Virtual reality; a question of subjectivity and authenticity?

Sigmund Trageton

University of Stavanger, Norway

Abstract

With the rise of the new medium like Virtual reality, it is important not only to understand the technology and marketing hype, but also to look closer on what differentiate this medium from others. Discussions around the medium's characteristics is an important way to understand the narratological power and how this can be used creatively.

This paper, will through a comparative analysis of four different VR-productions, try to shed light on the intricate play of narration and focalization and how this is differently exploited in each of the productions.

The matter of subjectivity and authenticity is two of the strong key aspect of Virtual reality. This has led to a form of kinetic excitement with a lot of production that utilizes the attraction of movement in the VR films (bungy jumping, roller coaster etc, etc). However, subjectivity is much more than just the perceptual cognition of the story diegesis, it is a central story exposition tool that need a close narratological analysis to be understood.

We will base the analysis on Edward Braningan's concept of 8 levels of narration (Braningan, 1992) where he distinguishes between the narration and the focaliziation. Furthermore, we will look on how the focalization fits, or not, with the intrinsic perceptual cognition characteristic of the VR medium.

Keywords:
Virtual Reality, Narrative theory, Analysis,

Cinema is Frame. Cinema is length of the lens, cinema is editing.

The position of the images that creates time and space.

VR, even when it is visual, is exactly all that cinema is not

Alejandro González Iñárritu

Introduction

To create emotional response and engagement with an audience has always been a goal for film directors, and the film medium was considered to be the "strongest" medium (meaning that it could create the strongest emotional response). With the birth of a technological capable virtual reality with mesmerizing resolution, head tracking and ambisonic sound, many journalists, technologist, producers and scholars have stated that this new medium has a stronger immersive power.

Sergei Eisestein stated that "Emotion is completely identical with the primary phenomenom of cinema. [In cinema] movement is created out of two motionless cells. Here, a movement of the soul, i.e. emotion (from the Latin root motio = movement), is created out of the performance of a series of incidents." (Eisenstein, 1991). Properly structured as a series of uncompleted incidents, montage calls on us to finish the actions mentally, and for Eisenstein this internal movement of filling in the gaps is emotion, a movement of the soul. (Smith, 2004)

In this article, I will try to compare this strong intrinsic immersive power of VR, with the challenges concerning the grammatical difference between film and VR. To understand this, we first have to look at how the brain processes information.

Process strategies

Top-down and bottom-up are both strategies of information processing and knowledge ordering. These strategies are used in a variety of ways inside humanistic and scientific theories. More specifically, they have become an important part of cognitive filmtheory. David Borwell defines the strategies as:

"Bottom-up" processing refers to those fast, mandatory activities, usually sensory ones, that are "data-driven." "Top-down" processes are more concept-driven; they are more deliberative, volitional activities like problem solving and abstract judgment. (Bordwell, 1989)

And then he goes on to explain why this is a key term for film analysis.

> The crucial assumption is that both bottom-up and top-down processes execute inference making; both "go beyond the information given" in determinate ways. For one thing, top-down processes can shape and steer bottom-up activity. Reading a text is not simply registering letters, adding them up to make words, adding them up to make phrases, and so on. Selected chunks of text cue us to extrapolate far ahead of the words that we next encounter; we start to build a semantic structure that guides our samplings of data. (Bordwell, 1989)

I will later show how these processes are crucial also for the narration in the VR documentaries and how it can help to analyze the narratological perspective.

Scholar Mandler shows that the top down processing is connected to earlier experiences and the schemata that a viewer already have in his mind (Mandler, 1984). Mandler and her colleagues propose a prototypical schemata that are characteristic of narratives. These "canonical stories" consist of certain elements in a standard order: an initial description of time and place; a delineated episode that undergoes development; a development, which consists of either characters' simple reactions that trigger immediate action, or characters' complex reaction that causes a "goal path" to be initiated. Such a schema can be shown to facilitate understanding and recall of a wide number of stories. Stories that do not follow the schema, such as tales lacking causal connections between episodes, are demonstrably more difficult to follow and remember (Mandler, 1984).

Julian Hochberg has proposed that physical movements and spatial configurations have distinctive visual features that mediate recognition. When the filmmaker seeks to represent a building as having a curved colonnade or to suggest that a person crosses a room, the onscreen cues should present the distinctive features that will prompt the spectator to make the proper perceptual inference (Hochberg, 1986, 44-59). And also from seeing films we tend to create conventions from what we have seen and this forms our processing of new films. Bordwell describes this as following:

> I have argued elsewhere that not only do narrative films utilize mental representations for their depicted events, but they draw on historically developed conventions that involve schemata and heuristics. For example, the classical Hollywood narrative is in many ways similar to Mandler's "canonical story," and it delegates to the spectator the task of assembling events into a coherent causal whole. By contrast, the tradition of "art-cinema" narration encourages the spectator to perceive ambiguities of space, time, and causality and then organize them around schemata for authorial commentary and "objective" and "subjective" realism (Bordwell, 1985)

This is a key aspect when we look at a new medium like Virtual reality because the viewer have not yet formed medium specific schemata but rather draw on experiences from film. Especially when it comes to the perception of subjectivity, this can both become an obstacle but also a blessing. Let me try to simplify with comparison between film and VR

The most obvious characteristic of VR and 360 video is the perceptual sensation that you are actually in a place that you are not. Because of the visual ques, the bottom up process tells you that you are inside a world surrounded by the diegesis. From a narrative point of view, you have a very restricted perspective on the diegesis, because if the director wants to stay faithful to the visual perception he can only

show you events in that specific place and from that specific perspective.

In film, on the other hand, the director can change the narrative perspective with every new shot. He can distribute the knowledge and thus make the spectator engage with the characters in different ways. I will now try to show this intricate process by using Edward Braningan theory on levels of narration and use this when I later analyze the four different VR productions.

Narration in film

Edward Branigan is also considered to be among the cognitivist movement and he describes narration as a dialectical process taking place between the spectator and a film and realizing a narrative (Branigan, 1984, 38). He has also done extensive work on focalization. I shall now present his main ideas and relate them to those of earlier theorists and discuss what can be problematic about them. First, it is necessary to see how narration is defined:

Narration is the overall regulation and distribution of knowledge, which determines how and when the spectator acquires knowledge, that is, how the spectator is able to know what he or she comes to know, in a narrative. A typical description of the spectator's position of knowledge includes the invention of speakers, presenters, listeners and watchers who are in a position to know and to make use of one or more disparities of knowledge. Such"persons" are convenient fictions, which serve to mark how the field of knowledge is being divided at a particular time. (Branigan, 1992, 76)

Branigan then proposes a model with up two eight levels of narration/focalization that tries to explain both the relation between the narration and the narrator/focalizer as well as the more cognitive effects. (Branigan, 1992, 87)
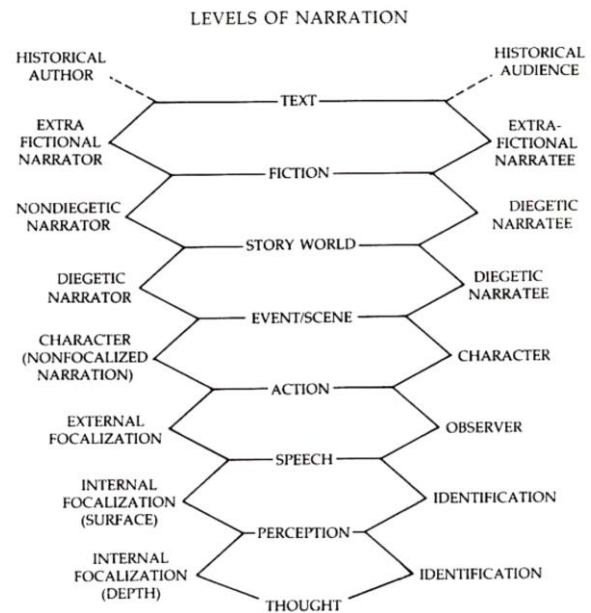

Fig 1 (Branigan, 1992)

As we see in this model, Branigan makes a strict division between narrator and character. Of course, these different levels will often be operating simultaneously and sometimes competing with each other, and a spectator may interpret the text in different ways, but these levels can help us to understand how the choices made by a director will influence the spectator's perception. I will now focus on Branigan's four level of focalization. He describes them as the following:

> The last four levels recognize that characters also provide us information about the story world, but in ways quite different from narrators. A character who acts, speaks, observes, or has thoughts is not strictly telling or presenting anything to us for the reason that spectators, or readers, are not characters in that world. Characters may "tell" the story to us in a broad sense, but only through "living in" their world and speaking to other characters. Indeed, one might almost say that these conditions, or restrictions, define what we mean by the concept of a character. (Branigan, 1992, 100)

This statement can be problematic in VR documentaries because from the perceptual point of view we are a character inside the diegesis. Some of the VR documentaries that I will analyze addresses this fact in a narrative way, whereas others just overlook it. In a way, they treat the new medium on the bases of film conventions. Let us now look closer on the different levels of focalization.

"Levels may multiply but there still exists a primary character-agent defined by actions and events. I will refer to this primary level of actions as a neutral, or non-focalized, narration of character" (Branigan 1992, 102). The next level is external focalization and is defined as: "A measure of character awareness but from the outside the character. It is semi-subjective in the manner of an eye-line match" (Branigan, 1992, 103). This is an interesting description because it implies a dialectical relationship between the film and the spectator. The inferences made by the spectator decide whether the narration is non-focalized or externally focalized. These kind of focalizations are the same in VR and film.

Another development from earlier focalizaion-theory models like Gerard Genette's model (Gennette, 1988) is the division between internal focalization, surface, and depth. 'Internal focalization is more fully private and subjective than external focalization. No character can witness these experiences in another character. Internal focalization ranges from simple perception (e.g. the point-of-view shot), to impressions (e.g. the out-of-focus point-of-view shot depicting a character who is drunk, dizzy, or drugged), to'deeper thoughts' (e.g. dreams and memories')' (Branigan1992, 104).

I will argue that these two last levels separates the inherent characteristics between film and virtual reality. In VR if we consider the spectator as a character in the diegesis these two levels should only be reserved to the viewer. Then the experience of presence would be enhanced and the storyworld would be narrative coherent, but this presupposes that the other characters addresses the camera directly and it can also imply certain amount of interactivity.

The border between non-focalized and external focalized can be hard to determine in film. When exactly is a shot becoming externally focalized? Moreover, what does semi-subjective imply? This border is even more blurred in VR as long as the director follow conventions of film. I would suggest a more deliberate use of internal focalization, as explained above, where the characters addresses the camera and thus get the emotional bounding that equals external focalization in film.

Methodology for the selection of VR documentaries

I have reviewed around 50 VR documentaries of different length, narrative structure and form of distribution. It is a competition going on between different technology providers to have the best content. Google, Facebook, Samsung, HTC Vive, Gopro and others try to make their own distribution platform with exclusive content. This is in addition to the traditional big media companies like New York Times, The Guardian, Spiegel, El pais and others. The main selection criteria has been to find productions that in different ways execute the narration and the grammatical VR language. It has been a clear choice to choose productions that differs in the use of the levels of narration and especially how they treat the focalization. First, I will present the four productions with a short plot followed by an analysis of each. Then I will conclude with a comparative discussion.

The fight for falludja
Produced by the New York Times and published august 2016. Length 11:08.

Image 1 (New York Times © 2016)

We follow journalist Ben C. Salomon on a news reportage about the final stages of fight around the Iraqi city of Falludja. Salomon follows the Iraqi forces as they try to fight the last Isis soldiers and liberate the city. It consist of two part

1. Taking the city

2. The aftermath (6.05)

We start on a rooftop outside the city with some Iraqi soldiers and Ben introduces where we are. Next scene is on a hummer vehicle where we get more information about the conflict. Then there is a scene outside a wall where we experience a gunfire shootout. This leads over to a new scene where Ben tells us that they prepare for the final move into Falludja. We

move into a building and the theme is about all the waiting in wars, but also the reporter anxiety of sudden rocket sound. Then we go with the Iraqi forces street by street inside the city. Part two begins with all the destructions of the city and then we are shown prison cells that Isis used to torture people. The next scene shows a dead Isis soldier and how the Iraqi mocks the dead body. In the last part, we are inside the refugee camp and Ben tells us about all the loss and pain the war has created for the civilians, and we get to meet some of them through interviews.

In general, this documentary follows a standard journalistic approach to tell the story with the weight of information through the voice-over. However, some interesting narrative techniques needs a closer look. In the opening the voice-over takes a stance as a non-diegetic narrator who describes the scene we see: "On a rooftop above Falludja, an Iraqi sniper takes careful aim at an Isis soldier…". Then at time code 01.05 The journalist introduces himself as a diegetic narrator and we also see him as a non-focalized character. On the VO he recognizes the audience as a diegetic narratee because he talks directly to us about the story world (diegesis). At the same time the character Salomon does not recognize the camera in the same way, in other words a visually non-focalized situation thus leading to a narrational mismatch between the oral and visual stimulus. This continues through the documentary as the VO is present almost all the time. Nonetheless, at TC 03:46 there is a moment of story action without commentary when an outgoing rocket is fired and journalist Solomon is shocked whereas the Iraqi soldiers laugh at him and his reaction. In this scene we are close to him and I would argue that here, the focalization is external instead of non-focalized. We experience the scene with him and therefore this becomes one of the most emotional point in the story.

After Solitary
Produced by Frontline PBS April 2017.
Directed by Cassandra Herrman and Lauren Mucciolo. Length 09:07

Image 2 (Frontline © 2017)

This is a story about Kenneth More that has spent more than 20 years in prison and many of them in solitary confinement. The documentary starts with Kenneth at home and he tells us on a voice over about his prison time and that he looks fine, but he is not. The next scene is in a cell and Kenneth is superimposed on the background while telling us about his time in solitary confinement. A superimposed text on the wall tells the backstory of why Kenny was put in jail. Next scene is outside the cell with Kenny behind the bars. He recounts how he provoked the distraction team and they were forced to do a security check on him. To illustrate this, a filmed action of a generic security check is shown on the window of the cell door. After a fade to black a new informative text tells us that, the prison started with more rehabilitation classes and Kenneth explains how he found motivation to change and finally managed to get out of prison. The last scene is in his bedroom where Kenny says he spends most of his time because it is the only place he feels safe. An ending text informs us that an ex-offender is more likely to go back to prison if he has spent significant time in isolation.

This production differs from the first one by the fact that the authors are not visible in the story. Cassandra Herrman and Lauren Mucciolo are thus extra fictional narrators. The main character Kenneth functions both as a character and a diegetic narrator. At TC 01.25 there is superimposed texts with information which belongs to the non-diegetic narrator lever. This is a quite complex narratological moment in the story. In the start of the film we see Kenneth in his bedroom and on the voice-over he tells us that he looks fine, but he is not. In the next scene we are inside a cell and Kenneth is

superimposed on this background and he describe how it was to be put in the cell for the first time. Here we have an external focalization where the character is speaking to us as an observer. At the same time there is a sound ambience from the past with other prisoners shouting and bashing the walls. This auditory belongs to the non-diegetic level since it is not a part of the scene at that time in the story. Then at TC 01.25 the text arrives and in a way brings the audience up to a non-diegetic level. At TC 04.40 is another interesting scene where we are outside the cell and Kenneth is inside looking at us through the glass. He is describing how he covered up the window to get attention because then they had to do a cell extraction on him. As he is telling us this, a generic video of a cell extraction is shown on the door. This video is thus on the non-diegetic level, the background can be read as internal focalization (us) and Kenneth voice-over is going from external focalization to diegetic narrator when he disappear visually. In general, this production tries to take into consideration that in the VR medium, the psychological perception per default, is internal focalized.

Nomads; Massai
Produced by Felix & Paul Studios for the 2016 Sundance film-festival. Later published on Gear VR (Samsung) Length 12:00.

Image 3 (Felix &Paul © 2016)

The production opens with a text about nomads since this piece is one in a series of four, about nomadic people. To enter the Massai episode you have to tap on the thumbnail and then a new introduction text appear. This text explains the background of the culture where they live and that it is a patriarchal society. The last paragraph says that nowadays they struggle with the rights of the land and have to look for other ways of income. To start the film you have to tap a thumbnail and a new text appears; God created Massai. He gave us the gift of being. We love who we are and are proud to share it. The rest of the film consists of single scenes with dip to black in between and can be described as follows:

1. On the savannah, a kid is slowly coming toward you and stop in front looking at you.
2. You sit together with a Massai group chanting in rytmh.
3. A group of young Massai is competing in javelin throwing.
4. An old man is telling a story to the rest of the family in a hut.
5. Outside a hut, an old woman sits and looks at you. In the distance, you can hear chanting.
6. In the village, everybody is sitting outside. The women are doing needlework and the kids are running around
7. You are inside a ring of man chanting rhythmically.
8. Some Massai are milking cows who seem malnourished.
9. Outside a hut, a man sits and looks at you.
10. An evening scene where the entire village is gathered together. There is a campfire and the kids learn to dance and sing from the adults.

This film is in the tradition of early travelogues and anthropological films where you do not have a voice-over or a clear story-line. The style is observational and poetic, and should therefore fit well with the VR-medium. The directors here is not visible in the diegesis thus they are on an extra-fictional level. It is also difficult to define a narrator at all since the scenes seems to appear in a random order with no clear narratological steering. An interesting aspect of this production is that the directors try to put you, as an audience, into the story-world without a strong narrative. Normally, in traditionally films, is often the narrative drive and manipulation that makes us forget ourselves, and just experience the film universe. The opening scene is emotionally strong because first, we are very alone on the savannah and we have time to look around, and then a Massai boy appears and stares directly at us. In a way, this is the first encounter with this tribe as it could have happened if we actually had travelled there in the real world. Another difference with this production is that there is no main character or plot to follow, but

still we bound with the tribe because of their recognition of us as a real character in the story world. This is particularly present in scene five and nine where our gaze are met from consecutively an old woman and a young man. From a narratological point of view this production is simple and coherent and the focalization is always on the internal (surface) focalization level.

Notes on Blindness
Produced by Ex Nihilo, ARTE France and AudioGaming for the 2016 Sundance film-festival. Later published on Gear VR (Samsung). Length ca 30.00.

Image 4 (ARTE France © 2016)

After losing sight, John Hull knew that if he did not try to understand blindness it would destroy him. In 1983, he began keeping an audio diary. Over three years John recorded over sixteen hours of material, a unique testimony of loss, rebirth and renewal, excavating the interior world of blindness.

Together with the premiere of the documentary feature film, Sundance 2016 saw the unveiling of Notes On Blindness: Into Darkness, an immersive virtual reality (VR) project based on John's sensory and psychological experience of blindness. The project won the Storyscapes Award at Tribeca Film Festival and the Alternate Realities VR Award at Sheffield Doc/Fest.

Each scene addresses a memory, a moment and a specific location from John's audio diary, using binaural audio and real time 3D computer graphics animations to create a fully immersive experience in a 'world beyond sight'.

First, there is an introduction with text and then 5 different chapter that can be summed up as follows:

1. Description with the lake (5 min)

John describes the beauty of the lake and its' surrounding and how every sound is connected to an activity: A man walking; Children playing; a car arriving. Without the activity, the image disappears.

2. Notes on weather. (6 min)

Description on how a blind person perceive the wind, sun, and thunder. Interactive elements like tapping to get the wind blowing.

3. Notes on panic (6.min)

Moving camera. A story about losing control and panicking. John tells a story about one time when he left the house and ended up having a panic attack. Interactivity where the viewer has to look on shoe steps to move the story forward.

4. Cognition is beautiful (6min)

A story about how beautiful it could be if it rained inside because then all the things in the room would be visible through the hearing.

5. The choir (6 min)

A depiction of how beautiful a choir can sound and how little important seeing is in such a context. John concludes that the privations of the sight has diminished through the years and in the end, he has adopted to a world without sight. Instead, he uses his ears as an interpreter of the world.

This production is the longest of the four and I have chosen to focus on one part since the style, form and narrative strategies are similar throughout the experience. Three aspects makes this production different from the others. The use of computer graphics instead of 360video, the use of binaural sound-design (the soundscape changes when the viewer rotate the head) and the use of interactivity. However, all this does not necessarily change the narrative possibilities.

The part I have chosen to analyze is "notes on panic". It starts with a CGI image on a veranda and writer John Hull tells us an episode where he tried to leave the house, but ended up having a panic attack because he felt that he lost control over the environment and the blindness overtook his entire body. Here the VO starts in a non-focalized manner where he tells us when and where this episode started. Then he goes to an internal focalized mode where he describe

exactly how he felt during the experience. This transition is emphasized by an interactivity action. The viewer has to look on the shoes to advance the story thus creating a perfect timing in the narrational level shift. Then we move forward in the garden as we are inside the head of John Hull with the visual CGI image to match. This prepare us for another transition when John on the voice-over says: "I had this need to break through this wall off blackness, this vail….". At the same time, the dot/pixelated universe we have gotten used to as the diegesis now fall apart and is replaced by expressive darkness with reflection trough vail-like tissue. This is the climax of the scene and I would argue that here we have an internal deep focalization because we experience in a flashback what John is describing. We are not only inside his head, but also inside his thoughts. To sum up, this VR experience use mostly internal focalization to give us a subjective feeling of how it is to be blind, not only to meet a character who is blind, but to give the viewer an experience of being blind.

Conclusion

Through the analysis of these examples, I have shown how different and complex the issue of different narration levels is. I would argue that the conventional film narrative uses the higher narration levels, and combine the bottom up and top down processing in a complex way. The viewers mind is intrigued both by the narrative schemata and by the perception psychological bottom up process of experiencing the film universe. Whereas the medium specificity of VR is prone to a more internal focalized story-world where it can be problematic to shift between higher narration levels without losing the suspension of disbelief.

Earlier in this article, I discussed how history and convention plays a big part in how an audience is interpreting a film. I would claim that the convention and grammar from film-history is also implemented in the new medium of VR. In the four VR-productions I have presented here, one can differentiate them in how much the narration levels are adopted towards the new medium and how much they lean towards film grammar and convention. The Fight for Falluja I would place very close to traditional film narrative with some moments of external

focalization. After solitary is presenting Ken as the main character and he functions both as a narrator and as character. Also here, there is an overweight of higher levels of narration that makes the suspension of disbelief difficult and it seems that the directors tend to use traditional film grammar.

The two last is more facilitated to the new medium, but they are also quite different. Massai has found a traditional style from film history that easily can be adopted in VR. They make a clear choice in the opening that the camera is recognized by the other characters in the story world. This leads to a high level of presence as defined by Barfield and Weghorst (Barfield & Weghorst, 1993). The directors here also deliberately place us very close to the Masssai tribe so we get the proximity effect (Brenton, 2015) where we are in a personal or intimate space with the characters. "Notes on blindness" uses a slightly different strategy since we do not see directly other characters or situations. This production rely heavily on the synergies between the auditory comments and the CGI animated visuals. What maybe the most interesting is the way the director handles the shift in narration level. It seems always in the combination with interactivity and thus creating both a mental and narrative tension for the viewer. This means that we get a natural narrational flow in the experience and therefore also a deeper engagement with the story.

Virtual reality is still a new medium and producer and audience have to negotiate an understanding of how the grammar of the medium should be understood. Indeed, the influence of film convention is present, but as I have shown, there is also producers that make a new kind of stories that is "Virtualistic" (meaning production that is genuine VR and employ the intrinsic characteristic of the medium). The future will show how the language of VR will be; we are still just in the beginning.

Bibliography

Bordwell, David. 1985. *Narration in the Fiction Film.* Madison: University of Wisconsin Press.

Bordwell, David, Janet Staiger, and Kristin Thompson. 1985. *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960.* New York: Columbia University Press

Branigan, Edward 1992. *Narrative Comprehension* London: Routledge.

Eisenstein Sergei 1991, *Towards a Theory of Montage.* Trans. Michael Glenny. Ed. Michael Glenny and Richard Taylor. London: BFI Publishing,

Genette, Gérard. 1988. *Narrative discourse revisited.* Ithaca: Cornell University Press

Hochberg, Julian. 1986. *Representation of Motion and Space in Video and Cinematic Displays.* In Kenneth R. Boff, Lloyd Kaufman, and James P. Thomas, eds. Handbook of Perception and Human Performance. Vol. 1: Sensory Processes and Perception.

Mandler, Jean Matter. 1984. *Stories, Scripts, and Scenes: Aspects of Schema Theory.* Hillsdale, New Jersey: Erlbaum.

Articles

Barfield W. and Weghorst S. 1993. *The Sense of Presence Within Virtual Environments: a conceptual framework.*

Bordwell, David. *A case for cognitivism.* Iris no. 9 (Spring 1989): 11-40.

Smith, Greg M. *Moving Explosions: Metaphors of Emotion in Sergei Eisenstein's Writings.* Quarterly Review of Film and Video 21.4 (October-November 2004) 303-315

VR productions

The fight for falludja (2016)
Produced by the New York Time.

After Solitary (2017)
Produced by Frontline PBS. Directed by Cassandra Herrman and Lauren Mucciolo

Nomads; Massai (2016)
Produced by Felix & Paul Studios

Notes on Blindness (2016)
Produced by Ex Nihilo, ARTE France and AudioGaming