my shots are framed around one character. (This is not to say that they only have one character in the frame; rather, the camera decision is made with respect to a single camera, and other actors are in the shot "accidentally".) Allowing the cinematography system to choose between shot framed around more than just one character will be an ever more powerful tool as this cinematography system continues to develop.

6.1.4 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.

6.1.5 Montage

Including some techniques for influencing shot choice based on montage theory would increase the cinematic potential of this system. 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.

6.1.6 Timing

Currently, the emotional state of the CameraCreature does not change the rate of turnover among actors and angles. Implementing this would add another channel of emotional expressivity. By applying emotional input to the gains and growth factors of the CameraCreature's motivations to choose various Actors and Angles, the emotional mood of the CameraCreature could come out in the rate of switching between shots.

6.1.7 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.

6.1.8 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.

6.2 Long-Term Ideas

6.2.1 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.

6.2.2 Adaptation to users

Three-dimensional visualizations are not a panacea for society's ills. Many applications are more easily understood without a perspective that zips all over the place. However, just as the web is embracing animation and graphics, I imagine that some descendent of the web will embrace the third dimension. And when that happens, cinematography will cease to be the realm of film makers and video game designers, and will become a more germane concern in a wide assortment of endeavors.

In that world, there are many directions that cinematography could go. It could follow the Netscape/Internet Explorer paradigm, where participants have a browser that is a fairly transparent window into which other people put content. Or it could go another direction, where people merely make their data structures available, and the participant's own, personalized cinematography system shows it to them in an appropriate fashion, given their interest, time constraints, preferences, and feedback. Most likely it will be some combination of the two, where participants have programs that are more personalized than the current browsers, that work out an agreement with the cinematography systems of the other organization about the best way to satisfy the desires and needs of the participant and the provider.

For example, considering the current trend toward advertising as the main funding model for the Web, the host site might say to the cinematography system, "Here's that 3D model of Mt. Everest your person wanted, but it'll cost you 10 seconds of quarter-screen viewing time." And the participant's computer will decide, "My person prefers airplanes, so every now and again, as he's interacting, I'll stick an airplane dragging a banner ad in the sky around Mt. Everest, and have the camera pan up to look at it." The cinematography system would have knowledge of you, and custom-design an interaction style that will get you the information you need in a way that you will enjoy. This would require an adaptive or trainable cinematography system, that could figure out how you like to see all the different things that you can see in 3D. But that is a topic for another thesis.

So, in the evaluation of this thesis, and its usefulness as a basis for future work, it is important to decide in what light it should be viewed. Is it a concrete example of how to do cinematography for a specific, game-like virtual environment? Or is it an exploration into the issues that make cinematography difficult and wonderful and complex in a range of applications? I hope that it is viewed as a hybrid of the two – a specific implementation that sheds light on the larger issues.

7 CONCLUSION

In this thesis, I have presented a specific cinematography system that I created, that balances interactivity and emotion in an interactive virtual environment. Throughout this document, I have tried to present this system's relevance to certain larger issues: what it means to be interactive, how emotion can be put on a screen, and which directions autonomous cinematography could go beyond this specific implementation. Cinematography and its descendent art forms will flourish with the upsurge in digital technology that appears to be marking the turn of this century. No longer can autonomous cinematography systems simply show the events that are happening; now it is time for them to show those events with style and sensitivity to emotion.

***(void*)* action shot**

By delving in depth into the process of building this cinematography system, and by developing it as an autonomous behavior-based creature, I hope to have provided some insight into the nature of the hybrid media that are arising from currently-existing art forms and technologies. My research focuses on emotion and interaction as the important elements in arranging the camera and lights in a scene. Emotions are central to human interactions. Interactions with people are the cause of most emotions. These two are closely linked. Media that play to both interaction and emotion will resonate doubly strong with humanity.

A world full of unpredictable characters demands intelligent cinematography. Autonomous actors choosing their behavior in real time cannot be put on a screen effectively by anything less than another autonomous behavior system.

My system advances the field of interactive cinematography by fully integrating an emotion system into our camera and lighting control system. Previous interactive cameras have dealt well with interactivity or emotion; this system is the first to address the two as equal contributors to the shot-selection process. The behavior-based approach is novel as well.

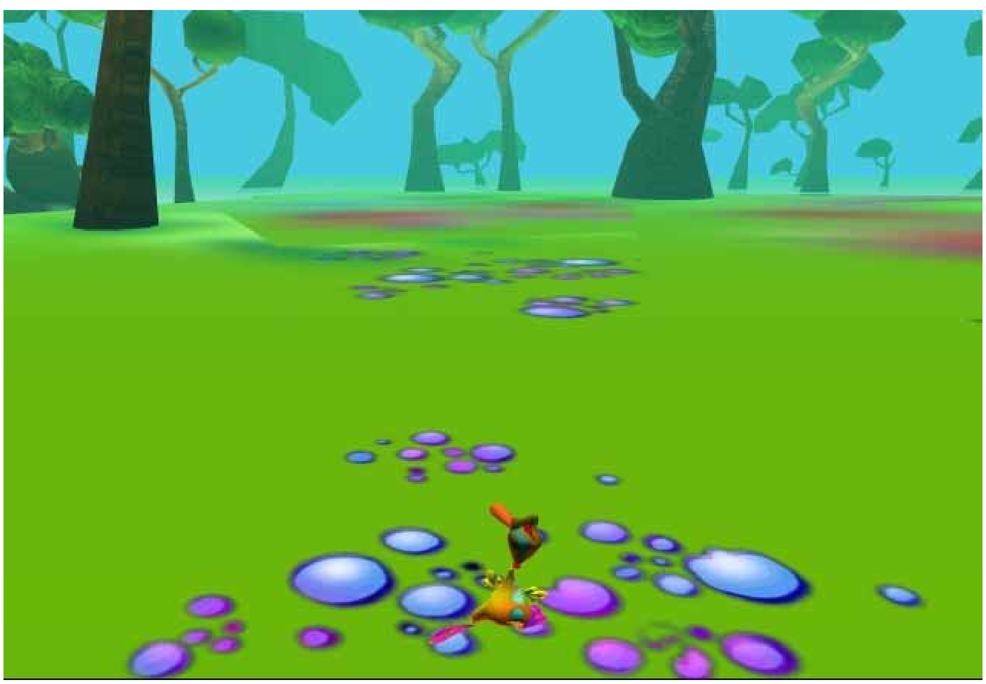
As advancing computing power opens the technological door ever wider onto three-dimensional virtual environments, cinematography must keep pace with those changes. As ever more interesting things happen in those worlds, by means of complexification of characters, improved rendering techniques, and advances in interactive storytelling, the art of cinematography will allow them to bring their wonders to the screen.

8 ACKNOWLEDGEMENTS

This thesis would not have been possible without the creative energy and support of a lot of people. Bruce Blumberg, my advisor, has been invaluable, tolerant, and encouraging. My two readers, Glorianna Davenport and Tom Porter, provided valuable suggestions and comments throughout the process.

I would like the thank the members and friends of the Synthetic Characters group at the MIT Media Laboratory, who created the virtual environments that my cinematography systems photograph. They are: Bruce Blumberg, Michael Patrick Johnson, Michal Hlavac, Christopher Kline, Kenneth Russell, Song-Yee Yoon, Marc Downie, Joseph Paradiso, Andrew Wilson, Teresa Marrin, Ari Benbasat, Jed Wahl, Zoe Teegarden, Dan Stiehl, Delphine Nain, Agnieszka Meyro and Julie Chasse. Special thanks to Delphine Nain, the undergraduate who has worked with me for the last few months to revise, improve and clean up the code in this cinematography system.

Thanks also to Kristin Wissig, who helped convince me to write this thesis. And of course, thanks as always to my family – Ma, Da, Lynn, Craig, Sam, Uncle Bill, and all four grandparents.



9 REFERENCES

Bares, W. H. and J. C. Lester (1997). "Realtime Generation of Customized 3D Animated Explanations for Knowledge-Based Learning Environments." Proceedings of the 14th AAAI National Conference on Artificial Intelligence.

Blumberg, B. (1996). "Old Tricks, New Dogs: Ethology and Interactive Creatures." Ph.D. Dissertation. MIT Media Laboratory.

Callahan, S. (1996). "Storytelling through Lighting." SIGGRAPH 96 Course Notes, Pixel Cinematography.

Christianson, D. B., S. E. Anderson, L. He, D. S. Weld, M. F. Cohen, D. H. Salesin (1996). "Declarative camera control for automatic cinematography." Proceedings of AAAI '96 (Portland, OR), pp.148-155.

Cruz-Neira, C., D. J. Sandin, T. A. DeFanti (1993). "Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE." ACM SIGGRAPH '93 Conference Proceedings, Anaheim, CA, July 1993.

Davenport, G., T. Aguierre Smith, N. Pincever (1991). "Cinematic Primitives for Multimedia." IEEE Computer Graphics and Animation special issue on multimedia, July, pp. 67-74.

Davenport, G. (1994). "Seeking Dynamic, Adaptive Story Environments." Visions and Views, IEEE Multimedia, Fall, Vol. 1, No. 3, pp. 9-13.

Drucker, S. M. (1994). "Intelligent Camera Control for Graphical Environments." Ph.D. Dissertation, MIT Media Laboratory.

Eisenstein, S. (1960). *Film Form and Film Sense,* edited and translated by Jay Leyda; Meridian Books, New York.

Galyean, T. (1995). "Narrative Guidance of Interactivity." Ph.D. Dissertation, MIT Media Laboratory.

Gianetti, L. (1993). *Understanding Movies: Sixth Edition*, Prentice Hall, Englewood Cliffs, New Jersey.

He, L., M. F. Cohen, D. H. Salesin (1996). "The virtual cinematographer: a paradigm for automatic real-time camera control and directing." Proceedings of SIGGRAPH 96, in Computer Graphics Proceedings, Annual Conference Series, 217-224, August 1996.

Johnson, M. P., A. Wilson, C. Kline, B. Blumberg and A. Bobick (1999). "Sympathetic Interfaces: Using a Plush Toy to Direct Synthetic Characters." Proceedings of CHI '99. To appear.

Kline, C. and B. Blumberg (1999). "The Art and Science of Synthetic Character Design." Proceedings of the AISB 1999 Symposium on AI and Creativity in Entertainment and Visual Art, Edinburgh, Scotland.

Malkiewicz, K. (1989). *Cinematography*, Simon & Schuster, New York.

Malkiewicz, K. (1986). *Film Lighting*, Simon & Schuster, New York.

Perlin, K. and A. Goldberg (1996). "Improv: A System for Scripting Interactive Actors in Virtual Worlds." Computer Graphics; Vol. 29 No. 3.

Picard, R. (1998). *Affective Computing*, MIT Press, Cambridge, MA.

Russell, K. B. and B. M. Blumberg. (1999). "Behavior friendly graphics." Proceedings of Computer Graphics International 99, Alberta, Canada, June 7-11, 1999. To appear.

Thomas, F. and O. Johnson. (1981). *The Illusion of Life*, Hyperion, New York.