https://doi.org/10.2352/lim.2022.1.1.08

The Art and Science of Displaying Visual Space

Robert Pepperell & Alistair Burleigh
Fovotec/Cardiff Metropolitan University
Cardiff, CF5 2YB, UK

rpepperell@cardiffmet.ac.uk

Abstract

This paper considers the problem of how to display visual space naturalistically in image media. A long-standing solution is linear perspective projection, which is currently used in imaging technologies from cameras to 3D graphics renderers. Linear perspective has many strengths but also some significant weaknesses and over the centuries alternative techniques have been developed for creating more naturalistic images. Here we discuss the problem, its scientific background, and some of the approaches taken by artists and computer graphics researchers to find solutions. We briefly introduce our own approach, which is a form of nonlinear 3D geometry modelled on the perceptual structure of visual space and designed to work on standard displays. We conclude that perceptually modelled nonlinear approaches can make 3D imaging technology more naturalistic than methods based on linear perspective.

1. Introduction

The question of how to represent visual space naturalistically in images has been a matter of controversy for several centuries [1,2]. What was originally a problem for artists and architects in the early Renaissance period has since become an issue for photographers, cinematographers, computer games designers and 3D graphics artists [3,4,5,6]. In this paper we provide a broad overview of the problem in the context of visual perception science, art history, and computer graphics.

One way to pose the problem is to think of visual space—defined as the space we experience perceptually—as a kind of internal display that helps us to navigate the world outside. The image that appears in that internal display need not necessarily be isomorphic to the world outside, nor need it conform to the data that our visual systems collect from the world [7]. In fact, as will be discussed, it can deviate markedly from both. The challenge then for those seeking to represent visual space naturalistically is to make an image of a scene that when presented on an external display— such as an electronic screen or a photograph—matches what we would expect to see in our internal display if we were looking at the same scene.

2. The Problem of Visual Space

Imagine that a person opens their eyes and fixates on a row of columns in front of them for one second. A stream of light passes through their ocular lenses and projects an inverted image onto the retina of each eye. These images stimulate arrays of light sensitive cells in the eyes, which triggers a cascade of neural activity that arrives at the visual cortex and is then integrated with neural activity elsewhere in the brain. The result is that the person almost instantaneously experiences a detailed, spatially wide and deep visual impression of the columns before them [8]. As noted, we can think of this visual impression as a kind of internal display that presents the scene in the viewer's mind as it is in the external world.

What appears in that internal display, however, can be very different from what is in the world itself and indeed from the optical data that our eyes collect. For example, the apparently detailed impression we get of the scene has relatively little to do with what is projected on the retinae; visual acuity is only high in the central foveal region of vision, which is around 2º of the total area of the visual field, while most of what we see during any fixation has very low spatial resolution; much of the apparent detail we perceive is generated by visual processes higher in the brain [9].

In addition, the space that appears to us in our internal visual display seems, on first inspection, to be structured according to the same Euclidean geometry that structures the world. So, for example, the columns in the scene that have equal width appear to have equal width in visual space. But it has been known since the early twentieth century that visual space is non-Euclidean in that its geometry varies across the visual field depending on where one is looking [10,11]. Related to this is the fact that while the columns, ground, and lintel may appear straight when viewed in the central visual field they are perceived as curved in the peripheral visual field. This effect has been known for several hundred years [12] but was first reported in the scientific literature by Hermann von Helmholtz [13] and illustrated with the image shown in Figure 1 [14]. The reverse effect can be seen if you look at a checkerboard pattern at proximity with one eye

In sum, the problem for the person wanting to represent what appears in our internal display with image media on an external display, such as an electronic screen, is that they have to first determine what actually appears in that internal display, which as noted here is not straightforward, and then work out how to emulate that on the external display in a way that feels natural—in other words, appears as expected—in the internal perceptual display when the person perceives the screen.

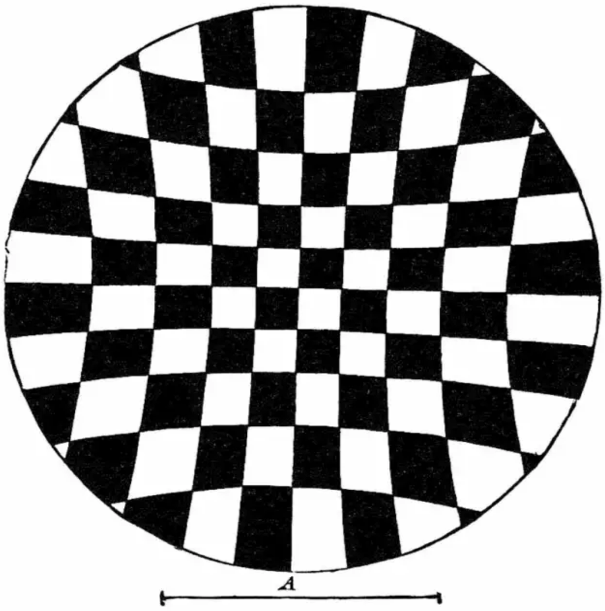

Fig. 22.

Figure 1. *The curvature of visual space as reported by Helmholtz in the nineteenth century. When displayed at a suitable size and viewed with one eye from a suitable distance the curvature of checkerboard squares disappears, and the vertical and horizontal lines appear straight.*

3. The Linear Perspective Approach

For several hundred years this problem was largely addressed through the expedient of linear perspective [15]. This form of projective geometry, discovered by artists and architects in fifteenth century Italy, is mathematically simple and, in strictly limited circumstances, perceptually natural. When set up and viewed correctly, i.e. with one fixed eye at the centre of projection of the image, a linear perspective image showing a narrow field of view can be almost indistinguishable from a real visual scene [16]. This is because, under those circumstances, the patterns of light entering the eye from the linear perspective image and projected to a plane closely match those that would enter the eye from the real scene. Linear perspective was a remarkably successful discovery that has underpinned many indispensable technologies such as photography, cinematography, and more recently 3D computer graphics. But as the pioneers of linear perspective such as Leonardo da Vinci knew, there are situations where it is unable to represent visual space naturalistically [17].

For example, the human visual field, even with one fixed eye, can span around 130º horizontally, and with two eyes is closer to 180º [18]. Projecting wide angles of view (>120º horizontally) to a flat surface using linear perspective results in an image that is highly distorted, with the central area appearing abnormally small and the peripheral areas abnormally stretched. This is so unless the image is viewed from its centre of projection with one fixed eye, which in the case of a 180º projection coincides with the projection surface. Add to this the fact that we normally see with two mobile eyes that are continuously moving with respect to our bodies and to the world then the inherent limitations of the linear perspective method start to become major drawbacks.

The problem would not be so acute, and may not exist at all in theory, if we had technology capable of exactly replicating the patterns of light that enter each eye as we move through space, e.g. a full field of view binocular virtual retinal display [19]. Head-mounted displays offer a partial solution, but fields of view are still limited and they pose health risks [20]. In the meantime, we remain highly dependent on conventional display formats such as computer screens, televisions, projections and print media. These typically occupy a relatively small portion of the visual field when viewed and the images they contain, which are predominantly linear perspectival, are rarely viewed from the centre of projection [5]. Therefore, the problem persists of how to naturalistically represent visual space—as it appears in the internal display—on standard external displays given the constraints that apply.

4. Alternative Nonlinear Approaches

Being aware of the limitations of linear perspective, artists, mathematicians, and technologists have developed alternative methods for representing visual space on external displays, often based on forms of curvilinear geometry, particularly when wide fields of view are required. It has been suggested that Leonardo himself devised a form of curvilinear perspective, now lost, that conformed more closely to the 'natural perspective'—as he called it—of visual space [17]. Overt examples of curvilinearity in paintings of the period certainly exist (see *The Arrival of the Emperor at St Denis* by Jean Fouquet, c. 1470, Cliché Bibliotheque, Paris).

In the nineteenth century, several European artists and mathematicians developed alternatives to linear perspective designed to better represent visual space [2]. Prominent among these was the "natural perspective" system developed by an artist from Liverpool, William Herdman, who specialised in expansive cityscapes of tourist locations. He claimed his curvilinear geometric method, which he published in some detail [44], produced images of "perfect accuracy according to vision". The German mathematician Guido Hauck launched a vigorous and rigorous attack on linear perspective later in the century with his publication of *Die subjektive Perspektive* [21]. Here he set out a form of curvilinear projective geometry based on psychological optics that, unlike mechanical optics, is intrinsically curvilinear, as noted above.

The period since has seen a number of further developments in the field of perceptually inspired perspectives, including notable contributions by artists such as Albert Floçon [22], Robert Hansen [23], Rackstraw Downes [24] and the scientist Boris Rauschenbach [25]. Perhaps most prominent, and most vocal in his denunciation of the limitations of linear perspective, has been the British painter David Hockney who has developed a range of innovative techniques for representing visual space naturalistically [26].

Researchers in the field of computer graphics have also taken an active interest in this problem and have broadly adopted two approaches to tackling it, often inspired directly by art history. One uses various methods of reprojecting, warping or otherwise deforming an existing linear projection and the other involves manipulating the three-dimensional geometry of the scene or model to create non-linear projections or composite projections from multiple camera viewpoints.

A prominent example of the first approach includes the Panini projection developed by Sharpless et al. [27], which is a form of cylindrical projection derived from the techniques of the eighteenth-century Italian artist Giovanni Panini who was famed for his naturalistic-looking wide vistas of buildings. Carroll et al. [28] presented several methods for artistically warping linear perspectival images by manipulating converging lines and vanishing points. Among early examples of non-linear multi projection rendering techniques is that of Agrawala et al. [29] in which single images were composited from several computer camera views to create more 'artistic' impressions of 3D space in the manner of painters such as de Chirico and Cézanne. Picasso's cubist style was an important source for the work of Singh [30] which also synthesised multiple computer camera views of 3D objects to create novel nonlinear renders that could be controlled interactively with several parameters. This technique was later developed into a plugin for Maya [31]. More recently, Liu et al. [32] have shown a depth-based editing tool for layering, resizing and recompositing photographs to overcome some of the restrictions of linear perspective-based projections and produce images that match visual space more effectively.

5. A Perceptually Modelled Approach

The approach briefly introduced here is the product of a long-standing interdisciplinary research project that draws on art practice, art history, vision science, perceptual psychology, geometry, and computer science. It began in 2011 as an exercise in recording the structure of visual space through drawing and painting, the aim being to capture the entire scope of the binocular visual field when fixating on a single point in space [33, 34]. Through this process of empirical observation, it became apparent that the human visual field during fixation has a particular nonlinear and non-Euclidean structure. In brief, objects under fixation—that is, objects in the middle of the visual field—appear larger than equivalently sized objects do in the periphery. Moreover, this apparent diminution in size by eccentricity also entails a change in shape that varies across the axis of the visual field. Taken together with the apparent curvature of visual space, noted above, the overall geometrical structure is somewhat akin to that produced by a non-uniform fisheye lens projection.

These observations were subsequently corroborated by art historical sources and by the perceptual psychology literature. The same essential geometrical structure appears in many paintings and drawings throughout art history, albeit in various forms and guises. It appears, for example in early mediaeval religious paintings, where a figure of importance is located centrally in the image, flanked on either side by less important figures that are also depicted as much smaller in size [17, 42]. Peripheral curvature and object minifications appear prominently and regularly, for example, in the landscape paintings of Paul Cézanne [35], and frequently in the paintings of several British artists of the twentieth century who were interested in capturing the effects of visual sensation [36]. Meanwhile, diminution of size and shape deformation across the visual field has been experimentally observed by Newsome [37] and Bedell & Johnson [38], among others, as well as in our own laboratory experiments [39, 40].

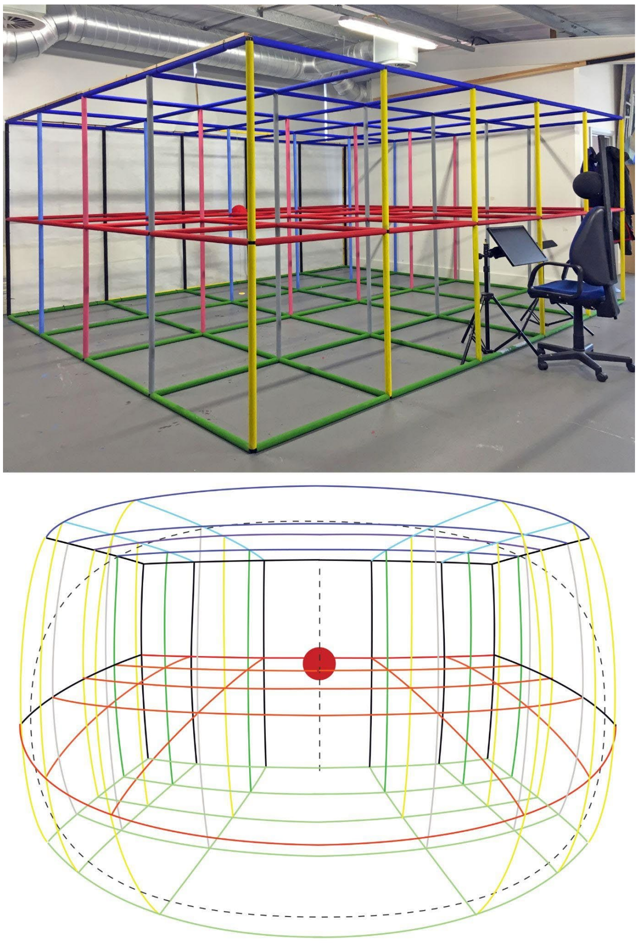

Figure 2. One of the experimental setups used by the authors to gather data on the structure of visual space. The top image shows the apparatus in which one of the authors viewed a red fixation object and made electronic drawings (as below) that conformed to his visual experience. These drawings were then used to derive computationally generated versions that were then presented to participants in the same space so they could judge how well they matched their own visual experience [43].

Our early attempts to emulate the nonlinear perceptual structure of the internal display on computer graphic displays were based on perceptual data that we gathered experimentally. We constructed various novel apparatus that we used to measure the perceived changes in size and shape of objects across the full visual field of participants (Figure 2). We then attempted to map this structure graphically by applying 2D image deformations to a linear perspective image to produce images that approximated the structure of visual space [41]. We were able to experimentally verify, albeit in limited cases, that these perceptually modelled images were indeed judged as more natural compared to a range of standard projection methods, such as linear perspective, stereographic, fisheye and others.

As we tried to generalise this approach to create a computer graphics technology that could automatically generate perceptually mapped representations by deforming 2D pixel arrays we met unforeseen challenges. First, we discovered many interacting variables that mitigate against a single geometrical structure to describe visual space in all viewing conditions [42]. Variables included the size of the external display and its viewing distance, whether the participant was viewing the scene monocularly or binocularly, the angle of the eyes with respect to the fixation object, and the scale and structure of the scene itself. For example, fixating on an object that is close to the viewer in a small space requires a different angle of convergence between the eyes compared to fixating on a distant object in an open space, such as a landscape, with a very wide field of view. This has a perceptible impact on the structure of the resulting visual space. We also found that in the absence of reliable eye tracking technology we were unable to determine the location of the viewer’s fixation, and so image effects that were applied to the periphery of the perceptually mapped image, and which would normally only be seen in the peripheral visual field, could be viewed directly, which undermined the naturalistic effect we were seeking.

To accommodate these variables within our computational model of visual space it became necessary to engineer a nonlinear rendering process that could produce an optimally naturalistic projection under a variety of conditions when shown on standard displays such as electronic screens and photographs. This process, which we call FovoRender, replaces the linear perspective mathematics on which most rendering engines are based with a different set of mathematics. There are two versions of FovoRender, one that runs in the vertex shader of a rasterization rendering pipeline and one that runs in the ray generation framework of a full path tracing rendering pipeline. Both versions were originally prototyped in Unity, later implemented in Unreal 4 and most recently in Cinema 4D through Open Shading Language with support for renderers including Octane and Arnold.

We briefly note some of the features and limitations of FovoRender with reference to two illustrations. Figure 3 shows a comparison between a 3D scene rendered in linear perspective (top) at 140º horizontal field of view and the same field of view in the real time rasterizer version of FovoRender (bottom). FovoRender provides a toolbox of sliders that allow the image designer to adjust a range of parameters that affect both the 3D geometry of a scene volumetrically, i.e. by changing occlusion paths, and the 2D pixel array in order to adjust the sizes and shapes of objects independently of the rest of the scene. The presets use values derived from both the artistic images created by direct observation described above and from experimental data gathered when measuring the structure of visual space for wide fields of view. User feedback from studies carried out in our laboratory and in collaboration with independent researchers, suggest that people perceive images rendered in FovoRender as more naturalistic than equivalent linear perspective renders [42, 43]. However, given the many interacting variables noted above, a challenge remains to find a universal set of parameter values that are optimal under all viewing conditions and for all kinds of scenes.

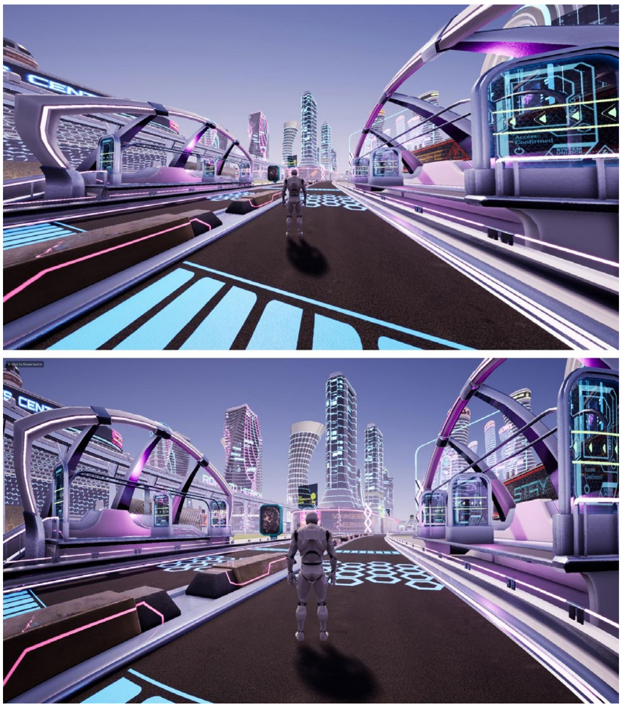

Figure 3. A comparison between (top) a 3D scene rendered with a wide horizontal field of view (140º) in linear perspective in Unreal 4 and (bottom) the same horizontal field of view rendered in the rasterizer version of FovoRender. Note the significant changes in occlusion paths between the two renders, which were necessary to match more closely the perceptual structure of visual space.

Figure 4 shows a comparison between a 3D scene rendered in linear perspective (top) at 110º horizontal field of view and the same scene rendered in the path tracer version of FovoRender (bottom). Both shots are made from the same camera position, but the FovoRender version displays more space and has less unnatural stretching of objects in the periphery. In the offline rendering mode, which cannot be explored in real time, there is more scope for nonlinearly adjusting the scene geometry to achieve a more naturalistic appearance. However, this flexibility puts more onus on the skills and judgement of the artist to decide what looks natural or not compared to using the presets provided in the rasterizer version, and there is more risk of creating unpleasant or unnatural looking deformations. A promising direction we are now investigating is to automate this process by intelligently modifying the image composition in 3D using rules generated through machine learning techniques with inbuilt scene understanding.

6. Conclusion

In this paper we considered the structure of visual space and the difficult challenge of representing it naturalistically in image media. We employed the analogy of the internal and external display to pose the problem and have defined the challenge as that of how to emulate the perceptual structure of visual space that we experience in our internal display on an external display. This problem is made more complex by the lack of isomorphism between visual space and the world and by many perceptual factors. Having considered the conventional solution of linear perspective, its limitations and some of the alternative solutions proposed by artists and computer graphics researchers, we briefly introduced a solution developed by the authors designed to allow 3D artists to create more naturalistic views of 3D scenes that show wider fields of view than would be practical using a standard linear perspective projection. Although there is no 'one size fits all' projection that captures the diversity of visual space yet, we conclude that the perceptually modelled approach outlined here, along with others in development, could lead in future to 3D computer graphic image experiences that are increasingly naturalistic compared to those made with linear perspective projections.

Figure 4. *A comparison between (top) a 3D scene rendered with a wide field of view (110º) in linear perspective in Unreal 4 and (bottom) a much wider field of view rendered in the path tracing version of FovoRender. These shots were taken from the same camera position but the space in the FovoRender version has less distortion (Images designed and rendered by Ben Walker at CreateCG).*

References

[1] Gombrich, E. H. (1972). The 'What' and the 'How': Perspective Representation and the Phenomenal World, R Rudner, and I Scheffler (eds.) Logic and Art: Essays in Honour of Nelson Goodman, pp.129-49.

[2] Kemp, M. (1992). The science of art: optical themes in western art from Brunelleschi to Seurat. New Haven: Yale University Press.

[3] Pirenne, M. H. (1970). Optics, painting, and photography. Cambridge: Cambridge University Press.

[4] Kingslake, R. (1992). Optics in Photography. Bellingham: SPIE Press.

[5] Banks, M.S., Cooper, E.A. and Piazza, E.A. (2014). Camera Focal Length and the Perception of Pictures. Ecological Psychology, 26(1-2), pp.30–46.

[6] Roettl, J. and Terlutter, R. (2018). The same video game in 2D, 3D or virtual reality – How does technology impact game evaluation and brand placements? PLOS ONE, 13(7), p.e 0200724.

[7] Hershenson, M. (1999). Visual space perception: a primer. Cambridge, Mass.: MIT Press.

[8] Wade, N. and Swanston, M. (2013). Visual perception: an introduction. New York: Routledge.

[9] Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. Journal of Vision, 11(5):12, 1–82.

[10] Luneburg, R. K. (1947). Mathematical analysis of binocular vision. Princeton, NJ: Princeton University Press.

[11] Wagner, M. (2006). The Geometries of Visual Space. Hove: Psychology Press.

[12] Elkins, J. (1994). The Poetics of Perspective. Ithaca: Cornell University Press.

[13] Helmholtz, H. von (1866). Handbuch der Physiologischen Optik, Hamburg: Voss.

[14] Oomes, A.H.J., Koenderink, J.J., van Doorn, A.J. and de Ridder, H. (2009). What are the Uncurved Lines in Our Visual Field? A Fresh Look at Helmholtz's Checkerboard. Perception, 38(9), pp.1284–1294.

[15] Alberti, L. (1991, originally published 1435). On painting. London: Penguin Books.

[16] Ten Doesschate, G. (1964). Perspective: Fundamentals, Controversials, History. Nieuwkoop, Neth.: B. De Graff.

[17] White, J. (1972). The birth and rebirth of pictorial space. London: Faber & Faber.

[18] Howard, I. P., & Rogers, B. J. (1995). Binocular vision and stereopsis. Oxford, England: Oxford University Press.

[19] Pryor, H.L., Furness, T.A. and Viirre, E. (1998). The Virtual Retinal Display: A new Display Technology using Scanned Laser Light. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 42(22), pp.1570–1574.

[20] UK Government (2020). The safety of domestic virtual reality systems A literature review. BEIS Research Paper Number 2020/038, RPN 4527.

[21] Hauck, G. (1879). Die Subjektive Perspektive und die Horizontalen Curvaturen des Dorischen Styls. Eine Perspektivisch-Ästhetische Studie, Wittwer, Stuttgart, Germany.

[22] Floçon, A. & Barre, A. (1988). Curvilinear perspective: From visual space to the constructed image. Berkeley, CA: University of California Press.

[23] Hansen, R. (1973). This curving world: Hyperbolic linear perspective. Journal of Aesthetics and Art Criticism, 32(2), 147–161.

[24] Schwartz, S., Storr, R. and Downes, R. (2005). Rackstraw Downes. Princeton, N.J.: Princeton University Press.

[25] Rauschenbach, B. (1982). Perceptual perspective and Cézanne's landscape. Leonardo, 15(1), 28–33.

[26] Gayford, M. (2022). Hockney's Eye: The Art & Technology of Depiction. London: Paul Holberton.

[27] Sharpless, T., Postle, B. & German, D. (2010). Pannini:a new projection for rendering wide angle perspective images, in Proceedings of the Sixth international conference on Computational Aesthetics in Graphics, Visualization and Imaging. Eurographics Association, 2010, pp. 9–16.

[28] Carroll, R., Agrawala, A., and Agrawala, M. (2010). Image Warps for Artistic Perspective Manipulation. ACM Trans. Graph. 29.4 (July 2010). ISSN: 0730-0301. DOI: 10 . 1145 / 1778765.1778864 2.

[29] Agrawala. M., Zorin, D. & Munzer, T. (2000). Artistic Multiprojection Rendering. Rendering Techniques, DOI:10.1007/978-3-7091-6303-0_12.

[30] Singh, K. (2002). A Fresh Perspective, in Graphics Interface, vol. 2002, pp. 17–24.

[31] Coleman, P. & Singh, K. (2004). Ryan: rendering your animation nonlinearly projected. NPAR '04 7 June 2004.

[32] Liu, S., Agrawala, M., DiVerdi, S., & Hertzmann, A. (2022). ZoomShop: Depth-Aware Editing of Photographic Composition. EUROGRAPHICS 2022, eds. R. Chaine and M. H. Kim, Vol. 41 (2022), No. 2.

[33] Pepperell, R. (2012). The perception of art and the science of perception. In B. E. Rogowitz, N. P. Thrasyvoulos, & H. de Ridder (Eds.), Human vision and electronic imaging XVII. Bellingham,WA: SPIE Press.

[34] Pepperell, R.(2015). Egocentric Perspective: Depicting the Body from its Own Point of View, Leonardo, 48 (5), pp. 424-429.

[35] Pepperell, R. and Haertel, M. (2014). Do artists use linear perspective to depict visual space? Perception, (43), 395-416.

[36] Pepperell, R. & Hughes, L. (2015). As Seen: Modern British Painting and Visual Experience. Tate Papers, Spring 2015.

[37] Newsome, L. R. (1972). Visual angle and apparent size of objects in peripheral vision. Perception & Psychophysics, 12, 300–304.

[38] Bedell, H. E., & Johnson, C. A. (1984). The perceived size of targets in the peripheral and central visual fields. Ophthalmic and Physiological Optics, 4(2), 123-131.

[39] Baldwin, J., Burleigh, A., & Pepperell, R. (2014). Comparing artistic and geometrical perspective depictions of space in the visual field. i-Perception, 5, 536–547.

[40] Baldwin, J., Burleigh, A., Pepperell, R. & Ruta, N. (2016). The Perceived Size and Shape of Objects in Peripheral Vision, i-Perception, July-August.

[41] Burleigh, A., Pepperell, R., & Ruta, N. (2018). Natural Perspective: Mapping Visual Space with Art and Science. Vision, 2(2), 21. https://doi.org/10.3390/vision2020021.

[42] Koenderink, J. & van Doorn, A. (2008). The Structure of Visual Spaces. Journal of Mathematic Imaging and Vision, 31: 171.

[42] Ruta, N. (2019). Visual perception in far peripheral visual space and its artistic representations. PhD Thesis, Cardiff Metropolitan University.

[43] Pepperell, R., Burleigh, A. & Ruta, N. (2021). Art and the Geometry of Visual Space, in Space-time Geometries in the Brain and Movement in the Arts, in the series "Lecture Notes in Morphogenesis" eds. T. Flash, A. Berthoz & A. Sarti, Berlin: Springer.

[44] Herdman, W. G. (1853). A treatise on the curvilinear perspective of nature; and its applicability to art. London: John Weale & Co.