An Autonomous Real-Time Camera Agent for Interactive Narratives and Games

Alexander Hornung[1], Gerhard Lakemeyer[2], and Georg Trogemann[1]

[1] Laboratory for Mixed Realities, Academy for Media Arts Cologne
Am Coloneum 1, 50829 Köln, Germany
hornung@lmr.khm.de, georg@khm.de

[2] Department of Computer Science V, Aachen University of Technology,
Ahornstr. 55, 52056 Aachen, Germany
gerhard@informatik.rwth-aachen.de

Abstract. Virtual reality environments provide the possibility to create interactive stories with the audience being an active part of the narrative. This paper presents our work on transferring cinematographic knowledge with respect to dramaturgical means of expressions of cameras to the domain of interactive narratives. Based on this formalisation we developed an autonomous real-time camera agent implementing this cinematographic knowledge with the goal of incorporating the camera as an active part into the storytelling process. The system was integrated into an interactive narrative environment to demonstrate the practicality of the system.

1 Introduction

The visual presentation of a narrative, like in movies, has a very strong influence on how we perceive and interpret a scene or situation of a story. These effects based on the camera position, image composition, choice of colours, etc., are well known from the theory of cinematography. But in contrast to classical cinematography, research in the field of interactive virtual reality narratives and the application of cinematographic concepts to this field is still rather young.

A variety of ideas for camera handling in 3D-based, virtual reality environments exist, ranging from first person views to complex, scripted camera movements. But as virtual environments evolve into a platform for interactive storytelling, these mostly geometrically oriented techniques fail to actively emphasise narrative content. Figure 1 shows two different examples of how the camera position actively contributes to the perception of an image.

We investigated classical cinematographic concepts applicable to this field, with a strong focus on dramaturgical principles of cameras and the visual emphasis of narrative content. The goal was to find a basic formalisation of the narrative expressiveness of cameras and cinematographic rules, and to implement

Fig. 1. Examples of visual emphasis of narrative content. a) A dark, low-angle camera shot emphasising the evil, terrifying nature of the vampire Nosferatu. b) A vast, calm landscape, with slow camera movement creating a lyrical, epic impression

this knowledge in form of a camera agent for interactive narrative applications. The agent autonomously chooses appropriate camera shots for a given situation within the current narrative context. The prototype of this virtual camera system was integrated into an interactive narrative environment to provide an example for the practicality of the system.

Possible target applications for an autonomous camera agent following cinematographic concepts are *Computer Games*: Current 3D-based computer games with a strong focus on narrative content rather than pure action, e.g., adventure games. *Storytelling Authoring Tools*: In tools like alVRed [Lab] for authoring interactive non-linear stories, an autonomous camera agent supports the author by preventing him from having to manually define camera positions. *Virtual Reality Environments*: VR-applications or e-learning systems supporting narrative content like a walk through a virtual museum. *Interactive Movies*: The capabilities of modern TV and DVD applications already point in the direction of interactive stories.

For further information and details about this work, and a playable demo of the camera agent see [Hor03].

2 Related Work

Publications dealing with models or implementations of virtual cameras are commonly based on [Ari76] and [Kat91]. In [BGL98], a system based on constraint satisfaction is described, [HCS96] creates a camera working with film idioms implemented as hierarchically organised finite state machines. Another approach for declarative camera control is presented in [CAH$^+$96]. [DZ95] presents a method of encapsulating camera tasks into well-defined camera modules, [HHS01] deals with a camera system designed for games, focusing on predictive camera planning and frame coherence. Most of these works deal primarily with geometrical

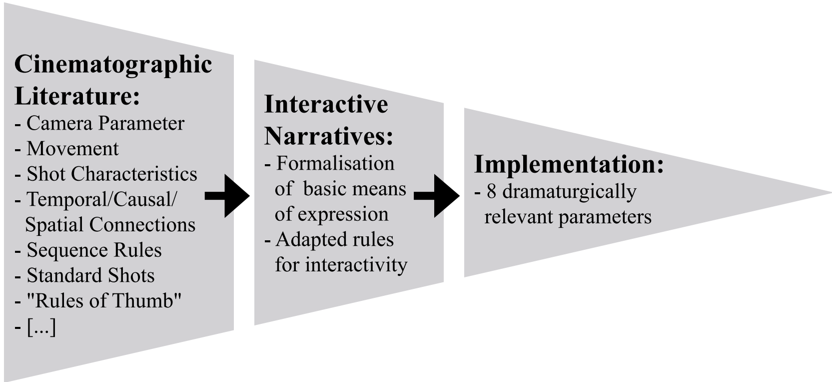

Fig. 2. Transfer and formalisation of cinematographic concepts to the domain of interactive narratives, resulting in an interface based on eight dramaturgically relevant parameters to communicate between a narrative application and the camera agent

constraint-satisfaction. [TBN] and [KM02] present cinematographic systems focusing with emotional content of digital scenes. The mathematical aspects of camera positioning are found in books like [AMH02], [Bli96].

3 Cinematographic Concepts

The inherent idea of visual storytelling is that the interpretation of a picture by the audience is based on a process of identification of the spectator with the camera standpoint and view (see Figure 1). This identification strongly influences interpretation of an image and must be considered during image creation.

Camera views depend on a large number of parameters. For example, the basic shot characterisations used in the cinematographic literature are based on the shot size, camera angle, camera movement, image composition, and shot duration. Furthermore, shots normally cannot be considered as single, atomic entities, but have always to be seen within the context of preceding and following shots. Within such sequences, temporal, causal, or spatial relationships are established, which significantly influence the narrative interpretation, and even can completely change the interpretation of a single shot [Kat91]. There exist a lot of cinematic rules to deal with such sequences, like the “Line of Action” rule, which forbids to cross the virtual line between two acting objects during a cut to preserve a consistent orientation for the viewer [Ari76] and [Kat91]. These rules of classical cinematography are often not (directly) transferable to interactive narratives, because future narrative events are usually not known until they occur. This makes narratively consistent planning ahead of shots very difficult.

We identified a set of basic dramaturgical principles, which describes the narrative expressiveness of cameras on a narrative level rather than on a camera level. Rules for camera placement were analysed and modified to meet the demands of interactive environments. The idea of this transfer of cinematographic concepts is depicted in Figure 2. Based on these principles we use eight parameters to describe dramaturgically relevant information in a story (Table 1). For continuous parameters $\in [-1.0 \ldots 1.0]$ like *Object Might*, positive values stand for a high power of the respective subject of the action (with respect to other objects for instance), zero for neutral values or equilibrium, and negative values for low power, or in general, the opposite meaning. These parameters are the basis to communicate story-events from the narrative application to the camera agent for proper visualisation.

Table 1. Eight parameters describing the dramaturgically relevant information of a narrative event, with respect to the expressiveness of cameras

Parameter	**Values**	**Description**
Given by the narrative application		
Scenery Variant	Action|Dialogue|Both	Classification of the event
Action Type	Physical|Mental|Predicate	Classification of the action
Radius of Interaction (RI)	$\in [-1.0 \ldots 1.0]$	Amount of space covered by the action
Object Might (OM)	$\in [-1.0 \ldots 1.0]$	Power of subject or relationship between objects
Emotional Involvement (EI)	$\in [-1.0 \ldots 1.0]$	Narrative climax, involvement of the spectator
Hectic / Dynamic (DY)	$\in [-1.0 \ldots 1.0]$	Hectic or calm, static or very dynamic situations
Excitement / Stress (EX)	$\in [-1.0 \ldots 1.0]$	Dominating inner emotion of characters
Computed by the agent		
Event Coherence	$\in [0.0 \ldots 1.0]$	Similarity of objects between subsequent events

4 Narrative Events

Our interface to communicate narrative information and to build a representation of the story within the camera agent are *narrative events*. Every narrative application can be assumed to have some form of representation of story events, be it in the form of explicit graphs like in [Lab], or as instantly occurring decisions or actions like in [VS98] for example. A narrative event encapsulates a single event of the story. The relevant narrative information about events and situations in a story can be formalised by a sentence-like structure consisting of the subject, the action, and the objects. Figure 3 gives an example of a narrative event. The only ‘carrier’ of dramaturgically relevant information is the

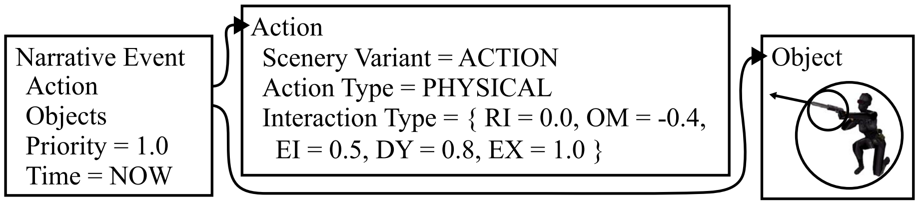

Fig. 3. An example of a narrative event, representing an attack of a wounded character. The interaction type describes an action with medium range, a low object might because of an injury, emotional involvement of the audience, high dynamics, and even higher stress of the fighting character

action itself. We characterise the action of a narrative event by the parameters representing the basic dramaturgical principles of cameras as introduced in Table 1. Objects like characters are described only by their respective geometrical information. This way we can keep out application specific knowledge from the camera agent. The coherence between subsequent events is actually computed by the agent, based on the similarity of objects. This knowledge enables the camera agent to detect connected sequences of events.

In addition, each event contains temporal information and a prioritisation assigned by the narrative application, so that the agent can build its own consistent representation of the story. The prioritisation and the coherence help to further distinguish important events of the story from unimportant ones, and to create visual continuity. In some forms of interactive narrative environments, like virtual reality systems based on independent agents inhabiting the world, there is seldom an explicit representation of the most important events, but the narrative emerges from interaction between these agents. In such cases it is necessary to enable the agent to distinguish between connected, story-driving events, and 'environmental' events. Knowledge about the coherence between events is also needed for cinematographic concepts like establishment shots to introduce new situations.

5 The Camera Agent

One of the strengths of agent-based software design is the construction of complex systems of interacting, autonomous modules. These approaches are more and more considered to model and implement autonomous characters or entities in virtual reality-based interactive narratives. Motivated by concepts from multiagent systems, we developed an agent-based software with the goal to provide intuitive software design and understanding of the decision procedures as well as an easy integration into agent-based, interactive narrative applications.

Figure 4 shows the overall system and the flow of information between the camera agent and the application, it is embedded in. The application sends relevant information about its current or future state as narrative events to the

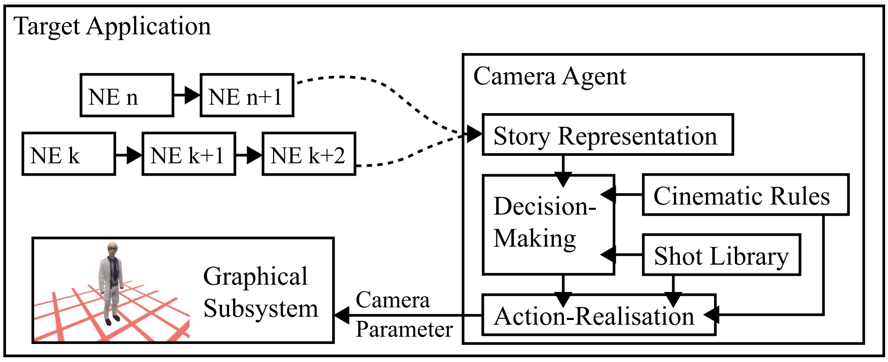

Fig. 4. System structure and flow of information

camera agent. These events contain geometrical information about the participating objects, and a description of the main action by the parameters found in Table 1. The camera agent then investigates and recomputes its internal representation of the narrative based on these events. It uses a hybrid method of rules, e.g., for the scenery type, and perceptron-based decision-making for continuous parameters to classify narrative events and to assign matching camera shots.

Based on the history of the narrative, the coherence between subsequent events, the event priority, and other factors like minimum shot durations, it chooses an active event for visualisation. The decision-module then adds potentially matching shots for this event from a user-defined shot library to a priority list. The shots in this list are ordered by a motivation value to apply this shot to the given event. This motivation is based on the activation of perceptrons associated with the different shots. The user can train different types of perceptrons with examples of narrative events to return a high activation only for specific actions. The inputs for each perceptron are the continuous parameters described in Table 1. For instance, we trained a perceptron to respond to a high event coherence and neutral interaction types, and associated this neuron with a close-up shots for dialogues (Figure 5). Using this approach, one does not have to hand-tune complex rules for action-classification. During the finding of matching shots, the decision-module ensures that cinematic rules like the 'Line of Action' are not violated.

In the final step, the module for action-realisation tries to carry out the shot with the highest priority for the chosen event. It computes the camera position, orientation, and other parameters like the field-of-view, and makes this information available to the narrative application for visualisation. If a shot cannot be realised due to geometrical constraints for instance, the next matching shot is chosen. Cinematic rules are considered also during this phase.

The user can easily add new shots to the shot library. Currently, we provide the possibility to describe shots based on spherical coordinates, and on the final on-screen position of the objects. Using these descriptions, we could easily imple-

a) b)

Fig. 5. A dialogue between four characters. a) Original first-person view. b) Camera agent: the spectator is guided through the conversation by appropriate camera shots.

ment all standard shots found in cinematic literature. We additionally provide key-frame animation for camera movements. Direct interdependencies between shots can be modelled by specifying positive or negative predecessor shot classes, similar to approaches based on state-machines like in [HCS96].

6 Experimental Results: Half-Life

The computer game *Half-Life* [VS98] was published in November 1998, and is considered to be one of the first games combining the elements of pure 3D action games with narrative elements, like dialogues, intermissions, and actions scenes, which are embedded into the game-play. We modified the software to generate a narrative event every time a character changes its internal state. For instance, if a character switches from an idle state to a conversation, it sends a corresponding event to the camera agent. The experimental evaluation was done by letting unprofessional audiences as well as professional cinematographers experience both views, the original first-person view of the game, and the shots created by the camera agent. Consensus was that the narrative, dramaturgical content of the game could be emphasised significantly by choosing cinematographically appropriate shots. For example Figure 5 compares the original first-person perspective during a dialogue to a sequence of camera shots created by the camera agent. In the first-person perspective, the player is not forced to concentrate on the dialogue and can miss potentially story-relevant parts of the conversation. By guiding the player through the conversation with appropriate shots of the camera agent, the player's focus can be manipulated to intensify the narrative experience. More examples for diverse situations can be see online at [Hor03].

7 Conclusion

We presented our work on transferring cinematographic concepts to interactive narratives, and an autonomous camera agent implementing this knowledge. The agent-based approach enabled us to integrate the system easily into an existing application. Our experimental results were very convincing and significantly enhanced the narrative experience of Half-Life for the spectator. Future research is necessary to understand and formalise more complex cinematic concepts, and to integrate planning techniques to allow for sophisticated reasoning about the visual outcome of camera shots to support the narrative. Reasoning about geometrical constraints within a scene is necessary as well as considering such complex concepts as 'visual metaphors'.

References

[AMH02] T. Akenine-Möller and E. Haines. *Real-Time Rendering, Second Edition*. A K Peters, 2002.

[Ari76] D. Arijon. *Grammar of the Film Language*. Focal Press, Boston, 1976.

[BGL98] W. H. Bares, J. P. Gregoire, and J. C. Lester. Realtime constraint-based cinematography for complex interactive 3d worlds. In *AAAI/IAAI*, pages 1101–1106, 1998.

[Bli96] J. Blinn. *Jim Blinn's Corner: A Trip Down the Graphics Pipeline*. Morgan Kaufmann, 1996.

[CAH+96] D. B. Christianson, S. E. Anderson, L. He, D. Salesin, D. S. Weld, and M. F. Cohen. Declarative camera control for automatic cinematography. In *AAAI/IAAI, Vol. 1*, pages 148–155, 1996.

[DZ95] S. M. Drucker and D. Zeltzer. Camdroid: A system for implementing intelligent camera control. In *Symposium on Interactive 3D Graphics*, pages 139–144, 1995.

[HCS96] L. He, M. F. Cohen, and D. H. Salesin. The virtual cinematographer: A paradigm for automatic real-time camera control and directing. *Computer Graphics*, 30(Annual Conference Series):217–224, 1996.

[HHS01] N. Halper, R. Helbing, and T. Strotthotte. A camera engine for computer games: Managing the trade-off between constraint satisfaction and frame coherence. In *Computer Graphics Forum: Proceedings EUROGRAPHICS 20(3)*, pages 174–183, 2001.

[Hor03] A. Hornung. Autonomous real-time camera agents in interactive narratives and games. Master's thesis, Department of Computer Science V, Aachen University of Technology, 2003. Download at http://www.cocoonpage.com/hlcam.

[Kat91] S. Katz. *Film Directing Shot By Shot: Visualizing From Concept To Screen*. Micheal Wiese Productions, 1991.

[KM02] K. Kennedy and R. E. Mercer. Planning animation cinematography and shot structure to communicate theme and mood. In *Proceedings of the 2nd international symposium on smart graphics*, pages 1–8. ACM Press, 2002.

[Lab] Laboratory for Mixed Realities, Cologne. alVRed. Nonlinear Dramaturgy in VR-Environments, http://www.alvred.de/.

[TBN] B. Tomlinson, B. Blumberg, and D. Nain. Expressive autonomous cinematography for interactive virtual environments. In *Proceedings of the Fourth International Conference on Autonomous Agents*, Barcelona, Catalonia, Spain. ACM Press.

[VS98] Valve-Software. Half-life. Computer Game used as Interactive Narrative Platform for the Camera Agent, http://www.sierra.com/games/half-life/, 1998.