Games, Frames, Screens and Seams. Spatial and Temporal Editing in Video Games

RUBEN MEINTEMA

Introduction

There exist a number of myths in academic circles about video games. The Russian media theorist Lev Manovich believes that New Media in general and video games in particular avoid editing and attempt to provide a seamless experience. (Manovich, 141-144) Likewise, the Danish game scholar Jesper Juul believes that video games are essentially non-narrative cultural forms, which do not deploy narrative editing techniques like the flashback and the flash-forward. (Juul, 223)

These 'myths' are grounded in an outdated anti-narrative ideology that emphasizes the uniqueness and difference of the new medium of video games. This is a common reaction to any new medium, guided by the wish of making the medium transparent. (Bolter and Grusin, 19) However, very thoughtful observations tend to be stretched too far in order to emphasize the 'newness' of the new medium. Hence neglecting that editing is very common in video games, though not in the same form as we are used to from other media.

It is the purpose of this article to show that despite the fact that some scholars have denied it, and most scholars have neglected it, editing is a common feature of video games. In addition, I will create a categorization of editing techniques in video games, which could also well be implemented onto other media. The insights gained from this could be of relevance for game designers, but also students of games, and for students of editing techniques in other media.

First I will start with a theoretical overview, in which I discuss the spatial and temporal characteristics of various media – including games – and introduce the concepts of spatial and temporal editing. In this section I will also argue against the view of such media theorists as Manovich and Juul that video games are smooth forms that resist editing. Secondly I discuss six types of editing; single frame, multiple frame, and successive frame editing, each with a simultaneous and a non-simultaneous variant. In coming to a conclusion this typology is made visual in the form of a table (figure 1, see page 110) including examples of each type of editing.

Space and Time across Media

The Arts have traditionally been divided into spatial forms and temporal forms, at least since Lessing's famous analysis of Laocoön.[1] Spatial forms are static objects that could be displayed in a museum, like sculpture, painting, photography, and architecture. Temporal forms on the other hand are dynamic forms that do not have a stable spatial extension. Examples are theatre, music, and dance. They are performed and therefore they do not 'exist' as such before and after their performance. They could thus not be displayed in a museum. Several modern art movements have tried to break with this distinction, most notably 'Performance Art.' Other art forms do not fit the typology exactly, because they have both spatial and temporal characteristics, like film, literature and poetry. (Leyssen and Vermeir, 46-62)

In literature and poetry the distinction between written and oral forms determines to a great extend the implementation of temporal or spatial editing. Spoken words are divided

1 The German Enlightenment thinker G. E. Lessing (1729-1781) argued in his essay *Laocon* (1766) that sculpture and poetry could not be analysed in the same way, because sculpture has a spatial extension while poetry has a temporal extension.

from each other in time, while written words on a page are separated by spaces. Respectively, this is due to the visual or auditory aspects of written and oral literature. Since the work of the German philosopher Arthur Schopenhauer (1788-1860) it has been held that perceptions of the eye are essentially spatial, and perceptions of the ear temporal, and this view is being adhered to until the present. (Kubovy, 318)[2] Although film images are predominantly visual phenomena, they act more like spoken instead of written words, because they are separated from each other in time. In this respect conventional film is said to be a temporal art form, because the film's structure is *temporally* edited.

A typical feature of New Media, according to Manovich, is *spatial* editing. (40, 157-158) Microsoft Windows is based on the possibility of opening several 'windows' on one screen, and of working simultaneously in Word, Internet Explorer, or any other program. These windows are an example of spatial editing. Films are usually temporally edited, but the film *Time Code* from 2000 is a reference to the 'Window thinking' of New Media, as it divided the screen into four frames, in which events are shown that take place at the same time.

In video games, as the most important contemporary cultural form of New Media, spatial editing plays a significant role. In early multiplayer games (before the widespread implementation of the Internet), the screen was divided into two or four frames. Examples of this abound in shooting games, platform games, and racing games. For instance, when two players race against each other, one player's car would be depicted on one frame, and the other player's vehicle on the other frame. The events that take place on both frames always

2 Modern cognitive research has shown that visual observations have to be divided by at least 4 milliseconds in order to be perceived as non-simultaneous, while auditory perceptions have to be divided only by 2 milliseconds, which implies that the ear is indeed better at perceiving time than the eye. (Pöppel, 713-729).

have to be simultaneous, otherwise a race or any other contest would be impossible. This is indeed the same editing logic as in *Time Code*.

In multiplayer games it could be clearly observed that the events on the screen logically take place in the present. But also in single-player games this is true, because otherwise the player would not be able to interact with and influence the events. The events on the film screen on the other hand, logically take place in the past. The film apparatus 'captures' the scene in front of the camera lens in order to show it at a later moment.

Manovich distinguishes three paradigmatic types of screens. The first one is the painting, which is a static representation of the past. Film is the second type: a dynamic representation of the past. The third paradigmatic screen is the radar, which is a dynamic representation of the present. (Manovich, 95-99) The radar image has to represent the present, otherwise the pilot can not use it to track his enemies down and the captain can not use it to steer his ship past perilous obstacles. So the possibility to act on the basis of the information displayed on the screen is always linked to the 'presentness' of this third type of screen. In video games this is not different, because the player needs to *act* on the information that the screen displays, so video games belong to the third type of screen.

The temporal reference of the video games to the present also has repercussions for the narrative structure of the medium. Because there is no temporal editing, there would be no flashbacks and flash-forwards. The semiotic conception of narrative (the plot as *signifiant* and the story as *signifié*, as described for film by Christian Metz and for literature by Gerard Genette) could not be applied to video games. According to Jesper Juul, the *Sa* and the *Sé* collapse into one single time. The video game would consist of one long take without cuts, flashbacks, and flash-forwards. (Juul, 222-223)

This is reason for Manovich to believe that the video game is seamless. Seamlessness is according to Manovich a feature of the 'Postmodern' aesthetic in general, and of video games in particular. In new media music, the DJ mixes together the separate tracks of his composition seamlessly, and the highest virtue of a mix is that one cannot determine where the previous track ended and where the next one begins. Because of the 'presentness' of video games, the temporal experience and perception of games would be closest to our everyday experience. (Manovich, 144, 142-143)

But Manovich's theoretical statements seem to contradict each other. Spatial editing is editing nonetheless, even though there are no temporal cuts like in film. Spatial editing through windows creates seams as well, which are perhaps even more visible than temporal cuts

(with continuity editing). Furthermore, the windows and frames on the screen are quite far from our everyday experience. Therefore, it is the aim of this article to criticize the myth of seamlessness of New Media and games by looking closely at the editing techniques of video games.

Spatial and Temporal Editing in Video Games

The large variety of editing techniques in video games can be systematically discussed by making distinctions between the possibilities of spatial and temporal editing along two axes: temporal reference, and frame configuration. There are three frame configurations: single frame, multiple frames, and successive frames. 'Single frame' refers to the situation in which the frame exactly coincides with the screen. With 'multiple frames' the screen is divided into two or more frames. 'Successive frames' means that several frames are displayed on the screen one at a time, so there is never more than one frame on the screen. Along the vertical axis I will distinguish between a simultaneous and non-simultaneous situation. 'Simultaneous' means that the time onto which a frame opens is the same time as displayed on the other frames. 'Non-simultaneous' means that different frames refer to different times.

Single Frame Editing

Single frame editing comes down to the situation of seamlessness that Manovich has described. It is the conventional type of editing in video games. A modern video game usually consists of one long take without cuts (although there are usually cuts between levels). This long take is always 'live' and could not be temporally edited because it has already been played. This means that unsuccessful game play attempts are also an integral part of playing, although they would be edited out of the work in film, or relegated to the status of bloopers.

But not every event depicted on the screen in a 'single frame editing' situation is simultaneous with the moment of playing. There are also instances of non-simultaneous editing in a single frame on a single screen. In many racing games, like *Mario Kart* for example, the player could choose to play a single-player time-attack race (that is without opponents) to attack a certain record time. The course of the player who had set the record time is represented by a transparent vehicle, and his course is displayed on the same race track, in the same frame, and on the same screen as the current race. This is a convenient device because the player can keep close track on the record time course. The player thus not only races against the abstract opponent known as 'time,' but also against

3 This aspect is described in more detail by the American game scholar Mark J.P. Wolf (1997).

a 'concrete' ghost car of the record time course. Note that the player who has set the fastest time could also be the same player that tries to break this record. So in that case the player races against himself as a ghost of the past.

This type of editing of non-simultaneous events in one frame is most relevant for racing games. It does not really apply to adventure games and platform games. But in the racing game genre it is conventional. It is a striking example of the depiction of events in the past and events in the present, shown on one and the same screen.

Multiple Frame Editing

But game play is not completely seamless. Spatial editing is editing nonetheless. In many games a small radar screen is displayed on the screen as a separate frame. This is convenient to navigate through the elaborate spaces of games, and to locate enemies. But this radar is not in any way similar to our everyday perception. Still, the radar is a conventional device in video games, in all genres from adventure and role-playing games to shooting and racing games.

The radar could be used in single-player as well as in multi-player mode. But in multiplayer games the division of the screen in one or four equal sized frames is a common device that is used mostly in the era of off-line multi-player mode. In this kind of gaming situation, the other players are located in the same space, and they use the same screen. This screen is divided into two or four frames, much like *Time Code*, and each player is assigned one screen in which he or she can follow his avatar or car. The actions of the other players are visible for any other player, so the

player can see himself in the frames of the other player from the perspective of the other player.[3]

The radar and the split-screen multi-player mode are both conventional examples of simultaneous multiple frame editing. But with the advent of the Nintendo DS multiple frame editing has become increasingly complicated. The Nintendo DS is a handheld console with two screens (Dual Screen), of which the bottom one is also a touch screen. These innovative features have revolutionized game editing, by distributing multiple frames over two screens. Games developed for the Nintendo DS in many cases use the bottom screen to display the radar map, instead of editing it into one screen.

But the two screens could be used in more innovative ways. In *Avatar: Burning Earth* for the DS, the character accidentally breaks the portrait of the general when he is away for a moment. The player then is challenged with the task of piecing the portrait back together like a jigsaw puzzle before the general comes back to see the damage. The general walking back is shown on the top screen, while on the bottom screen the player tries to piece the puzzle back together. This is a creative way to use the two screens, because it shows events in two places simultaneously, and it adds to the experience of suspense by the player. The proximity of the general, and thus the time left to finish the puzzle, is knowledge that the player has, but the character does not have. This knowledge has the effect of making the puzzle more interesting, and involving the player more in the character and the fictional world.

Multiple frame editing in games is usually simultaneous, because otherwise the information on the screen would be unreliable, and the player would not be able to act on it. But this type of editing could also be non-simultaneous. In graphic novels non-simultaneous multiple frame editing is the conventional situation. A story, which consists of events that have taken place in succession to each other, is told through multiple frames on one page. Each frame opens up another time frame. When reading the frames, the story is being reconstructed by the reader. Therefore it is no surprise that the video game that makes use of this type of editing is based upon a graphic novel. In the first-person shooter game *XIII* there is a normal video game screen. But this screen is spatially edited with multiple frames that open up to events in different times, much like a graphic novel. The American game scholar Michael Nitsche has shown how there could be four different time frames on one screen, and argues that players have no difficulties understanding the temporal make up of this screen. (146-147)

Successive Frame Editing

Successive frame editing consists of showing frames in succession, while there is no more than one frame at a time on the screen. The frames are divided temporally instead of spatially, so this is a type of temporal editing. In simultaneous successive frame editing the events displayed after each other are really taking place at the same time, while in the non-simultaneous variant the events that are displayed after each other also take place after each other. The latter is the conventional type of editing in film.

But also video games often include 'cinematic sequences' between game play sequences, in which cinematic editing techniques are used. In the 2001 PlayStation 2 game *Final Fantasy X* for example, cinematic sequences abound, and they are seamlessly integrated into the game play. In the beginning of the game, there is an attack on a city. The depiction of the attack starts out as a cinematic sequence, after which there is a seamless transition to a game play sequence in which the player has to repel the attack.

It would a misconception, however, to attribute the seamlessness of this example to a 'Post-modern aesthetic of New Media,' as Manovich has done. At least since the ceiling fresco of the Sant'Ignazio church in Rome by Padre Andrea Pozzo, in which the fresco is seamlessly integrated into the church' architecture, it has been an aesthetic preference. Also since Classical Hollywood cinema, the system of continuity editing is aimed at hiding the cut and providing a seamless viewing experience. (Bordwell and Thompson 231-240) As Bolter and Grusin have argued in *Re-mediation* (1999), seamlessness is not an exclusive or essential feature of New Media, it is rather a strategy deployed to achieve immediacy and transparency of the medium in all or most arts. (19)

Simultaneous successive frame editing is rather uncommon in cinema. It is used in several action films, when for example the same explosion is shown from different angles in successive shots to produce a spectacular effect. This editing device is also transplanted onto action sequences in video games. In the 1998 fighting game *Tekken 3* for example, the final knock-out punch or kick is repeated three times, each time from a different perspective. But also in the racing game *Burnout* (a game that is not only about racing but also about crashing), a big crash is repeated quickly through successive cuts, and this adds up to the spectacular effect of the crash. So while this type of editing does exist in games, it is used mostly in action sequences to achieve a cinematic effect.

Simultaneous successive frame editing is also used in cinema in the form of 'parallel editing,' in which events that are taking place

at the same time are shown after each other. In video games this type is predominantly used in cinematic sequences to show events that are taking place somewhere else. But in some cases also game play sequences are edited in this way. This is not so common, because the player usually only controls one character, and events involving other characters are shown as non-playable cinematic movies. For example, in the previously discussed game *Avatar*, there are three protagonists, and they are all playable characters. When two of them leave a fortress behind to search for a medicinal herb, the third protagonist stays in the fortress to take care of the injured. When the party of two characters has found the medicine, there is a cut back to the fortress, and the player can control the third character who is healing patients. After the player has healed a number of patients, there is another cut, and the two other characters are playable again to bring back the medicine to the fortress. This type of editing is used mostly in adventure games and role-playing games which have a high level of narrativity and multiple playable characters.

frame configuration: / temporal reference:	**Single Frame**	**Multiple Frames**	**Successive Frames**
Simultaneous	conventional game editing	multiplayer games, radar, *Time Code*, *Avatar (DS)*	parallel editing, *Avatar (DS)*
Non-Simultaneous	'ghost player'	conventional graphic novel editing, *XIII*	conventional film editing, *Final Fantasy X*

Figure 1: Typology of Spatial and Temporal Editing

Conclusion

In this article I have shown that editing is not uncommon to video games, as some of the most influential theories about contemporary media suggest. Video gaming shows a broad variety of editing techniques, some of which are closely related to film, others are more akin to illustrated literature, in particular the graphic novel. Instead of being a seamless entity video games consist of elements that are pieced together spatially or temporally. I have shown how video games

could implement at least six types of editing, making use of a typology consisting of varieties of frame configurations and temporal references. The three frame configurations (single frame, multiple frame, or successive frame) could have the same temporal reference or various temporal references, so they are simultaneous or non-simultaneous. In this article I provided each type of editing with examples from video games, as well as other media.

Not only have I created a typology of editing techniques in video games, and placed them in perspective vis-à-vis editing techniques in other media, I have also linked these editing techniques to effects that they might have on the player. This could be of relevance not only for academic game scholars, but also for game designers and developers who are interested in how editing techniques could influence the playing experience. Experiences of suspense, spectacle, narrative coherence, continuity, of stylistic pleasure, etcetera could all be produced by making use of editing techniques. Game designers and scholars alike would do good not to miss this considerable aspect of video gaming.

BIBLIOGRAPHY

Bolter, J. David, and Richard Grusin. *Re-mediation: Understanding New Media*. Cambridge: MIT Press, 1999.

Bordwell, David, and Kristin Thompson. *Film Art. An Introduction*. New York: McGraw-Hill, 2008.

Juul, Jesper. "Games telling stories?" *The Handbook of Computer Game Studies*. Ed. Joost Raessens and Jeffrey Goldstein. Cambridge and London: MIT Press, 2005. 219-226.

Kubovy, Michael. "Should we resist the seductiveness of the space:time::vision:audition analogy?" *Journal of Experimental Psychology: Human Perception and Human Performance* 2/14 (1988): 318-320.

Lessing, Gotthold Ephraim, *Laocoon: or, The Limits of Poetry and Painting*. London: Ridgeway, 1836.

Leyssen, Sigrid, and Koen Vermeir. "De film en zijn museum. Over de esthetiek van het verleden." *Transgressie in de kunst*. Ed. Bart Vandenabeele and Koen Vermeir. Budel: Damon, 2004.

Manovich, Lev. *The Language of New Media*. Cambridge: MIT Press, 2001.

Nitsche, Michael. "Mapping Time in Video Games" *Situated Play. Proceedings of DiGra 2007 Conference*. Ed. Akira Baba. University of Tokyo, 2007. 145-151.

Pöppel, Ernst. "Time Perception." *Handbook of Sensory Physiology. Volume VIII: Perception*. Ed. H. Autrum et al. Berlin, Heidelberg and New York: Springer Verlag, 1978. 713-729.

Wolf, Mark J.P. "Inventing space: A taxonomy of on- and off-screen space in video games." *Film Quarterly* 1/51 (1997): 11-23.

SUMMARY

In this article I criticize the view that video games resist editing and are 'seamless' forms. As opposed to this myth of seamlessness I state that a great deal of editing is involved in video games. Apart from temporal editing, which is well-known in cinema, video games make heavy use of spatial editing, which is also used in graphic novels. I will systematically discuss six types of editing, based on their spatial or temporal configuration, and the simultaneity and non-simultaneity of the depicted frames. In addition to this typology I will continuous debate how the editing technique works for the viewer to create interaction, suspense, narrativity, and spectacle.

Ruben Meintema (1983) is currently finishing the Research Master of Literary and Cultural Studies at the University of Groningen, where he has focused on the study of video games. He has contributed before to *De Game-industrie. Een inleiding.* (2007) and his article 'Beneath the Skin of Digital Characters: Race and Body in Japanese Fighting Games' is forthcoming in *Race and Racism in Modern East Asia: Western Constructions and Local Reactions* (2010).