**Deakin University**

**Rose Woodcock**

# Predatory vision: 3D imaging and the transformation of screen-space

Abstract:

Despite Wheatstone's academic interests in the device, the stereoscope languished somewhat as an optical toy. Yet the advent of 3D screen-spaces for home and mass entertainment suggests today's consumers and practitioners of screen culture hold the view that screen culture will be 'improved' through 3D imaging technologies. Like cinema and photography, stereoscopic 3D imaging has the potential to transform visual culture. But what is transformed, as optics and electronic imaging techniques deliver *Alice in Wonderland* in 3D? This paper links the advent of 3D cinema and TV to the notion that vision is itself a 'technology of the visual'. As such, our innate binocular stereoacuity is ripe for exploitation by developers of 3D imaging technologies. I argue that contemporary 3D imaging marks an epistemological visual-perceptual shift: toward screenspaces becoming spaces for potential action. Such a shift entails seeing as *doing* rather than seeing as thinking. 3D imaging exploits binocular vision's spatial acuity (stereopsis), but is effective only for objects within near distal space. The 3D effect tapers off dramatically for objects only some metres away, because the two retinal images lack significant lateral disparity (difference) to trigger stereopsis: the imagery flattens out and becomes 'monoscopic'. Information available from conventional 2D media entails a peculiarly unspecified spatiality. Perceptually, the contents of a conventional cinematic screen are like those of a painting: they are situated neither near nor far, and constitute a shared and ambiguous visual space. Our own eyes are like those of a cat: frontally placed for predatory action. The visuality of 3D screen-spaces assumes a perceptuality of the near-by and close at hand, since this is the structure of the visible information to which stereopsis is adapted to respond. Noting the binocular acuity of predatory animals, as well as some etymological links, this paper examines the implications of perceptually 'capturing' the sensation of visually solid objects in one's immediate space. Stereopsis is about decisive action within an immediate environment: but it also presupposes the single viewpoint of an active observer toward which the 3D imagery is targeted.

Biographical note:

Rose Woodcock lectures in Animation and Digital Culture in the School of Communication and Creative Arts, Deakin University. Areas of focus include stop-motion animation, compositing, and motion capture. Rose's teaching emphasises

experimentation through exploration of digital and analogue techniques. Of particular interest is the connection between observation, drawing and the visual representation of form and motion. Rose is currently completing PhD research into experimental stereography, investigating pictorial perception as a methodology to critique and understanding the possibilities of stereo-immersive Virtual Reality (VR). This research is the culmination of a long interest in the 'presence-and-absence' of depicted objects in visual imagery, from high end VR to the humble landscape painting.

Keywords:

Stereoscopy – Visual perception – Aesthetics – Epistemology

## Introduction

Charles Wheatstone (1802-1875) was elected a Fellow of the Royal Society of London in 1836. As a scientist and inventor, Wheatstone is best remembered for having devised a means to create the illusion of three-dimensional 'stereoscopic' pictures. Wheatstone's academic interest in the stereoscope centred on the information it revealed for understanding binocular vision. Noting in a 1838 lecture at the Royal Society that the lateral displacement of retinal images in binocular vision (stereopsis) is used by the visual system to detect depth in a given scene, Wheatstone proposed calling the device a stereoscope 'to indicate its property of representing solid figures' (Wheatstone 1838: § 2); a reference to the perceptual solidity of the three-dimensional information uniquely available from stereoscopic vision. Thereafter, the stereoscope languished somewhat as an optical toy. In the meantime, studies of the mechanisms and affordances of human binocular vision were consolidated, and the workings of visual perception are now well understood in contemporary vision research. Among such findings is that the ability for a perceiver to judge distances with accuracy (binocular stereoacuity) drops off quickly over distance. While there is no consensus on the exact limits, some empirical studies suggest that beyond six metres away—across the length of an average lounge-room—the two retinal images lack significant lateral disparity to trigger stereopsis (Gregory, 1966: 53)[1]. Binocular vision is therefore only effective for seeing stereoscopically in 3D, when the object of attention is viewed within near distal space.

Nevertheless, the advent of 3D screen-spaces for home and mass entertainment suggests that today's consumers and practitioners of screen culture subscribe to what Friedel calls a 'culture of improvement': the belief that new imaging technologies will improve screen-space 'sensibilities', making them 'better than they were' (Friedel 2007: 2). Indeed, why watch something 'flat' on a screen (so the sales pitch typically goes), when you have the 'better' alternative of watching it fully rendered in three-dimensions? Like photography[2], stereoscopic 3D imaging has the potential to transform visual culture by bringing novel (i.e., transformative) experiences of the possibilities of visual representation to a broader public, through cinema, TV, and video games. But what exactly is transformed as new imaging techniques deliver (say) *Alice in Wonderland* (Dir. Tim Burton 2010) in 'solid' three-dimensions? Leaving video games aside, on a point for point comparison between the narrative cinematographic structure of *Alice* as 'conventionally' projected, and that of *Alice* in 3D (the same applies to football on TV), the diegetic content remains largely the same: the drama unfolds, the characters are animated, *dénouement* comes about eventually.

However, an additional body of content is introduced in the stereoscopic screening of a film (or football match). One might call this content 'non-diegetic': delivered surreptitiously through the delivery system of 3D projection, the additional content goes largely unnoticed. I refer here to the three-dimensional spatiality itself, the '3D effect'. For to *notice* it as a separate property of the experience of seeing *Alice* in 3D, would entail reflecting upon the film's particular 'spatiality' in relation to the film's holistic contents. Such a dialectical response would be atypical and run counter to expectations of what it is like to 'enjoy' a film. Moreover, the purpose of introducing

stereoscopic 3D screen technologies to the cinematic experience, aside from the sheer novelty, is to make the experience of screen-space depth more real. The 3D effect may be engaging and even hyper-real, yet the technique for bringing cinematic hyper-reality through 3D imaging is grounded in the technology of 'ordinary' binocular vision. Sensory real-world visual experience is diaphanous:

> Experience itself is transparent. There's no experiencing it. There's only encountering the world—content—as you experience it. It would seem, then, that we cannot reflect on experience itself (Noë, 2004: 175).

This brings to mind a further point of consideration: namely, that the principle of producing scenes of startling realness in 3D operates through the 'perceptual illusion of non-mediation' (Lombard & Ditton 1997). Of course, watching *Alice* in the flat is already a 'mediated' visual experience, and like paintings, photographs and other 2D representations, films provide indexical references *to*, rather than the 'real' dimensions *of* spatial depth. Such classical representational indices as linear perspective, foreshortening and interposition are essentially 'pictorial cues' across a functionally flat screen. While the relative movements of figures and perspectives across screen-space provide important signifiers of depth through the operations of motion parallax, figures and scenes on the conventional cinema screen have no perceptible three-dimensionality *per se*, whether nearby or far away. Seeing something—a figure, a scene—'in 3D' on a screen, thus presupposes a very different visuality from conventional screen media, because the visual-perceptual conditions (the status of vision as a knowledge-bearing faculty) have shifted.

## Predatory vision

In a legitimate effort to advance their craft, developers of stereoscopic 3D screen technologies exploit binocular vision's innate spatial stereoacuity. The technique produces scenes in which the surface layout of objects and scenery appears in three dimensions: objects appear stereoscopically three-dimensional, with the effect accompanied by a sensation of volumetric spatial depth—transparency and diaphanousness—in the visual relations between foreground and background objects. Yet, in binocular perception of an ordinary environment, stereopsis is effective only for objects close to hand. For mid-ground and far distances, the imagery flattens out and becomes 'monoscopic'. In the vision science literature, monoscopic information is termed 'pictorial' (Palmer 1999: 248), which is suggestive of the way information available from conventional 2D media (films, TV, paintings), entails a peculiarly unspecified and underdetermined spatiality. Perceptually, the visible contents of the 2D cinematic screen are situated neither near nor far away.

A cat's eyes are frontally placed like ours for the stereoacuity of predatory vision. All predators take advantage of two eyes oriented for the rapid and accurate response to where things are located in three-dimensional space: for hunting in general, and for grasping and catching specifically. We may not consider ourselves 'predatory' on the basis that we have two eyes frontal, and we may not recognize this innate capacity in our day-to-day view of the world. However, with seeing comes knowing, and all

visualities carry an epistemological load. For mice and rabbits (whose wide-spaced eyes give an in-the-round view of the world) seeing brings forewarning of a predator; seeing means the possibility of escape. For predators, seeing is power. Binocular visual perception is, among other capabilities, a potential for structuring decisive actions. Coupled with 3D screening technologies, our innate visuality becomes instrumental, an extension of the projection technology.

## Hunting and capturing

Stereoscopic 3D screenspaces are often vaunted as affording a 'virtual' perceptual experience of objects and spaces. It is instructive here to consider the meanings associated with 'stereoscopic' and 'perceive' respectively. Classical Latin gives us *percipere*: to take possession of, seize, get, obtain, receive, gather, collect; to apprehend with the mind or senses, to understand. Thus, 'perceive' (*per-* + *capere*) also means 'to take, seize': *to capture*. Its Greek origins reveal 'stereo-scopic' to mean '*solid sight*'. Virtual 3D imaging technology of the kind that gives us *Alice* in 3D thereby 'arms the eye' and 'gives the eye a hand of its own, propelled (apparently) by the gaze itself' (Penny 1994: 203). In the light of such observations it is pertinent that in 1857, 'stereoscope' was a generic term used to describe 'an instrument for detecting a calculus in the bladder, and foreign bodies in the soft parts' (Dunglison 1857).

More pertinent is terminology found in the contemporary 3D imaging industry itself, wherein the stereoscopic 'comfort zone' for viewing 3D screens is a much sought-after commodity, and 'hunting time' is described in the technical literature of 3D imaging as:

> [T]he time it takes for a viewer's eyes to converge on the focal elements of a shot. If the transition between shots entails large convergence changes, it will take time for eyes to adjust. Dramatic convergence changes can be uncomfortable. Further, if the shots are too short, hunting time can exceed shot length and viewers will not be able to track the action. Good 3D entails leaving the viewer's convergence at the end of a shot near the same stereo distance of the focal elements of the next shot (Seigle 2010).

This is tantamount to applying preconditions upon the cinematic and televisual content at the pre-production (visual-aesthetic, *thinking*) stage, to avoid such problems of convergence and tracking. It is suggestive also of the power invested in the laboratory method (which parallels the media production studio): going beyond the study and representation of real things, such methodologies become 'powerful enough to define *reality*' (Latour 1987: 87). If dramatic camera movement and rapid intercutting—shots that go from near to far too quickly for 'hunting time' to catch up—are strategies for cinematographers to *avoid*, what happens to the 'cut', the montage and other epistemic (knowledge-bearing) cinematic forms?

My concern here is that the 'reality' of existing screen-space (of film/animation in terms of the aesthetic and sensory experiences it already affords), will be over-determined by an unthinking shift toward what can be done 'better' in 3D. Schematically speaking, this entails a collapse of the manifold of visual-spatial

possibilities of such experiences, into a peculiar singularity: the cinematic screen is no longer a shared space of indeterminable distance available equally to all (because its space is nowhere in particular). Rather, in 3D the content is more specifically targeted to individual pairs of eyes. The two-eyedness of binocular stereoacuity, rather than suggesting an opening out of visuality into the generalised screen-space, instead construes the cinematic experience as an array of personally directed televisual trajectories. The resulting visuality is characterized by the perceptual 'capturing' of those parts of the film's contents that protrude out of the screen and into the immediate space of each individual viewer: there is a more *tangible* sense in 3D cinema that what I am seeing from here, cannot be the same object that you are seeing from there, because my seeing this object depends on the lateral displacement of just my eyes. Each separate pair of eyes is involved in a process of range-finding as it triangulates its own portion of screen-space to capture the 3D effect. The shift into 3D not only emphasizes what I have described metaphorically as 'predatory vision', but also privileges the individual viewer's singular point of view.

It is not clear that any shift away from 'uncomfortable' dramatic convergence is a move toward 'better'—more interesting or transformative—cinematic experiences. This over-determination of screen-space visuality by the push toward stereoscopic imaging so as to minimise the inconvenience of 'hunting time', is akin to colour grading a film pink, on the grounds it has a calming effect on viewers. Making a pink film in this case is less about filmmaking than about pragmatism. This raises the broader issue of who is best qualified to 'design' screen-spaces and their contents. For the majority of our more important needs, we turn to experts in the disciplinary field. We would not let just anybody with a small sharp knife perform our cataract operation. Why would we turn to technicians to prescribe the cinematographic possibilities of our existing screen-spaces?

## Image space as composition

Various theorists of VR have noted the parallel with Bazin's dream of 'total cinema' in which 'the technology of cinema would be effaced in favour of a 'complete illusion of reality.' (Rutsky 1999: 111) Similarly, Eisenstein recognised the interrelated histories of art, science, and technology, stating that the 'ultimate synthesis of all art genres would culminate in the imminent realization of Stereokino, stereoscopic cinema' (Grau 2003: 154). When Grau preferences the term 'image space' in his discussion of stereographic immersive VR, his emphasis is similarly on 'the practice (by developers of 3D imaging technologies) of using all available technical means to remove the boundary between observer and image space' (Grau 2003: 143). If the future of 3D imaging is doomed to fail in transforming screen-spaces in positive ways, it is because it succeeds so well in achieving what it sets out to do technically. No art form is ever complete as an Art, but there are epistemic issues (losses) involved in the introduction of the third dimension to an art form that is already 'complete' in its orchestrations across two-dimensional space. In this context, the 1972 'Manifesto' by stereo-enthusiast Roger Ferragallo reads as a cautionary tale.[3]

In his essay 'Art as Technique', the Russian Formalist, Viktor Shklovsky develops the concept of 'estrangement' (defamiliarisation), suggesting that the 'aim of art is to convey a sense of the object, to make us see it, not to [merely] recognize it'. For Shklovsky, 'the process of perception is an end in itself and must be lengthened' (Shklovsky 1916). William Hogarth expresses a similar sentiment in *Analysis of Beauty*: the beauty of '*composed intricacy of form*; and how it may be said, with propriety, to lead the eye a *kind of chace'* (Hogarth 1755: 28). In the design of 3D screen-space, the compositional pleasures and peculiarities of form seem overlooked in the rush to exploit the commensurability of human stereoacuity to the technical foundations of 3D imaging.

Sergei Eisenstein's claim that the ultimate visuality of cinema will be of the '*stereokino*' is thus a curious one, because many of the key passages in his *Battleship Potemkin* (1925) and *Ivan the Terrible* (1944) are exceptionally well articulated as visual spaces across the screen. While the cinematography portrays both shallow and deep spaces, what is also perceptible when watching *Potemkin*, is a space configured through flat compositional devices: dramatic figure-ground relations of positive and negative spaces. The tense silhouetted sailors anticipating battle, and the foreboding views across the sea at night under a heavy sky are exemplary in their artfulness. Such devices are essentially compositional, and their spaces are resolutely two-dimensional: it is difficult to imagine how *Potemkin's* iconic scene of the crowd streaming down the stairways at Odessa would be any 'better' by appearing in 3D.

A stereoscopic rendering of the elements within these particular scenes would add nothing 'spatial' that is in any way meaningful cinematographically. Compositionally, these passages in *Potemkin* work compositionally (their study as still frame-captures confirms this) because they are also, inherently, already spatial. The contents of *Battleship Potemkin* appear across a flat screen, but we do not *sense* that we perceive the scenes or shots as explicitly 'flat'. As with any other film, or painting or photograph, *Battleship Potemkin* represents both its shallow and deep spaces through a range of visual devices that are not themselves three-dimensional (stereoscopic). Yet is there any sense in which we feel a film, painting or photograph is visually *lacking*, because the spatial contents are merely represented, rather than configured to appear solidly in three dimensions?

Frank Popper states that immersion in a visual display is 'characterized by diminishing critical distance from what is shown and increasing emotional involvement in what is happening' (Popper 2007: 286). The distinction Popper makes here between two ways of responding to content, reinforces the key point of this paper: namely, that the mode by which the eyes are engaged through screen-projection technologies, will significantly determine the epistemic possibilities of the experience. The link between seeing and knowing is a crucial one in a culture bent on improving itself through technologies of the visual. Insofar as vision is 'one of our chief resources for discovering the content and character of our surroundings' (Dretske 1969: 78), it is important that we pay attention to how vision can be structured—over-determined—by those technologies which initially set out to mimic its innate affordances. If binocular stereoacuity is only effective for objects close to hand, then much of the technical effort in rendering scenes depicting wide-open

spaces and distant vistas is wasted. Mountains, clouds, and figures in the middle distance remain inexorably flat. This perceptual redundancy is not in itself a problem, since we are accustomed to experiencing our environment visually in much the same way. To maximise the effect of 3D imaging—to get what you paid extra for—requires that a significant amount of screen-space be dedicated to objects in immediate near space. This is not necessarily a cinematic problem. However, if the capacity for film to work as a mode of seeing-as-thinking is conditional upon issues that are not in themselves about film and filmmaking, then it is a problem for cinematic experience: to decide in the pre-conceptual stage of production that a film will not have too many cuts because the audience will be disorientated suggests the tail is wagging the dog.

## Conclusion

Technically, it is not difficult to produce imagery in 3D. The technique is well understood, and further development will be rapid given intense competition for market share. Although some technical problems are non-trivial, they are easily solvable through standard technical methods: in many respects, 3D screens today are an elaboration of Wheatstone's stereoscope. Eisenstein's *Battleship Potemkin* depicts revolutionary uprising, but *Potemkin* is also about the cinematic possibilities of screen space. Through the use of montage and other compositional devices, movement and action comes from watching, not doing. One can only wonder what would happen, compositionally, to this film if it were projected in 3D.

The important question is not about how to solve technical problems, but how to understand the epistemological shifts that accompany the new visuality of 3D screen spaces. For if, as Deleuze suggests, the definitive visuality of cinema is one that presents an opportunity for a different mode of *thinking* about meaning and representation (Deleuze 1983 & 1985), then a '3D turn' toward a cinema of solid spaces suggests a philosophically retrogressive step for visual culture. Developing scholarship around 3D screen spaces, rather than perfecting the technique, is the urgent issue. We should consider the existing heuristic function of film *as film*, before we make assumptions about it needing to be 'improved'. Pictures, as Margaret Hagen reminds us, are 'adequate representations of selections of the visual world, not substitutes for it.' (Hagen 1999: 80). The same argument applies to the 'moving picture'.

Insofar as there is a 'culture of improvement' and that filmmaking over the past one hundred years has contributed to that culture in ways that are significant, how does 'being in 3D' constitute an improvement? The short answer is that it does not necessarily improve our experience of a film *as film*, if (as is typical) adding the third dimension merely projects objects along the z-axis, toward and away from the observer's viewpoint in predictable and unnecessary ways. In its application to an essentially two-dimensional form of representation, the visuality of 3D stereoacuity may over-determine the experience of watching a film, the brute form of which is a screen-space characterised by an array of projectiles: solid forms eliciting a perceptual response better adapted to predatory reaching and grasping (film as sport). Yet there are reasons to be optimistic precisely because the 3D shift is itself new and

undeveloped. In the hands of innovative, thoughtful filmmakers, new visualities of the film form are always possible. We may be productively transformed—'estranged' afresh—by different materialist 3D screen practices that are yet to emerge.

## Endnotes

1. More recent research suggests eighteen metres as the absolute limit of stereoacuity; see, for instance, Allison, RS, Gillam, BJ, & Vecellio, E 2009, 'Binocular depth discrimination and estimation beyond interaction space', *Journal of Vision* 9.1(10): 1-14.

2. From Niepce's first photograph (1826), through Oscar Rejlander's layered composites (1857), to Laszlo Moholy-Nagy's 'photograms' of the 1930s, the beginnings of photography were experimental and knowledge-seeking.

3. Ferragallo asserted, vehemently, that not only was it possible to render all art stereoscopically, but that it was 'necessary' to forge a 'new space aesthetics' for painting, with retinal disparity at the centre. Ferragallo, R 1972 *A Manifesto Directed to the New Aesthetics of Stereo Space in the Visual Arts and the Art of Painting* (inactive web document; no pagination). Ferragallo republished his *Manifesto* in 1974, retitled 'On stereoscopic painting' in *Leonardo*, 7.2: 97-104. All quotations in the passage above are from Ferragallo's original 1972 *Manifesto.*

## Works cited

Allison, RS, Gillam, BJ & Vecellio, E 2009 'Binocular depth discrimination and estimation beyond interaction space', *Journal of Vision* 9(1).10: 1-14

Deleuze, G 1983 (1986) *Cinema 1: The Movement-Image*, London: Athlone Press; 1985 (1989) *Cinema 2: The Time-Image*, London: Athlone Press

Dretske, FI 1969 *Seeing and Knowing*, London: Routledge & Kegan Paul

Dunglison, R 1842 *Medical Lexicon. A dictionary of medical science; containing a concise explanation of the various subjects and terms of anatomy, physiology, pathology, hygiene, therapeutics ... with the accentuation and etymology of the terms, and the French and other synonyms* (1874), Philadelphia: HC Lea Publishers

Ferragallo, R 1972 *A Manifesto Directed to the New Aesthetics of Stereo Space in the Visual Arts and the Art of Painting,* at http://www.ferragallo.com/newmf4.html (accessed 21 March 2011)

Friedel, R 2007 *Culture of Improvement: Technology and the Western millennium*, Cambridge, Mass.: MIT Press

Grau, O 2003 *Virtual Art: From illusion to immersion*, trans. G Custance, Cambridge Mass.: The MIT Press

Gregory, RL 1966 *Eye and Brain*, London: World University Library

Hagen, M 1999 'How to Make a Visually Realistic 3D Display', *Computer Graphics*, 25.2: 76–81

Hogarth, W 1753 *Analysis of Beauty*, London: W Strahan

Latour, B 1987 *Science in Action*, Cambridge: Harvard University Press

Lombard M & Ditton, T 1997 'At the heart of it all: The concept of presence', *Journal of Computer-Mediated Communication*, 3.2, at http://jcmc.indiana.edu/vol3/issue2/lombard.html (accessed 21 March 2011)

Noë, A 2005 *Action in Perception*, Cambridge, Mass.; London: MIT Press

Palmer, SE 1999 *Vision Science: Photons to phenomenology*, Cambridge, Mass. and London: The MIT Press

Penny, S 1994 'Virtual Reality as the Completion of the Enlightenment', in Loeffler, CE & Anderson, T (eds) *The Virtual Reality Casebook*, New York: Van Nostrand Reingold: 199-213

Popper, F 2007 *From Technology to Virtual Art*, Cambridge, Mass.: MIT Press

Rutsky, RL 1999 *High Techné: Art and Technology from the Machine Aesthetic to the Posthuman*, Minneapolis and London: University of Minnesota Press

Seigle, D 2010 'Glossary of Terms used in 3D Post Production', In-Three Inc., at http://www.in-three.com/glossary.html (accessed 2 November 2010)

Shklovsky, V 1916 *Art as Process*, cited in Lovink, G 2006 'Catalogue of the 1st international video festival at the Bauhaus, Dessau', *Mediamatic* online journal, 8.1: *Storage Mania,* at http://www.mediamatic.net/page/8421/en (accessed 15 June 2006)

Wheatstone, C 1838 'Contributions to the Physiology of Vision—Part the First. On some remarkable, and hitherto unobserved, Phenomena of Binocular Vision', originally published in the *Philosophical Transactions* of the Royal Society of London*,* 128: 371-94, at http://www.stereoscopy.com/library/wheatstone-paper1838.html (accessed 15 April 2010)