Darshak – An Intelligent Cinematic Camera Planning System

Arnav Jhala

Liquid Narrative Group, Department of Computer Science
North Carolina State University
890 Oval Dr, Raleigh, NC - 27606
ahjhala@ncsu.edu

Abstract

A virtual camera is a powerful communicative tool in virtual environments. It is a window through which a viewer perceives the virtual world. For virtual environments with an underlying narrative component, there is a need for automated camera planning systems that account for the situational parameters of the interaction and not just the graphical arrangement of the virtual world. I propose a camera planning system called Darshak that takes as input a story in the form of sequence of events and generates a sequence of camera actions based on cinematic idioms. The camera actions, when executed in the virtual environment update a list of geometric constraints on the camera. A constraint solver then places the camera based on these constraints.

Introduction and Related Work

A 3D narrative-based system (e.g. video games and training simulations) must not only create engaging story-world plans, it must use its media resources to tell the story effectively. In the work that I describe here, I will focus on one aspect of the effective creation of cinematic discourse: automatically determining the content and organization of a sequence of camera shots that film the action unfolding within a story world. Strategies for determining shot content in 3D virtual environments fall into one of three categories. Several research systems use camera constraints that are pre-specified relative to the subjects being viewed (Bares & Lester 1997). These approaches are of limited value in dynamic domains where constraints need to be dynamically generated. Other approaches, like many commercial computer games, provide dynamic camera positioning based on the viewpoint of the user's character. However, these approaches limit the information about that story world that is conveyed to the user to just those elements of it that the user chooses to view.

I propose a technique that falls along a third line of research – that of the automatic determination of shot composition based on the dynamics of the scene and the story events occurring within it. The ability of my narrative system to generate camera control directives automatically and dynamically is important for two reasons. First, automatic generation of camera shots can ensure that those objects in the environment that should be visible in shots are not obscured by scene geometry. Second, and more central to the current discussion, automatic composition of shots allows the system to select shot sequences that exploit cinematic knowledge relating shots to the unfolding action in order to more effectively communicate aspects of the plot. Cinematographers have identified patterns of shot sequences that define stereotypes for ways to film certain types of action or action sequences (Arijon 1976). These stereotypes are called film *idioms*; their use is central to the creation of a cinematic experience for a virtual world's user.

Our Approach

Generation of camera shots for conveying a story can be seen as *planned intentional communication* from the director and cinematographer to the audience. This approach parallels both the film production process and the natural language discourse generation process (Jhala 2004). I see three main requirements for generation of coherent cinematic discourse. At an abstract level, the director extracts the *salient elements* of the discourse from the given events in the virtual world. These events are then *organized* into a rhetorical structure for the coherent telling of the story. Finally, camera shots are chosen that set up *constraints* on the camera to satisfy certain film idioms during the geometric placement of the camera. These requirements only differ from natural language discourse generation at the realization level where syntactic constraints on sentences in natural language are replaced by geometric constraints on camera shots. Another difference in generation of cinematic discourse is that camera shots need to be synchronized with actions/events happening in the virtual world. Duration of shots also affects viewer's perception of the context in terms of parameters like the *tempo/pacing* of the narrative. Thus a camera planning system needs support of durative actions and temporal reasoning.

For satisfying the aforementioned requirements, I have adopted the representation of film idioms as hierarchical

Copyright © 2006, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

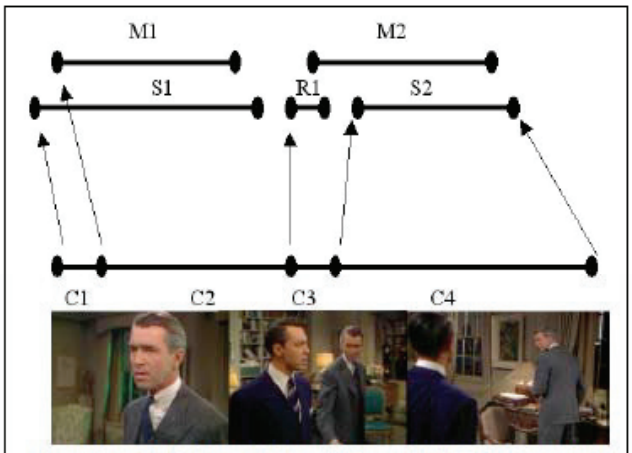

Figure 1: Relationship between camera actions and story world actions. The arrows indicate the start times and end-times of camera actions.

plan operators. I have extended an existing discourse planning algorithm (Young et. al. 1994) to incorporate temporally qualified expressions (Ghallab et. al. 2004) as constraints. I use interval temporal relations described by Allen (1983) for specifying constraints on execution of actions as well as for representing temporal links between the execution of actions and events in the story and the camera shots.

Implementation

My action representation is based on the formalism developed by Young et. al. (1994). I have extended the planning problem to include additional temporal constraints on the execution of actions. Details of the representation can be found by Jhala et. al. (2005).

I have modified the DPOCL algorithm (Young et. al. 1994) with the following extensions based on our representation of actions mentioned above.

Causal Planning: Causal Planning takes place in the planner as described in (Young et. al. 1994), with temporally quantified conditions. A precondition *p@[t)* is required to be true at the start of the action (t=tstart). Thus if effects of an action A1 satisfies the precondition of another action A2, then a temporal constraint of type is added for $tend_{A1} < tstart_{A2}$ such that no $p@[tend_{A1}, tstart_{A2})$.

Temporal Planning: In our implementation, each camera action C_i is associated with a (starts-at C_i ?t_p) , and a (ends-at C_i ?t_p) marker where t_p is a temporal marker to a story world action (of the form (during ?s)). These temporal markers are added to relate the camera actions with the corresponding story actions that they film during execution. For instance, in Figure 1, each camera shot C is temporally marked for execution with the respective speech act S.

Temporal Consistency Checking: After a causally consistent camera step is added to the plan, the temporal consistency module checks the start and end times of the newly added camera steps against the temporal constraints on a constraint list. A Temporal Flaw is added for the steps that do not satisfy the constraints. A temporal flaw is handled by adding additional temporal constraints or binding constraints on variables.

Discussion

I am working towards the following goals through this work.

1. Formalizing film idioms as plan operators to utilize planning algorithms for automated generation of camera shots.
2. Identifying requirements for cinematic discourse actions and extending the existing work in discourse generation in natural language to a new medium.
3. Implementing semi-automated tools for story-authors and directors to help generate branching narratives and for pre-visualization.

I have made progress in 1. and am evaluating the systems that have been developed based on 2 above. I am also currently in the process of developing the applications mentioned in 3.

References

J. F. Allen. *Maintaining knowledge about temporal intervals*. Communications of the ACM, 26(11), 1983.

Bares William, Lester James, *Cinematographic User Models for Automated Realtime Camera Control in Dynamic 3D Environments*, UM, 1997.

Christianson David, Anderson Sean, He Li-wei, Salesin David, Weld Daniel, Cohen Michael, *Declarative Camera Control for Automatic Cinematography*, Proceedings of AAAI, 1996.

Jhala, Arnav, Harish *Intelligent cinematic camera planning system for dynamic narratives*. Masters' Thesis, North Carolina State University, 2004.

Jhala, Arnav, Young R M. *Discourse Planning Approach to Cinematic Camera Control for Virtual Environments,* Proceedings of the 20th AAAI, Pittsburgh, PA, 2005.

Ghallab M, Nau D, Traverso P. *Automated Planning: Theory and Practice.* Morgan Kaufmann, May 2004.

Young R M., Moore, J. *DPOCL: A Principled Approach To Discourse Planning*, INLG workshop, ME, 1994.