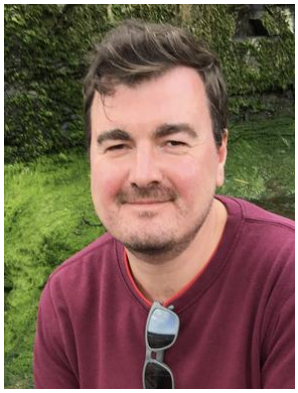
Regulating Observers: Visuality and the influence of scopic regimes on the experience of virtual space

Dr Christopher Fry, BA(Hons), MA, PhD, FHEA
Westminster School of Art, University of Westminster
e-mail: c.fry@westminster.ac.uk



Abstract

This paper examines visuality and the 'scopic regimes' that influence how we experience digital spaces such as those offered by 'Virtual Reality' and 'Augmented Reality'. These 'regimes' are socially constructed ways of seeing which lie outside the mechanics of vision and affect not only what we see but also how we understand our relationship with visual media.

This paper considers the particular influence of the camera and how alternative conceptual models might offer a different approach to creating experiences of virtual space. This includes the ancient 'extramissionist' understanding of vision and 'emptiness'.

1. Introduction

In the influential collection of essays 'Vision and Visuality', Hal Foster and others describe how 'visuality', distinct from the mechanics of vision, has shaped not only how we see but how we are made or allowed to see [1]. Certain conventions and devices, sometimes referred to as 'scopic regimes' [2] or 'perceptual schemas' [3], shape how we receive and understand media. Perhaps among the most well known and widely critiqued of these 'schemas' is perspective. While often seen as enabling an objective depiction of reality, Panofsky and others have shown how perspective is simply one 'reality effect', and that far from being definitive, several models of perspective exist [3]. Norman Bryson likens the role of visuality to a screen inserted between the retina and world consisting of "all the multiple discourses on vision built into the social arena" [4, p92]. This paper will look at how some of these 'scopic regimes' may regulate our experiences of virtual spaces such as those found in Virtual Reality (VR) and Augmented Reality (AR).

Although there are many influences on the development and experiencing of VR and AR, including references to prior media, this paper will examine the influence of the camera and lens media. This includes the camera as a metaphor, the related concept of the 'viewer/camera' and the potentially constraining nature of Flusser's photographic 'gesture'. The camera is chosen not only for its widespread

occurrence but as it can be seen to influence both the making and reception of virtual spaces and experiences.
Not all VR aims to create lifelike representations of the 'real world' or employ lens-based approaches. This paper will not aim to address how experiences can be made more 'real' or 'immersive'. It is not intended to offer a comprehensive survey of work in this area. Nor does it suggest that the camera is the only influence on the way virtual experiences are made and experienced. Instead, it uses the camera's influence as a way of thinking about other models and metaphors and what they might offer.

A range of terms and acronyms have emerged to describe virtual spaces and technologies including VR and AR. Terms such as Mixed Reality and XR are sometimes used as umbrella terms. They are also used to describe hybrid and complex combinations of VR, AR and other elements such as live performance and installation. Slippages between these different categories often occur, especially where they are employed as marketing tools. Here the abbreviations of VR, AR and XR are suggestive of a technological avant-garde, frequently used alongside terms such as 'immersive', 'interactive' and 'experience'. Such slippages are unhelpful as they imply a particular experience and draw attention away from more meaningful consideration of the qualities of the experience. The common term is generally 'reality'. This paper will not hope to address what constitutes 'reality' or the 'real'. However, what the term reality does suggest is an apparent link to previous media and 'reality effects'.

2. VR as a Continuation of Reality Effects

It would be easy to see VR as a continuation of several technologies used to capture and depict reality. Such a lineage might include the use of perspective in painting, photography, cinema, television and so on. Jonathan Crary notes how a continuity is often traced in which "Renaissance perspective and photography are part of the same quest for a fully objective equivalent of 'natural vision'" and in which the camera obscura and cinema are seen as "enduring apparatus of power" that continue to "define and regulate the status of the observer" [5, p30].

Each 'new' media seemingly fixes an imperfection in what came before through a process of remediation that brings us closer to the 'transparent immediacy' of pure and unmediated reception [6]. While clearly an oversimplification, the links to prior media and particularly lens-based media are regularly noted. Histories of the development of VR typically include reference to the panorama, stereograph and photographic camera [7], while histories of photography and cinema typically begin with the camera obscura [8, p26].

As Crary points out, rather than there being a continuous link between the camera obscura and later lens media such as photography and cinema, new understandings of the mechanics of vision in the 18th and 19th century dismantled the camera obscura as a model for our understanding of vision. While the camera obscura provided a model for accessing an 'objective truth', key to undoing the model was an

acceptance of our ability to misperceive [5, p39]. Crary points to the example of experiments which showed how electricity applied to the optic nerve could produce the experience of light [5, p39]. Nevertheless, an assumed connection between the camera obscura, photography and cinema persists based largely on certain formal qualities [8].

Not all VR, AR or virtual experiences aim to faithfully reproduce the 'real' world. What is useful to note is the connection to other lens media and to consider what these media bring with them, including assumptions about things such as image quality and fidelity. Much time and effort are given to refining VR technology, correcting perceived flaws. A typical research paper from this area describes how "In a perfect virtual world, each eye would be presented with exactly the visual geometry that it would be exposed to when viewing a real scene" [9]. What this seems to suggest is that there is a perfect 'real scene' that can be worked towards. This brings with it a danger that we assume all issues can be resolved technologically, through greater resolutions, faster refresh rates and correcting perceived imperfections. As Baudry and Williams note, imperfections and limitations such as 'depth of field' imply a particular conception of reality in the first place [10]. The idea that reality can be perfectly and definitively represented is far closer to the camera obscura model than later understandings of vision.

It is important to note that lens media are not the only influence on the development of VR and virtual experiences. The immersive theatre experience 'Draw me close' by the National Theatre in the UK is an example that combines elements of theatre, drawing and animation [11]. It creates a complex and compelling hybrid that makes significant use of haptic and tactile feedback. Haptic and other sensory feedback are not considered here. However, as this example shows, other models and media disciplines can and do inform practice in this area.

3.The Scientific Character of VR

If the link between the camera obscura and VR may be largely superficial, what they do seem to share is their scientific flavour.

Michael Punt has noted the relationship between science, technology and early cinema in describing the technological imaginary of early cinema [12]. As an apparent descendant of cinema, VR has a similar character and its own technological imaginary. I have previously written about the influence of computational bias on our experiences of media [13], a subject that has been widely noted by James Bridle among others [14]. This includes how we are inclined to accept 'auto focus' and 'colour correction' for their convenience, as well as our privileging of algorithmic judgement over our own. As computational media VR and AR are framed by this wider context. However, they also have a particular character as an optical technology which may have its own implications for their visuality.

In reflecting on the cinematic apparatus, Baudry and Williams ask: "Does the technical nature of optical instruments, directly attached to scientific practice, serve to conceal not only their use in ideological products but also the ideological effects which they may

provoke themselves?" [10, p40], adding, "Their scientific base assures them a sort of neutrality and av oids their being questioned" [10, p40]. Arguably the scientific and technological character of VR and A R, seen in the way that such experiences are marketed, may have a similar effect. The Saatchi Gallery website describes how in the hugely successful immersive experience 'We Live in an Ocean of Air' "cutting-edge technology illuminates the invisible- but fundamental- connections between human and natural worlds" [15]. Meanwhile 'Draw me close' is described as "Weaving theatrical storytelling with cutting-edge technology" [11]. The link between 'cutting-edge technology' and new or extended ways of seeing seems clear. To what extent are we prepared to defer our judgement about how these experiences are constructed in the same way that we may defer judgement to other computational assistants such as auto focus?

Optics and scientific understanding of how the eye works is clearly of benefit to creating virtual spaces and experiences. Textbooks such as 'Practical Augmented Reality' often include chapters on optics and the working of the eye [16]. However, this scientific explanation of vision alone cannot account for the visuality of the experience of seeing. Perhaps because of the scientific and technological character of the processes involved, assumptions are made about the objectivity of the experiences offered and their ability to replicate how we see. A mistaken assumption that we see like a camera, or perhaps that the camera sees like us.

4. The Viewer/Camera

Cameras are involved in many aspects of making virtual spaces, from photogrammetry used to capture and create 3D models, to the cameras on devices that enable AR. Meanwhile virtual cameras can be found in many of the software tools used. The documentation for the widely used game engine Unity describes how: "cameras are your eyes into an interactive experience" [17]. The concept of the virtual camera is an i mportant one in creating virtual 3D environments but also brings with it and implies a certain visuality. It blurs the distinction between the camera and the viewer creating a new entity, the 'viewer/camera'.

In many software tools virtual cameras are positioned and controlled to provide a viewpoint, or as the Unity quote suggests, a window or portal through which the viewer can look onto the created space. The intertwining of camera and viewer often occurs through an intricate nesting of objects and components. Complex hybrids of what are referred to as 'actors', 'avatars' and 'cameras' create an i nterlacing of the viewer, camera and n avigation controls. These allow for the kind of first person or POV experiences with which we are familiar.

These are all clearly needed for the creation and functioning of certain types of experience. However, they also imply a particular way of thinking about the viewer and is far more constrained than the illusion of autonomy that they create within the experience. The 360-degree field of view seems to break from the constraints of the camera or cinema frame and allow the viewer to select their viewpoint. However, this is an i dealised or intended view sometimes referred to or defined by parameters such as 'targetview'.

This borrowing of concepts is an example of the encoding of existing tools described by Lev Manovich's computer as 'metamedium' [18]. Understood this way, game engines such as Unity and Unreal incorporate earlier mediums including, animation, cinematography, typography and photography bringing with them tools, workflows and concepts. Although as Manovich notes, as well as incorporating and digitising existing media and processes, the 'metamedium' can create new hybrid forms. The 'viewer/camera' could be seen as one such hybrid, albeit one that is often overlooked by its apparent familiarity.

The 'viewer/camera' is not only found in software tools. The concept of the 'viewer/camera' is used to describe and plan the production of Cinematic Virtual Reality (CVR) experiences [19]. This model places the viewer at the centre within the 'frame', the director arranging the scene around them.

Cinema is often described as a passive experience as the audience submit to the images which provide them with a 'God's eye view' that they cannot themselves control. CVR seemingly offers a more active role by allowing the viewer to look around and control what they see, but at the expense of the God's eye view since they may look away and miss things [19].

Viewing a film involves drawing on a learnt understanding of the grammar of film, allowing us to turn the cuts and camera movements into a unified experience [19]. The challenge for the makers of CVR is to develop new grammars and techniques to guide the viewers' attention. This raises the question of how the viewer understands their role and how learnt grammars inform the experience. Turning our head in the 360 environment when understood as a cinematic experience may lead the viewer to search for the 'correct' or significant shot. It has turned us, if not into a director, then perhaps into a cinematographer. The viewer becomes a camera operator.

5. Seeing Like a Camera

Where the 'viewer/camera' is often found in the enclosed experiences of VR, AR employs the camera in a different way and more commonly sits between the viewer and the virtual experience. While technologies such as the HoloLens or Google Glass may promise their own form of the 'viewer/camera', AR is typically experienced using a camera phone or other handheld device.

The Google AR website describes how AR lets you "experience digital content in the same way you experience the world. It lets you search things visually simply by pointing your camera at them" [20]. Here we can see a tacit assumption that to experience the world is to do so through the camera, or at least that to experience through the camera is normal. Given the ubiquity of cameras and camera phones it is easy to imagine how we might be allowing the camera to mediate our experiences. The camera phone is used to augment our memories, to record our existence in the world and share it. It is important to note that these are networked devices and so the images that we create and receive are understood as part of a network of images [21]. AR appears to offer us the ability to extend our experience, summoning from the cloud virtual objects and manifesting them. In this way seeing through a camera phone is not a discreet experience between subject and object

but understood as entering into a shared digital space. However, this is also the logic of the camera, framing and extending its logic into real space.

In the software tools used to create VR and AR experiences, the virtual camera is typically positioned within a grid space and coordinate system. This regularly divided and seemingly infinite space has its own visuality connected to that of perspective techniques. The grid space is a dominant schema that occurs not just in 3D game engines and 3D modelling software such as Blender but has extended even to tools more associated with 2D imagery such as Photoshop. Lev Manovich in describing the 'automation of sight' and 'computer vision', refers to the 'perspectival machines' and 'geometry engines' of computational media [22]. Meanwhile, Damjan Jovanovic refers to the 'ground grid', marking out a uniform grid space, homogenising space and suggesting a certain 'total visual empowerment' [23]. These grid spaces are empowering due to their suggestion of limitless and granular control. However, they also constrain, relying on a single viewpoint and typically requiring us to think according to the logic of the grid space.

To understand the virtual object inserted into our space is to understand our surroundings according to the regulating grid space projected out into the world. AR experiences will often begin with a calibration process in which the grid space of the AR world is mapped onto and aligned with the space the viewer occupies (figure 1). Rather than the viewer being able to determine exactly where a virtual object can be placed, it is the device that has the final say on what is acceptable.

To understand the tension between the empowerment of AR and the control asserted by the camera through the grid space we could turn to Flusser's description of the 'gesture' made with the camera 'apparatus' [24, p198]. For Flusser the photographic gesture sees us search for a viewpoint through the camera view finder. This seemingly offers limitless possibilities but, Flusser argues, actually takes place within the "possibilities offered by the apparatus" [24]. When viewed this way, AR requires and is limited by the projection of the grid space reducing our understanding of the world to one of flat planes.

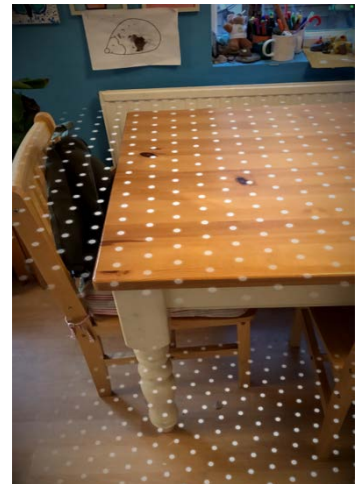


Figure 1. Screenshot of AR app calibration

It is interesting to note that the patents for AR technologies seem to show a f ar more active and controlling role than Flusser's gesture. The technical drawings submitted in support of patent applications for VR and AR technologies such as those shown in figure 2, show the eyes of the user connected to the proposed device or technology by straight lines as though they are projecting their will into reality. The patents aim to show the need for such a

technology rooted in an active human intention. As Baudry and Willams wrote of the cinematic apparatus, an extension seemingly "grafted on to replace his own defective ones" [10, p46]. On the one hand we have the technological imaginary of humans commanding these technologies as extensions of ourselves. On the other we have Flusser's warning that we may mistake the limits of the apparatus as the limits of possibility.

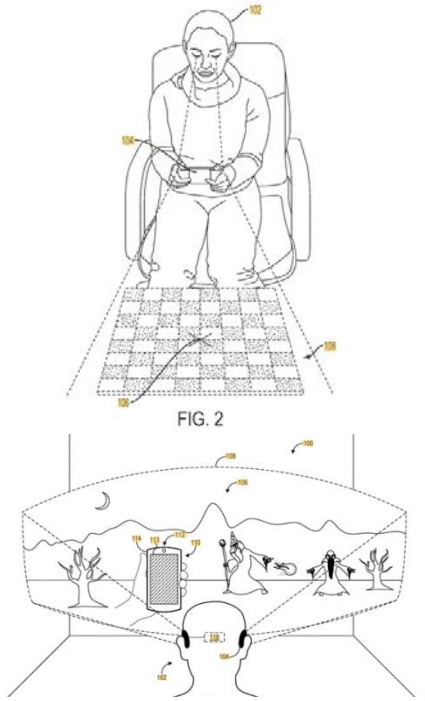


Figure 2. Images from patent applications for VR and AR technologies

The lines between the user's eyes and devices the patents show could also be interpreted as intention, an important aspect of vision that more scientific and technological explanations may overlook. Iain McGilchrist notes the importance of intention for vision and how this changes the way that we see, what we see and how we understand seeing arguing: "attention is modified by the intention that lies behind it" [25, p169]. In dismissing the camera metaphor for vision, McGilchrist argues "we never just 'see' something in the sense that a photographic plate receives rays of light. In the real world we bring a lot of ourselves to the party" [25, p165].

This would seem to reinforce the need to consider visuality and the contexts brought to the experience. However, the question remains, to what extent is the intention handed to us and shaped by notions of technology, relationships with the camera and understandings of prior media?

6. Extramissionist Drawing

Considering these issues and de bates led to the question: What are the alternatives to some of the more prevalent metaphors and scopic regimes?

An alternative understanding of vision that seems to resonate with a more active and i nterpretive seeing can be found in the ancient extramissionist theories of sight. Extramissionist explanations of vision describe rays of light emanating from the eye and landing on the objects we see [26]. Thinking about light 'emanating from the eyes' rather than simply falling into them suggests a d ifferent role for the viewer. The rays might be i magined in much the same way as the lines that appear in the patent diagrams, suggestive of the intentions and will of the viewer. It also casts seeing as an act of discovery and interpretation. McGilchrist notes Plato's description of a s tream of gentle light from within merging with the light from objects to form one body [25]. Meanwhile Bryson describes how the signifier operates on light and y et has no light of

itself only “the light it borrows from my eye” [4, p91]. While bearing no relationship to the mechanics of vision, it may be a metaphor that better describes how it feels to see.

This understanding of seeing as a tactile sensory experience also resonates with descriptions of the haptic qualities of drawing. The inherently ambiguous and felt nature of drawing offers another way of exploring virtual space, and one that may bear closer relation to our actual experience of seeing. Nicolas Mirzoeff describes seeing as more like a rapidly drawn sketch than a photograph and how certain effects in painting such as ‘papillotage’ or ‘blinking’ accepted the constructed nature of sight including the role of eye movement [27]. Drawn virtual spaces already exist and can be seen to successfully exploit these qualities, including the previously mentioned ‘Draw Me Close’.

To explore how an extramissionist model might be applied, a set up was devised to produce ‘drawings’ of spaces and objects using the ‘emitted light’ of a depth camera. This employed a depth camera, such as the RealSense or Kinect, and custom code written using Processing. It is interesting to note that depth cameras such as the RealSense and other 3D scanning tools send out their own rays of light, albeit infrared light, in order to measure depth.

By holding the depth camera and moving it through space and around objects, depth data was captured and turned into a motion sequence that could be viewed in 2D or VR. Rather than capturing all the depth data, only a focused section was used. This created a point of focus that could be traced over the environment or objects, synonymous to the tracing of light from the eye. Multiple depth cameras could be used simultaneously in an attempt to move away from the notion of a single viewpoint. The depth threshold and the limits of the hardware’s capabilities were pushed such that it created anomalies and visual artefacts that could be considered as errors or imperfections (figure 3). Parallels might be drawn here with the ‘papillotage’ effect, mimicking, or at least reflecting the mechanic imperfection of sight.

These ‘extramissionist drawings’ are intended to be records of seeing. They capture a process which combines the framing action of the camera, with a tactile notion of sight and the observation and attention of drawing. They aim to show that the capturing of space does not need to adhere to the apparently ‘objective’ logic of the camera, their constructed nature being very apparent.

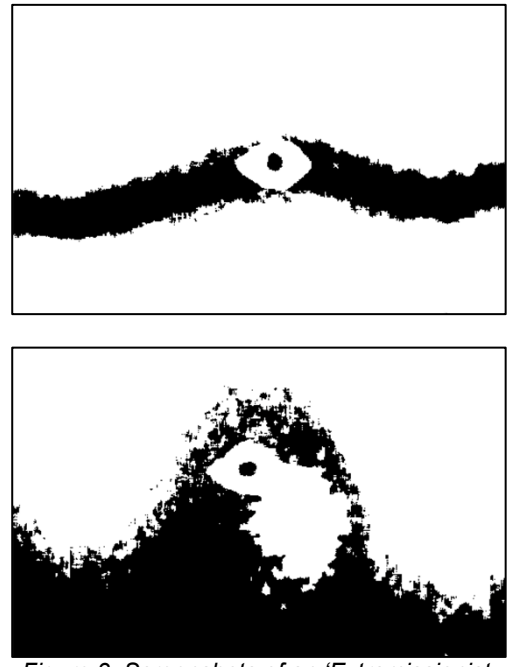


Figure 3. Screenshots of an ‘Extramissionist Drawing’ of an oil lamp

7. Emptiness and Immeasurable Space

Drawing at its simplest might be defined as a mark on a ground [28]. To this end a white background was used to provide a contrast to the 'marks' produced by the 'drawing' process. However, this seemingly 'empty' space also suggests an alternative to the constraints and predictability of the 'grid space'.

Norman Bryson describes how Keiji Nishitani's critique of the centring of the viewer found in Lacan and Sartre, put forward the concept of 'emptiness' [4]. This countered the framing process commonly found in painting and photography and which resonates with many forms of VR and AR. Bryson describes how "only that which appears within the framing apparatus – perspective, picture frame, camera – exists; the viewer on one side, the subject on the other" [4, p100].

Nishitani's approach is "to dissolve the apparatus of framing" [4, p100]. In doing so "The viewer is pulled away from the aperture of the viewfinder or lens and redefined as radically dis-framed." [4, p100]. As an example, Bryson refers to the 'Flung Ink' drawings of Sesshū Tōyō (figure 4). The cast ink creates forms that float on a " field of nihility or emptiness" breaking out of the tunnel vision that fixes object and subject [4, p95]. The rapidity and ambiguity of the marks deny the 'framework of control' that more regular marks might suggest.

Figure 4. Haboku-Sansui (1495) by Sesshū Tōyō (Wikimedia)

Dissolving the frame may seem to be equivalent to VR's removal of the screen in favour of viewer control and choice. However, this would overlook the centred nature of the viewer and the 'viewer/camera' metaphor, tethered to a subject. Far from breaking free from the frame, the VR and AR camera operator is in a constant act of framing.

Rather than using the grid space to construct a total and apparently full or continuous visual scene, emptiness can allow for a more ambiguous space. Bryson describes how when grids are absent "space itself becomes immeasurable" [29, p43].

An immeasurable space runs counter to the regular and regulated spaces of VR that rely on a quantification of space to be rendered. It draws attention to the mutability of our perception of space. A similar effect can be seen in the example of the Roman painted interiors such as the Black Room from Boscotrecase which play with the "ontological thresholds" between the represented and

reality [29]. These murals set hints of architectural and figurative elements on a field of black. The effect is to cause the eye to “move from the impalpable and insubstantial blackness […] to images set at some indeterminate distance within it” [29, p43]. Such ‘fictive spaces’ can create a sense of limitless space precisely because of the removal of scale [29].

8. Conclusion

The technological nature of virtual experiences can overshadow the consideration of other qualities. While undoubtedly capable of offering radically new experiences, they inevitably draw on conceptual models and ways of working received from prior media. These shape both the making and the reception.

The approach outlined here is not intended as a replacement for the dominant ‘perceptual schemas’ or ‘scopic regimes’. Instead, it aims to open up discussion about what approaches might be developed at a time when the new languages and grammars of VR and AR are still being negotiated. The breadth of work in the field suggests there cannot be a unified approach. As a result, it pays to be aware of the options available rather than risk defaulting to existing approaches and allowing the audience to do the same. New possibilities will be created by looking beyond the most obvious and established models and metaphors. Only this way will it be possible to take full advantage of the ‘metamedium’ and create new hybrids and new ways we are ‘allowed’ to see.

References

[1] Foster, H. (1988) ‘Preface’ in Foster, H (ed.) *Vision and Visuality*, Dia Foundation, Bay Press, Seattle

[2] Jay, M. (1988) ‘Scopic Regimes of Modernity’ in Foster, H (ed.) *Vision and Visuality*, Dia Foundation, Bay Press, Seattle

[3] Panofsky, E. (1997) *Perspective as Symbolic Form*, translated by Christopher S. Wood, Zone Books

[4] Bryson, N. (1988) ‘The Gaze in the Expanded Field’ in Foster, H (ed.) *Vision and Visuality*, Dia Foundation, Bay Press, Seattle

[5] Crary, J. (1988) ‘Modernizing Vision’ in Foster, H (ed.) *Vision and Visuality*, Dia Foundation, Bay Press, Seattle

[6] Bolter, Jay David, and Richard Grusin. *Remediation: Understanding New Media*. Cambridge, MA: MIT Press, 1999

[7] Tricart, C (2017) *Virtual Reality Filmmaking* Routledge

[8] Crary, J. (1992) *Techniques of the Observer*, MIT Press

[9] Jaekl, P.M. et al, (2002) “Perceptual stability during head movement in virtual reality” *Proceedings* IEEE Virtual Reality 2002, pp 149-155

[10] Baudry, J.L. and Williams, A. (1975) ‘Ideological Effects of the Basic Cinematographic Apparatus’ in *Film Quarterly*, Winter, 1974-1975, vol. 28, No. 2, pp39-47, University of California Press

[11] National Theatre (2020) *Draw Me Close.* Website: https://www.nationaltheatre.org.uk/immersive/projects/draw-me-close accessed November 2022

[12] Punt, M. (2000) *Early Cinema and the Technological Imaginary* The Postdigital Press

[13] Fry, C. (2019) 'Visuality and the haptic qualities of the line in generative art', *XXII Generative Art Conference Proceedings*, available at: https://www.generativeart.com/GA2019_web/10_C_Fry_2_168x240.pdf accessed November 2022

[14] Bridle, J. (2018) *New Dark Age: Technology and the end of the future* Verso: London

[15] Saatchi Gallery (2018) *We Live in an Ocean of Air*. website: https://www.saatchigallery.com/exhibition/salon_009__we_live_in_an_ocean_of_air accessed November 2022

[16] Aukstakalnis, S. (2016) *Practical Augmented Reality: A Guide to the Technologies, Applications, and Human Factors for AR and VR*, Addison-Wesley Professional

[17] Unity (2022) *Creative Core: Cameras* Online course. Website: https://learn.unity.com/project/creative-core-cameras accessed November 2022

[18] Manovich, L. (2013) *Software Takes Command* Bloomsbury Academic

[19] Dooley, K. (2021) *Cinematic Virtual Reality: A critical study of 21st Century approaches and practices*, Palgrave Macmillan

[20] Google (2022) *Augmented Reality* website: https://arvr.google.com/ar/ accessed November 2022

[21] Rubinstein, D and Sluis, K. (2013) 'Concerning the Undecidability of the Digital Image' in *Photographies*, 6:1, 151-158, Taylor and Francis

[22] Manovich, L. (1997) *Automation of Sight from Photography to Computer Vision*, available at: http://manovich.net/index.php/projects/automation-of-sight-from-photography-to-computer-vision accessed November 2022

[23] Janovich, D. (2016) 'Fictions: A Speculative Account of Design Mediums', in *Drawing Futures*, conference proceedings, The Bartlett School of Architecture

[24] Flusser, Vilém (2012) [1980], 'Towards a Theory of Techno-Imagination', *Philosophy of Photography (POP) 2* (2), p. 198.

[25] McGilchrist, I. (2009) *The Master and His Emissary: The divided brain and the making of the modern world* Yale

[26] Bielfeldt, R. (2016), Sight and Light: Reified gazes and looking artefacts in the Greek cultural imagination', in Michael Squire (ed.) *Sight and the Ancient Senses*, Routledge

[27] Mirzoeff, N. (2015) *How to See the World* Pelican

[28] Van Alphen, E. (2008) 'Looking at Drawing: Theoretical Distinctions and Their Usefulness' in Garner, S. *Writing on Drawing: Essays on Drawing Practice and Research*, Intellect, Bristol

[29] Bryson, N. (1990) *Looking at the Overlooked: Four Essays on Still Life Painting* Reaktion Books