

Albertian Perspective and Augmented Reality: Lessons from Panofsky

Tom McGuirk, Alan Summers
Department of Art and Design,
University of Chester
Chester, United Kingdom
, t.mcguirk@chester.ac.uk, a.summers@chester.ac.uk

This paper addresses the ubiquity of Albertian perspective as the dominant paradigm in the production of certain diagrams. Panofsky recognized the cultural specificity of perspective as, “a systematic abstraction from the structure of ... psychophysiological space.” He considered it essential to ask with regard to artistic periods, not only whether they have perspective, but also what kind of perspective they have. This paper asks a similar question with regard to the employment of such perspective in augmented reality technologies.

In East Asian culture an alternative use of floating perspectives has developed, this is recognised by cultural psychologists as indicative of the different sensitivities to contextual information. Differences in the interpretation of the visual field between Western and East Asian subjects further call into question the universal application of Albertian and Cartesian models in the design of the diagrammatic environment.

Augmented reality technologies are now capable of overlaying diagrammatic information directly upon the user’s visual field. Therefore the perspectival conventions of three-dimensional visualisation techniques might potentially come to reinforce Cartesian principles, and thereby be regarded as the unjustifiable imposition of a culturally specific worldview. This paper addresses the psychological, philosophical and indeed cultural ramifications of this phenomenon.

Keywords—augmented reality; perspective; symbolic form; Panofsky;

I. INTRODUCTION

The development of augmented reality (AR) technologies that overlay diagrammatic information upon a user’s field of vision presents many technological challenges, some of which will be touched upon in this paper. Over and above these challenges however, there are psychological and even philosophical concerns with regard to those perspectival systems that underpin these technologies. Panofsky [1] revealed that Albertian/Cartesian perspective is not a ‘natural’ or inevitable mode of engaging with space, but rather a construction and indeed a peculiarly Western instance of what he describes as “symbolic form”. However, in Western cultures, linear perspective has been remarkably successful; Jay

describes it as “the dominant scopic regime of the modern era” [2]. Rather than being a neutral, ‘natural’ vision, as is often supposed, linear perspective carries quite significant psychic implications, which bear upon how we encounter and interpret the visual field and – particularly with regard to AR – potentially how we engage with space. On a profound psychological level, it can moreover, condition and indeed prescribe how we interpret and even how we encounter our environment.

Panofsky correlates the Renaissance development of linear perspective, representing geometric space on a two-dimensional plane, with Descartes’s development of Cartesian space. Panofsky also considers that “‘aesthetic space’ and ‘theoretical space’ recast perceptual space in the guise of one and the same sensation: in one case that sensation is visually symbolized, in the other it appears in logical form” [3].

There are a number of philosophical implications that flow from this. In the above formula Panofsky identifies a conceptual juggernaut that, at least in the West, has swept all before it. However philosophical objections, to this reification of Cartesian perspectivism, emerge from divergent sources. Bourdieu, guided by Panofsky, sees in the Renaissance development of pictorial perspective, the camera obscura and “scientific” cartography, the emergence of a “distant lofty gaze”, characterized by an attitude of detachment, surveillance and control. Furthermore he identifies this with the “scholastic disposition” and the supposed objectivity of Western science, and he argues that it is Albertian perspective that first presents us with, that “point of view on which no point of view can be taken” [4]. Bourdieu tells us that the “only way”, to avoid this “blind spot” is to do as “Panofsky does”, by putting “perspective into historical perspective.”

Panofsky understood that perspective “mathematizes this visual space”, in what he terms “an ordering of the visual phenomenon,” just as Cartesian method (indebted to the same worldview) orders all phenomena. However this does not come without a price, as Panofsky explains:

Whether one reproaches perspective for evaporating “true being” into a mere manifestation of seen things, or rather for anchoring the free and, as it were, spiritual idea of form

to a manifestation of mere seen things, is in the end little more than a question of emphasis. [5]

The mechanics of “detachment” that Bourdieu recognizes in both Cartesian perspective and Cartesian method is outlined by Schrodinger, who describes such “objectification” as; “the ‘hypothesis of the real world’ around us.” And he explains the considerable “price” paid for what he witheringly refers to as merely “a fairly satisfactory picture for the time being.” [6] As he puts it:

We step with our own person back into the part of an onlooker who does not belong to the world, which by this very procedure becomes an objective world. [7]

Two points of note here are Schrödinger’s recognition of the psychic price paid for this “fairly satisfactory picture” and the extent to which it reflects a particularly Western, individualistic vision. And Schrödinger also notes that part of this price is the removal of the body (that of the observer and of others) from the world picture as presented by Albertian/Cartesian perspective.

II. HISTORICAL CONTEXT

In “Perspective as a Symbolic Form,” Panofsky [1] details consecutive systems of perspective from Antiquity, to the Middle Ages, to the emergence of Albertian perspective in the early Renaissance, these are described as a succession of worldviews. Panofsky thereby counters the positivist narrative of progression toward ultimate mathematical certainty and geometric purity.

Taking his lead from Cassirer, Panofsky analyses these worldviews in terms of symbolic form, or “symbol systems that structure an aspect of reality according to certain organizing principles ... [something] perspective does for the representation of the visible world.” [8] Panofsky repudiates the notion that linear perspective is the unique and inevitable outcome of some natural progression, rather, as Neher puts it, it is merely “an artifice suited to a certain Western manner of apprehending the world.” [9] As though to emphasise this, Panofsky insists that in studying “artistic periods and regions” it is therefore essential “to ask not only whether they have perspective, ... but also which perspective they have” [10].

If linear perspective is but one of a number of culturally specific ways of representing space, it is rather an odd one. It makes, as Panofsky points out, two implausible assumptions:

... first, that we see with a single and immobile eye, and second, that the planar cross section of the visual pyramid can pass for an adequate reproduction of our optical image. [11]

As Panofsky recognizes these are no small concessions but “rather bold abstractions from reality” and he warns us against conflating “‘reality’ as an “actual subjective optical impression” with what he calls the “structure of an infinite, unchanging and homogeneous space”, because, as he points out, “a purely mathematical space - is quite unlike the structure of psychophysiological space”[12].

Panofsky’s core insights concern the artificiality and culturally specificity of linear perspective, more usefully regarded as a system of symbols than a somehow privileged or ‘natural’ way of perceiving reality. It is a system moreover that relies on certain “assumptions”. Indeed, this assumption that a natural vision of the world can be attained from a single fixed point of view is one that many commentators, as we shall see, have queried. As Berger observes:

The inherent contradiction in perspective was that it structured all images of reality to address a single spectator who, unlike God, could only be in one place at a time. [13].

III. CULTURAL CONTEXT

Panofsky objects to the notion that a contrived artificial mathematical space can substitute for psychophysiological space. His work, in the art historical realm, is starkly echoed in the findings of recent research that sheds light on the cultural specificity of systems of perspective in more contemporary contexts. In a comparative survey [14], which compared the responses of Western, compared with East Asian participants, to particular visual displays and scenarios, stark differences were observed in the reactions of participants, reflecting culturally specific responses to perspectival displays. Amongst other findings a number of their conclusions are particularly relevant to our current discussion. The authors concluded that; “the convergence of evidence across multiple measures would favour the conclusion that East Asians are more sensitive to context in artistic and representational domains than Westerners.” Moreover, the authors conjecture that perhaps:

... individualistic Europeans developed a linear perspective (in which the spectator sees the social world from his or her point of view) and collectivistic East Asians developed a floating perspective (in which the spectator sees the social world, including himself or herself, holistically) may contribute to ... [these] culturally shaped patterns of perspective. [14]

This insight tallies with a number of commentators on the topic, Jay speaks of the “multiple vantage points” of medieval Western perspective being replaced “by one, sovereign eye.”[15] If we consider that East Asian perspective systems encompass a similar “multiple” viewpoint perspective, this strengthens Panofsky’s argument that Albertian/Cartesian perspective might even be regarded as an aberration, as opposed to somehow being the default, ‘natural’ system for mapping reality and certainly not a universal system.

IV. TECHNICAL DEVELOPMENT

The production of diagrammatic perspective is easily facilitated by the digital toolsets at a designer’s disposal. The use of computer-generated imagery (CGI) to construct two-dimensional renderings of three-dimensional space has become ubiquitous. However as Panofsky points out, the ubiquity of linear perspective and our habituation to it in the West plays a major role in its dominance, therefore the role of the tools used to communicate perspective should also be addressed.

The development of three-dimensional CGI is inextricably indebted to the Cartesian system of logic, space and geometry.

Descartes may not have envisaged the development of CGI, but his thinking enabled algebraic equations to generate geometric shapes. It is through algebraic equations that computers can model three-dimensional geometry within a virtual space to be experienced by users on two-dimensional computer screens. This process can be discussed in terms of logical or theoretical systems generating the visual or aesthetic form of perceptual space.

The digital Cartesian bodies produced are initially devoid of colour and texture as these are calculated and visually rendered by the computer using the process of ray tracing. It is this process that enabled certain design fields to set out on a quest for hyper photo-like imagery in the rendering of three-dimensional geometry. Mitchell notes the striking similarity between the development of computer rendering and the development of linear perspective during the Renaissance [16]. The use of Cartesian geometry and ray tracing — tracing rays of light to form a two-dimensional computer-generated image of a three-dimensional scene — clearly reflect the principles of linear perspective. Alberti's thinking of lines of sight as "threads" stretching between the eye and the object observed, are reconfigured within ray tracing as rays, calculated by Cartesian logic to construct the computer-generated image on screen.

In films, and in digital games produced using CGI techniques, the viewer passively or actively interacts with 3D environments representative of 'real' space. The common element within these early formats was the viewing of these representations through the frame of a two-dimensional screen. The technological development of immersion or removing the frame is a further significant development. Immersive spaces such as the 3D IMAX for films or virtual reality headsets require different technical solutions to those provided by Alberti's frame. With these technologies the frame of the cinema or computer screen as a method for organising and structuring perception is obsolete. Grau [17] discusses these visual forms as immersive because they "seal off the observer from external visual impressions". The relocation of the observer *into* the image as opposed to looking through the frame *at* the image, from outside, is the recurring model. In order to intensify the illusion and initiate immersion, the observer is only able to see the virtual reality, a reality rendered in the perspective form chosen by the designer. Grau goes on to state that, "Over the last years of the twentieth century, the incidence of immersive image media has spread to all industrial nations, and thus, has helped the Euclidean representation of space to achieve dominance, in Asia as well." [17]

The subsequent development of Augmented Reality (AR), a combination of the real world and CGI representations of space, raises the most interesting questions about the dominance of Euclidean space. This technology superimposes a mathematically generated perspectival interpretation onto a view of the observer's field of vision and thereby imposes a geometric structure onto what Panofsky termed the viewer's "psychophysiological space," potentially orchestrating the viewers interaction with their environment, with all the psychic and philosophical implications that entails.

Using mobile devices to view 2D or 3D information diagrams, rendered on screen as a live camera image, thereby augmenting the user's view of the world, is now commonplace. In these instances any immersion is reliant on the on-screen device's perspective rendering matching the observers field of vision. The observer needs to hold the device in a consistent central position, ensuring the image frame on the device matches the users field of view, in order that Alberti's frame synchronises with the actual real-world space. AR headsets or glasses resolve the dual view issue as the point of view is fixed when projecting information onto a lens firmly set in front of the observer's eye. Systems with dual projection — an image projected onto each lens — produce stereoscopic rendering of the augmented information. This enables three-dimensional perspective to be projected directly over the observer's view of the real world, Cartesian space is thus superimposed directly into the observer's psychophysiological space, such that the habituation to linear perspective is reinforced without ever being questioned.

V. CURRENT AR CONTEXT

Panofsky argues that we cannot reduce psychophysiological spaces to purely mathematical spaces. To elucidate this let us examine the two implausible assumptions that Panofsky considers essential for central perspective "that we see with a single and immobile eye" and "that the planar cross section of the visual pyramid can pass for an adequate reproduction of our optical image" [11]

A user of AR glasses does not use a single eye but two eyes in constant motion and the resulting spheroidal field of view is projected onto their concave retinal surfaces. This is not equivalent to the fixed view of a camera with a flat image sensor or the linear perspective model based on projection onto a flat plane. The human spheroidal field of view results in a curved perspective, where vertical and horizontal lines that are not central in the field of view are actually perceived as curved.

In the pre-Renaissance period, both artistic and philosophical theories are known to have observed that straight lines are seen as curved and curved lines are seen as straight. The curvatures of Doric pillars demonstrate that the Greeks understood the practical implications of this observation. The ancients recognised that the projection of an image onto a curved surface is determined by angles not distances and the relationship of the width of the angles of vision. Euclid's eighth theorem states that it is the ratio of angles, not the ratio of distance between two different magnitudes at differing distances that provides a more factual representation of natural vision.

Euclid's model is based upon projection onto a curved surface and Panofsky implies it might be more than mere accident that this model was suppressed during the Renaissance; where artists may have looked more favourably upon a system for projection onto flat surfaces, in terms of wall decoration and easel paintings. Few people actually perceive these curvatures and Panofsky cites the habituation of linear perspective being reinforced through visual formats such as photography — to which can now be added the many forms of computer-generated imagery.

In consideration of the assumption of the immobile eye, Chua [18] observed cultural differences when tracking the eye movements of Western and Asian subjects. Western participants quickly concentrated on central objects, and for longer than Asian participants, who made more eye movements in general, which are more directed toward the background. This supports theories of more contextual awareness in Asian subjects but also raises implications for the binding of any perspectival rendering to a real world background. Subsequent research has found that Japanese participants recognised objects better when they are bound to their original context, than American participants, where contextual manipulation had much less effect [14]. This indicates the binding of diagrammatic perspectives to the actual real worldview in AR systems may be more problematic for Asian than for Western subjects.

It can be argued then in relation to Panofsky's assumptions, that with regard to the rendering of diagrammatic perspectives for AR; firstly habituation and technological ease might be considered to have strengthened the dominance of linear perspective over more relevant ocular models. Secondly there is no immobile eye, and moreover cultural differences in eye movement may prove problematic for linear perspective rendering for diagrams in AR environments.

VI. CONCLUSION

In conclusion, this paper raises questions (just as Panofsky did in 1924-25) concerning the appropriateness of an unquestioning acceptance of Albertian/Cartesian perspective as the gold standard of perspectival systems. It does so in the context of increasingly rapid developments in augmented reality and other technologies and it does so from a number of viewpoints.

With the development of AR and other similar technologies it can be argued that unless we question, as Panofsky did, the kind of perspective we want by not allowing our habituation to one dominant form to reign supreme, then the perspectival conventions of three-dimensional visualisation techniques could potentially reinforce the dominance of Albertian/Cartesian perspective and thereby enforce and indeed reinforce a Cartesian worldview. There are, as this paper has argued, good practical technical reasons to eschew such a course, for example the aspirations of areas of the CGI industry for a 'photo-real' finish often conflict with the diagrammatic perspective renderings used in augmented reality. However, there are other more philosophical reasons to oppose this dominance. One of these is the danger of the unjustifiable imposition of a culturally specific worldview onto all our retinas. Just as pre-Renaissance European cultures had valid ways of representing and engaging with space through specific perspectival systems, so too do non-occidental contemporary cultures, as evidenced in the work of Masuda et al.

Habituation, technological convention and 'herding' could potentially hamper consideration of other alternative forms of perspective as candidates for development, where ocular, floating-point or multi-perspective forms of CGI may offer more appropriate representational forms that bind to the viewers "psychophysiological space".

As Panofsky, Bourdieu, Berger, Schrödinger and Jay assert; Cartesian/Albertian perspective presents us with a singularly individualistic vision of the world. Indeed one where, as Schrödinger puts it, we become onlookers who do "not belong to the world, which by this very procedure becomes an objective world". [7] This is a vision apparently quite alien to East Asian perspectival systems.

In discussing Renaissance perspective Panofsky wrote the following:

When work on certain artistic problems has advanced so far that further work in the same direction, proceeding from the same premises, appears unlikely to bear fruit, the result is often a great recoil, ... Such reversals, are often associated with a transfer of artistic "leadership" to a new country or a new genre, ... they do this precisely by turning back to apparently more "primitive" modes of representation. These reversals lay the groundwork for a creative reengagement with older problems, precisely by establishing a distance from those problems. [19]

Perhaps we have another lesson to learn from Panofsky, perhaps a "recoil", is precisely what is called for, a recoil from the dominance of Albertian/Cartesian perspective whereby the new genres of Augmented Reality and related technologies can be freed from the dominance of a certain kind of prescriptive vision that binds us to a singular, some might even say tyrannical worldview.

VII. REFERENCES

1. Panofsky, E. (1997) *Perspective as Symbolic Form*. New York, NY: New edition, Zone Books.
2. Jay, M. (1994). *Downcast eyes: The denigration of vision in twentieth century French thought*, pp. 69–70. Berkley, CA: University of California Press.
3. Panofsky, E. (1997) *Perspective as Symbolic Form*, pp. 45. New edition, Zone Books.
4. Bourdieu, P. (2000), *Pascalian Meditations*, pp. 22. Cambridge: Polity Press.
5. Panofsky, E. (1997) *Perspective as Symbolic Form*, pp. 71-72. New York, NY: New edition, Zone Books.
6. Schrödinger, E. (1992). *What is Life?: With Mind and Matter and Autobiographical Sketches*, pp. 119. Cambridge University Press.
7. Schrödinger, E. (1992). *What is Life?: With Mind and Matter and Autobiographical Sketches*, pp. 118. Cambridge University Press.
8. Neher, A. (2005). How perspective could be a symbolic form. *The Journal of aesthetics and art criticism*, 63(4), 359-373.
9. Neher, A. (2005). How perspective could be a symbolic form (pp.362). *The Journal of aesthetics and art criticism*, 63(4), 359-373.
10. Panofsky, E. (1997) *Perspective as Symbolic Form*, pp. 41. New edition, Zone Books.
11. Panofsky, E. (1997) *Perspective as Symbolic Form*, pp. 29. New edition, Zone Books.
12. Panofsky, E. (1997) *Perspective as Symbolic Form*, pp. 30. New edition, Zone Books.
13. Berger, J. (2008) *Ways of Seeing*, pp. 16. Penguin UK.

14. Masuda, T., Gonzalez, R., Kwan, L. and Nisbett, R. E. (2008) *Culture and Aesthetic Preference: Comparing the Attention to Context of East Asians and Americans*, pp. 1273. *Personality and Social Psychology Bulletin*, vol. 34, pp. 1260–1275.
15. Jay, M. (1994). *Downcast eyes: The denigration of vision in twentieth century French thought*, pp. 54. Berkley, CA: University of California Press.
16. Mitchell, W. J. (1994) *The Reconfigured Eye: Visual Truth in the Post-Photographic Era* New Ed edition. MIT Press.
17. Grau, O. (2004) *Virtual Art: From Illusion to Immersion* New Ed edition. MIT Press.
18. Chua, H. F. (2005) *From The Cover: Cultural variation in eye movements during scene perception*, *Proceedings of the National Academy of Sciences*, vol. 102, pp. 12629–12633.
19. Panofsky, E. (1997) *Perspective as Symbolic Form*, pp. 47. New edition, Zone Books