Interaction and Expressivity in Video Games: Harnessing the Rhetoric of Film

Laurent Cozic, Stephen Boyd Davis, Huw Jones

Lansdown Centre for Electronic Arts, Middlesex University, Barnet EN4 8HT, UK
<l.cozic, s.boyd-davis, d.h.jones>@mdx.ac.uk

Abstract. The film-maker uses the camera and editing creatively, not simply to present the action of the film but also to set up a particular relation between the action and the viewer. In 3D video games with action controlled by the player, the pseudo camera is usually less creatively controlled and has less effect on the player's appreciation of and engagement with the game. This paper discusses methods of controlling games by easy and intuitive interfaces and use of an automated virtual camera to increase the appeal of games for users.

1 Introduction

The film-maker benefits from more than a hundred years' development of expressivity in the medium, using viewpoint, editing and other means to affect the way the action of the story is perceived. Video Games designers have attempted to imitate some of these devices, but several difficulties arise in trying to adapt the language of film to an interactive context. It might seem that the best solution for games is simply to show the action as simply and straightforwardly as possible; we therefore discuss the benefits that a more inventive approach potentially confers. Our interest particularly is in seeing how far the demands of interactivity and the expressive possibilities of film-making can be reconciled in a game. This work is relevant not only to video games design but also to the visual presentation of virtual environments and to the automated construction of extended CGI sequences.

Bares *et al.* [2] describe a constraint-based approach to camera planning. The camera settings are generated from a set of visual constraints on the objects of a scene in the same way a camera operator move the camera according to the constraints of the story board. Drucker and Zeltzer encapsulate camera tasks into "camera modules" that "can be used as the underlying framework for controlling the virtual camera in widely disparate type of graphical environment". [6]

He *et al.* use "idioms" to describe the behaviour of the camera for each type of scene. The rules of cinematography "are codified as hierarchical finite state machine" which "controls camera placements and shot transitions automatically". [9]

Halper *et al.* discuss the importance of having a good balance between constraint satisfaction and frame coherence in order to have "smooth transition and appropriate cuts". They also propose a new method to deal efficiently with occlusion based on projective shadow casting algorithm. [7]

De Loor *et al.* [10] describe a virtual camera that has many features controlled by human interaction but with some drawbacks when it is set into automatic activity.

In this paper, we discuss our approach to implementing more cinematic features and 'seamless' interaction in a system for 3D video games. In section 2, we explain what we mean by expressive cinematography and outline the aspects of camerawork that have implication in game design. In section 3 we review the different type of third-person viewpoints in adventure and action games. Section 4 discusses briefly a game engine, devised to illustrate the features discussed elsewhere through creation of a prototype game, 'The Intruder', written in an OOP scripting language. We sum up in section 5.

2 Cinematographic Concepts in Games

Two aspects of camerawork in film-making have implications for the deployment of virtual cameras in a video game: the target of a given shot and the attributes of the shot such as angle and

framing. The techniques discussed in this paper deal with aspects of the first and second issues, which are closely interconnected, and the extent to which practices derived from film-making may be useful in games design. We also touch on the third issue of the relationship between shots, in our discussion of camera selection.

2.1 Information and Affect

Film-making works at two levels. It provides visual (and aural) information, and it does so in such a way as to impact psychologically on the viewer. For these reasons it has been called a form of *rhetoric*, a mode which aims not just to inform but also to influence and persuade [8]. More than a hundred years of film-making have seen the development of a mature 'language' for this purpose. In the design of video games, the designer similarly needs not just to inform the player about the game events but also to afford some particular perception of them. Different games may require different approaches but a first step is the investigation of techniques which, once developed, might be customised to the requirements of particular game genres.

In the simplest model of film-making, the camera needs to be pointed at the place where the action occurs. Since actions may occur in several places, some rule needs to determine what should be shown. The depiction will often need to show the context of an action, which may be provided by adjacent shots, or it may be provided within the shot itself, for example by a medium or long shot. As described so far, the camera might be thought of as depicting a pre-existing drama. This is the position taken for example by He *et al.* [9] in devising an automated cinematography system for use with virtual environments. They aim to ensure that the significant event is always in view. This is the simple principle of what we have called the *optimal view* [3]. But the camera has a crucial role not just in presenting, but also in articulating the drama. Armes [1] remarks that the slippage between the events and the depiction has in itself significant interest for the viewer. This slippage may take the extreme form of denying viewers the view that they feel they need, a technique dating back at least to *The Birth of a Nation* [15] [12]. Or in framing a shot a viewpoint may be chosen which partly conceals the object of interest. In *Rosemary's Baby* [16], Ruth goes to use the phone in an adjoining room. Polanski uses the doorway as a frame for this action, so that Ruth's face is concealed from view – the viewer yearns (and expects) to see her expressions, but cannot. The resulting frustration leads to intense engagement with the action. Such examples suggest that there are limits to the 'optimal view' approach for virtual cinematography. Significant benefits may arise from choosing the target more subtly and framing the shot in other ways.

2.2 Shot Selection, Virtual Cinematography

All we have said up to now concerns continuous camerawork. The other aspect potentially useful to the games designer is shot selection. Whereas in film-making this is divided into two phases of activity, the capturing of action to film and the subsequent temporal organisation of that material by editing, in the live environment of the game we might think of this as a single process. It is in some ways therefore more akin to the selection of cameras in televising a live event. However both the requirements and the opportunities differ. The requirements differ for the reason already set out above: the aim is not always, as it might be in televising a football game, to choose cameras merely to achieve the 'optimal view' – the one which affords the clearest picture of the event – but also to affect how the events are perceived. The opportunities also differ, since the game designer has greater foreknowledge of what will occur under a given set of circumstances and of its significance in the development of the 'story'. Cameras can be positioned with a view to affecting the relationship between the viewer and the events.

3 Third Person Views in Video Games

Two of the most widespread third person view games are the adventure and the action game. By adventure game we mean those based on the resolution of puzzles and on an elaborated plot. The possible actions and movements of the player's avatar are in general very limited as this kind of game mostly expects reflection from the player. Typical examples include the *Resident Evil* [21] or the *Monkey Island* [20] series. Action games by contrast appeal to the player's dexterity and

reflexes. The main character has many more possible movements. The plot is poor or nonexistent as it is not the main concern of the game. Typical examples include *Super Mario Sunshine* [24] or *Tomb Raider* [25]. Finally we will review the way games that mix these two genres handle the camera.

3.1 Adventure Games

Most of the camera systems of these games track the character using static cameras with a position and an orientation set during the making of the game. The main advantage of this technique is that it allows precisely describing each scene by showing the player what he is supposed to see under an angle that is coherent with the story. For example, in *Resident Evil 2* [22], the first encounter with one of the creatures is prepared by a series of carefully chosen shots that aim at creating suspense. First, as the player moves towards the place where the creature is located, a shot shows a window through which the player can see a form that appears briefly. The next shot shows the character through a window in a subjective view. The wooden window – which can visibly be easily broken – is at a few centimetres from the character; thus making it very vulnerable. The shot just before the encounter shows blood drops falling from the ceiling in the foreground. This example shows that predefined viewpoints can be used to create a dramatic impact in games without having to use cut scenes.

However it is often difficult with this technique to make views that are both functional and dramatically relevant. A viewpoint should not hinder the player's actions because, as he does not control the camera, he will not be able to choose a better one. The risk is of making the game artificially difficult by masking a hazard he should have seen. Thus, the narrow shots of *Resident Evil* [21] sometimes do not show the player the enemies that can nevertheless be very close; thus forcing the player to shoot randomly in the direction of the footsteps.

Another problem with this type of display is that it may be difficult to understand the link between successive shots. Thus in *Fear Effect 2* [17] the game cuts from shot A to shot B without any visual link between shots except for the presence of the central character. This can give the strange feeling of arriving suddenly in a completely new place although the character is still in the same room.

3.2 Action Games

The most commonly used camera in this type of game is the tracking camera. As well as the default following mode, in some systems the player can freely zoom or rotate to choose the most useful view. The level of control depends directly on the level of precision required by the game play. In *Super Mario 64* [23], the player has four buttons to control the camera, to allow precise positioning. The game can indeed require great precision to avoid obstacles or to jump from one platform to another. In comparison, only one button is used to control the camera in *The Legend of Zelda: Ocarina of Time* [18]: it resets its position so that it points in the same direction as the character.

Although giving a lot of freedom to the player, the free camera has the disadvantage of being difficult to control in narrow spaces. It tends to "knock" against the walls or to be obstructed by objects of the scene and, finally, it does not allow seeing the scene with as much efficiency as with predefined views.

Finally, as the camera is only driven by the player's movement and never by the plot or by the characters' emotions, this type of camera system only provides dramatically neutral viewpoints. Thus in *Ocarina of Time* [18], the fight between the main character and his worst enemy will be "filmed" in exactly the same way as any other common fight.

3.3 Mixing the Genres

Games that mix action and adventure generally have two alternatives: whether they use a camera system with several modes − one for each type of situation, or they leave the player complete control over the camera.

In *Ocarina of Time* [18], the player has to explore the virtual world, talk with characters and collect various objects to resolve puzzles; but to progress in the game he also has to defeat enemies

using swords, bows or other weapons. These different phases of the game each use a different camera style. When the player is exploring the houses or the town, the views are predefined and cannot be changed by the player. However, when the player is on the outside world or in a castle infested with enemies, the game switches to a tracking camera. As this example shows, mixing genres can simply mean mixing the corresponding camera styles.

Another solution found in some games is to give the player complete control over the six degrees of freedom of the camera as it has been done recently in *The Legend of Zelda: The Wind Waker* [19]. This could be considered as a universal camera system as it could theoretically fit any situation. Indeed the player can always translate and rotate the camera to find the 'optimal view'. However, we think this solution should be avoided because it forces the player to consider too many parameters at the same time which can lead him to lose concentration on the game-play itself. Besides the task of controlling the camera is arguably not as entertaining as controlling the character. Steven Poole remarks that "controlling the camera should not be the player's job" [11] but it should be the developer's goal to devise a fully automated camera system.

4 The Intruder's Camera System

In this section we briefly discuss features of our camera system used by our prototype game 'The Intruder'. Greater detail of implementation will form the core of a later paper to be submitted for publication. [4]

4.1 Shot Selection

Each scene comprises several cameras with a predefined position. The game switches from one to another according to the position of the character. When the character is occluded by an object for more than a second, the system attempts to find a camera under which the character is visible. If it is visible under several cameras the system will select one under which the character is not too big or too small on screen.

Continuity is handled by positioning the camera in such a way that it is not possible to cross the line of action. Each shot is also selected so that it shares some visual features with its previous and successive shots. This helps the player to make sense of the virtual world easily.

4.2 Tracking Modes

Each camera can track a character in three different modes: #simple, #anticipate and #watcher. In the #simple mode (cf. Fig. 1.) the camera is locked on the character and re-frames without delay as soon as it moves. This view shows the player too much useless information as it uses half of the screen to show the player what he has already seen, i.e. the space behind the character.

The second mode (cf. Fig. 1.), commonly used in films, anticipates the motion of the character so there is more space in front than behind [14]. In this respect this view provides more useful information to the player than the previous mode. As the character is on the left or the right of the screen, the player's centre of interest is also more likely to be on the centre of the screen. Therefore by showing the player what he wants to see and by getting rid as much as possible of the useless information, we achieve an 'optimal view' as described previously. It is also the most playable view as it allows the player to anticipate his movements. However this view can be dramatically weak: by showing the player everything he wants to see, potential dramatic effects such as those described previously are lost. It is to deal with these issues that a last mode has been implemented.

The #watcher mode (cf. Fig. 1.) makes the camera follow the character as if seen by an external observer who cannot anticipate character movements, so adjusts the shot with a small delay. The fact that it leads to more space seen behind the character than in front can be used to support the dramatic purpose of a scene. For example we can imagine a scene where the player just heard a call for help coming from somewhere near his character. He decides to run towards this scream. Here by activating the #watcher mode, the camera will very slowly follow the character in such a way that the player will see what he is looking for only when he will be right in front of it. By playing on the contrast between the emergency of the situation and the slowness of the camera, some suspense could be achieved. This view, as its name shows, can also be used to simulate an

external observer: for example a journalist followed by his cameraman. In The Intruder, which unfolds in a kind of ghost house, this view can also enhance the narrative by being used to draw the player's intention to the invisible presence that observes the character.

The different character of the latter two modes is crucial to the distinction we wish to make. The #anticipate mode provides the so called 'optimal view' – that view which provides the best information – but the #watcher mode offers the more expressive view, chosen for its rhetorical effectiveness. The more expressive view, as we have implemented it, does not conflict with the demands of usability.

Fig. 1. Examples of shots generated by the #simple (left), #anticipate (middle) and #watcher (right) modes.

5 Conclusions

We have reviewed different ways of 'filming' a character for dramatic effect. We consider a free camera under control of the player to be dramatically poor in most circumstances, though such a system can be suitable for action games. Predefined viewpoints can be used to add more meaning to a given scene and are seen as dramatically more interesting. However, they can create usability problems because they do not always show the player what he or she wants to see. In particular in this paper, we have discussed what is needed for an autonomous virtual camera system to generate more expressive shots.

We do not believe that videogames should become an ersatz cinema. Boyd Davis and Jones say, 'it is not a matter of simply transferring existing spatial practices to the newer medium … There is a need to rethink inherited pictorial and spatial practices to suit the demands made on the artefact. It is useful to remember ways in which digital interactive media are unlike their antecedents'. [3]

We have presented work in progress that will continue for at least a further year with the aim of producing an engine that will support games that are intuitive to play, engaging and dramatically illustrated. Davenport enunciates our aims, 'Gradually, these smart tools will gain the expertise and knowledge of their human masters, and their utility (and complexity) will increase.' [5]

References

[1]   Armes, R., 1994, Action and Image: dramatic structure in cinema, Manchester University Press, Manchester, UK

[2]   Bares, W. H., Thainimit, S., and McDermott S. A model for constraint-based camera planning. In Smart Graphics. Papers from the 2000 AAAI Spring Symposium (Stanford, March 20--22, 2000), pages 84--91, Menlo Park, 2000. AAAI Press.

[3]   Boyd Davis, S. and Jones, H., Screen Space: Depiction and the Space of Interactive Media, in J.A. Jorge, N.M. Correia, H. Jones and M.B. Kannegai (eds.), Multimedia 2001, Springer, Vienna, 2002, pp. 165-176.

[4]   Cozic, L., The Intruder: Expressive Cinematography in Videogames, Working paper, Lansdown Centre for Electronic Arts, Middlesex University, 2003.

[5]   Davenport, G., "Smarter Tools for Storytelling: Are They Just Around the Corner?", IEEE Multimedia, 3(1), 1996, pp. 10-14.

[6]   Drucker, S. M., and Zeltzer, D. CamDroid: A system for implementing intelligent camera control. In 1995 Symposium on Interactive 3D Graphics, pages 139--144, Apr. 1995.

[7]   Halper, N., Helbing, R., Strothotte, T., Computer games: A camera engine for computer games, Computer Graphics Forum 20 (2001)

[8]   Harrington, J., The Rhetoric of Film, Holt, Rinehart and Winston, New York, 1973.

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

[10] de Loor, P., Favier, P.A., and Tisseau, J , Programming Autonomous Entities with Purposes and Trends for Virtual Storytelling, International Conference of Virtual Storytelling, Avignon, Sept 2001, pp. 40-43.

[11] Poole, S., Trigger Happy, Fourth Estate, London, 2000.

[12] Reisz, K., and Millar, G., The Technique of Film Editing (2 ed.), Focal Press, 1982.

[13] Thompson, R., Grammar of the Edit, Focal Press, 1998.

[14] Thompson, R., Grammar of the Shot, Focal Press, 1998

Computer Games and Films

[15] The Birth of a Nation, D.W. Griffith, 1915.

[16] Rosemary's Baby, R. Polanski, 1968.

[17] Fear Effect 2, Kronos Digital Entertainment, 2001.

[18] The Legend of Zelda: Ocarina of Time, Nintendo, 1998.

[19] The Legend of Zelda: The Wind Waker, Nintendo, 2002.

[20] Monkey Island, LucasArts, 1990.

[21] Resident Evil, Capcom, 1996.

[22] Resident Evil 2, Capcom, 1997.

[23] Super Mario 64, Nintendo, 1996.

[24] Super Mario Sunshine, Nintendo, 2002.

[25] Tomb Raider, Core Design Ltd., 1996.