FOVOGRAPHY: A NATURALISTIC IMAGING MEDIA

Alistair Burleigh, Robert Pepperell & Nicole Ruta

Fovolab
Cardiff School of Art & Design
Cardiff Metropolitan University
Cardiff, CF5 2YB, UK
Correspondence to: rpepperell@cardiffmet.ac.uk

ABSTRACT

This paper broadly describes a novel approach to representing the three-dimensional world in a two-dimensional image. Rather than just relying on the physical behavior of light rays projected onto a surface, as is the case with most current imaging technologies, we are attempting to create a new imaging technology that emulates the subjective appearance of the world as perceived via the human visual system. We do so by computationally modelling a number of complex psychological processes occurring in visual perception. This approach results in a new non-linear projection framework called Fovography that we have shown has higher ratings than standard projections when measured across a range of psychological factors, including sense of presence, comfort, and ecological validity. Moreover, Fovographs are rated as being equally immersive when compared to a large format 180º cylindrical projection screen, and on a par with a virtual reality system, despite needing no glasses, headsets, or specialist display hardware to view.

Index Terms— Linear perspective, depth perception, 3D imaging, human visual perception

1. INTRODUCTION

The aim of our research is to improve the way visual experience is depicted in imaging media by increasing the *naturalism* of the media. Numerous technical methods for depicting the visual world currently exist. Most of them rely on optical laws that determine the way light rays project onto a plane through an aperture [9]. Physical cameras and computer graphics rendering systems are the most common examples. They are designed to accurately capture or computationally simulate the patterns of light emitted or reflected by objects in the world. The images they produce can be thought of as objectively *realistic*. However, such images to do not necessarily represent what a human viewer would experience when viewing the same objects; the human visual system is not a camera. While there are some similarities between eyes and cameras, much of what we visually perceive is the result of complex psychological and neurophysiological processing occurring in the visual areas of the brain, for which there is no parallel in current imaging technology [12]. This paper briefly outlines our approach to creating naturalistic imaging media based on the emulation of human visual processes, and how this can improve the way we depict the visual world.

2. NATURAL VISUAL EXPERIENCE

Humans with healthy vision have two eyes that together create a field of view spanning about 180 degrees horizontally and 130 degrees vertically [7]. This visual field is roughly oval in shape, and is constituted by the images received on the retina from both eyes that fuse to form what we (mostly) perceive as a single image [5]. Eyes, heads and bodies can move relative to each other with varying degrees of freedom, and eyes themselves have several mechanisms for adjusting to visual conditions, including vergence, accommodation, saccadic motion, and pupillary dilation. It has been argued that visual space is non-Euclidian, meaning that there is a non-uniform distribution of perceived object distances, sizes and shapes across the visual field [6]. The peripheral areas of the visual field are much less distinct than the central area, which is served by the fovea [18].

Few of these properties of visual perception are incorporated into current imaging technology. Instead, cameras and most common computer graphics rendering systems are designed to capture or computationally simulate light rays travelling through a pinhole and intersecting with a plane. Linear perspective, discovered by artists and architects in the fifteenth century, was the first method of formalizing this process [8]. It held the promise, in theory, of accurately representing the three-dimensional world on a picture plane. But its limitations soon became apparent. It proves useful only for depicting relatively narrow angles of view (normally <60 degrees horizontally); it lacks the binocular properties of natural vision; it is static, and relies for its greatest effect on the viewer adopting a motionless, one-eyed viewing point that in practice is almost impossible to obtain [19]. Moreover, linear perspective does not discriminate between the differing appearance of the central and peripheral visual fields. Nor is it able to accommodate the non-Euclidean structure of visual space [4, 14].

3. THE FOVOGRAPHY APPROACH

Our general approach is to model imaging media as closely as possible on perceived visual experience. We aim to record what a human being sees rather than what a camera sees [13]. This requires an analysis of both the structure of three-dimensional human visual experience and how that experience can be depicted in two-dimensional images. This we have studied through various methods, including artistic observation and scientific investigation [1]. Having observed the general features of this structure, we are now mathematically modelling them in order to simulate them technologically. The resulting images, which we label ‘Fovographs’ (field of view drawings), are then tested experimentally to validate the model.

Fig. 1. The upper image is a photograph taken with a 24mm wide-angle lens. The lower image is a computer-generated Fovograph depicting the same viewpoint and showing a different geometric distribution of space and other perceptual features such as binocular diplopia and reduced acuity in the periphery. Note that due to the size of reproduction not all features of the process are clearly visible here. Images © Robert Pepperell/Alistair Burleigh, 2014.

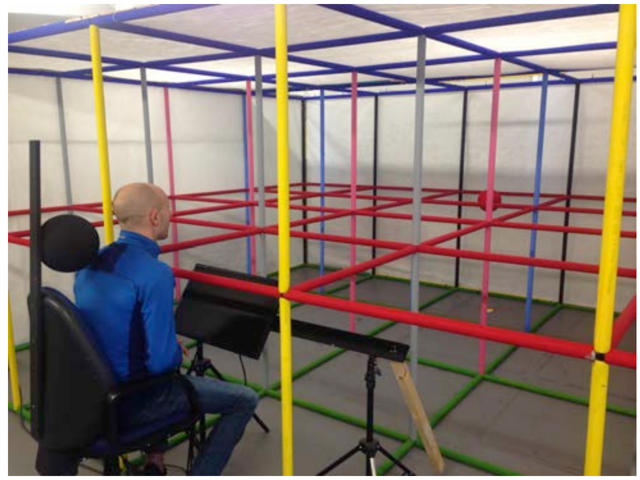

Fig. 2. A photograph of the experimental setting of the study in [17].

Through this process we have identified a set of key features of natural visual experience that, when digitally synthesised in an image, show an improvement across a range of objective measures when compared to images created with conventional projection methods (see Figure 1). These key features include a novel gaze contingent non-linear projective geometry; gaze contingent variations in acuity; simulation of binocular diplopia, gaze contingent modification of luminance, contrast, depth of field, and other image variables. These imaging processes can also be combined with bespoke methods of physical display, projection, and interaction to improve perceptual impact [3].

4. EVIDENCE FOR THE APPROACH

There is converging scientific evidence that Fovography images are preferred to conventional images when measured on a range of criteria. For example, eye tracking studies have shown significantly improved depth perception in Fovographs versus photographs [2]. A recent study [17] showed that Fovographs obtained higher ratings on four critical psychological variables over fisheye and linear perspective depictions of wide fields of view. The experimental setting consisted of a room-sized grid that fully encompassed each participants’ visual field. Three projections of the room made from the participants’ viewpoint were displayed on an LCD screen, including standard wide angle rectilinear, fisheye, and a Fovograph (see Figure 2). Participants were asked to rate the images on a range of measures.

Results showed that Fovographs obtained statistically significant higher ratings in the following psychological variables: liking ($t(15) = -3.049$; $t(15) = -3.25$, $p < .05$) spatial presence ($t(15) = -5.895$; $t(15) = -5.813$, $p < .05$), ecological validity ($t(15) = -3.802$; $t(15) = -2.129$, $p < .05$) and comfort ($t(15) = -4.672$; $t(15) = -3.569$, $p < .05$), compared to fisheye and linear perspective respectively. Overall, Fovographs were preferred, judged as having a better sense of presence, being more ecologically valid [10] and more comfortable to look at compared to the other two projections. This suggests that Fovography offers an improved representation of visual space compared to the standard geometric projections tested.

A pilot study [15] compared Fovographs presented on a 55’’ screen (Sony Bravia, KD KD-55X9005A) with the following immersive media technologies: wide-angle linear perspective (presented on the same screen as Fovographs), virtual reality (VR) head-mounted display (Oculus Development Kit2) and a large format rear-projected curved screen covering 180 degrees of the visual field (4m diameter, 1.45m high; facility available on site at Cardiff Metropolitan University). All media used in the study represented the same computer generated interior scene, as illustrated in Figure 3. We collected ratings for liking, comfort, immersion, spatial presence and ecological validity from 7 participants [10].

The comparison revealed that Fovographs were preferred ($t(5) = -3.081$, $p < .05$), rated as more comfortable to look at ($t(5) = -4.914$, $p < .05$) and more ecologically valid ($t(5) = -3.0003$, $p < .05$) compared to the wide-angle rectilinear counterpart. No significant difference was found between Fovographs and the other two technologies, VR and the curved screen, for liking, comfort, immersion and ecological validity ratings. Results show that Fovographs were judged being as equally powerful on the rated measures as VR and the curved screen. Spatial presence ratings revealed Fovographs scored significantly higher compared to wide-angle rectilinear pictures ($t(5) = -3.846$, $p < .05$), but significantly lower than VR ($t(5) = 2.940$, $p < .05$). No significant difference between Fovographs and the curved screen ratings was found, meaning that participants judged these two technologies having an equal sense of spatial presence.

These preliminary results suggest that Fovographs offer a comparable sense of immersion to a large format rear-projected curved screen, but using a much smaller display. Moreover, Fovographs are liked as much as, and judged as ecologically valid as VR, but with the advantage that participants did not need to wear the head-mounted display, thus avoiding its potentially adverse effects [20,11].

5. FOVOGRAPHY TECHNOLOGY

As noted above, our approach to creating technology based on the principles of Fovography departs from the widely accepted conventions of optically based image making, in which straight rays of light pass through a single aperture to be projected on a flat single surface as an inverted image. Whilst the behaviour of light is, of course, key to visual perception and imaging technology, rigidly adhering to the Physics-based rules of linear or curvilinear projection, particularly for wide fields of view, imposes well-known limitations [8].

As with any image technology, Fovography technology is based on modelling the behaviour of light, and the interaction between light and surfaces, but does so virtually using a set of computational processes that allow various novel manipulations of the resulting image. These processes can be used to

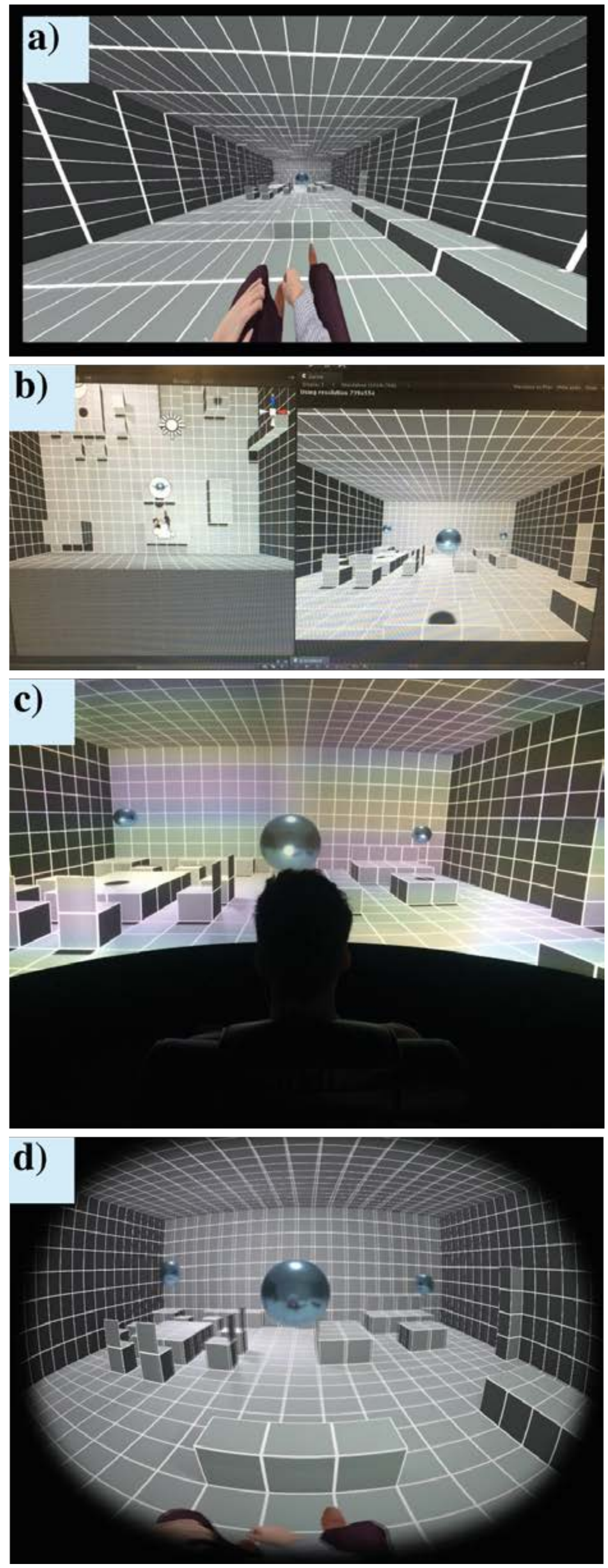

Fig. 3. Stimuli used for the experiment [15] showing a) wide-angle linear perspective, b) VR head-mounted set, c) 180º curved screen, and d) a Fovograph projection.

produce images that are more perceptually naturalistic than common alternatives, or have other useful geometric properties not achievable using conventional methods [16].

The Fovography process begins with a set of data points representative of a three-dimensional scene, with their various properties such as 3D coordinate and colour. The data can be captured from an optical device or devices, or generated or modelled computationally. The software then applies a series of mathematical transformations and coordinate system conversions to the 3D data in order to produce a two-dimensional image on a screen or display device. These transformations differ in five key ways from conventional linear projection based graphics pipelines:

1. The data points can be transformed along any arbitrarily shaped trajectory or path prior to hitting a virtual camera sensor in order to make an image;

2. The virtual camera sensor can be non-planar, curved or arbitrarily shaped;

3. There is no single aperture in the virtual camera, multiple or individual per-data point convergence;

4. The virtual camera sensor can be a volume through which light passes, rather than just a plane;

5. The resultant virtual camera sensor, whether layered or volumetric, can be unwrapped in various further ways to create a final 2D image on a flat surface.

An example image generated using these principles is shown in Figure 3d. In comparison to Figure 3a, a conventional linear perspective projection showing approximately the same angle of view, there are a number of geometrical differences. Notably, the central area of the image is less minified than in Figure 3a, and there is less stretching of space in the left and right regions. Note also the elliptical image boundary, which mimics that of the human binocular visual field [5].

The five key features of the technology broadly outlined above have now been integrated into a real-time interactive computer graphics pipeline running on a Graphics Processor Unit (2560-core Geforce GTX 1080). When applied with specific settings, this pipeline can automatically output a Fovography image modelled on human visual perception for any given combination of viewing position and fixation point in a virtual scene at frame rates over 120 frames per second at Ultra High Definition resolution. The technology offers a versatile set of functions and features in a real-time platform unconstrained by the physical laws of optics and cameras.

6. FUTURE DEVELOPMENTS

Fovography began as an artistic method of capturing the full scope of the human visual field in painting and drawing. It has since evolved into a general-purpose computational method of representing visual space more naturalistically. Our current goal is to turn this into a commercially available computer graphics technology for a wide range of possible applications, including medical imaging, entertainment, communications, photography and cinematography, and robot vision. The benefits to users we anticipate will be images that have greater sense of depth, breadth, presence, and immersion, and that offer a more engaging first-person perspective than can be obtained with current media technology.

7. REFERENCES

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

[2] Baldwin, J. (2016) *Can artistic methods be used to improve the perception of depth in pictures?* PhD Thesis, Cardiff Metropolitan University.

[3] Burleigh, A. & Pepperell, R. (2014). Improvements in and relating to image making. *International Patent Treaty Cooperation,* WO2014122477 A1.

[4] Flocon, A. and Barre, A. (1987). *Curvilinear Perspective: From Visual Space to the Constructed Image*. Berkeley, CA: University of California Press.

[5] Gibson, J. J. (1950). *Perception of the Visual World*. Boston: Houghton Mifflin.

[6] Hecht, H., Doorn, A., & Koenderink, J. J. (1999) Compression of visual space in natural scenes and in their photographic counterparts. Perception & Psychophysics, 61(7), 1269-1286. doi:10.3758/BF03206179

[7] Howard, I. P., & Rogers, B. J. (1995) *Binocular Vision and Stereopsis*. Oxford University Press.

[8] Kemp, M. (1990). *The Science of Art*. London: Yale University Press.

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

[10] Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. Presence: Teleoperators and virtual environments, 10(3), 282-297.

[11] Nichols, S. and Patel, H. (2002). Health and safety implications of virtual reality: a review of empirical evidence. *Applied ergonomics*, *33*(3), pp.251-271.

[12] Palmer, S. (1999) *Vision Science*. Cambridge, Mass.: Bradford.

[13] Pepperell, R. (2012) The Perception of Art and the Science of Perception, in *Human Vision and Electronic Imaging XVII,* ed. Bernice E. Rogowitz, Thrasyvoulos N. Pappas, Huib de Ridder, Proc. of SPIE-IS&T, Elec. Im. Vol. 8291, 829113.

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

[15] Pepperell, R., Ruta, N. and Burleigh, A. (2016). Exploring and evaluating a new artistic natural perspective and its possible

applications to mindfulness practice. [Conference Presentation] In: *Symposium of New Technologies for Mindful Awareness and Wellbeing*. London, 24th November, 2016.

[16] Pepperell, R. & Burleigh, A. (2017). *Improvements to and relating to mage making*. U.S. Patent 9,684,946.

[17] Ruta, N., Burleigh, A., Vigars, R., Barratt, E. & Pepperell, R. (2016). Evaluating an artistic method for depicting human visual space, *Applied Vision Association meeting*, London, 19th December 2016.

[18] Strasburger, H., Rentschler, I. & Jüttner, M. (2011) Peripheral Vision and Pattern Recognition: A review. *Journal of Vision* 11(5):1-82.

[19] Ten Doesschate, G., (1964). *Perspective: Fundamentals, Controversials, History*, Nieukoop: Hes & de Graaf.

[20] Wann, J.P., Rushton, S. and Mon-Williams, M. (1995). Natural problems for stereoscopic depth perception in virtual environments. *Vision research*, *35*(19), pp.2731-2736.