

Film and the Development of Interactive Narrative

Andy Clarke, Grethe Mitchell

Kinonet, 78 West Kensington Court, Edith Villas, London W14 9AB

andy@kinonet.com

Dept. Of Innovation Studies, University of East London,
Docklands Campus, 4-6 University Way, London E16 2RD

grethe@kinonet.com, grethe@uel.ac.uk

Abstract. This paper explores narration in film and in videogames/virtual environments/interactive narratives. Particular attention is given to their use of the continuity of time, space and action and this is used as a means of classifying different types of work. The authors argue that the the creators of these videogames etc. need to have more authorial presence and that this can only be done through abandoning their traditional reliance on the continuity of time, space and action.

1 Introduction

There is a clear crossover between film and interactive computer-based entertainment. This crossover occurs in a variety of forms and at a number of levels, and is appropriate given film's central role as a storytelling medium and the clear formal similarities that exist between films and computer-based media - the main ones being that they are both screen-based, they are both time-based and they both convey most of their information visually.

If we take videogames as an example, we can see that they borrow the established conventions and iconography of film. The guns in Unreal Tournament look like those in Aliens because doing so provides the designers of the game with a shorthand to describing the characteristics of each weapon. Similarly, the lighting, camera angles, and music of Resident Evil are like those in George A. Romero's Living Dead series of movies because drawing upon the conventions of the horror movie genre (and this subgenre) provides the game with a shortcut to creating a sinister atmosphere.

But what is more significant than this is the way in which a game such as Soul Blade uses its "virtual camera" - tracking and zooming to follow the fighters, and slowly tracking back and up to give the impression of the life draining from your character's body when you die - or the way that Resident Evil uses editing. In both of these examples, the videogame is using the language of film, but in a way that is subtly different to how it is used in film.

In other papers, we have looked in detail at various formal aspects of the videogame, with a particular interest in the relationship between viewpoint and identification, and between immersion and narrative in videogames[1, 2]. Although

there is some benefit in choosing another formal aspect of film and virtual environments to explore here or to examine in fine detail the way in which a single one of these environments presents its narrative, both of these tasks are probably best left to a book-length examination of the subject. Instead, we intend to pursue another option in this paper, to draw these - and other - threads together to explore a single theme: narration.

Narration is something very distinct from narrative. The narrative is the story: “what happened”. The narration, on the other hand, is the storytelling. As we will see in the next section, there are elements of film narration that are common to all films, no matter what they are about (though each film may use these narrative techniques in a different way, possibly using some not at all). Are there a similar set of narrative techniques in interactive narratives, and if so, what are they? Are they complete, unified and coherent, and if not, where are the lacunae and what do these absences and omissions tell us about the nature of the medium and/or the stories that people choose to tell in it? Are there better ways to present interactive narratives or to present narratives in interactive environments?

It is worth pointing out that for reasons we explain later in the paper, we make little or no distinction between the various types of interactive narratives, videogames, arcade games, immersive environments, and so on. As a result, we will use these terms somewhat interchangeably, and our examples will be drawn mainly from the world of videogames as these are the most widely available, should the reader wish to view the work for themselves.

2 Narration in Film

Film has, over the years developed a wide range of formal techniques (formal in the sense of dealing with film form, rather than in the sense of being a precise, coherent, complete or logical system). These conventions are so familiar to us that they become “invisible” - they are helped in this respect by the fact that film normally engages us so strongly at an emotional level, as a narrative, that it is difficult to identify at the same time the techniques by which it is presenting this narrative or having this emotional effect.

It is useful therefore to list the most important of these formal components (particularly for the benefit of those who have not studied film theory).¹ They can be summarised as follows: the *mise en scene* (the choice of actors, costumes, props and setting; the blocking of the scene; the use of special effects; etc.); the shot (the choice of camera position, focal length, film stock, aspect ratio and framing; the use of a moving camera or zoom lens; the choice of lighting style, setup and level; etc.); the juxtaposition of shots (the speed and style of editing; the use - or non-use - of continuity editing; the use of the long take and/or a moving camera as an alternative to editing; the use of optical effects such as dissolves between shots; etc.); sound (the use of sound, sound effects, music and silence; etc.).

¹ Good introductory texts are Film Art: An Introduction [3] and The Cinema as Art [4].

This list is not meant to be complete, definitive or exhaustive, but is intended merely to give some indication of the type of elements involved and highlight some of the more important. There is clearly a great deal of overlap between categories and ongoing discussion about how to divide even the partial list of properties above between them - one could, for instance, easily class lighting as part of the *mise en scene*, rather than as part of the shot as the lighting is often motivated by the choice of location and the lights which are present in the shot. The use of a moving camera or zoom lens is also interesting in that it is normally the use of this within a long take that replaces editing, rather than the long take per se.

One can likewise easily add to this list of categories or create subtle distinctions within them. One example of this is within sound - between sound that comes from things which are (or could be) present in the scene (diegetic sound) and sound that is not motivated in this way (non-diegetic). Even this can be further subdivided between diegetic sound from onscreen people or things, and that from offscreen ones.²

But these formal elements of film cannot work without the filmmaker and their audience having an agreed understanding of what they “mean”. There is an infinite number of ways that the filmmaker can choose to take a shot (different camera positions, focal length, etc.) but these choices are meaningless unless there are conventions that give meaning to the various possibilities - i.e. that says that a wide shot from high up looking down on a figure means one thing, and a low tight shot looking up at the same figure means something else.

In fact, it is not as clear cut as this. Each of these shots can have a whole range of different meanings according to its context (both on the basis of its position within the narrative and depending on what shots precede and follow it). Indeed, the filmmaker can create new meanings through the context of the shot and these can be incorporated into the evolving body of film grammar.

What all of the above should indicate is that when we use terms such as “film language”, “the conventions of film” and “film grammar”, we are not talking about a single, fixed, consistent, coherent set of rules that can be written down - indeed, attempts such as Metz’s *Grande Syntagmatique*³ have failed to define an internally consistent set of rules of film grammar even within a single film.

It should therefore be obvious that when we talk about applying the lessons learnt from film theory and practice to the creation of virtual worlds and interactive narratives, we are not talking about distilling the conventions of film into a neat set of axioms and formulae which then can be programmed into the computer - neither films nor virtual environments represent bodies of work which are consistent and coherent enough in terms of form, content, aims, resources or techniques to allow this.

But in spite of this, we still regard the theories and techniques of film as being potentially far more useful than those of other media. There are a number of reasons

² This is further complicated by the fact that the same sound may be diegetic in one shot and non-diegetic in the next.

³ See Stam, Burgoyne and Flitterman-Lewis [5] for an overview of Metz’s theories and their weaknesses.

for this, but the key one is that they can often be applied both at a “macro” level and at a “micro” level. To clarify what we mean by this, compare, for example, Propp’s theories of character and plot structure [6] to those of character, genre and *mise en scene* in the Western. Propp’s theories describe the role that the villain has in the narrative, but offer very little assistance in how to implement them in an interactive narrative. There is no advice as to how this villain should look, how he behaves (beyond the broadest description), and so on.

Film theory offers far more concrete guidance in this respect. The conventions of genre sketch out the role that the villain has in the narrative (as in Propp), but there is a whole host of other advice that it (and other film theories) can give us to “flesh out” this character. They will tell us that the villain will wear a black hat and will be badly shaven. In addition, it will provide a set of generic locations for the story to take place around (jail, saloon, main street, etc.) and generic set-piece events for the characters to engage in (crooked card game, quick-draw shoot-out, and so on).

But it doesn’t stop here. The theories of film also provide some guidance as to how to best convey the information above to the player. The conventions of the Western are that the action is often shown in a wide shot, so this can be our default way to render a scene, but the fact that the villain is badly shaven is best conveyed by showing him in a close-up. Likewise, other objects - such as the sheriff’s star - are known to be significant for this genre and will likewise be worthy of close-ups.

This has hopefully given some brief indication of how a (silent) interactive Western system could be assembled from what we know about the Western from film theory.⁴ Such a system is far more concrete than the idealised storytelling engines in Aaseth [7] and Laurel [8].

3 Narrative and Narration In Interactive Entertainment

Many writers including Laurel [9], Murray [10], Aarseth [11], Poole [12], and others have written about interactive narrative, and although different writers may use different terms, most come to a common conclusion - that the more freedom the player/user has to intervene in the narrative or choose their own path through the narrative, the weaker the voice of the author becomes.

We would argue that the problem is not the player’s intervention in the narrative, but rather their intervention in the narration. The clear and undeniable problems that current interactive narratives of all types have - a lack of empathy with the characters, a lack of engagement with the events of the story - has more to do with the narration of these works than with their narratives: the most exciting of stories can be made dull when presented in an uninspiring way, while the plainest of events can be made interesting and exciting through its presentation.

Nowhere is this more true than in videogames. At a surface level, these are the most seductive of stories to immerse oneself in - they tell action movie stories with

⁴ Creating characters that speak and respond to speech is a separate AI issue, though again this task may also be simplified by the conventions of genre.

glossy computer graphics and put you in control of the hero of the story - but as you play them, there is a sense of disconnectedness: even though you may have the sense of “being there” and jump (in the real world) when you are ambushed (in the game), there is little or no sense of engagement with the narrative or with the characters within it.

Film narration, as outlined above, is the interplay of a set of conventions regarding the presentation of people, objects and events onscreen. We must therefore ask whether a similar, parallel or equivalent set of conventions exists - or is developing - within interactive narratives, and whether the use of these techniques can solve the problems mentioned above and increase the emotional impact of these works.

As we have indicated before, there are clear formal similarities between film and computer-based forms of entertainment that have allowed many aspects of film to already be adopted by these media. Most of these similarities are so obvious that they are barely noticed and only a brief run-through of them is needed: they are both screen-based media; they are both time-based media; they both use technology in their production and presentation; they both tell their stories predominantly through image, rather than through dialogue.

There are, however, also differences. The most fundamental of these - if one leaves aside, for the moment, issues such as the scale of the cinema screen or the fact that films are projected - is the nature of view and viewpoint in the two media. A film, on the whole, uses a variety of shots to tell its story and while some of these shots may coincide with the viewpoint of one of the characters, most of the time they don't. The filmmaker decides what object, person or event to show on screen - and how to show it - and the viewer of the film can only choose where to focus their attention within this limited view given to them (or to look away).

In interactive environments, however, the player tends to have a continuous view of the action, typically through the “eyes” of their character or from behind them. Taken individually, neither of these points - having a continuous view or seeing through the “eyes” of the character - is a problem. The problem comes when they are used exclusively and thus impose a continuity of time, space and action on the videogame: the designer of the world is left with little opportunity to control what the user sees or how they see it.

Under such conditions, the designer of a virtual environment cannot prioritise an object through showing it in close-up as a filmmaker would do - they must do it through the design of that object, its placement in the world, the way that it is lit, and so on. They are likewise forced to create mood through the design of the world, the lighting, the props and decor, etc. Essentially, they are reduced to one narrative technique: that which we refer to in film as *mise en scene*.⁵

But there is a limit to how much the designer can prioritise an object solely through its design, placement and lighting. Virtual worlds therefore tend to have every non-essential item removed so that the few objects that you can interact with stand out enough to be noticed. Significant objects are also often constructed in a deliberately non-realistic fashion so that they stand out from their environment - keys in Quake,

⁵ Genre is another element, though through providing a ready-made set of characters, objects, locations, etc., it contributes greatly to the *mise en scene*.

for instance, are oversized, floating in mid-air and rotating so that they cannot be missed. In effect, these videogames use the symbol of a key, or something that stands for the function of a key, rather than a key *per se*.

Virtual environments are sparsely decorated and furnished, present few objects to interact with, present them in a deliberately unreal way; they are also - for different reasons - sparsely populated (usually with characters that you can't talk to). The worlds themselves often have a maze-like structure, with paths blocked off until certain tasks have been done, as this is the easiest and safest way to guide the user.

The design of the world is, to a very real extent, the design of the narrative, and there are limits, therefore, to how subtle and sophisticated this narrative can be. As we have mentioned elsewhere [13], these are not problems that can be solved solely through achieving photorealistic quality in the rendering or greater realism in the animation.

4 The Continuity of Time, Space and Action

In the previous section, we said that the continuity of time, space and action presented a problem in interactive narratives, (limiting the narration to a sparse *mise en scene* and the problematic design and placement of significant objects), and it is now necessary to explore this issue, its origins, and its implications in greater detail.

By "continuity of time, space and action", we mean that the user experiences the environment as a space that exists as a whole independent of their presence or actions, and that the user's actions in this world are both presented and experienced as a single continuous event.

This is a complicated explanation for what is a very simple concept - that the world of the game is a world that we experience like the real world. At first, it may be difficult to imagine any alternative to this, but consider the following examples. In a film, for instance, scenes are made up of shots - taken from a variety of angles - edited together. This allows the film to jump backwards and forwards in time, or from one location to another. A play will likewise include jumps in time and between places not only between scenes, but also within them - indeed such jumps can take place mid-sentence. In addition, one can see how cubist paintings present multiple viewpoints in a single image, or different moments in time, or both.

What this should indicate is that there is not one form of presenting time and space that is "natural" and others that are "conventions" - they are all conventions. The fact that virtual environments tend to obey the continuity of time, space and action is not "natural" - it is merely a convention (and one that is particularly strange given the discontinuity of the player's experience: the fact that they will die within the game and be "reborn", that they will pause and resume the game, and so on).⁶

⁶ One cause of this is the way in which virtual reality was "colonised" by architects at a very early stage - see Benedikt [13] for a prime example of this.

5 “Showing” and “Telling”

Interesting patterns emerge if one groups together works - both interactive and non-interactive - that use the continuity of time, space and action and ones that don't. There should be no surprise that this division (between works that use/obey the continuity of time, space and action and those that don't) means that most immersive virtual environments - whether they were produced as arcade games, research projects, simulations, interactive art, etc. - can be grouped together as they all obey the convention of continuity of time, space, action.

Also included this group are videogames such as Tomb Raider or Quake. We ignore the so-called “cut scenes” that may appear between levels in games such as these or their very occasional cut-away shots (such as to show what effect pulling a lever has had). The reason for this is that these are used so rarely that they do not interrupt the overall continuity of the player's experience to any significant degree: a player may spend an hour completing a level and only get a cut scene at the start and end of it; they will likewise only get a handful of cutaway shots.

A videogame such as *Myst*, however, doesn't fall into this grouping. This is because when you click to “walk” ahead in *Myst*, you don't see your viewpoint change as you take every step through the environment - you cut to what you would see from this new position. This contrasts with the experience of playing *Quake*, for example, where your viewpoint changes fluidly as you walk forward, or *Tomb Raider*, where the camera is constantly and continuously on your character.⁷

Also outside the grouping are text-based MUDs and MOOs. You might think that MUDs and MOOs would be grouped with the other virtual environments above, but they are not. This is not because they allow actions such as “teleporting” or “emoting” (although both of these are interesting in that they are like cutting out your movement from one place to another - as you would do in a film - or stepping out of the narrative to deliver an aside - like an actor on the stage). In stead, it is because one can imagine that in a MUD, the world could be built so that when you go up a flight of stairs, you type “up” once to go halfway up and once more to reach the top, but when you go down, you can do so in a single “down”. Here we are achieving the same sort of fluidity of time and space - and ability to compress and expand time and distances for narrative or emotional effect - that we have in film.

By dividing works of all types into those that obey the continuity of time, space and action and those that don't, one ends up with the following broad categorisations. In the category of those that obey continuity of time, space and action, we would place most modern videogames (including *Quake*, *Tomb Raider*, etc.), immersive virtual environments, multi-user worlds (such as *Alphaworld*), etc. The list of those that do not would include most old videogames, MUDs, MOOs, films, novels, plays, etc.

The pattern that emerges as a result of this distinction is easy to see, and echoes the terms used by other writers on (non-interactive) narrative: i.e. the distinction between

⁷ *Myst* is a prime example - and in many ways the apotheosis - of the sparseness typical of virtual environments.

“showing” and “telling”, “mimesis” and “diegesis”, “imitation” and “presentation. Essentially, what we are talking about here (with regard to interactive narratives) is the ability to structure the way that the events within the narrative are viewed or experienced by the user, rather than seeing the limit of one’s role as “author” of the world as simply to decide what happens within it.

6 The “Birth” of Immersion

An activity such as the one outlined above allows us, on one hand, to identify similarities between various works that might, at first glance, seem to be very different and, on the other, to perceive differences between those which might initially appear to be similar.

It has other benefits too. If one examines the history of videogames in the light of the distinction between works that obey the continuity of time, space and action and those that do not, other interesting points emerge. One comes to realise that it is only relatively recently that the videogames obeying this continuity of time, space and action have come to dominate as they do now - before this the majority did not obey this convention.

If one looks at 1980, for example, one can see that virtually all of the top arcade games - Defender, Missile Command, PacMan, Space Invaders, Asteroids - consisted of separate, distinct levels of stylised play within a game area that is typically no larger than the screen. Only Battlezone stands out as an exception to this rule: here the action - tank combat - takes place in a “real” space, rendered with realistic 3D perspective (albeit with only vector graphics).⁸

If one looks at the current state of computer, console and arcade games, however, one sees the situation reversed: they are all dominated by games which *obey* the continuity of time, space and action. The crossover point is difficult to place precisely, but 1992/93 is a critical moment, with the release of first-person “shoot-em-ups” such as Doom. Before Doom, immersion was not essential; after it, immersion has become - to a very great extent - the *sine qua non* of a videogame.

7 Conclusions

From this brief examination of narration in film and interactive narratives, it is possible to draw some conclusions about the way that interactive narratives can develop, and to do this, we will return to the distinction that we made between works that obey the continuity of time, space and action and those that do not. It is important to remember, however, that we are not making predictions about the

⁸ PacMan etc. do not class as obeying the continuity of time, space and action as their levels are short - we are talking about the *sustained* continuity of time, space and action.

narratives of these works - we are engaged in predicting developments in the *narration* of these works.

As we have pointed out earlier in this paper (and elsewhere), the continuity of time, space and action offers a quick and easy sense of immersion, but places severe limitations on the type of interaction that can take place and the level of narration that the creator of the game can impose on the events within the world and on the player's actions.

During the early nineties, at around the time that Doom was released, there was a fundamental shift in the design of videogames - from the games before this, which tended to not to obey the continuity of time, space and action, to those after it, which did. We can therefore speculate as to what videogames may have been like if they had continued along their original path and had taken advantage of the phenomenal increases in processor speed and computer graphics that have taken place since 1993.

Games were, at that point, already establishing conventions of their own for bridging space and time. A game such as Elite, for example, allowed the player to skip forward in time and cut out the boring bits of space travel (thereby progressing rapidly from one dogfight to another or one planet to another). Streetfighter, likewise, was establishing conventions of how to use montage - in a very comic book-like fashion - to convey the power of the fighter's blows, etc.

We believe that, over time, these games would have borrowed more extensively from the narrative techniques of film as these offer a ready-made and highly familiar set of conventions that they would have been able to exploit fully through their increasing ability to perform real-time rendering.

We can imagine, therefore, that these games - and the players of them - would have been comfortable with a whole range of filmic techniques that break up the continuity of time, space and action for narrative ends. These would include: changing viewpoint in mid-action; bridging time or space with a cut; cross-cutting between separate locations, characters or threads of the narrative; incorporating flashbacks and flashforwards; and so on.

We therefore believe that the designers of videogames and interactive narratives should abandon their obsession with immersion and return to exploring the freedom that breaking the continuity of time, space and action gives them as narrators.

It is important to realise that what we are arguing for is a clear and distinct break with the dominant form of videogames and interactive environments (i.e. those that obey the continuity of time, space and action). It is not sufficient to keep producing this type of game, merely with a little film language added on.

Games such as Quake or Unreal Tournament do what they do extremely well - they offer a very strong sense of immersion and are very exciting to play. Film language has only a limited role to play in videogames such as these, and adding it inappropriately would detract from their strengths. But there is, however, a separate but equally valid form of videogame for which the narrative techniques used in film are ideal: those which aim for a sense of engagement in the narrative and/or the characters but which are currently unable to produce this because they obey the continuity of time, space and action.

We therefore believe that there needs to be a rediscovery of the “path not taken” in the design of videogames, virtual environments and interactive narratives - a return to the freedom that breaking the continuity of time, space and action provides. Essentially this provides the opportunity to *author* a narrative, rather than just design a world - the opportunity to choose how the events that happen in the world will be experienced, rather than just deciding what events to show.

We are talking here about reclaiming the virtual world as a storytelling medium, rescuing it from architects and those who would make it just a perfect, yet pale, imitation of the real world. We can currently only speculate on what exciting form these works will take.

8 References

1. Clarke, A., Mitchell, M.: Playing with Film Language. In: Coulter-Smith (ed.): The Visual-Narrative Matrix. Southampton Institute (2000)
2. Clarke, A., Mitchell, G.: Screen Play: Film and the Future of Interactive Entertainment. In: Earnshaw, R. A. (ed.): Digital Content Creation. Springer-Verlag (2001)
3. Bordwell, D., Thompson, K.: Film Art: An Introduction. Alfred A. Knopf (1986)
4. Stephenson, R., Phelps, G.: The Cinema as Art. Penguin (1989)
5. Stam, R., Burgoyne, R., Flitterman-Lewis, S.: New Vocabularies in Film Semiotics. Routledge (1998) 33-68
6. Propp, V.: Morphology of the Folktale. University of Texas Press (1968)
7. Aarseth, E.: Cybertext. John Hopkins University Press (1997) 104
8. Laurel, B.: Computers as Theatre. Addison Wesley (1993) 138
9. Laurel, B.: Computers as Theatre. Addison Wesley (1993)
10. Murray, J.H.: Hamlet on the Holdeck. M.I.T. Press (1997)
11. Aarseth, E.: Cybertext. John Hopkins University Press (1997)
12. Poole, S.: Trigger Happy. Fourth Estate (2000)
13. Clarke, A., Mitchell, G.: Screen Play: Film and the Future of Interactive Entertainment. In: Earnshaw, R. A. (ed.): Digital Content Creation. Springer-Verlag (2001)
14. Murray, J.H.: Hamlet on the Holdeck. M.I.T. Press (1997) 48-49,120-121
15. Benedikt, M. (ed.): Cyberspace: First Steps. M.I.T. Press (1994)