

# Exploring Spatial Augmented Reality With Holographic Projection



Author: **Agnes Cameron**

Supervisor: **Prof. Tim Wilkinson**

Date: **31/05/2017**

I hereby declare that, except where specifically indicated, the work submitted herein is my own original work.

*Signed* \_\_\_\_\_ *date* \_\_\_\_\_

# Technical Abstract

Spatial Augmented Reality (SAR) encompasses a range of projection techniques to cast images onto non-uniform and variable-depth surfaces. SAR is used to augment and change the perception of real objects within a 3D scene, and has applications ranging from urban planning to performance art. This project explores the application of holographic projection techniques to SAR scenarios. It encompasses both 2D and 3D projection, extending beyond traditional occlusion-based technologies and exploiting the unique transplanar properties of holographic projection.

Practical modulation schemes using the projection setup are reviewed in section 3, and Binary Phase Modulation is determined to be ideal for the SLM used. A range of holographic optimisation schemes are discussed, and the preferred generation method is established as One Step Phase Retrieval. Greyscale testing is suggested as a good qualitative perceptual metric for the limitations of different holographic generation schemes, with the Mean Squared Error proposed as a suitable quantitative metric in both 2D and 3D holography.

Section 4 of this project builds on the investigation of Fresnel slicing techniques for 3D holography developed in previous projects. The projection system is fully calibrated using the SLM as a virtual lens, and focal errors in the system are corrected for using linear regression. The fundamental limitations of holographic techniques are reviewed, and the additional effects of space-multiplexed 3D images are discussed. A demonstration of time- and space-multiplexed Fresnel techniques is investigated using CT scan data, exploring the limits of the current projection setup.

Section 5 develops SAR techniques on non-uniform surfaces using 2D holograms. A key application of SAR is the mapping of real and imaginary forms to 3D scenes. Previous projects have shown that the use of depth cues in holographic projection can be used to alter the perception of depth in an image. The use of textural and perspectival depth cues is investigated in constructing stereograms, examining the effect of varying the number of vanishing points in a scene. The effect of perspectival depth cues on the perception of '3D' animations is discussed, using both an orthographic and perspective-projected animation to test the combination of motion parallax with other cues.

The investigation of SAR is then extended to custom surfaces, demonstrating both static and animated projection mapping to non-uniform 3D scenes. Moderate and extreme false-perspective illusions are investigated in the context of architectural spaces, and a qualitative analysis of the effectiveness of false perspective projection is provided. Finally, animated SAR is demonstrated on a non-uniform 3D scene, and possible extensions to projection mapping approaches are suggested.

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Background</b>	<b>2</b>
2.1	Augmenting Visual Perception . . . . .	2
2.2	Depth Cues . . . . .	5
2.3	Perspective . . . . .	6
2.4	Fundamentals of Holographic Projection . . . . .	7
<b>3</b>	<b>Experimental Setup and Calibration</b>	<b>12</b>
3.1	Generating optimal Fourier holograms . . . . .	13
3.2	3D Holography . . . . .	16
<b>4</b>	<b>Information Limits in Holographic Projection</b>	<b>20</b>
4.1	Fundamental 2D losses . . . . .	20
4.2	Quality Metrics in 3D Holography . . . . .	21
4.3	Representing CT Scan Data . . . . .	27
<b>5</b>	<b>Spatial Augmented Reality on Nonuniform Surfaces</b>	<b>31</b>
5.1	Ray tracing 3D scenes . . . . .	32
5.2	Flicker stereograms . . . . .	34
5.3	Animation . . . . .	37
5.4	Custom surfaces . . . . .	39
5.5	SAR on a 3D scene . . . . .	40
<b>6</b>	<b>Conclusions</b>	<b>42</b>
6.1	Future Work . . . . .	43

# 1 Introduction

Spatial Augmented Reality (SAR) is a technique proposed by Raskar et. al [17] that uses surface-mapped projections to enhance, modulate and transform physical environments. It encapsulates possibilities for headset-less augmentation: both static and dynamic, inhabiting the gap between fully-tangible media, where physical objects are the subject of user interaction, and ‘see through’ AR, where the user wears a headset, or sees the real scene through a head-up display. This project explores the potential for applying holographic projection techniques to SAR scenarios, exploring the limits and potential for 2D, 3D, static and dynamic projection.

Digital holography represents a unique and efficient solution to many projection problems. However, holographic projection has yet to see a mainstream commercial implementation, in part due to the rapid rate of development of occlusion-based Silicon projection schemes. As Moore’s law slows and there is a plateau in the development in LED technologies, it is not infeasible to imagine a diversification in projection media, driven by an interest in SAR applications. Holographic projection has been used to create micro projectors [7], head-up displays (HUDs) [11] and fully-3D volumetric displays [12]. In a previous project, Freeman [11] demonstrated a mixed reality HUD system for pilots using holographic projection, which utilised eyeball tracking to exploit binocular effects. The development of 3D holographic HUDs is currently being investigated by startup VividQ – a Cambridge-based company looking to holograms as the next development in headset-based mixed reality applications [21]. Holograms have the unique property of being able to modulate the depth of focus of an image without changing the optical setup [5, 24], enabling the possibility for transplanar optical effects that reduce the nausea associated with depth cues used by traditional mixed reality applications [21].

It is thus of great interest to explore the aesthetic potential of holographic projection techniques, emphasising the unique focal properties. Projection mapping to non-uniform surfaces to create depth illusions has been demonstrated in previous 4th year projects by Bheemireddy [2] and Senanayake [19], whose work this project seeks to build upon. This project chooses to focus on the production of effective optical illusions and projection maps using holography, and to provide a detailed analysis of fully-3D holographic projection.

# 2 Background

## 2.1 Augmenting Visual Perception

We see a 3D world with a 2D eye, and are reliant on the blanket tools of visual perception to form a reliable realisation of the physical dimensions of the scene. The brain relies on a combination

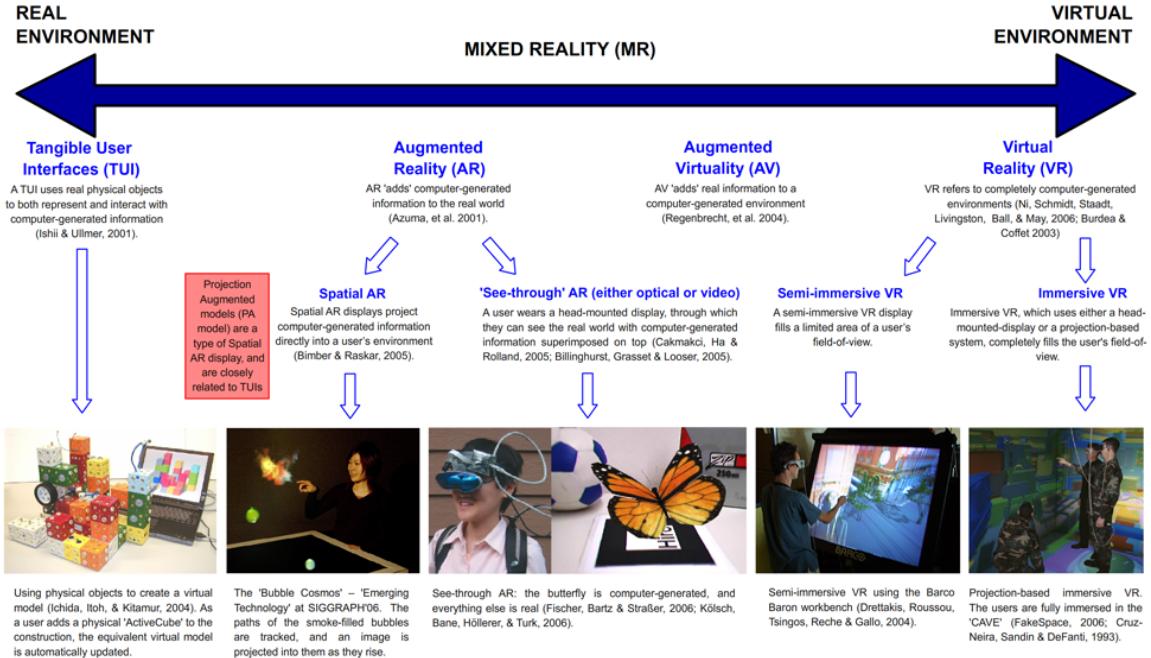


Figure 1: *The 'reality-virtuality' continuum [22]*

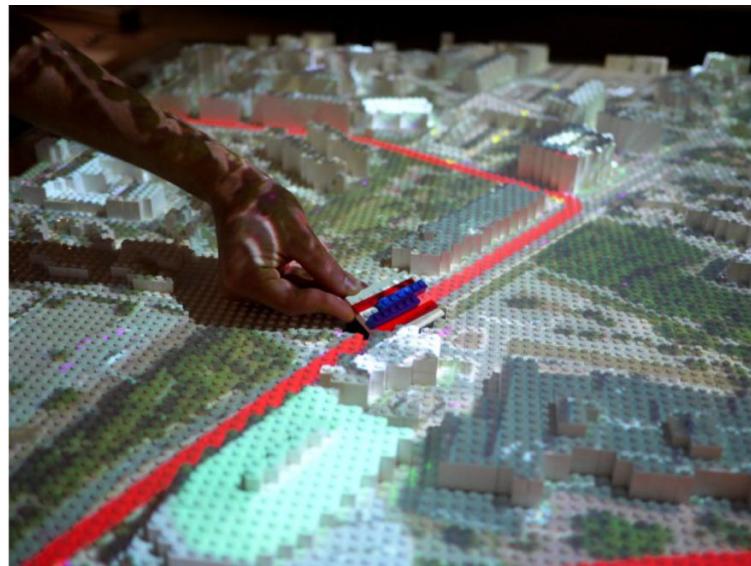


Figure 2: *Use of spatial AR in urban planning project, MIT [16]*



Figure 3: *The ‘sea monster’ illusion – an example of Amodal perception [4]*

of contextual cues and prior knowledge (a ‘schema’) to interpret the 2D image falling on the retina as that of a 3D world [18]. By manipulating depth cues within the scene we can fool the brain into decoding flat or near-flat objects as possessing depth, forming the basis for depth-based illusions.

3D computer graphics rely heavily on the principles of depth perception, rendering digitised 2D images from a mathematical description of a 3D scene. The description of the scene is constructed using a set of primitives – simple geometric forms that act as the building blocks for structures and surfaces [8, 13].

SAR relies both on the manipulation of images to modulate depth cues, and a mathematical understanding of the projection surface [17]. By augmenting 3D surfaces with additional images, the environment can be powerfully modulated. For example, the MIT project Materiable [16] utilises an aerial projection linked to a dynamic surface to provide a morphable tool for urban planning.

### 2.1.1 Gestalt Psychology

Gestalt psychology refers to the tendency of the brain to ‘fill in’ additional information needed to interpret an image. This is useful in the production of 3D illusions, particularly in ‘impossible shapes’ where the brain’s autocomplete is used to create an illusion of a shape that would be geometrically impossible. By distorting the context of a projection, the way an object is perceived can be changed. Amodal perception – the perceptive illusion that an object has volume based only on one visible surface – is a useful Gestalt effect that can be applied to suggesting depth in SAR applications [4].

## 2.2 Depth Cues

Depth perception may be based on either binocular cues (which pertain to the construction and position of the eyes) or monocular cues, which tend to rely on psychological effects [18]. For the purposes of this project, all of the effects used exploit monocular depth cues.

### 2.2.1 Binocular depth cues

Binocular cues for depth perception are derived from the use of stereo vision: the translation of the image plane between eyes provides a differential viewpoint: by examining the relative displacement of objects in a scene, it is possible to extract clues of relative depth.

**Binocular disparity** Distant objects – which subtend virtually the same angle at each eye – will show little disparity between the image seen by each eye, whereas objects in the near visual field show a large displacement. By combining these two offset images, the brain constructs a 3D image of near objects, using the binocular parallax in the near visual field. This effect is also known as *stereopsis*.

**Convergence** Both eyes focus on same object, rotating in order to do so. This rotation extends the extraocular muscles, contributing to a physical perception of depth.

### 2.2.2 Monocular depth cues

Monocular depth cues assume that both eyes see approximately the ‘same’ image, and can be divided into kinetic and pictorial cues. These cues are the most appealing for use in projection systems, and as such form the basis for the investigation of illusory effects in section 5.

**Kinetic cues** These rely on a moving image; the relative motion of objects in a scene can reveal information about their relative depth. This motion parallax effect differs from binocular vision in that the ‘stereo view’ is built up from the relative motion of a monocular camera within the scene, rather than the relative distance between two static views. However, the calculation involved in deriving these two results is the same. This forms the basis for the traditional stereogram, which takes 2 displaced scenes to reconstruct structural depth cues from the relative rotation of objects.

**Pictorial cues** These are perhaps the most diverse avenue of exploration for potential illusion, and the most practicable for a static projection setup. The eye may be tricked into the assumption of depth through a range of possible cues, listed in fig. 4. Of particular interest to this project are depth cues involving perspective, gradient, shadow and elevation, as all these effects are reproducible in monochromatic holographic projection [2, 19]. In particular, as Fourier holograms are in focus at all depths, it is possible to create high quality projections using slanted and even curved projection surfaces.

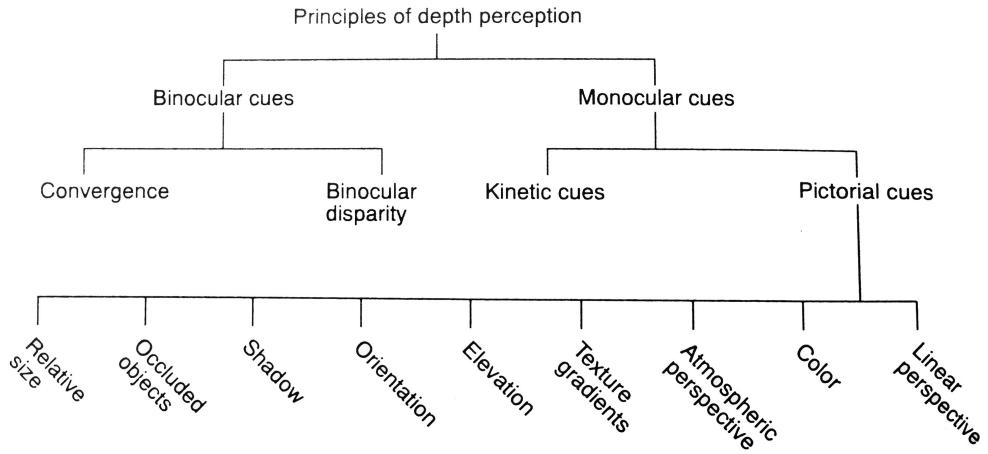


Figure 4: *Classification of depth cues [18]*

## 2.3 Perspective

One of the most powerful depth cues is perspective, the principle by which the apparent size of an object will decrease with the orthogonal distance from the viewer. Perspective has been used in representing 3D imagery since the Fifteenth century – developed by the painter Brunelleschi [1], and has been key in the development of art and architecture.

### 2.3.1 Linear Perspective

Linear perspective is most commonly used, and gives the impression of looking at a scene in front of the viewer. Parallel lines in the scene converge to ‘vanishing points’, which are collinear, and form the horizon line. Architectural projections by Jan Vrederman de Vries (see fig. 5) are an early example of linear perspective projection.

### 2.3.2 Curvilinear Perspective

Curvilinear perspective is somewhat more complex to calculate, and gives impression of immersion within a scene. It is used extensively in VR and game design, where the viewer should feel surrounded by an environment. The human eye has a curvilinear perspective [18], the effects of which are most pronounced in peripheral vision.

### 2.3.3 Forced Perspective

Forced perspective is an illusion based on the properties of linear perspective cues. By creating a greater artificial variation in size than would otherwise exist in an image (for example, narrowing

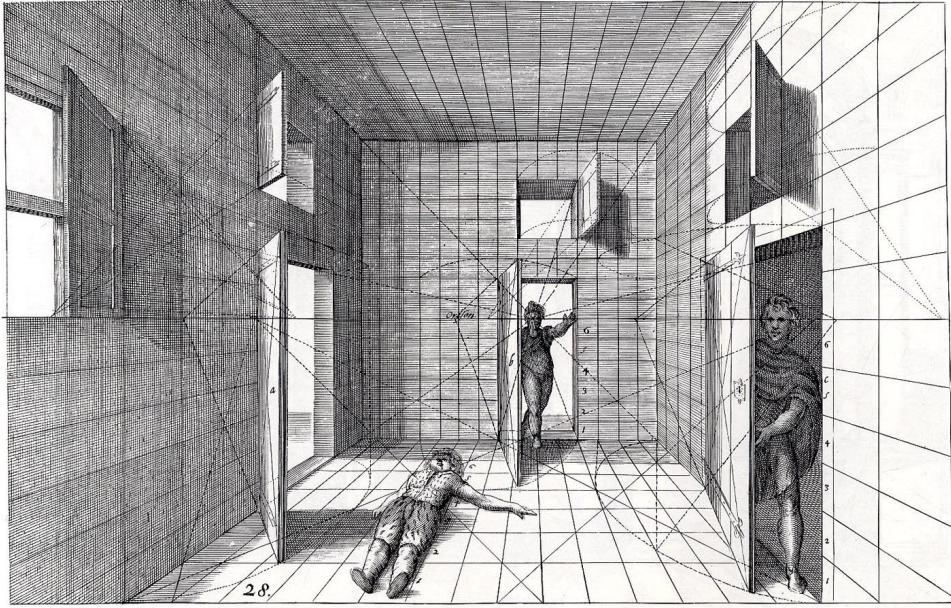


Figure 5: *Depth cues from linear perspective, in Jan Vredeman de Vries' Perspective, 1604 [10]*

the proportions of a set of steps as they become higher, thus appearing to stretch into the distance), it is possible to trick the eye into seeing a far deeper scene than is actually present. This is a technique used extensively in Renaissance art and architecture, and forms the central feature of the Palazzo Spada (see fig. 6), which has an apparent depth of 37 metres compared to an actual depth of 8 metres. Instead of a straight-sided parallelepiped (as is assumed from a view along the optical axis), the walkway is a truncated pyramid, narrowing and shortening at the far end. The tiles on the floor are not square, and instead trapezoidal. [18] The more extreme the projective illusion, the narrower the viewing angle from which it may be effective. In SAR, this is a trade-off: the more an environment is manipulated, the more limited the range of views from which the effect can be upheld.

## 2.4 Fundamentals of Holographic Projection

Holograms form an attractive projection scheme due, in part to their superior efficiency and size in comparison to traditional occlusion-based projection systems, which block around 80% of incoming light [7]. Instead of blocking light, the hologram bends it, using the principles of diffraction to form an image in the ‘far field’ of the optical system. A comprehensive analysis of both Fraunhofer and Fresnel computer generated holography is discussed in [7], in which a lightweight holographic projection system is developed.

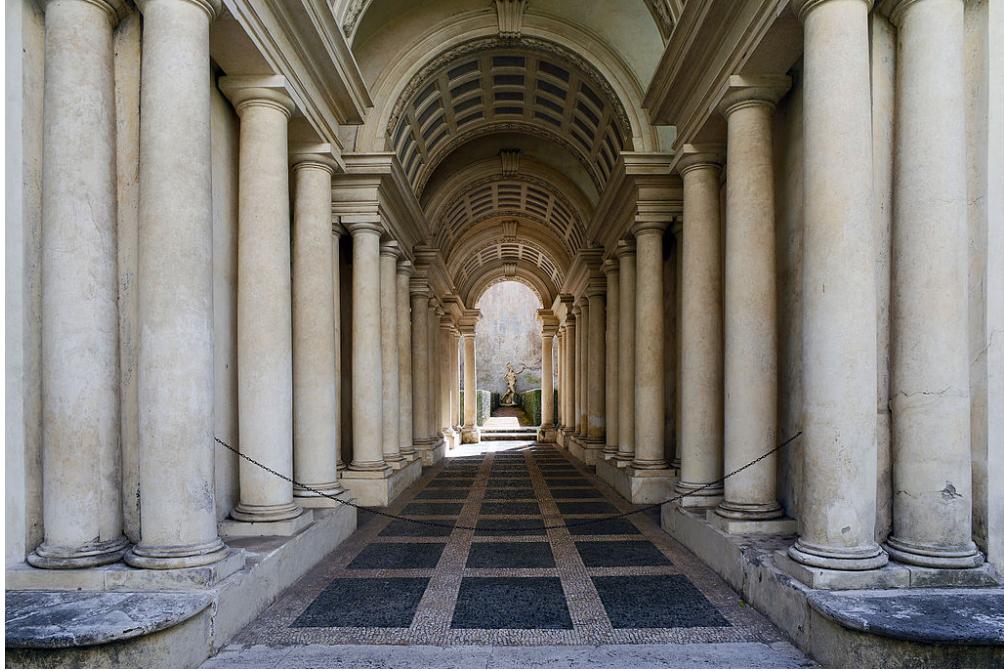


Figure 6: *The gallery of the Palazzo Spada, an example of false perspective [18]*

#### 2.4.1 Diffraction Regions

When light is passed through an aperture, the Huygens wavelets that comprise the flat wavefront are deflected, forming an interference pattern. For a single aperture and a flat incident wavefront, the pattern of light observed after diffraction through the aperture varies with distance (see fig. 7).

It may be shown, through the paraxial and Taylor series approximations that ‘far field’ or Fraunhofer image of this wave forms the sinc function (e.g. the Fourier Transform of a rectangular wave) [2, 19]. Likewise, through the paraxial approximation (the assumption that points in the object field are approximately coincident with the optical axis), the Fresnel region may be represented by the Fourier Transform of the aperture multiplied by a parabolic term.

#### 2.4.2 Computer generated holography

As the Fourier Transform is linear, for an array of apertures, it may be shown through linear summation that the far field diffraction pattern is the Fourier Transform of that array. This principle forms the basis for computer generated holography. By changing a pattern of apertures on the hologram to form the Inverse Fourier Transform of an image, that image may be replicated in the far field. The computer generated holograms (CGHs) described in this project are created using a ferroelectric liquid crystal (FLC) device known as a Spatial Light Modulator (SLM). The

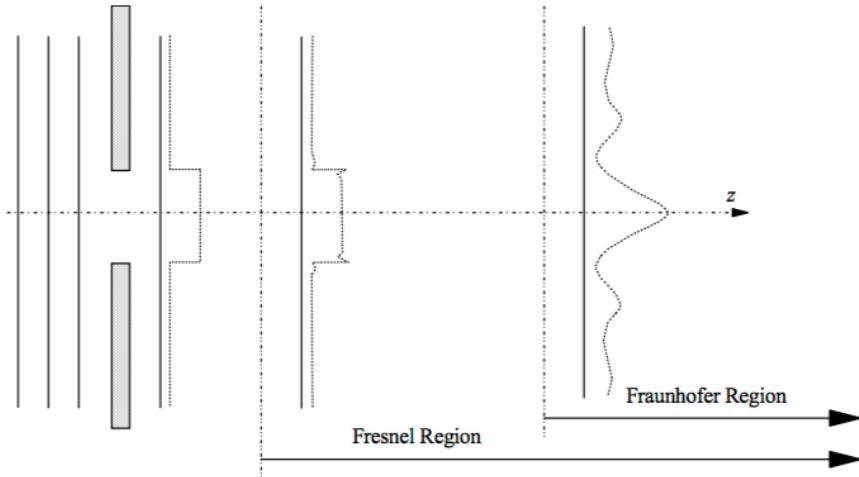


Figure 7: *Showing diffraction regions after aperture - general case can be derived through linear summation [23]*

SLM is constructed from a rectangular array of individually-addressable pixels, the phase and opacity of which may be changed between binary values.

#### 2.4.3 Conventions for co-ordinate systems

Here we define the replay field  $\mathbf{I}(x, y)$  as the visible region of the projection, and the hologram  $\mathbf{h}(u, v)$  as the image on the SLM. For Fraunhofer (far-field) holography, these are related by:

$$\mathbf{I}(x, y) = \mathcal{F}[\mathbf{h}(u, v)] \quad (1)$$

This transform must be discretised over the finite-pixel SLM, thus, the Discrete Fourier Transform is given by:

$$\mathbf{I}(x, y) = \sum_{x=\frac{N}{2}+1}^{\frac{N}{2}} \sum_{y=\frac{N}{2}+1}^{\frac{N}{2}} \mathbf{A}(u, v) e^{\frac{-j2\pi}{N}(ux+vy)} \quad (2)$$

Where  $\mathbf{A}(u, v)$  represents a single aperture (e.g. a pixel) on the SLM, and  $N$  is the number of pixels on the SLM.

#### 2.4.4 Modulation schemes

Due to the properties of FLCs, FLC devices can only represent real, binary values of amplitude and phase. Whilst multiphase devices are achievable using Nematic Liquid Crystal (NLC) systems,

these are slow and unstable, due to lower viscosity, and are thus unsuitable for many projection applications. [23] As such, any display of a desired image will be lossy, as it is not possible to display the true value of the IFT of the replay field [7]. SLM holograms may be formed by modulating either the amplitude or the phase of the FLC pixels: these schemes are called Binary Amplitude Modulation (BAM) or Binary Phase Modulation (BPM) respectively.

**Binary Amplitude Modulation:** If the amplitude of the real component of the IFT at a pixel is less than the mean of the IFT amplitudes, that pixel is assigned amplitude 0 – if greater than the mean, it is assigned 1

**Binary Phase Modulation:** If the amplitude of the real component of the IFT at a pixel is less than the mean of the IFT amplitudes, that pixel is assigned a phase shift of  $-\pi/2$  – if greater than the mean, it is assigned  $+\pi/2$  (for the ideal case)

Previous work has shown that BPM is a preferable modulation scheme [2, 7, 19], due to the presence of a ‘zero order’. This lies at the centre of the replay field, and corresponds to the average energy of the hologram. As, for an amplitude scheme, this average will always lie between 0 and 1, the zero order forms a central bright spot that distracts from the desired image. Phase modulation schemes are ‘DC balanced’, with an ideal scheme ( $\pm\pi/2$  switching angle) suppressing the zero order entirely.

As the modulation scheme is purely real, there is no distinction between the desired image and its 180 degree rotation, thus both are displayed in the replay field (see fig. 8). This restricts the display of holographic images to the top half of the SLM, otherwise there is overlap.

$$\mathbf{I}(x, y) = \mathcal{F}[\mathbf{h}^*(u, v)] = \mathcal{F}[\mathbf{h}(u, v)] = \mathbf{I}(x, y) \quad (3)$$

#### 2.4.5 Projective 3D Holograms

Chen [8] provides a broad overview of approaches to 3D holography, modelling two approaches to graphical construction of 3D scenes.

**Image order:** This is based on a ‘ray casting’ strategy, where a mathematical description of the scene is stored, and the displayed image adapted to the ‘rays’ cast by the viewer onto the scene. The main problem with image order techniques is the resolution required to generate a realistic image, which is prohibitively high for current SLM technologies, and a means of detecting the viewer’s position relative to the scene is also crucial.

**Object order:** These schemes deconstruct the objects in the scene into object primitives, and assemble the scene over multiple focal depths. This is more easily interpreted into a holo-



Figure 8: *Binary phase modulation reconstruction, showing 180 degree hermitian symmetry in the conjugate image*

graphic context, as quantisation along the optical axis provides an easy means of deconstructing 3D objects into simple geometric primitives.

Two main object-order techniques exist for the generation of 3D holograms: Fresnel slicing [7, 3, 19], and the ‘polygon method’ [24]. Fresnel slicing generates an image using a set of 3D slices (‘kinoforms’) to construct an image, multiplexed into a single hologram that forms a set of ‘slices’ in the replay field. The polygon method – which requires a high-resolution SLM to generate a replay field of any meaningful size – tiles together a polygon mesh through manipulating individual ‘tiles’ of the SLM.

#### 2.4.6 Fresnel Holography

Implementations of Fresnel holography are described in [3, 5, 7, 19, 24], who discuss both direct-view (where the image is focussed on the eye of the viewer) and projective methods that use Fresnel techniques to generate 3D images. Bheemireddy [3] proposes a direct-view method of live Fresnel slice generation from point cloud data, using geometric primitives to reconstruct an image captured with a Kinect.

Buckley outlines a method of Fresnel generation, considering the Fresnel region with an extra parabolic term applied to the Fourier Transform in the spatial frequency plane. The hologram  $\mathbf{h}(u, v, z)$  required to generate image  $\mathbf{I}(x, y)$  in the replay field at a depth  $z$  is given by:

$$\mathbf{h}(u, v, z) = \mathcal{F}^{-1}[\mathbf{I}(x, y)e^{\frac{-j\pi}{\lambda z}(u^2+v^2)}] \quad (4)$$

Thus, the value of the image  $\mathbf{I}(x, y)$  in the replay field is given by:

$$\mathbf{I}(x, y) = \mathcal{F}[\mathbf{h}(u, v, z)e^{\frac{j\pi}{\lambda z}(u^2+v^2)}] \quad (5)$$

The parabolic phase factor may be considered as a ‘virtual lens’ [7] of phase profile  $\mathbf{L}(u, v)$ , such that:

$$\mathbf{L}(u, v) = e^{\frac{j\pi}{\lambda}(\frac{u^2+v^2}{2f_v})} \quad (6)$$

Where  $f_v$  represents the focal length of the ‘virtual lens’, and thus the plane at which the 3D image is focussed. This consideration is very useful in defining Fresnel holograms as part of an optical system, as it allows standard optics results such as the Lensmaker’s equation to be applied [7, 19].

### 3 Experimental Setup and Calibration

The projection setup used in this project was developed by Freeman [11]. A coherent source is provided by a DPSS 532nm 50mW laser, with a collimator and beam expander used to achieve near-uniform illumination of the SLM. The SLM is a binary phase SXGA-R2 ForthDD ferroelectric LCOS microdisplay with a refresh rate of 1440Hz, and a resolution of 1280x1024, and a pixel pitch of  $13.6\mu m$ . The holograms uploaded to the SLM are 24-bit images, with a greyscale depth of 8 bits, and RGB components. The SLM displays each bit plane successively, with ones and zeros decoded as opposing phase modulations of each pixel.

Photographs of the replay field were captured with a Canon 5d mark III DSLR with a 24-105mm Kit Lens<sup>1</sup>, and with an iPhone 5 camera where the full camera was unavailable.

#### 3.0.1 Switching angle

The optimal switching angle for a Binary Phase Modulation scheme is  $90^\circ$ , giving a theoretical maximum transmission efficiency of  $\eta_{switch} = 100\%$ . The actual switching angle ( $\theta_{FLC}$ ) of the LCOS SLM is only  $33^\circ$ , meaning that a zero order is still present in the image, and the efficiency is

---

<sup>1</sup>with kind help from Gary Zhexi Zhang

lowered. Jones matrix analysis gives the overall transmission efficiency as  $\eta_{trans} = 1 - \sin^2(\theta_{FLC})$ , giving a theoretical  $\eta_{trans}$  of 70.3% for this scheme. Higher switching angle SLMs have been theoretically developed, but are not used in practice due to cost and stability issues [23].

### 3.0.2 Viewing Angle

The viewing angle of the projection system is given by the first order diffraction angle  $\theta$ , which is calculated using the relation:

$$\theta_{view} = 2 \tan^{-1} \left( \frac{\lambda}{\Delta} \right) \quad (7)$$

Where  $\lambda$  is the wavelength of the laser, and  $\Delta$  is the pixel pitch of the SLM. For this setup, the theoretical SLM viewing angle is  $2.24^\circ$ , though this is expanded by a 2-lens array (see fig. 9) to an experimental value of  $46.5^\circ$ . The demagnification of the second lens ( $f = 19mm$ ) means that the far field of the projection is achieved very close to the projector.

### 3.0.3 Pixel shape compensation

The shape of the microdisplay pixels contributes to an effect known as *apodisation*. As the pixels have a square (rather than the theoretical  $\delta$ -function) shape, a sinc envelope attenuates the image in the replay field. This places a limit on the size of the replay field that may be produced. The attenuation from pixel shape effects at point  $(x, y)$  in the  $N$ -pixel replay field is given by a sinc envelope:

$$\mathbf{A}(x, y) = \left( \frac{\sin \frac{\pi x}{N}}{\pi x} \times \frac{\sin \frac{\pi y}{N}}{\pi y} \right)^2 \quad (8)$$

This leaves a central field of 640x360 pixels (approximately 80% of the available replay field) with a reasonable attenuation [2].

## 3.1 Generating optimal Fourier holograms

The initial projection setup uses 2 lenses, with second (outer) lens chosen with a short focal length, allowing the far field image of the hologram to be formed very close to the projector, and thus the replay field to be effectively ‘in focus’ at all depths. This project uses a  $f = 19mm$  outer lens for its relatively low distortion, in comparison to the ball lens used by Freeman [11], which was shown to introduce large optical aberrations.

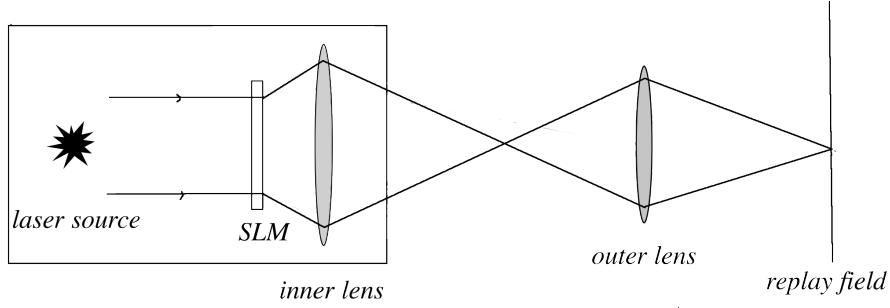


Figure 9: *Projection setup for 2D holography: 2-lens system*)

### 3.1.1 Naive Transform Algorithm

The simplest CGH algorithm uses a binary phase transformation based on the quantisation of the IFT of the target replay field image. In all CGH generation algorithms, the square root of the image intensity is used as the eye acts as a square law detector [5]. The stages of the basic transform are as follows:

#### 1. Square root intensity:

$$\mathbf{T}(x, y) = \sqrt{\mathbf{I}(x, y)} \quad (9)$$

#### 2. Fourier Transform:

$$\mathbf{h}(u, v) = \mathcal{F}[\mathbf{T}(x, y)] \quad (10)$$

#### 3. Binary Phase Quantisation:

$$\mathbf{h}_{quant}(u, v) = \begin{cases} -1 & \text{if } Re[\mathbf{h}(u, v)] < 0 \\ 1 & \text{if } Re[\mathbf{h}(u, v)] \geq 0 \end{cases}$$

This implementation results in an edge enhancement (see fig. 10b): energy in the target image is concentrated at low frequencies, as FT conserves energy, the energy becomes concentrated at high frequency in the replay field. This edge enhancement may be avoided by adding an initial random phase distribution to the target image. As the human eye is insensitive to the phase of light, this does not adversely affect the appearance of the replay field. This addition is performed in the first step:

$$\mathbf{T}(x, y) = \sqrt{\mathbf{I}(x, y)} e^{j\phi(x, y)} \quad (11)$$

### 3.1.2 Iterative Techniques

In order to improve on the basic hologram generation, a number of iterative techniques have been developed to optimise for a particular image. Thorough research into optimal generation techniques has been investigated in previous projects [2, 5, 7, 19], and as such only a brief comparative discussion is outlined here.

**Direct Binary Search:** Instead of using the IFT of the target image, a purely iterative technique is used. A random array of pixels are initialised as the hologram  $\mathbf{h}(u, v)$ , the structure of the replay field determined, and the error between the actual and desired replay field is calculated using a cost function such as the Structural Similarity Index or Mean Squared Error. Iteratively, random pixels in  $\mathbf{h}(u, v)$  are flipped; if the cost function is lowered by a pixel flip, that pixel value is kept: if it is increased, the value is rejected. This is repeated until the cost function falls below some threshold acceptable value. This produces a reasonable replay field, however, the algorithm is very slow (requiring around 200,000 iterations for good convergence [19]) and has a tendency to find local, rather than global minima.

**Gerchberg-Saxon:** This algorithm iteratively determines the phase profile of  $\mathbf{h}(u, v)$  required to generate a target image in the replay field. It loops between the hologram and desired image, applying constraints at each iteration. A local minimum may be reached within around 20 iterations [19], four orders of magnitude faster than Direct Binary Search, though it requires both a FT and IFT every step, reducing the computational saving somewhat.

**One Step Phase Retrieval (OSPR):** First demonstrated by Buckley [5], OSPR is a solution to high-quality hologram reconstruction that relies on time multiplexing of holograms, exploiting the response time of eye in order to reduce noise in the replay field. This time response – around 1/60 of a second – is due to the isomerisation of the protein opsin in the retina [2], and forms the basis for frame-based animation systems. Thus, by overlaying slices within this 1/60 second window, noise variance  $\sigma^2$  may be reduced by  $1/N$  factor [7]. Slices must be independent and identically distributed (IID) in order to produce good-quality images. The noise variance is considerably more significant in the perception of illusion than the noise mean [7], and thus the images produced appear to be of high quality. The perceived image is approximated by:

$$I(u, v) = \frac{1}{N} \sum_{i=1}^N N |\mathcal{F}[H_i(x, y)](u, v)|^2 = \alpha^2 T(u, v) + \frac{1}{N} \sum_{i=1}^N N \epsilon_t(u, v) \quad (12)$$

**Adaptive OSPR** This is an extremely high-quality method for hologram generation, combining iterative and time-domain techniques to use active compensation for existing noise in the current frame within the next frame. At each cycle of the algorithm, the reconstruction of the

Figure 10: Optimising modulation scheme

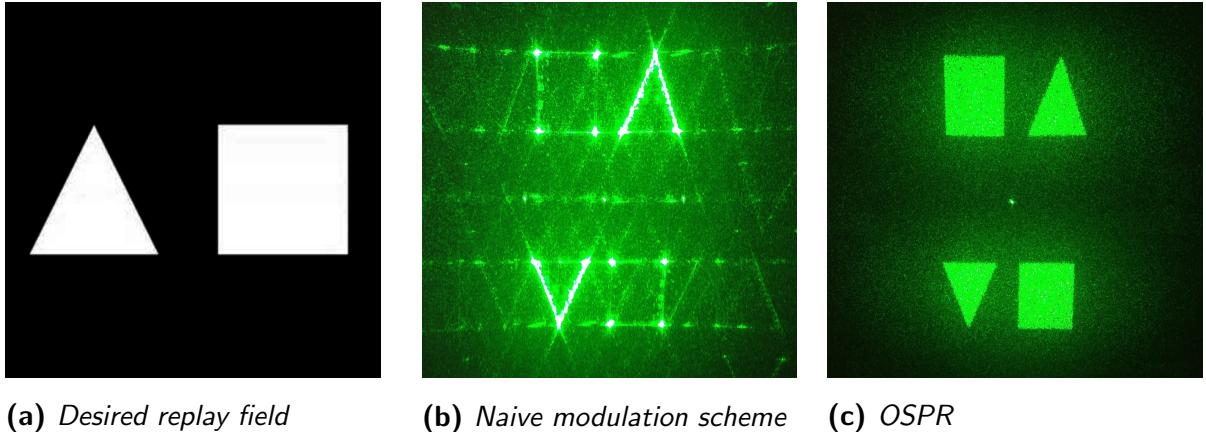


image is calculated, summed with prior distributions and normalised, finding the cumulative average of all reconstructed sub-frames. By making the new target image the difference between the time-averaged reconstruction and the target image, the noise is progressively compensated for, leading to faster convergence [2, 7, 19].

Cable [7] demonstrates that the noise mean of ADOSPR is given by  $\mu/N$  and the noise variance by  $\sigma^2/N^2$ , giving superior reconstructions to non-adaptive OSPR. In perceptual terms, this improvement is noticeable, but more slight on projective system (in comparison to computer reconstruction) as the improvement is obscured somewhat by optical non-idealities inherent in the system.

### 3.2 3D Holography

In order to calibrate the experimental setup for 3D holograms, it was first necessary to characterise the optical setup within the projector. The projector already contains an internal lens, the focal length of which must be determined in order to produce holograms at a known depth. This is the arrangement used for Fresnel holography in [6], which established an exploration of single-lens multifocal projection.

In order to calibrate this focal length, ‘zone plates’ were used to turn the SLM into a virtual lens. This is a method described in [20], and has the advantage of a sharp, clear focal depth (the point at which a central spot in the replay field is sharpest), making them ideal for this early calibration stage. This required a degree of trial and error, requiring a while to find an image that would focus, as even a small variation in the focal depth of the virtual lens can make a large difference in the replay field.

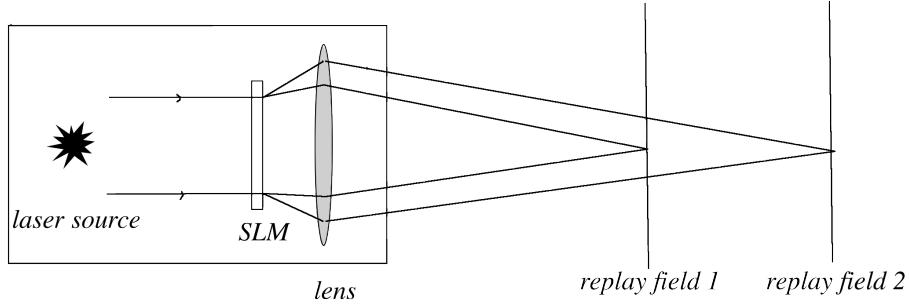


Figure 11: Variable-depth holographic projection with one lens

### 3.2.1 Zone plates

The zone plate – also known as a Fresnel lens – is a lensing technique used on reconfigurable optical devices to create a virtual lens of modifiable focal depth. of a phase-based lens transform [20]. Concentric rings phase-shift light from the laser to focus at a single point in the replay field. Modulating the width of the rings  $r_n$  changes the focal depth  $f$  with the relation:

$$r_n = \sqrt{n\lambda f + \frac{n^2\lambda^2}{4}} \quad (13)$$

$$\approx \sqrt{n\lambda f} \quad (14)$$

For multiple zones:

$$f = \frac{2r_N \Delta r_N}{\lambda} \quad (15)$$

Where  $r_N$  is the radius of the outermost zone. A zone plate generated with ring width  $r_n$  will thus produce a focal plane at a point at depth  $f$ , and can be used to determine the magnitude of the lens in the projector, using the lens maker's equation.

$$\frac{1}{f_{total}} = \frac{1}{f_1} + \frac{1}{f_2} + \dots + \frac{1}{f_n} \quad (16)$$

Thus, we may derive the inner lens focal length,  $f_{inner}$ , for a zone plate of focal length  $f$  that is focussed at a distance  $d$  from the projector:

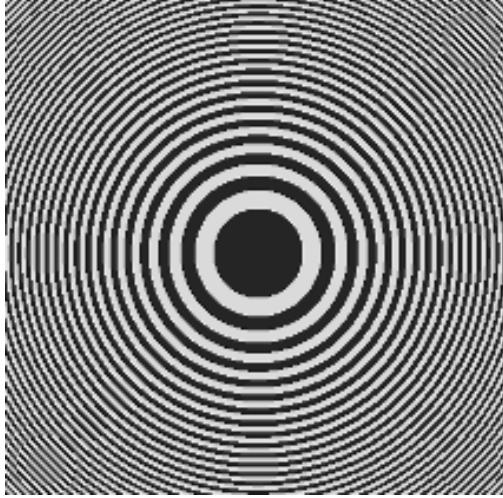


Figure 12: *The central order of a single-mode zone plate.*

$$d = \frac{f \cdot f_{inner}}{f + f_{inner}} \quad (17)$$

Thus,

$$f_{inner} = \frac{f \cdot d}{f - d} \quad (18)$$

### 3.2.2 Calibrating Depth

Previous investigations of Fresnel holography using this projection system [19] found that there was a moderate variation in the expected and actual focal depth of the Fresnel hologram. In order to investigate this, the depth error,  $\epsilon_d$  of the laser behind the aperture was determined using linear regression, and the setup recalibrated. This distance is slightly greater than the actual distance, as the rays that illuminate the SLM are not perfectly parallel, and SLM itself is not perfectly flat; both of these effects create an additional effective focal effect. When applied to the Fresnel generation algorithm using an  $\epsilon_d$  of 0.23m, this gave significantly better results fig. 13, with all planes in a range of a metre focussing within 1-2cm accuracy.

The most accurate calibration results for  $f_v$  were established behind the inner lens – thus producing an actual projection setup shown in fig. 14. It was not possible to calibrate a sufficiently-accurate value for  $\epsilon_d$  for the far-field focal depth shown in fig. 11, which could be due to additional artifacts in the optical system. This system was well-calibrated, however, as the focal plane of the virtual lens was in the extreme near field, it is subject to the defocus noise effects discussed in section 4.2.1.

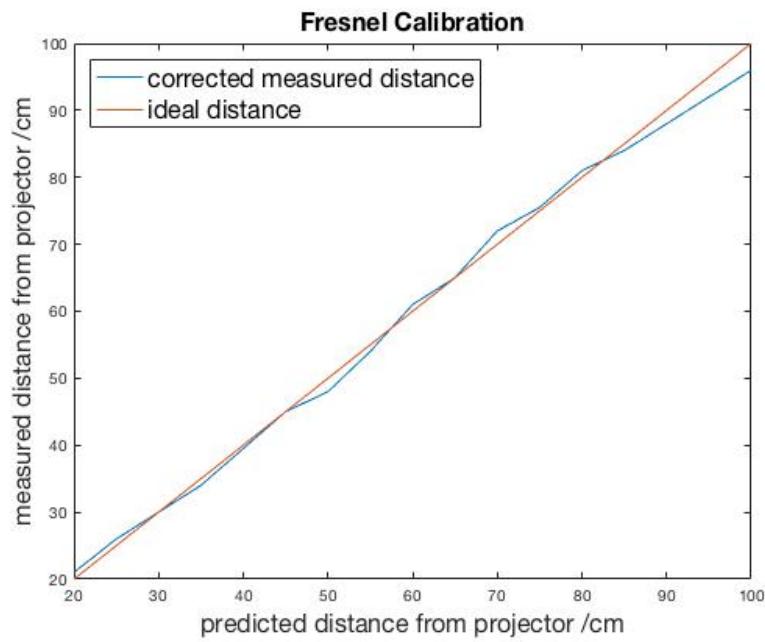


Figure 13: Desired vs. actual focal depth for the calibrated near-field setup

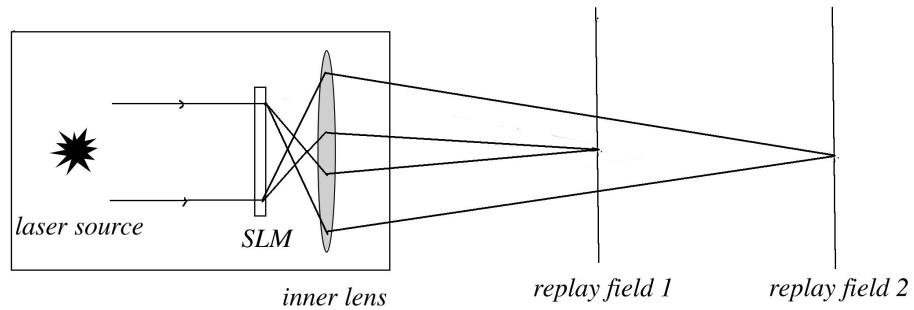
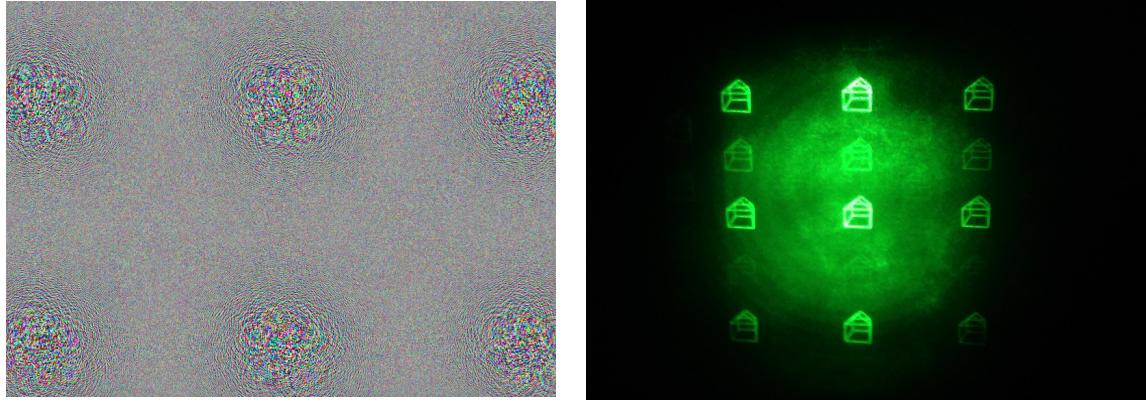


Figure 14: Final calibrated projection setup, with Fresnel holograms focussed in the near field

Figure 15: Fresnel Hologram and Replay Field



(a) *Image of a Fresnel hologram, showing characteristic ‘zones’*  
 (b) *Image captured at a depth of 400mm in the replay field*

### 3.2.3 Full Fresnel Holograms

A simple implementation of the variable-depth hologram is to XOR a generated hologram with a zone plate. However, higher harmonics mean that the image is repeated several times, with the energy evenly spread across them, making the efficiency low, and the resultant images are low resolution. Instead, the virtual phase term of the additional lens is introduced in the initial image transform – now a Fresnel, instead of Fourier transform, substituting  $f_v$  for  $z$  in equation (5). This adds a ‘zone plate like’ appearance to the hologram, as can be seen in fig. 15a, which clearly shows a radial pattern amongst the holographic material.

## 4 Information Limits in Holographic Projection

### 4.1 Fundamental 2D losses

According to the conservation of energy, the energy in the hologram and the replay field must be the same. Thus any information not represented in the hologram will be dispersed as noise in the replay field [7].

Information theory provides an absolute limit on the practical bandwidth of binary modulation schemes. The minimum loss generated for an 8-bit greyscale image is derived in [7]. For a generic implementation, the information content in the replay field must be the same as that in the hologram, as the Fourier Transform is a loss-less operation. Thus, for an  $M \times M$  bit IFT represented on the hologram, a  $M \times M$  bit image is theoretically reproduced in the replay field.

If we consider a binary modulation scheme, however, the informational content of the replay

field is reduced to  $M \times M/2$  bits, due to the conjugate symmetry inherent in binary schemes. Furthermore, considering the holographic reconstruction of an 8-bit greyscale  $M \times M$  bit image in the replay field, we further require  $8 \times M^2/2 = 4M^2$  bits of information on the SLM. As the SLM has an informational content of only  $M \times M$  bits, we are thus left with an inherent loss of  $3M^2$  bits, which may only be compensated for by using a multi-phase modulation (8 phase ‘bits’ sufficient), which is not possible using this projection setup.

#### 4.1.1 Effective Greyscale Resolution

Crucial in the display of greyscale images is the number of discernible grey levels that can be perceived in the replay field. In order to test the effectiveness of a particular modulation scheme, a ‘grey level’ test is used. The effective greyscale resolution provides an intuitive metric for noise effects on image resolution. The greater the overlap between individual pixel distributions, the more difficult it becomes to distinguish between levels. For the grey levels to be distinguishable,  $\Delta T \geq \sigma$ , where  $\sigma$  is the standard deviation of the noise.

Thus, the effective greyscale resolution is given by the ratio of the maximum pixel amplitude in the replay field to the standard deviation of the noise:  $\frac{T_{max}}{\sigma}$ , giving a bit depth  $\log_2 \frac{T_{max}}{\sigma}$ . This may be interpreted as an application of the Shannon-Hartley theorem for the maximum transmissible information rate over a noisy channel, defining the capacity of the holographic transform.

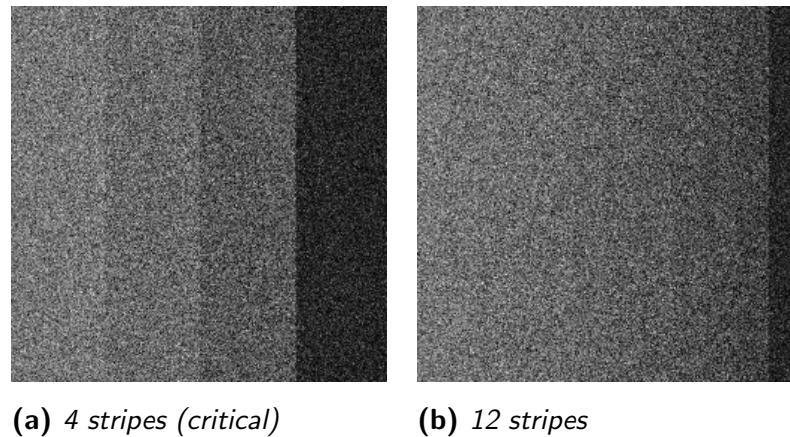
Fig. 17a demonstrates a subcritical case for the OSPR modulation scheme – the four grey levels are clearly distinguishable. By contrast, fig. 16b represents a supercritical case for six levels over the naive binary modulation scheme, demonstrating a noise term far in excess of  $\Delta T = 0.2$ . For the critical case shown in fig. 16a, the levels are barely-perceptible at a  $\Delta T = 0.33$ , giving an effective noise level of approximately  $\sigma = 0.33$  for this case. For OSPR, the critical noise level is significantly higher, with 12 grey-levels giving an approximate  $\sigma = 0.091$ .

These metrics may be used in real holographic projection systems, in which the perceptual noise can vary significantly from the theoretical noise, due to additional effects. Thus, these are taken as a means of evaluating the effectiveness of Fresnel projection schemes, which prove otherwise complex to characterise [3, 7, 19].

## 4.2 Quality Metrics in 3D Holography

Although a number of quality metrics exist for 2D holography, quality metrics for 3D holograms are considerably less well defined [3, 7]. In general, preferred 2D metrics (such as the Structural Similarity Index, Universal Image-Quality Index) favour image structure over individual pixel values as a measure of success, which ties in with human perception of an image [3].

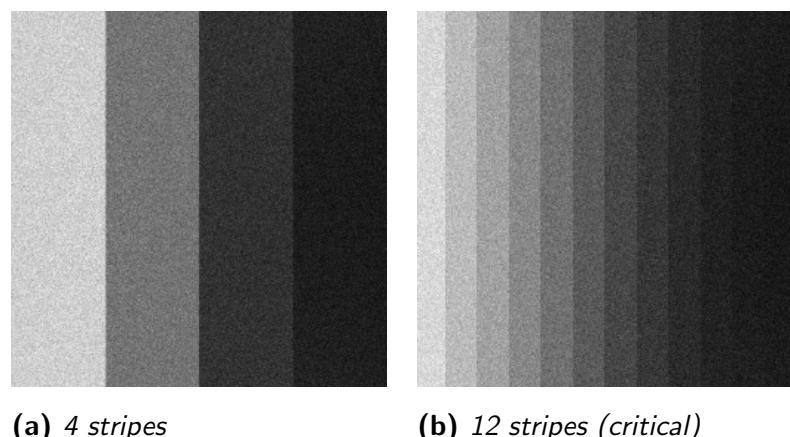
Figure 16: Greyscale Resolution, Naive Phase Modulation



(a) 4 stripes (*critical*)

(b) 12 stripes

Figure 17: Greyscale Resolution, OSPR



(a) 4 stripes

(b) 12 stripes (*critical*)

When considering Fresnel slicing techniques, each additional Fresnel plane introduces a defocus blur to all other planes. Blurring from defocussed planes cannot be considered a priori, and it is thus difficult to consider the quality of a 3D hologram in comparison to a ‘target’. Bheemireddy [3] suggests applying the mean squared error only over points in the desired replay field, a more selective metric is used.

The main issue with the MSE as a quality metric is that it accounts for the mean of the image noise, rather than the variance (as in the Structural Similarity Index) – however, as the holograms in this experiment are generated using the same algorithm (ADOSPR), of more interest is the relative, rather than the absolute noise distribution, so this issue may be overlooked. A slight change to all the pixels in the image produces a subtle perceptual change, but will drastically effect the MSE. Thus, normalised values for the MSE are used.

$$MSE = \frac{1}{N} \sum_{n=1}^N (T_n - R_n)^2 \quad (19)$$

#### 4.2.1 Depth effects on noise in Fresnel holograms

3D holography generates a volumetric, rather than a flat noise field, changing the statistical properties of noise distribution. Volumetric noise is not fully characterised, though it can be shown to increase as the focal length shortens. The signal energy variance remain relatively similar to their Fourier equivalents within  $f_v > 0.1$ ,  $f_v < -0.1$ . For schemes that vary out of this ‘Fresnel range’, additional noise effects are observed in the replay field [7].

This effect is due to the noise from the defocussed conjugate plane dispersing across the replay field. As the focal power increases, this plateaus, as the noise is more evenly dispersed. The noise power is constant in the replay field, and it becomes localised in the conjugate image as virtual focal length tends to infinity. A graph of the variation in Mean Squared Error with the log of the focal length is shown in fig. 18.

Fig. 19 shows the visual change in noise distribution with focal power. Whilst not so visible in printed form, the replay field of fig. 19c is subtly noisier than that of 19b and 19a, where the noise is concentrated in the conjugate image. This can be clearly seen in fig. 20, which compares holograms focussed in the near and far field using different lensing schemes.

#### 4.2.2 Different Lens Setups

In order to produce a hologram with the single-lens projection setup in fig. 14, it is necessary to focus the holograms in the extreme near field, introducing noise effects discussed in section 4.2.1. However, by focussing the images onto a second lens (at a variable depth) it is possible to

Figure 18: Plot of the Mean Squared Error against the log of the focal length

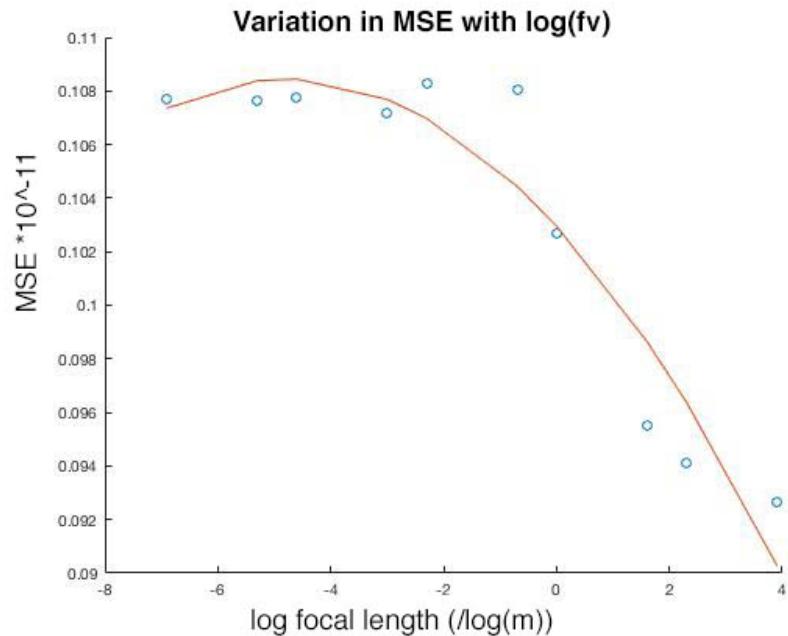
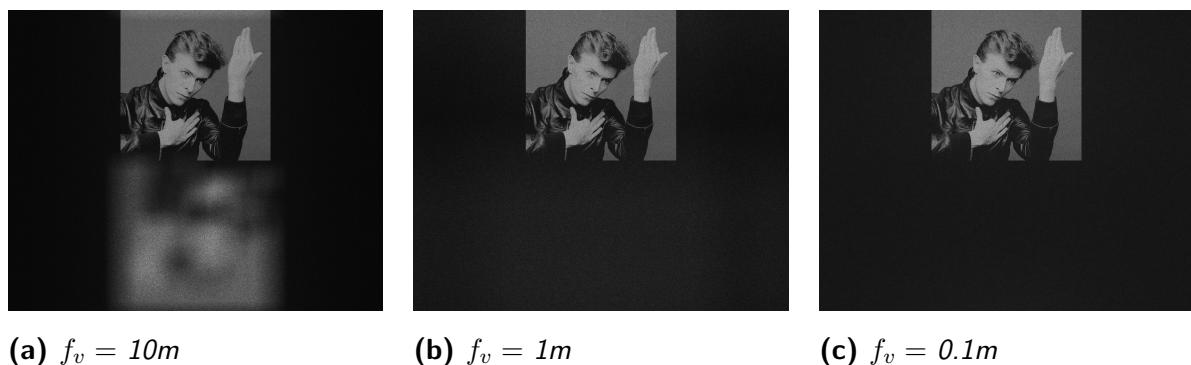


Figure 19: Reconstructions of Fresnel holograms with varying focal power

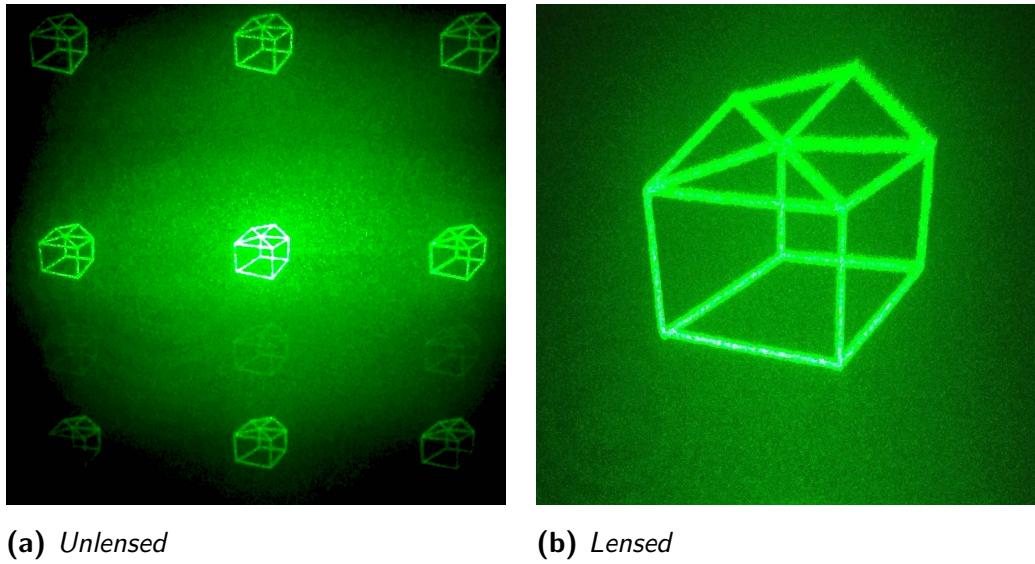


(a)  $f_v = 10m$

(b)  $f_v = 1m$

(c)  $f_v = 0.1m$

Figure 20: Comparing 60mmx60mm area of replay field of one- and two-lens Fresnel holograms at 500mm depth



use sufficiently larger values of  $f_v$ . The results of both of these schemes are pictured in fig. 20. The hologram was focussed onto the lens, and the Fresnel region was navigated moving the lens back and forth. This reduced the dynamic potential of the 3D hologram (as only one plane was in focus at a time, and for all of the replay field), but allowed larger and less noisy images to be projected.

The replay field produced in the single-lens setup (fig. 20b) forms an image of 0.175 times the size of the modified 2-lens setup (fig. 20a) at the same depth, with visible higher orders, and conjugate noise due to the extreme near field present in the replay field. Magnification of the image reduces the aperture, however, this has the effect of obscuring higher orders and does not impact the desired part of the replay field. As the quality of the 2-lens hologram is perceptually significantly higher, the 2-lens system is used to construct Fresnel-type images for the remainder of this report.

#### 4.2.3 Scaling Fresnel Holograms

As the focal plane of the hologram gets further from the projector, the relative size of the replay field increases. In order to counteract this, a scaling factor (relative to the nearest slice, which will have the maximal image size already), is introduced.

The image size increases due to the viewing angle of the projection: as this is constant, the transformation is a linear projection. This varies with the viewing angle  $\theta_{view}$  calculated in section 3.0.2.

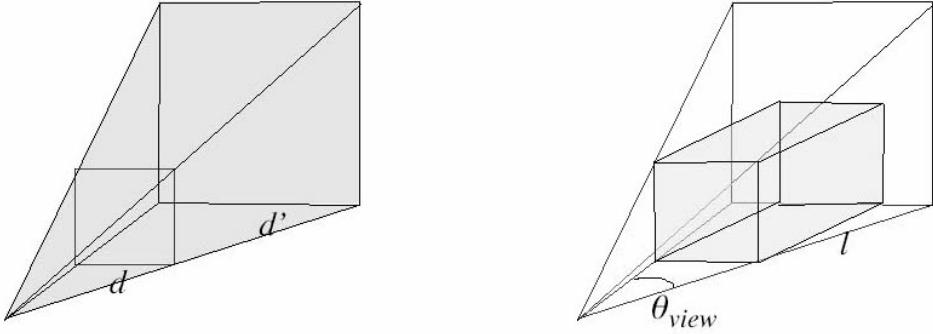


Figure 21: *Diagram of the scaling issues associated with viewing angle, where  $d$  is the scale of the nearest viewing plane,  $d'$  the scale of the plane at displacement  $l$  along the optical axis, and  $\theta_{view}$  is the viewing angle*

$$d' = d + 2l \times \tan\left(\frac{\theta_{view}}{2}\right) \quad (20)$$

Where  $l$  is the displacement from the Fresnel slice focussed closest to the projector, shown in fig. 21.

#### 4.2.4 Multiplexing Fresnel Holograms

Cable [7] describes a scheme for multiplexing several Fresnel slices into the same hologram, allowing multiple replay fields, focussed at different depths, to be projected simultaneously, forming a fully 3D image in the replay field. A diagram of the algorithm for slice-multiplexing is pictured in fig. 22. Each separate slice is generated using ADOSPR (to maximise quality) and the Fresnel transform, and these slices are then summed into a single hologram and phase quantised, producing a SLM image with multiple focal depths.

Less computationally-intensive schemes for Fresnel multiplexing are outlined in [3], which discusses the multi-quantised OSPR method for generating live 3D holograms. This uses a single FT over multiple slices (as opposed to one per slice), allowing rapid generation of multi-planar images. However, the quality is somewhat lower, and as this is a non-live application the ADOSPR step is retained.

#### 4.2.5 Noise effects of Fresnel multiplexing

By multiplexing several images into the same plane, the amount of resolution-per-slice available is reduced by a factor of approximately  $1/N$ , where  $N$  is the number of multiplexed slices, and the defocussed planes contribute significantly to unwanted noise effects (as the energy in the replay

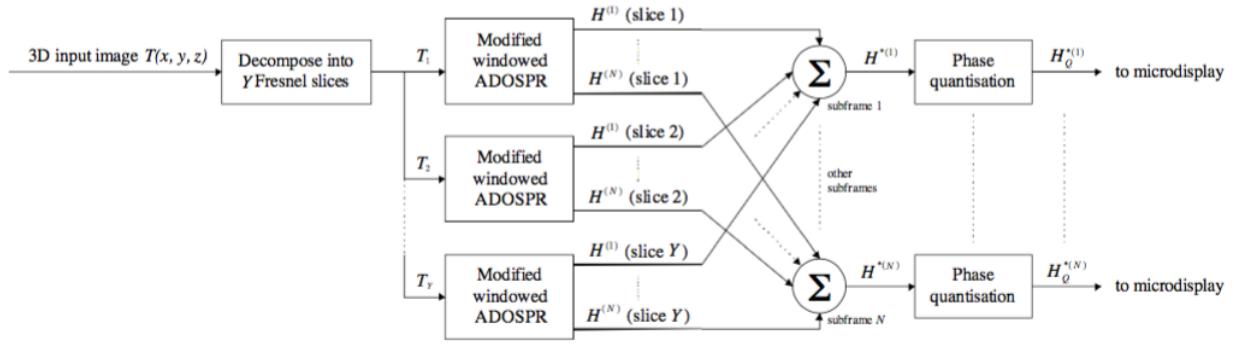


Figure 22: *Diagram of Fresnel slice merging algorithm, from [7]*

field remains constant regardless of depth). In order to investigate the variation in noise effects with the distance between Fresnel layers, a 2-slice image was focussed with layers at an offset of 1, 2 and 4cm respectively in fig. 24. The closer the multiplexed slices, the clearer the image in the defocussed plane becomes, forming a distracting artifact in the replay field.

The noise effects of multiple layers in comparison to a single layer are shown in fig. 26, showing a significant degradation in perceptual sharpness as more layers are added. A greyscale test for multiplexed Fresnel holograms (for 2 slices) shows a noise level critical at 2 levels (fig. 23) – indicating that only binary black and white schemes are perceptible in a multiplexed holographic scene.

Noise metrics were less helpful here, as the defocus noise is highly image-dependent, and the perceptual effect of the defocus plane as slices become closer is far greater than the measurable change in noise. Thus, for this scheme, limits are defined in terms of perception, using the greyscale test and images of variable spatial frequency to determine the constraints on the number and spacing of slices.

### 4.3 Representing CT Scan Data

One of the primary scientific applications of 3D computer graphics is in the field of medical imaging, which requires the use of 3D visualisations as a tool for diagnosis and teaching [14]. Medical imaging data is often detailed, high-resolution and over a wide dynamic range, thus making a challenging form for holographic projection. For this section, a 400-slice data set<sup>2</sup> of a human skull was used.

<sup>2</sup>kindly donated by Dr Graham Treece

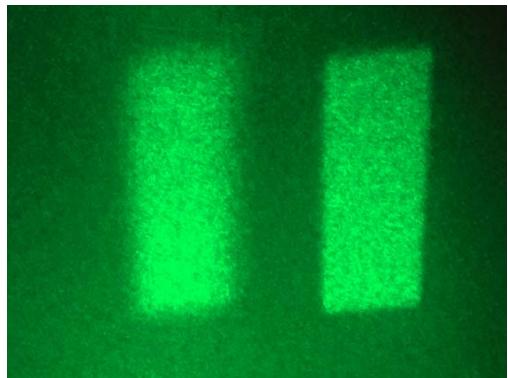
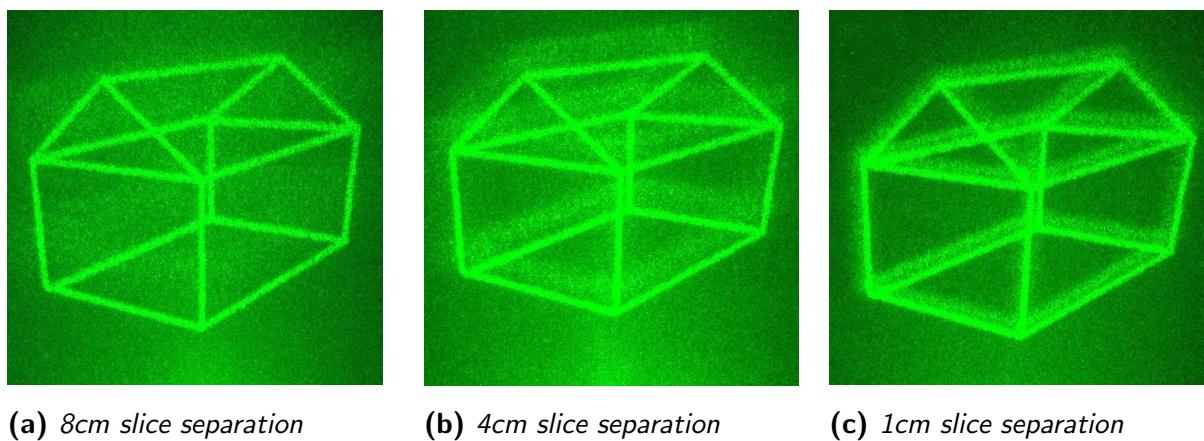


Figure 23: Greyscale test on 2-slice hologram (in-focus slice right, defocussed slice left), with slice separation of 2cm

Figure 24: Relative slice depth effects on noise



(a) 8cm slice separation

(b) 4cm slice separation

(c) 1cm slice separation

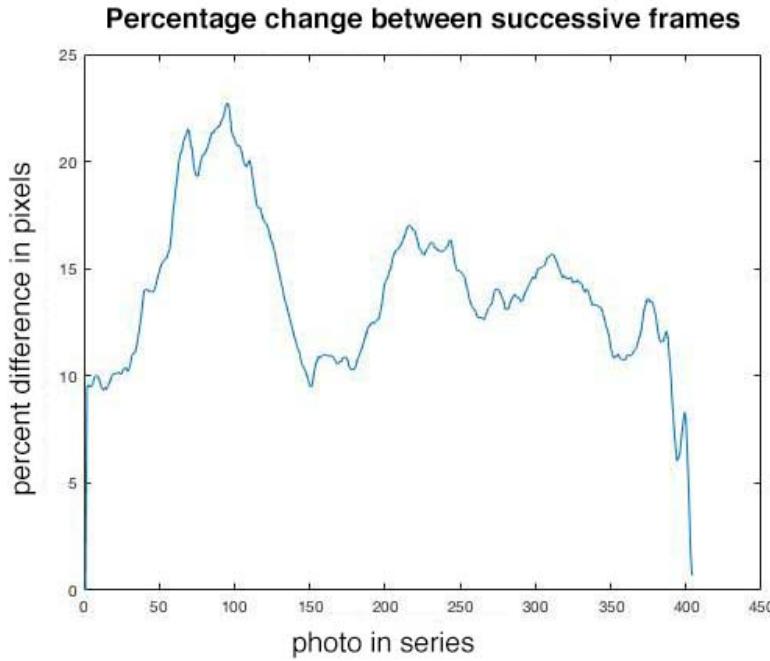


Figure 25: *Graph of the percentage pixel change between each image in the set*

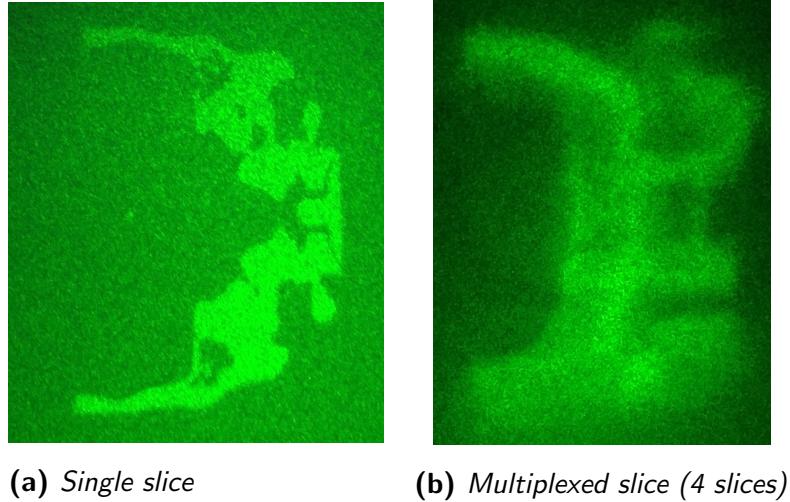
#### 4.3.1 Compressing image sequence

The first step in information reduction was to render the CT scan in black and white, using the Matlab image processing library. High spatial frequency components were then smoothed with a Gaussian, to prevent noise in the replay field. In order to compress the 3D component of the scan, a 'reverse interpolation' scheme was used to significantly reduce the number of slices, while retaining the essential progression of the scan.

This scheme measures the change in pixels between each successive frame, and plots the rate of change between frames in different parts of the scan. Sample 'keyframes' are then taken from the data according to the rate of change of the local set. For sets of frames with a high rate of change, the sampling rate is increased – for sets with a low rate of change, the sampling rate is decreased. Frames either side of the sampled keyframe are then averaged into a single slice to make the transition as smooth as possible.

Even with a significant reduction in the number of frames (e.g. 400 to 15), multiplexing 15 Fresnel slices into the same hologram introduces so much noise into each slice that individual images are illegible. 15 frames was decided as the lower limit on the number of keyframes required to prevent unnecessary jumping, as the sample was taken every time approximately 10% of the image had changed in intensity.

Figure 26: Comparing single and multiplexed CT slices



#### 4.3.2 Multiplexing schemes

**Spatial multiplexing:** Due to the noise introduced in combining slices through the mechanism discussed in 4.2.5, an upper limit of 4 spatially multiplexed slices was a limit for reasonable resolution in the replay field. This was determined by perceptual testing over a range of slice numbers (1 to 6). This is a far lower number of possible slices than the ‘house’ used in fig. 24, due to the comparatively high spatial frequency of the CT data. As this was significantly less than the 15 slices required for a reasonable CT representation, slice multiplexing in the time domain was also required.

**Time multiplexing:** Time-multiplexing of holograms is used in the OSPR algorithm, relying on the transfer function of the eye to switch between multiple images simultaneously. This is an effect used to create 3D images by both Zeng et. al [24] and Bheemireddy [3], who use sequential Fresnel frames to construct 3D projections. The refresh rate of the projector is 1440Hz; with the number of OSPR slices required for good resolution (8 or greater [2]), and the time response of the eye (60Hz), this gives a theoretical limit of the projector of 3 time multiplexed slices without visible distraction.

**Combining space and time multiplexing:** Two main methods exist for combining Fresnel slices in both space and time (see fig. 27), where the spatially multiplexed slices are either interleaved, maximising inter-slice depth, or each set of slices is separately time-multiplexed. The former approach is chosen for this experiment, as it allows individual slices to be packed closer without increasing inter-slice noise.

Combining both the spatial and time multiplexing limits of the projection system gives a theoretical limit of 12 CT slices without major visual defects: lower than the 15 required for the

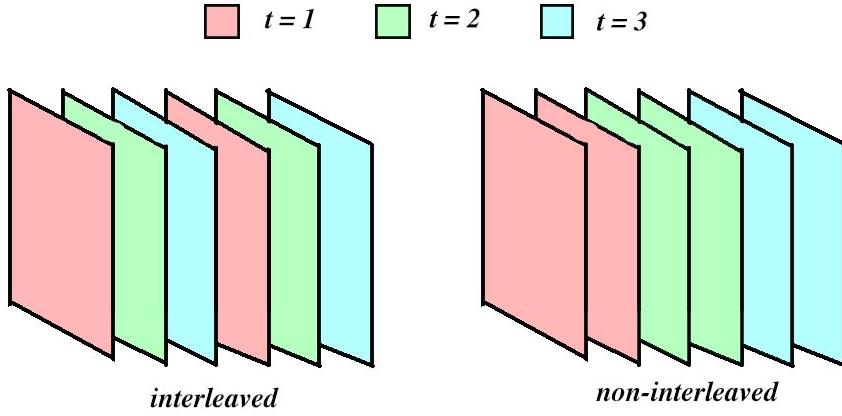
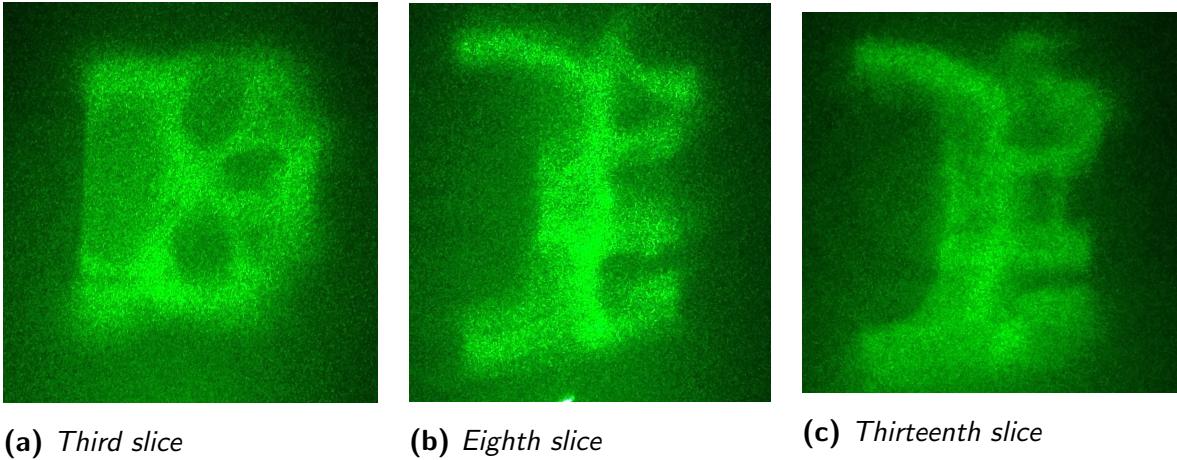


Figure 27: *Alternative space- and time-multiplexing schemes*

Figure 28: Slices from time- and space- multiplexed CT scan data



CT scan. A compromise was reached by spatially multiplexing five interleaved slices, however, latency in the upload code proved a limiting factor, with a visible ‘flicker’ when more than 2 slices were cycled through. A more computationally-intensive programme could form the basis for investigation by a future project, however, considering the already-poor resolution of the slice-multiplexing scheme, this was not investigated here.

## 5 Spatial Augmented Reality on Nonuniform Surfaces

Previous projects have explored the projection of holograms onto slanted and curvilinear surfaces [2, 19] and corners [2]. Here, this work is developed to include projection mapping and animation

to a range of non-uniform 3D surfaces.

## 5.1 Ray tracing 3D scenes

Scenes in nature are rendered visible by the presence of light sources. These sources emit rays, which, on intersecting with objects, are reflected, refracted or transmitted. The net incidence of light rays on the eye is used by the brain to construct a 3D scene. Thus, by reconstructing the light pattern created by a 3D scene in 2D – ‘ray tracing’ from a source to the eye – the illusion of 3 dimensions may be created.

The surface rendering pipeline is the multi-stage process used in 3D computer graphics to render a 3D ‘world’ object in terms of 2D device (‘image plane’) co-ordinates. This process consists of many stages – not all of which are relevant to this project – but the overall transformation forms the basis for effective projection onto 3D scenes [13]. In particular, perspective projection and the rigid body transform are needed for SAR mapping onto non-uniform surfaces.

### 5.1.1 The perspective projection matrix

The perspective projection matrix is used to map the 3D co-ordinates of an object in a scene  $\mathbf{X} = \begin{bmatrix} X & Y & Z \end{bmatrix}^T$  to the 2D co-ordinates of a camera  $\omega = \begin{bmatrix} u & v \end{bmatrix}^T$ . By transforming into a homogeneous co-ordinate set, perspective projection may be represented by a single transform:

$$u = \mathbf{P}X \tag{21}$$

Where the perspective projection matrix,  $\mathbf{P}$  is a  $4 \times 3$  matrix, and may be decomposed into the extrinsic camera transform  $[\mathbf{R}|\mathbf{T}]$ , which gives the position of the projector in the 3D scene, and the intrinsic camera properties (focal length, CCD co-ordinates), which determine depth perception:

$$\mathbf{P} = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} & p_{1,4} \\ p_{2,1} & p_{2,2} & p_{2,3} & p_{2,4} \\ p_{3,1} & p_{3,2} & p_{3,3} & p_{3,4} \end{bmatrix} = \begin{bmatrix} k_u f & 0 & 0 & u_0 \\ 0 & k_v f & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} [\mathbf{R}|\mathbf{T}] \tag{22}$$

The perspective projection is due to the depth of the 3D scene,  $f$ . By altering the effective value of  $f$ , the perception of the scene is altered. By modulating the apparent value of  $f$ , the depth of objects in the scene – rather than the perceived properties of the imaging medium – appear to change.

A separate perspective projection matrix is required for the viewing plane and the projection surface – the viewer is assumed to stand at  $30^\circ$  to the optical axis in this model.

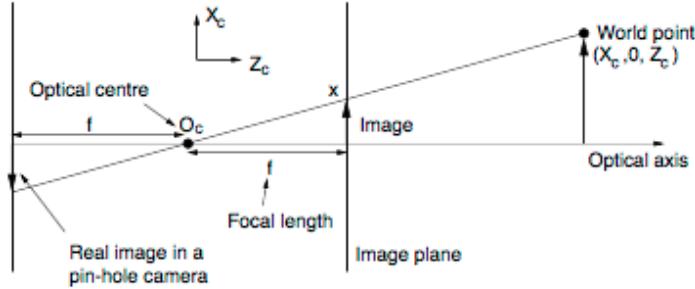


Figure 29: Diagram of 1D perspective projection, from [9]

### 5.1.2 Applying the ray tracing algorithm

The ray-tracing algorithm ‘pre-distorts’ the projected image to fit the inverse of the surface, thus compensating for distortion and appearing ‘flat’ from the correct viewing angle. The size of projected pixels in the replay field  $\Delta_{replay}$  is given by:

$$\Delta_{replay} = \frac{R\lambda}{N\Delta_{eff}} \quad (23)$$

Where  $\lambda$  is the wavelength of the source,  $N$  is the size of the hologram (along its shortest dimension), and  $\Delta_{eff}$  is the effective pixel pitch of the SLM. As the internal lens is used to magnify the image,  $\Delta_{eff}$  is calculated using the true viewing angle of the hologram:

$$\Delta_{eff} = \frac{\lambda}{2\tan(\theta_{view})}[2] \quad (24)$$

This gives a measured value of  $\Delta = 612\text{nm}$ . Perspective projection is then used to map from the projection plane to the image plane, defined at the SLM. For each pixel in the image plane, a ray is projected through the centre and the intersection with the projection surface taken as the R-component. Extending this to the viewing plane requires a second transform: a planar projection (homography) that scales the image to appear flat in the field of view.

Pixels in the distorted image are interpolated by taking the average of the original pixels mapped to that location, a method suggested in [19] to mitigate the appearance of ‘grid’ artifacts.

### 5.1.3 Vanishing Points

One side effect of perspective projection is the creation of vanishing points within a scene. The number of vanishing points depends on the number of ‘orientations’ seen by lines in the scene

– each set of parallel lines has a vanishing point associated with it. For a horizontal camera, it may be shown that vertical lines have a vanishing point at infinity, and the vanishing points of all the other objects in the image relate to their orientation, not position [9].

By manipulating the placement of vanishing points, the viewer is ‘tricked’ into seeing parallel lines where they do not exist [9]. False parallels are used extensively in Renaissance architecture, manipulating the perception of space. Vanishing points are also a strong depth cue for 3D scenes, and are used to reconstruct scenes from images. By varying the position of vanishing points in an image, the image appears to move, and an amodal perception of solid shapes may be established [4, 9, 18].

## 5.2 Flicker stereograms

The stereogram is a victorian illusion that takes offset images of some scene, either for the creation of a binocular (e.g. headset-based AR, viewmaster) or monocular (SAR, flicker stereogram) illusion of 3D. In combining two images of the same scene offset in the x-direction, motion parallax is exploited to produce an illusion of depth, giving the impression that the viewer is walking around an image.

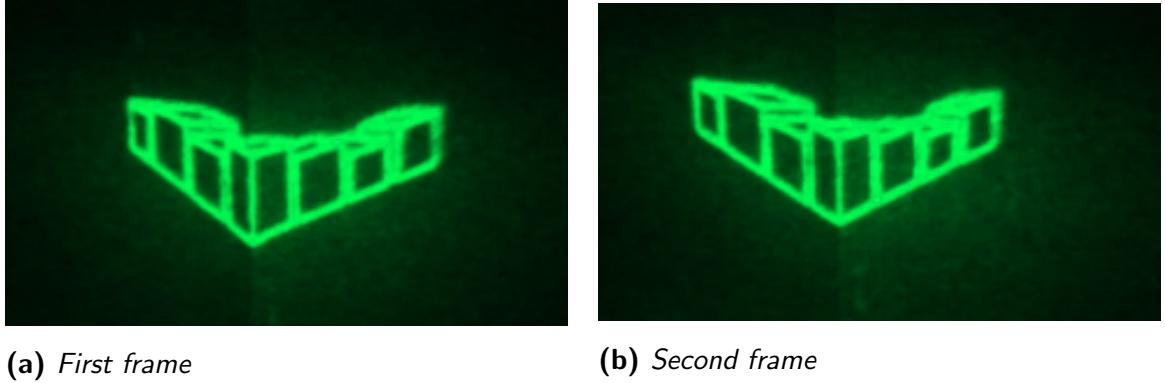
### 5.2.1 Comparing one- and two-point perspective

In order to test the effectiveness of flicker stereograms – alternating images projected onto a static surface – for different perspective projections, scenes were constructed with 1- and 2-point perspective. The scenes – showing a simple schematic of a street and a corner respectively – incorporate perspectival, textural and occlusion-based depth cues, and are simply-interpreted by the viewer. These allow for low-spatial frequency holograms to be produced (thus high quality in the replay field), whilst producing strong depth illusions. Architectural shapes are unambiguous, and rely on straight lines – thus giving the viewer a ‘schema’ that the image is expected to fulfil, and encouraging amodal perception [4].

The models were constructed in architectural drawing software Sketchup by rotating a scene by 10 degrees in the x-direction, and simplified in Photoshop, where the central point of each image was aligned with the placement of the corner. These scenes were projected onto flat surfaces, and ray-traced onto both concave and convex corner surfaces.

**Two-point perspective** The first stereogram tested was two-point perspective, projected onto a flat, convex (fig. 30) and concave (fig. 31) corner surface. The stereogram created a good illusion of a scene rotating, however, the motion parallax was unsettling. For the central axis to remain in place, the ‘virtual translation’ of the viewer with respect to the 3D scene is very

Figure 30: Mapping 2-point perspective stereogram to positive corner (45 degrees), images at 10 degree offset



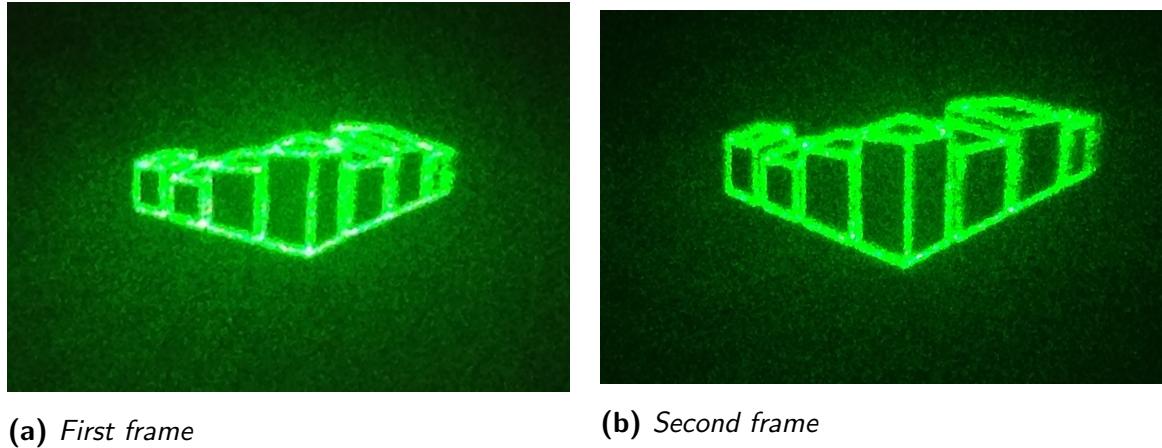
large, creating an unnatural and somewhat nauseating effect. The most effective projection was onto the concave corner: this confirms the hypothesis suggested in [2] that the most effective 3D illusion is constructed in surfaces that oppose the depth cue of the image. Removing the corner to project onto a flat surface, and allowing translation of the central axis in the x-direction created a far more convincing illusion, as this neutralised the motion parallax effect.

**One-point perspective** A second projection was constructed using a single, static vanishing point in the centre of the replay field. This was far more effective, with motion parallax aiding the perception of the scene, with the ‘nauseating’ effect neutralised. This worked most effectively projected onto the convex corner (see fig. 32), although the viewing angle was greatest on the flat plane. For stereograms projected onto a corner, the illusion only holds from a narrow viewing angle: thus, for SAR applications, a trade-off emerges between an effective 3D effect and a restrictive viewing environment.

### 5.2.2 Multiple viewing planes

An optical illusion commonly mistaken for a hologram is the lenticular. Patented by Walter Hess in 1915 [15], it uses a flat sheet printed with 2 interleaved images at some spatial frequency  $\Delta$  to create a stereographic illusion. The sheet is overlaid with transparent plastic or glass, ridged in the same spatial frequency as the interleaved images. As the viewing angle of the lenticular changes, the angle of reflection of incident light is modulated by these ridges, making one set of stripes visible from one angle, and the other visible from the opposite angle. This can be used to create depth and/or motion illusions in the image, by using either 2 offset views of the same object (creating a stereogram), or two different images that appear to ‘morph’ into one another. When seen from the correct angle, one viewing plane masks the other, showing one

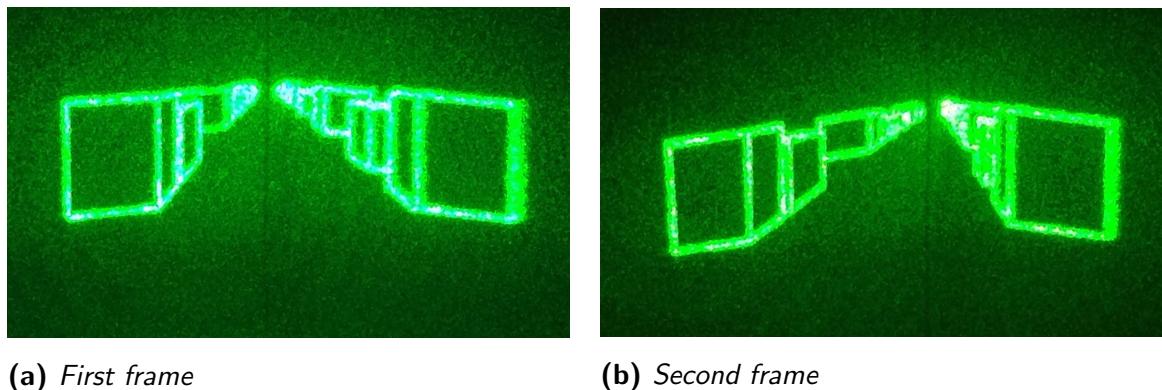
Figure 31: Mapping 2-point perspective stereogram to negative corner (-45 degrees), images at 10 degree offset



(a) *First frame*

(b) *Second frame*

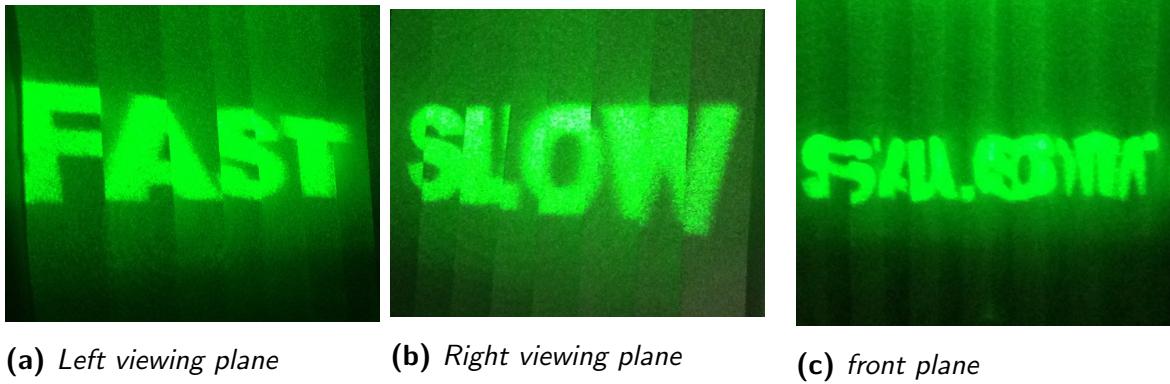
Figure 32: Mapping 1-point perspective stereogram to positive corner (45 degrees), images at 10 degree offset



(a) *First frame*

(b) *Second frame*

Figure 33: Multiple viewing planes – uncorrected version



'image' comprised from a set of strips whilst occluding the other.

In order to investigate this illusion with a hologram, the ridged glass was substituted for by a piece of paper folded in a regular  $45^\circ$  ridged pattern, and two interleaved images projected onto the surface at the same spatial frequency. Strips were measured out on pieces of paper, and scored with a craft knife for greater folding accuracy. The initial test version is shown in fig. 33: as can be seen from figs. 33a and 33b, there are two projective distortions: one due to the position of the viewer with respect to the projection plane (global effect), and one due to the angle of each ridge with the projector (local effect) (see fig. 34).

By aligning the viewing plane with the slope of the scene in the replay field ( $\alpha = \beta$ ), only a single transform was required to project each image onto opposing 'slice planes' of the projection surface. The corrected version produced with this method is shown in fig. 35, showing an approximately-flat image from both viewing planes. The illusion deteriorates somewhat at the edges of the replay field, in part due to sinc envelope compensation, as for this experiment it was chosen to maximise the horizontal size of the replay field.

An upper limit on the spatial frequency of slices was dictated by paper folding – more than 8 stripes from each image was affected by small inaccuracies in the measurement, and produced nonsensical images from the viewing plane. A better projection surface could be produced with machined wood or metal for a sharper edge (with the resolution of 3D printing also proving a limiting factor for a large number of stripes), and could form the basis for an investigation of 'spatial stereograms'.

### 5.3 Animation

The influence of different depth cues on the 3D perception of animated images is investigated for 2 different animation schemes. The first – the 'spinning ball' illusion – uses motion parallax and

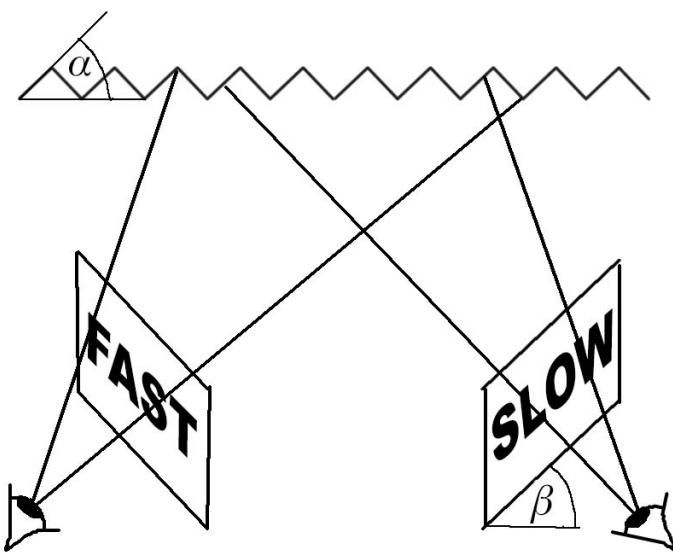
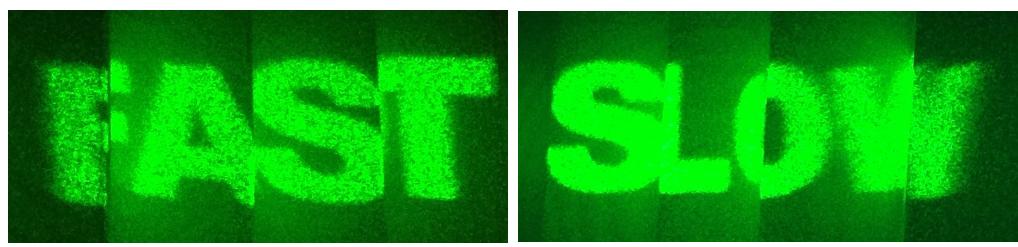


Figure 34: Diagram showing projection plane angle ( $\alpha$ ) and viewing plane angle ( $\beta$ ) for the simultaneous viewing plane setup

Figure 35: Multiple viewing planes – corrected version



(a) Left viewing plane

(b) Right viewing plane

Figure 36: Stills from holographic animation with motion parallax and orthographic projection

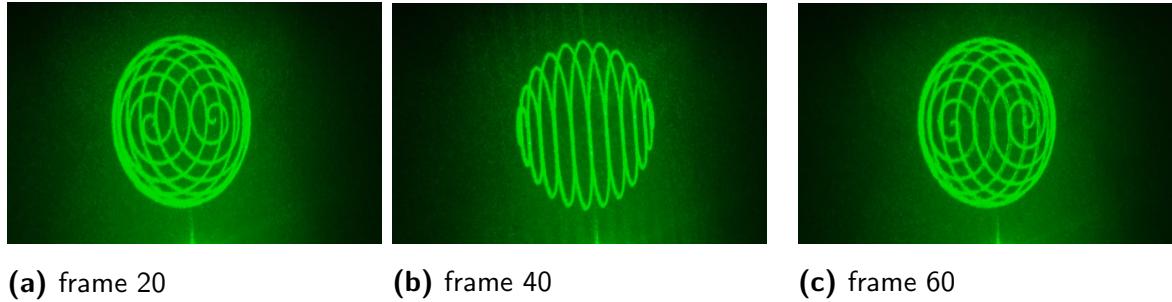
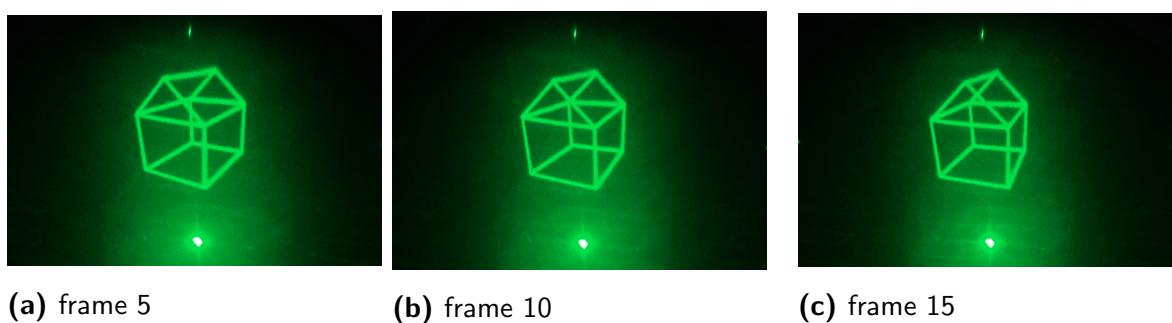


Figure 37: Stills from holographic animation with motion parallax and perspective projection



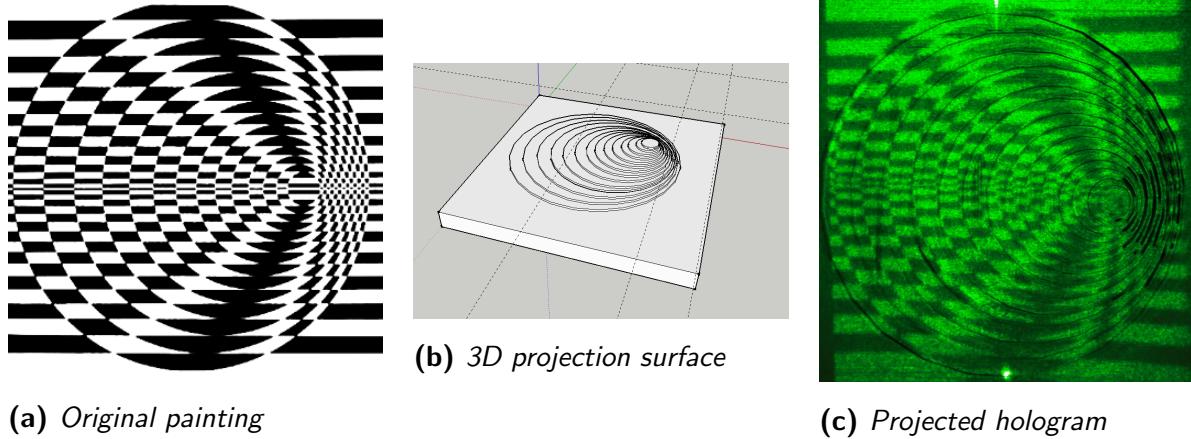
orthographic projection to generate the impression of a rotating sphere from a wireframe (fig. 36). This was generated from an animated gif, and consists of 140 frames cycled through the projector. The second animation used both motion parallax and perspective projection to create a 3D effect, using a simple architectural form to produce an impression of depth (see fig. 37). The perspective grids that map the images were created in Sketchup.

In both cases, the motion parallax observed between the lines gives the illusion of a 3D object rotating in space. Overall, the inclusion of a perspective in the second video contributed to a much more convincing depth effect, though the increased number of frames in the first image produced a smoother and more pleasing video overall.

## 5.4 Custom surfaces

3D printed surfaces were constructed using the Ormerod RepRap Pro, and sprayed with chalk paint to create a Lambertian projection surface. This both reduces the risk associated with reflection of laser light, and improves the quality of the projection, reducing the adverse effects of the bright zero order and undesirable high-energy points.

Figure 38: Op-art projection of Bridget Riley's Britannia (1961)



#### 5.4.1 Op-Art and Visual Illusion

Op-art refers to an art movement of the 1960's, that saw the development of a form of abstract art which exploited visual phenomena to create optical illusions. Stemming from effects such as false perspective and anamorphosis, op-art pieces are characterised by the impression of movement, warping or projective ambiguity suggested by abstract shapes.

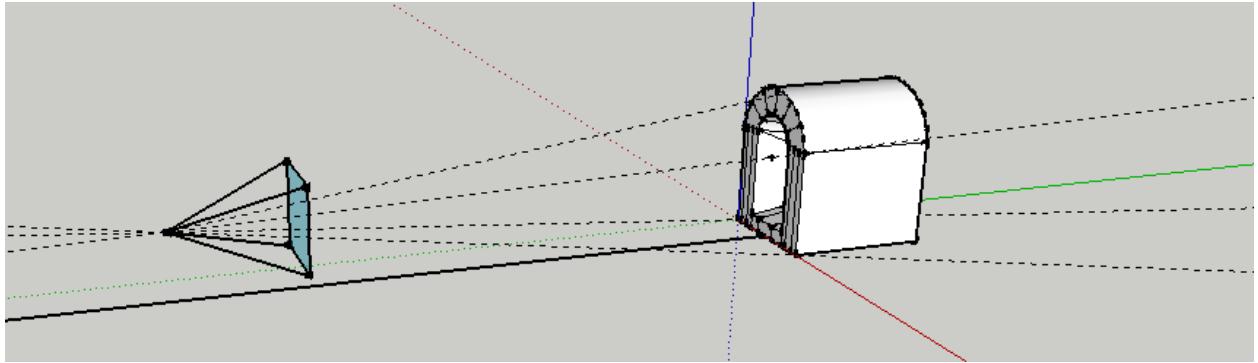
The painter Bridget Riley plays with both colour and optical frequency to modulate the viewer's perception of an image. Here, her painting *Britannia* is projected onto a 3D surface, investigating the changing effect of a 3D projection surface on the perception of depth cues in the scene. As the change in depth within the scene was small in comparison to the distance from the viewer, only weak perspective projection was used to project to the surface (used for any  $\Delta z/z < 1/10$  [9]). This proved a reasonable approximation, with only very slight distortion around the edges of the surface.

The projection formed a reasonable 3D effect with a wide viewing angle. Moving around the image, the sensation of motion was amplified from projection onto a flat surface, although the limited resolution and contrast of the projector attenuated much of the 'visual power' of the original painting, highlighting the spatial frequency limitations of the system.

#### 5.5 SAR on a 3D scene

In order to demonstrate projection mapping to non-uniform surfaces (e.g., ones without a well-defined mathematical description that could be implemented in Matlab), Sketchup was used to construct the perspective projection onto a 3D scene. The software aligned with the projector's image plane with the measured position of the object in the 3D scene (according to the extrinsic transform  $[R|T]$ ), using the value of  $\theta_{view}$  to determine the limits of the 'viewing frustum' used

Figure 39: Virtual projection setup – Sketchup still



to define a 3D scene. Placing the desired projection surface within the frustum, and projecting a grid outline onto the shape gave a simple black-and-white image with a low spatial frequency, producing a high-quality image in the replay field.

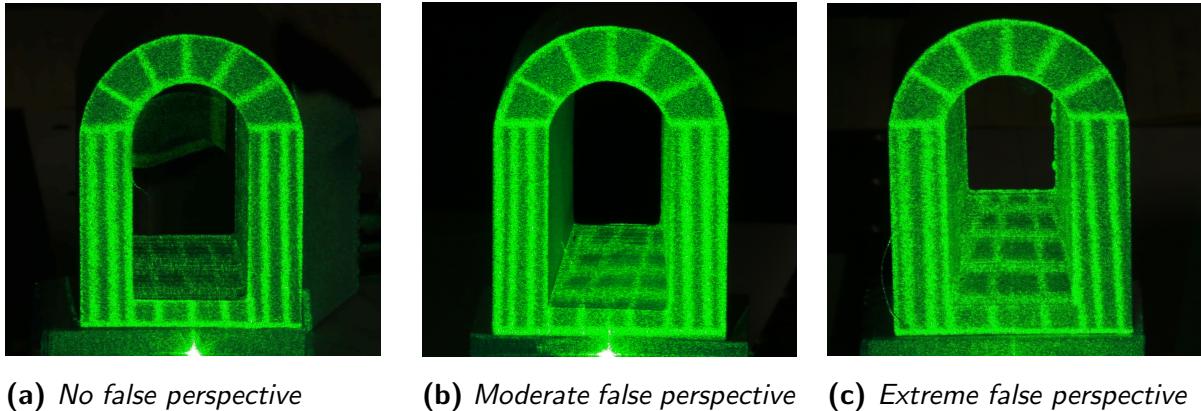
### 5.5.1 False Perspective – Palazzo Spada

In order to demonstrate false perspective projection using the holographic projector, the false depth illusion of the Palazzo Spada (se fig. 6) is recreated using SAR on a 3D scene. A simplified model of the archway was recreated using a 3D printer. Passages were printed to range between three levels of false perspective – from a straight passage, through to extreme false perspective, with all passages possessing the same physical length.

Two viewers were asked to estimate the length of a set of 3D printed passages illuminated with projections. Though this is not enough to provide a full data set, both described an increase in the perception of depth in the passage as the degree of false perspective was increased, with the greatest change in perception between the straight and moderately-sloped passage. Both reported viewing angles that decreased slightly with the increase in the degree of false perspective, though between the moderate and extreme cases this decrease was slight, implying that more extreme scenes can be projected without greatly enlarging the restriction on effective viewing angle. To avoid the introduction of bias, the viewers did not watch as the scene was changed, and viewing angles were only calculated once the initial estimates had been recorded.

The viewing angles for the projected surfaces were reasonably large (between  $\pm 30^\circ - 33^\circ$  for the moderately-sloped surface, and  $\pm 26^\circ - 30^\circ$  for the extreme surface), with the illusion only properly breaking down as the outer edge of the tunnel (which could not one projected onto) became fully visible. This gives a ‘visible range’ spanning around  $50^\circ$ , allowing for significant ‘travel’ before the illusion breaks down.

Figure 40: False perspective with 3D projection mapping



### 5.5.2 Animated SAR

Using the extreme false perspective illusion, animation on a 3D scene was demonstrated. By shifting the lines parallel to the x-axis in the image plane along the false optical axis, using 2 frames, an illusion of motion into the scene is generated. This provided a simple but effective notion of depth, creating a ‘void-like’ effect. Viewers reacted largely positively to this demonstration, though some also reported a slightly strange and nauseating effect. Slight distortion at the extreme end of the tunnel might have been the cause of this dissonance, an effect that could be mitigated in future projects with the use of dedicated projection-mapping software.

## 6 Conclusions

This project has demonstrated a range of holographic projection techniques, investigating both qualitative and quantifiable metrics for success in projecting 2- and 3D scenes. The Mean Squared Error was defined as a good quantitative metric for single-slice 3D holography, though it was not possible to define a reliable quantitative metric when multiple slices were multiplexed. The greyscale resolution was proposed to perceptually define the noise levels of an image, and showed it was not possible with this scheme to view greyscale information in a multiplexed Fresnel scheme.

Fresnel projection was successfully calibrated for both the one- and two-lens projection arrays, with the latter found to have superior perceptual and scaled effects. Spatial multiplexing of Fresnel slices was demonstrated, though it was shown that incorporating multiple Fresnel slices into a single hologram significantly decreases the quality of the images in the replay field. Even by exploiting both time- and spatial multiplexing techniques, it was not possible within the scope of this project to generate high-quality Fresnel sliced images for a large data set.

Augmenting textural depth cues in op-art with a complementary surface produced some par-

allax, which contributed to a 3D effect, although the perception of the image was limited by the resolution and colour of the projector. Op-art is characterised by high spatial-frequency cues, which do not translate effectively onto the projector. The intensity and sharp focus of the light is most powerful when used to generate images of low spatial frequency (e.g. grids and lines), as earlier demonstrations of op-art deformations, though interesting, did not have the same powerful visual effect.

Projection mapping of 2D holograms was demonstrated on non-uniform surfaces. Textural and perspectival depth cues were used with motion parallax to create stereograms in 1- and 2-point perspective on flat surfaces and on corners. It was considerably easier to successfully map a 1-point perspective stereogram to a cornered surface, whereas the 3D effect – coupled with an excessive sensation of motion – reduced the effectiveness of the 2-point stereogram. Static images were mapped to ridged viewing planes to create a ‘lenticular’ effect where the image seen by the viewer depended on the viewing angle. These were mapped to the correct viewing angle successfully, however, small inaccuracies in the projection surface meant the images were not smooth.

The use of depth cues in animations was investigated, with a combination of motion parallax and perspective projection producing an effective illusion of depth. False perspective projection was used with SAR to create an effective projection onto a 3D scene. The notion of depth increased with a higher degree of false perspective, although in the extreme case blurring at the edges distracted somewhat from the overall effect. Lastly, the application of animation suggesting translation in the z-direction was effectively applied to a false perspective projection, demonstrating the use of depth cues to generate convincing SAR on a non-uniform 3D scene.

## 6.1 Future Work

An application for a later project might be to examine the use of commercial projection mapping software using a live upload to the projector, to create animated SAR on more complex (or even moving) 3D scenes. This would also involve computational considerations if used for a ‘live’ generation application, and would form an interesting challenge to rapidly generate high-quality projections. Full colour holograms could also be used to investigate the effect of colour-based cues on depth, and create more naturalistic scenes.

A further investigation of Fresnel optimisation would also be an interesting project, perhaps using the more effective ‘on-eye’ projection techniques discussed in [3, 7] as these seem to have a superior perceptual effect over the direct projections used in this analysis.

## Acknowledgements

With thanks to Tim Wilkinson for his support over the year, to Vamsee Bheemireddy and SJ Senanayake for the use of their legacy code, to the CMMPE group for being kind and welcoming, to Graham Treece for CT scan data, and to Gary Zhixi Zhang for capturing high-quality photographs of the holograms.

## References

- [1] Bertol D., 1997, "*Designing Digital Space: An Architect's Guide to Virtual Reality*", Wiley
- [2] Bheemireddy V., 2015, "*Holographic Projection onto Non-Uniform Surfaces*", 4th Year Project, Department of Engineering, University of Cambridge
- [3] Bheemireddy V., 2016, "*An Architecture for Next Generation 3D Holographic Video*", MRes Project, Integrated Photonic and Electronic Systems CDT, University of Cambridge
- [4] Breckon T., Fisher R., 2005, "*Amodal volume completion: 3D visual completion*", Computer Vision and Image Understanding, Vol. 99, pp. 499-526
- [5] Buckley E., 2006, "*Computer-Generated Holograms for Real-Time Image Display and Sensor Applications*", PhD Thesis, University of Cambridge
- [6] Buckley E., 2010, "*Holographic projector using one lens*", Optics Letters, Vol 35, No. 20, pp. 3399-3401
- [7] Cable A.J., 2007, "*Real-time high-quality two- and three-dimensional video projection using the one-step phase retrieval (OSPR) approach*", PhD Thesis, University of Cambridge
- [8] Chen R. H.-Y., 2010, "*Computer-Generated Holograms for 3D Holographic Display*", PhD thesis, Department of Engineering, University of Cambridge
- [9] Cipolla R., 2016, *4F12: Computer Vision*, lecture notes, Department of Engineering, University of Cambridge
- [10] Crawford C., "*Perspective*", <https://professorcrawford.wordpress.com>, last accessed 15/05/2017
- [11] Freeman J., 2009, "*Visor Projected Helmet Mounted Display for Fast Jet Aviators using a Fourier Video Projector*", PhD Thesis, University of Cambridge
- [12] Freedman, D., 2002, "*Holograms in Motion*", <https://www.technologyreview.com/s/401697/holograms-in-motion/>, last accessed 26/05/17
- [13] Gee, A., 2016, *3G4: Medical Imaging and Computer Graphics*, lecture notes, Department of Engineering, University of Cambridge
- [14] Hackett M., Proctor M., 2016, "*Three-Dimensional Display Technologies for Anatomical Education: A Literature Review*", Journal of Scientific Education Technology, Vol. 25, pp. 641-654

- [15] Hess W., 1915. "Stereoscopic Picture", US patent 1,128,979, patentimages.storage.googleapis.com/pdfs/a205c4f373beee9adcd3/US1128979.pdf
- [16] Nakagaki, K. et. al, 2016, "Materiable: Rendering Dynamic Material Properties in Response to Direct Physical Touch with Shape Changing Interfaces", CHI 2016, San Jose, CA
- [17] Raskar R., Welch G., Fuchs H., 1998, "Spatially Augmented Reality", First International Workshop on Augmented Reality, San Francisco
- [18] Solso R., 1994, "Cognition and the Visual Arts", MIT Press, Bradford Books Series in Cognitive Psychology
- [19] Senanayake S. J., 2012, "Holographic Projection onto Non-uniform Surfaces", 4th Year Project, Department of Engineering, University of Cambridge
- [20] Takaki Y., Hitoshi O., 1995, "Liquid-crystal active lens: a reconfigurable lens employing a phase modulator", Optics Communications, Vol. 126, pp. 123-134
- [21] VividQ website, 2017, <http://www.vivid-q.com>, last accessed 26/05/2017
- [22] Glogger commonsWiki, 2002, "Mediated reality continuum", Image accessed via the Wikimedia Commons
- [23] Wilkinson T., 2016, 4B11: Photonic Systems, lecture notes, Department of Engineering, University of Cambridge
- [24] Zeng Z. et. al, 2015, "Dynamic holographic three-dimensional projection based on liquid crystal spatial light modulator and cylindrical fog screen", International Workshop on Holography and related technologies (IWH2014), Beijing, pp. 853-861

## **Appendix: Evaluation of Risk Assessment**

The main hazard present in this project was the low-power class 3B laser used in the holographic projection system. The key risk was exposure of the eye to laser light, which could occur by looking down the optical axis or at a reflective object in the path of the beam when the laser was active.

The projection bench was kept clear of reflective objects, and opaque curtains were placed around the projector when turned on. To use the projector, the curtain facing the wall was opened, facing away from other users of the lab. Analysis of Fresnel holograms was limited to the end of the bench (this proved adequate), as any further would have required holding screens in the path of the beam. No alignment changes to the optics were required, and thus the casing surrounding the laser was not moved. Care was taken not to knock or jolt the array, as this would warrant re-alignment.

Potentially specular surfaces were identified, and all 3D printed projection surfaces were rendered matte using chalk paint. The painting itself was conducted outside, to ensure no solvent was inhaled, and painted objects were ensured fully-dry before returning to the lab. Volunteers for the qualitative analysis of false perspective were taken from the lab, and instructed carefully to remain on the 'non-projective' side of the projector.

The use of a photographer was discussed with the laser safety officer in advance, and they were clearly instructed as to the risks associated with the laser. The camera was positioned on a tripod well behind the projector to ensure that they did not cross the beam. The initial risk assessment proved adequate, and no additional changes are required.