THE ART OF 3-DIMENSIONAL CONTENT CREATION - DO 3DS AND VR SHARE THE SAME VISUAL GRAMMAR?

C. F. Orme and M. J. Uren

Ravensbourne, UK

ABSTRACT

3-Dimensional Stereoscopic (3DS) and Virtual Reality (VR) productions are generating new artistic concepts and techniques to enable them to take full advantage of the emotive power of these media that incorporate the perception of depth. This paper looks at how the authors are using innovative 'visual grammar' in some 3DS film and VR productions. They are exploring how the creative aesthetics of these media can enhance the art of storytelling, and improve the user experience for stereoscopic displays and VR headsets.

INTRODUCTION

It is only through the use of screen technologies recently developed for smartphones that VR headsets have become a potential mainstream consumer product. Most content for this new VR medium is either computer-generated gaming graphics or single-shot live action. There is very little narrative live action content for this new experiential medium, and the grammar for storytelling in VR has not yet been established.

This paper compares the nature of the visual grammar needed for generating emotive and immersive content in 3DS and VR. The approach to this research is informed by both practice-based and traditional academic research methodologies for the production of film content, both in 3DS, and now VR. The film director Sergei Eisenstein, who was a pioneer of the theory and practice of montage, believed that connecting with audiences was at the heart of good filmmaking.

> 'In the last essay Eisenstein ever wrote, the master of montage characterized stereoscopic cinema as "a dream of unity between spectator and actor".' (1)

Understanding how to create this "unity" using 3-dimensional visual grammar is the central theme of this paper and the accumulated film work of the authors.

Anorak research

The authors have between them watched and reviewed over 300 items of 3D content, either seen in a cinema, film festivals, from Blu-Ray discs, 3D television broadcasts or YouTube.

STATE OF THE ART

The following sections describe some of the 3DS techniques that have been applied by the authors in their productions. The paper then looks in more detail at their latest research findings.

Walk On The Wild Side

The authors' focused firstly on placing the viewer at the centre of the work and how audiences can be taken on a physical and emotional journey. They documented a journey to the summit of *Snowdon – Yr Wydffa* (2012) that was filmed with moving camera shots and a point-of-view (POV) style (see Figure 1). Like a theme park ride this was to simulate the experience of a journey.

1 - Snowdon POV Shot

Sculpting In Space

Secondly their investigations made use of 3-dimensional scene *volume* as a creative tool. Using character theme-based depth they demonstrated how to 'sculpt in space' in their documentary *Pave The Way* (2012) (see Figure 2).

Figure 2 - Pave The Way Landscape Shot

Out Of The Box

Thirdly they explored the idea of breaking the 2D frame and getting 'Out of the Box' using chromakey in their promo for Basement Jaxx' *Back 2 The Wild* (2013) (see Figure 3)

As Jeffery Katzenberg quoted with reference to the film *Polar Express 3D* (2004), 'The Wall of the screen

had dissolved into an open window, drawing us into the film, while also releasing the characters into the theatre.' (2)

Figure 3 - Back 2 The Wild Chromakey

Viewer As Player

An example of constructing a bridge between actors and audience was used in *The Silver Ghosts* (2013), a multi-camera film-noir musical. This experimental piece explored the nature of the passive/active viewer by giving the viewer a key role to play. One of the most significant camera set-ups in the production (see Figure 4) was from the spectator's perspective; the viewer therefore became essential to the narrative, i.e. the protagonist.

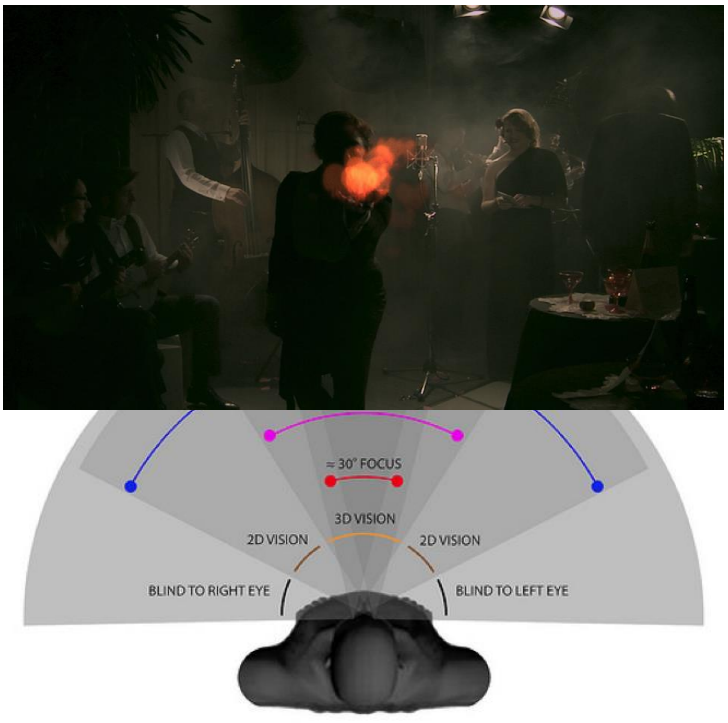

Figure 4 - The Silver Ghosts Master Wide Shot

Human 3D Vision

Figure 5 shows how stereoscopic vision is limited to a cone of about 60° in front of us. This is determined by the overlapping fields of view of our two eyes, limited by the bridge of the nose. The narrower 30° cone gives the region in which we perceive focused detail. The wider 120° cone defines the edges of 2D vision where we see shapes but no detail, and the outer edges of close to 180° contain areas blind to one eye and in which we only really see motion. We can only fuse into a 3D image those parts of a scene that are within a small angular range of the object of our attention. Objects outside of this zone will be seen as low-resolution double images that are out of focus.

Figure 5 - Human stereoscopic vision

> ‘By adding stereoscopic vision to a moving picture, we bolster the illusion of realism shown on screen considerably. Under ideal conditions, this leads to what is called ‘immersion in action’. After a few minutes of watching a 3D movie projected under optimal conditions, the viewer truly identifies with a part of the fictional world being presented. This is where ‘suspension of disbelief’ kicks in.’ (3)

> ‘The basic principle in creating an immersive experience is to fill the audience’s field of view... almost all efforts to create immersive motion picture experiences have involved increasing the amount of information presented, with larger film frames, higher frame rates, or both.’ (4)

This latter quote sums up the immersive nature of the IMAX experience, offering images that almost fill the viewer’s field of view. The average field of view in an IMAX cinema is 70° (5). This is much greater than the average 54° in a normal cinema, giving the impression that the image fills the viewer’s field of view. In the VR world, the Oculus Rift headset has a horizontal field of view of 100° (6), while the Samsung Gear VR’s is 96° (7).

Eye of the Viewer

One strand of the authors’ research centred on gaining an understanding of human visual perception (8); not just how we see the world stereoscopically (9), but how our brain interprets this into our visual sense of the world we inhabit. This can be considered to have three main components: the eyes, the visual pathways between the eyes and the third component, the visual cortex. Here the brain interprets the input it receives and projects in front of our eyes the view of the world that we perceive ourselves as being part of.

The eye is a vehicle to understanding the realities of the world and is our primary sense. Within different cultures it symbolizes and signifies: mirror, evil eye, third eye, eye of God, eye of authority and window to the soul.

> ‘As a privileged sense, the eye organ acts as mediator between interior and exterior, subject and world. Above all, second-degree vision - charged with reflexivity - was to guarantee critical access to reality. Its semantic multi-dimensionality is expressed clearly in the French language: voir (vision), savoir (knowledge) and pouvoir (power) have the same stem. The etymological relationship symptomatically reveals the dual structure of vision, a connection to reason as well as the control of power, to illumination but also the illumination of reality and therewith its surveillance.’ (10)

The Viewer’s Perspective

In both 2D and 3D filmmaking, the position of the viewer can change during a film and the viewer can move between multiple roles: observer, spectator, performer, voyeur, or witness. In the case of *Rear Window* (1954), an American suspense film, the main subjective viewpoint of ‘Jeff’ the protagonist, is from a wheelchair throughout the film. The viewer is placed into the protagonist’s role as a voyeur/spy who then becomes a witness to a crime. Another example is *The Lady In The Lake* (1946), where almost the whole film is shot as point-of-view with the viewer as the main character.

In many current examples of VR content, the narrative is constructed as either first or third person, and this perspective does not change. Laurel emphasizes the importance of encouraging the user of a technology to develop a first-person, rather than third-person,

relationship with his or her mediated environment. Engagement, which Laurel describes as a primarily emotional state with cognitive components, serves as a critical factor in engendering a feeling of “first-personness” (11). Engagement is likened to what poet Samuel Taylor Coleridge called the “willing suspension of disbelief” (12).

The question the authors are now asking is, ‘With VR, does the viewer have to be locked to one perspective?’ The authors are particularly interested in the possibility of changing a viewer’s perspective during a shot, e.g. from observer to participant. An example of this technique in 3DS is during the opening sequence of *Gravity* (2014), when the viewer is transported through the astronaut’s helmet. The viewer starts as an observer and then becomes the astronaut.

Where’s Your Head At?

Unlike other art forms, which are created in space – notably sculpture and theatre – we cannot physically move our heads or bodies in 3DS to see around an object. 3DS only reveals an image with depth as if seen from a single viewpoint. If the viewer moves their head, shear distortion (13) occurs that breaks the stereoscopic illusion, This restriction does not apply with VR where the relationship between screen and viewer is locked by the act of wearing a headset. This encourages the viewer to explore a scene by moving their head. There is a similarity here with 8K and Cinerama. Once the viewer has grasped the overall view, the image scale and detail encourages a more exploratory way of viewing a shot.

Against The Brain

Avatar (2009) was innovative with the 3D technology and to some extent challenged ideas of subject within the world.

> ‘The screen plane or the convergence point is a very important stereo tool that we have for storytelling. So with Avatar for example James Cameron locks focus and convergence, what this means is that the subject that’s in focus, always lies at the screen plane...This is a big point of contention amongst the 3D community...’cos we are not really used to seeing the world move around us. We are used to seeing subjects move around in the world which is fixed.’ (14)

The use of subject and shot convergence (15) is another creative tool the authors have used in their 3DS productions. With a virtual screen at around 20 metres in a VR headset (7), the question is whether any form of convergence is possible in VR filmmaking.

Orthographic Filming

The purest, and therefore perhaps the most realistic immersive S3D content is through orthographic 3D imaging, where all the viewing conditions of real-life stereoscopic vision are duplicated. These conditions are: camera separation (IA) = human eye separation (IO), camera image magnification = human magnification (by choosing appropriate lens focal length), matching distance to subject when imaged, and matching distance to displayed subject when viewed on a display.

> ‘An orthographic movie can duplicate the three-dimensional experience of watching a live stage show...There is no camera movement, no lens change and no editing.’ (16).

In *The Silver Ghosts* (2013) the master wide shot was close to orthographic (see Figure 4) with a slightly less than human image magnification (wider lens), and a less-than-human IA. The viewer is positioned at this camera position, simulating being seated at the performance.

> 'The definition of virtual reality is based on concepts of "presence" and "telepresence," which refer to the sense of being in an environment, generated by natural or mediated means, respectively. Two technological dimensions that contribute to telepresence, vividness and interactivity, are discussed.' (17)

In the presentation of this paper, the authors will explain and demonstrate the play-offs between IA, IO, image magnification and depth budget.

Mise-en-Scène

In *The Silver Ghosts* (2013), the scene is staged as a mise-en-scène. All of the cameras are locked off for the most part, with the action being staged through rehearsed choreography. In the last shot, the viewer finally uncovers their role as an adulterer when the singer throws a glass of champagne into audience space over them.

In the film *Dial M for Murder* (1954), Alfred Hitchcock manipulates depth as an effective storytelling device. When the heroine reaches back into audience space and grabs a pair of scissors the audience become complicit in the murder that follows. It is one of the most effective uses of 3D stereoscopic depth to date.

The authors are extending the mise-en-scène approach to VR filmmaking, with more emphasis on camera position and tracking than on cutting.

Depth of Field (DOF)

The established 2D filming grammar of narrow DOF, rapid cuts and shaky hand-held action shots is to a large extent unsuitable for generating an immersive 3D experience in 3DS. Narrow DOF is frequently used in 2D filmmaking to separate background and foreground from the main subject. Depth itself does this in a 3DS shot, so the use of DOF can be different. Slight background softening is appropriate where there is a foreground or mid-ground object of interest. This compensates for the viewer's eyes always being focused at the screen plane. However, any foreground softening can instantly break the 3DS illusion if fixated by the viewer, since soft foreground objects cannot be made sharp.

> 'The audience involuntarily looks first at the object closest to it, even if it's blurry.' (18)

Other monocular depth cues such as desaturation and reduced contrast of background elements, and effects such as mist and smoke, can be used to provide more visual separation in 3DS. Figures 1, 2 and 4 illustrate this.

Optimising Immersion

The senses we use to experience the world (touch (haptic), smell, sight, hearing, taste) would ideally all need to be stimulated in 'high-definition' to reproduce a completely realistic immersive experience. A crude binary puff of wind is probably worse than no haptic stimulation at all. Research is progressing to provide a more refined haptic sensation by using focused beams of ultrasound – UltraHaptics (19). D-Box cinema seating (20) with sophisticated motion actuation is also becoming available.

> A highly vivid medium can be considered “hot” in the McLuhanesque sense, as it “extends one [or many] sense[s] in ‘high definition.’” (21)

If we do not stimulate one or more of these senses then we are leaving it to the viewer/listener’s imagination to fill in the missing sensation(s).

> ‘This illusion [immersion] depends on successfully fooling several senses, principally sight and hearing, and eliminating or reducing any cues that would tend to break the illusion.’ (22)

Combining 4D with VR

Other areas of interest in the world of film and cinema experiences are HFR (High Frame Rate), object-based 3D audio and motion effects, with systems such as 4DX and D-Box using electronically controlled cinema seats that can move in sync with on-screen action (20). The impact of these technologies on the immersive experience of an audience in conjunction with S3D or VR needs to be the subject of further research.

Dreaming

Lucid dreams (23) may be the most immersive sensation we have. However, there is clearly a difference between a dream and VR.

> ‘Virtual reality can be distinguished from purely psychic phenomena (such as dreams or hallucinations), since these experiences require no perceptual input at all, and can be distinguished from the “real” reality as experienced via our unaided perceptual hardware, since virtual realities (unlike real realities) can be experienced only through a medium.’ (24)

Immersion and 8K UHDTV

Some coverage of the London Olympics was filmed in 8K Ultra High-Definition (called Super Hi-Vision by NHK) and one of the authors saw a demonstration in 8K resolution at BBC Broadcasting House. Zubrzycki (25) has published more technical details of the demonstration.

The cinematic-style presentation on an 8-metre wide screen was absorbing and immersive in a way not experienced with any other video-based sports content, including 3DS. The transparency of the images was breath taking. Watching such a detailed image, the viewer’s eyes are drawn to fixate on various detailed areas during a shot. But, if the camera pans while one fixates on an element in the shot, the element blurs as the viewer’s eyes try to follow it across the screen. This is because the frame rate is too low to capture motion successfully. It has been estimated (26) that at least 120 fps or possibly 150fps are needed to maintain sharpness with motion at 4K or 8K image resolutions. It also appeared to the author that close-up shots became invasive. It was like looking at someone’s face under a magnifying glass. Mid-shots felt far more natural to watch from this perspective.

Viewing 3DS

The display technology used to view 3DS content will affect the quality and comfort of the experience and the sense of immersion that can be achieved. Viewing 3D in the cinema requires the viewer to wear a pair of active or passive glasses. This is often seen as a barrier to further engagement with 3DS content.

> ‘Natural, full-parallax motion cues and consistent oculomotor cues of convergence and accommodation are likely to be the most decisive factors for a comfortable 3D viewing experience.’ (27)

The 3DS illusion relies on viewers breaking the normal linkage between convergence and accommodation. This can be done when the separation is kept within the Coupled Zone (The range of distances within which a viewer is able to maintain focus and convergence easily and comfortably due to the natural depth of field of the human eye). One question is whether differences in the coupled zone for a 3DS viewing experience at home and in a cinema impacts on the viewer’s comfort and the immersiveness of their experience?

Then there is roundness. A spherical object will only appear to be correctly spherical at one distance from a S3D display. If the viewer is closer than this distance then the object will appear to be flattened, if further away the object will be elongated. With VR headsets, the viewing distance is constant, so roundness should not be an issue.

Viewer discomfort

There have been several reports published on the potential damage to health that S3D viewing might represent. In 2012, a study at the Vision Performance Institute, Pacific College of Optometry in Oregon concluded that:

> ‘S3D viewing provides greater immersion, but it can also lead to heightened visual and motion sickness symptoms. Viewers with prior symptoms in viewing TV and computer screens are not more likely to have increased ocular and physical symptoms in 3D viewing. Young viewers incurred higher immersion but also greater visual and motion sickness symptoms in 3D viewing; both will be reduced if a farther distance and a wider viewing angle are adopted.’ (28)

Angelo Solimini’s observational study used a relatively large sample of 497 who watched a 2D and a 3D film. The study concludes:

> ‘Seeing 3D movies can increase rating of symptoms of nausea, oculomotor and disorientation, especially in women with susceptible visual-vestibular system. Confirmatory studies which include examination of clinical signs on viewers are needed to pursue a conclusive evidence on the 3D vision effects on spectators.’ (29)

However, what this and other studies fail to address is to what extent these reported levels of discomfort are caused by bad 3D filmmaking, and not to 3DS per se. While ‘Bad 3D’ can cause symptoms of discomfort and eye fatigue, ‘Bad VR’ may allegedly cause disorientation, loss of balance and physical retching. It is apparent that more research is needed, driven by a better understanding of the 3DS and VR phenomena. This is a key area of investigation for the authors, who are currently in pre-production on *Poundline* (working title), a docudrama for dual platforms of 3DS and VR.

CONCLUSIONS

The viewing experience for 3DS and VR content can be improved by considering what visual grammar techniques are applied at the creative stages of the pre-production, production and post-production. In their research, the authors have derived some innovative grammar rules, and these will be demonstrated with 3DS film clips during their presentation. They are investigating whether the immersive experience of 3DS and VR can be enhanced further with '4D' effects like synchronised seat movements, and apparent touch (haptics).

The optimisation of visual grammar rules and the use of new creative aesthetics will enhance the viewers' quality of experience, and open up the opportunity for '3D' to regain its acceptance in the consumer marketplace.

REFERENCES

1. Mendiburu, B., Pupulin, Y. and Schklair, S., 2012. 3DTV and 3D Cinema: Tools and Processes for Creative Stereoscopy. Elsevier, p.143

2. Pennington, A. and Giardina, C., 2012. Exploring 3D: The New Grammar of Stereoscopic Filmmaking. Focal Press. p.ix

3. Michel, B. J., 2013. Digital Stereoscopy: Scene to Screen 3D Production Workflow. Stereoscopy News. p.1

4. Hyder, J., 2008. What is “Immersive”? LF Examiner. Available at: www.lfexaminer.com/20090522a.htm

5. www.imax.com/community/blog/imax-101-theatre-geometry-video/

6. www.wearable.com/oculus-rift/how-oculus-rift-works

7. www.samsung.com/global/microsite/gearvr/gearvr_specs.html

8. Sekuler, R. and Blake, R., 1994. Perception. 3rd ed. McGraw-Hill

9. Mendiburu, B., 2009. 3D Movie Making: Stereoscopic Digital Cinema from Script to Screen. Focal Press. ch.2

10. Schmidt-Burkhardt, A, 2002. The All-Seer: God’s Eye as Proto-Surveillance. CTRL [SPACE]: Rhetorics of Surveillance from Bentham to Big Brother. MIT Press. p.18

11. Laurel, B., 1993. Computers as theatre. Addison Wesley. p.113

12. Coleridge, S. T., 1817. Biographia Literaria. The Watchman. ch.14

13. Woods, A., Docherty, T. and Koch, R., 1993. Image distortions in stereoscopic video. Stereoscopic Displays and Applications IV, Proceedings of the SPIE Volume 1915. Available at: www.andrewwoods3d.com/spie93pa.html

14. Uren, J. Interview 22/12/2011

15. Block, B. and McNally, P., 2013. 3D Storytelling: How Stereoscopic 3D Works and How to Use It. Focal Press, p.120-127

16. Block, B. and McNally, P., 2013. 3D Storytelling: How Stereoscopic 3D Works and How to Use It. Focal Press, p.16

17. Jonathan, S., 1993. Defining Virtual Reality: Dimensions Determining Telepresence. SRCT Paper #104, p.3

18. Block, B. and McNally, P., 2013. 3D Storytelling: How Stereoscopic 3D Works and How to Use It. Focal Press, p.151

19. Carter, T., Seah, S., Long, B., Drinkwater, B. and Subramanian, S., 2013. UltraHaptics: Multi-Point Mid-Air Haptic Feedback for Touch Surfaces. UIST 2013, October 8-11. Available at http://dx.doi.org/10.1145/2501988.2502018

20. Hazelton, J., 2013. The big attraction. Screen International, April 2013, pp.30-31

21. McLuhan, M. (1964). Understanding media: The extensions of man. Penguin. p.36

22. Hyder, J., 2008. What is “Immersive”? LF Examiner. Available at: www.lfexaminer.com/20090522a.htm

23. Post, T., 2013. Lucid dreaming. University of Twente. Available at: www.lucidipedia.com/lucid-dreaming/

24. Jonathan, S., 1993. Defining Virtual Reality: Dimensions Determining Telepresence. SRCT Paper #104. p.17

25. Zubrzycki, J., 2012. The Olympics in Super Hi-Vision. BBC Research & Development Blogs. Available at: www.bbc.co.uk/blogs/researchanddevelopment/2012/08/the-olympics-in-super-hi-visio.shtml

26. EBU, 2013. Subjective testing confirms the importance of frame rate for UHDTV. Geneva: EBU Broadcast Technology Futures (BTF) Group. Available at: http://www3.ebu.ch/cms/en/sites/ebu/contents/knowledge/technology/news/201308/subjective-testing-confirms-impo.html

27. Reichelt, S., Haussler, R., Futterer, G. and Leister, N., 2010. Depth cues in human visual perception and their realization in 3D displays. Three-Dimensional Imaging, Visualization and Display 2010, Bahram Javidi and Jung-Young Son, Editors, Proc. SPIE, Vol. 7690.

28. Yang, S., Schlieski, T., Selmins, B., Cooper, S., Doherty, R., Corriveau, P. and Sheedy, J., 2012. Stereoscopic viewing and reported perceived immersion and symptoms. Optom. Vis. Sci. Vol.89, Issue 7, pp.1068-80, July 2012

29. Solimini, A., 2013. Are there side effects to watching 3D movies? A prospective crossover observational study on visually induced motion sickness. Plos One. Available at: www.ncbi.nlm.nih.gov/pmc/articles/PMC3572028/

Films

Avatar (2009). Directed by James Cameron. 20th Century Fox. Lightstorm Entertainment

Back 2 the wild teaser (2013). Directed by Caroline Orme. Available at: http://www.youtube.com/watch?v=93h3GjbyTGE

Dial M for murder (1954). Directed by Alfred Hitchcock. Warner Brothers

Gravity (2013). Directed by Alfonso Cuarón. Warner Brothers

The lady in the lake (1946). Directed by Robert Montgomery. MGM

Pave the way (2013). Directed by Caroline Orme. Available at: www.youtube.com/watch?v=p2q7XlMTBBw

Polar express (2004). Directed by Robert Zemeckis. Castle Rock Entertainment

Rear window (1954). Directed by Alfred Hitchcock. Paramount Pictures. Patron Inc.

Snowdon – Yr wydffa (2012). Directed by Caroline Orme. Available at: www.youtube.com/watch?v=aPKORcamGm8

The Silver Ghosts (2013). Directed by Caroline Orme

ACKNOWLEDGEMENTS

The authors would like to thank colleagues at Ravensbourne for their continued support for their work.