DEVELOPMENT AND EVALUATION OF DEPTH ESTIMATION FROM PLENOPTIC IMAGING FOR APPLICATION IN RETINAL IMAGING

by

RICHARD JAMES MARSHALL

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

PSIBS Doctoral Training Centre
School of Chemistry
College of Engineering and Physical Sciences
University of Birmingham
August 2016

Abstract

Plenoptic imaging is a technology by which a three dimensional representation of the world can be captured in a single image. Previous research has focused on the technology itself, with very little focusing on applications of the technology. This thesis presents an investigation into its potential application to the field of retinal imaging, with the aim of producing three dimensional images of the retina at a cheaper cost than the current gold standard of retinal imaging, optical coherence tomography. Both a theoretical and practical approach have been utilised through the means of computational simulations and plenoptic imaging through the use of a commercial camera. The key novel contributions presented throughout this thesis are:

- Computational simulations of light transport in the retina to investigate the potential of plenoptic retinal imaging. Initial investigations focused on using ballistic photons to quantitatively measure the depth of different layers in the retina. However, as only 0.35% of light reached the retinal pigment epithelium without scattering, it was determined this would not be practical. By then looking at the angular distribution of photons reflecting from the retina, this indicated that the distribution was not a function of the thickness of the retina. Therefore, surface topography would be the most appropriate avenue to pursue.

- Through imaging surfaces with limited features with a Raytrix R11 plenoptic camera, it was noted that the depth recovered was not correct. By then applying texture, both inherent in the scene and by projecting texture, the recovered depth improved significantly.

- A work-flow has been developed to compute quantitative, metric depth values whilst imaging different scenes with a plenoptic camera, rather than the relative, virtual depth.

- Investigation of characteristics and concepts of plenoptic based imaging for application in retinal imaging. Developed work flows for imaging both flat planes and spheres, and the improvement of the depth recovery by projecting various patterns onto the scenes.

- Novel cost effective production of microlens arrays. Preliminary experiments to fabricate lenses of appropriate specifications for the manufacturing of a retinal plenoptic camera were performed, showing promise for future experiments to develop.

The computational simulations showing how retinal layered imaging was not possible with the use of a plenoptic camera, but surface imaging was, drove the direction of the project towards the presented focus on recovering quantitative depth of featureless surfaces. The study to fabricate novel, cost effective microlens arrays stemmed from the aim to produce a low cost alternative to optical coherence tomography. Although the conclusion of this work did not lead to a functioning microlens array for the purpose of building a plenoptic camera, significant steps were taken to build a foundation for future development.

Acknowledgements

I would firstly like to thank the EPSRC and the Physical Sciences of Imaging in the Biomedical Sciences (PSIBS) Doctoral Training Centre for funding my studies.

I would like to give a special word of appreciation to my supervisors Dr Iain Styles, Professor Kai Bongs, Professor Ela Claridge and Dr Alex Robinson, whose expert advise throughout has helped to make the project possible.

A huge thank you will also have to go to Massimo Turola and Chris Meah, whose both technical knowledge and encouragement to keep going was vital for the success of this work.

I'd like to thank both Steve Gruppetta at City University and Nigel Davies at Chelsea and Westminster hospital for the use of their facilities and the Raytrix R11 camera, without which many areas of this project just wouldn't have been possible to complete.

A thank you to everyone in PSIBS, the medical imaging group in the School of Computer Science, Dr Jochen Kronjaeger and Dr Nadine Meyer for the support and supervision throughout my studies.

Finally, I would thank my parents Ann and Trevor, who supported me throughout my entire studies, for which I will always be grateful.

Publications

R. J. Marshall, C. Meah, M. Turola, S. Gruppetta, E. Claridge, N. Meyer, K. Bongs, A. Robinson, and I. Styles, "Optimisation of patterned illumination to improve depth estimation from a plenoptic camera," in *Imaging and Applied Optics 2015, OSA Technical Digest (online)* (Optical Society of America, 2015), paper JT5A.41.

Contributions on this publication involved performing the research presented throughout and the writing of the paper.

Richard J. Marshall; Chris J. Meah; Massimo Turola; Ela Claridge; Alex Robinson; Kai Bongs; Steve Gruppetta; Iain B. Styles, "Improving depth estimation from a plenoptic camera by patterned illumination", *SPIE Optical Metrology*, pages 952815–952815, International Society for Optics and Photonics, 2015.

Contributions on this publication involved performing the research presented throughout and the writing of the paper.

R. Marshall, I. Styles, E. Claridge, and K. Bongs, "Plenoptic imaging of the retina: can it resolve depth in scattering tissues?," in *Biomedical Optics 2014, OSA Technical Digest (online)* (Optical Society of America, 2014), paper BM3A.60.

Contributions on this publication involved performing the research presented throughout and the writing of the paper.

Chris Meah, Massimo Turola, **Richard J Marshall**, Ela Claridge, Alex Robinson, Kai Bongs, Steve Gruppetta, and Iain Styles. "Towards plenoptic multi-view imaging." In *Propagation through and Characterization of Distributed Volume Turbulence and Atmospheric Phenomena*, pages JT5A–42. Optical Society of America, 2015.

Contributions on this publication involved an advisory role with the primary author on both the research and writing of the paper

Massimo Turola, Chris J Meah, **Richard J Marshall**, Iain B Styles, and Stephen Gruppetta. "Effect of wavefront aberrations on a focused plenoptic imaging system: a wave optics simulation approach". In *SPIE Optical Metrology*, pages 95260X–95260X. International Society for Optics and Photonics, 2015.

Contributions on this publication involved an advisory role with the primary author on both the research and writing of the paper

Meah, Chris, **Richard Marshall**, Ela Claridge, Kai Bongs, and Iain B. Styles. "A Simulation Toolbox For Biomedical Plenoptic Imaging." In *MIUA*, pp. 277-282. 2014.

Contributions on this publication involved an advisory role with the primary author on both the research and writing of the paper

Contents

List of Figures

List of Tables

Chapter 1

Introduction

According to the World Health Organisation, an estimated 253 million people are living with vision impairments, of which 36 million are blind [2]. However, 80% of all vision impairments can be prevented or cured. Retinal imaging could play a major role in helping to reduce this figure. Unfortunately, the high costs associated in purchasing the equipment necessary to image in three dimensions is prohibitive in many parts of the world. This thesis comprehensively evaluates the potential of imaging the retina with a plenoptic camera, the ability and accuracy of a plenoptic imaging system to recover actual depths of a variety of object geometries, and finally investigates microlens fabrication at a fraction of the cost of buying one commercially. As a plenoptic camera, which gains 3D information from a single image acquisition, contains the same components as a normal camera but with the only addition being a microlens array, it is hypothesised that 3D images of the retina would be achievable from a single image at a reduced cost of current retinal imaging methods.

The visualisation of the retina in three dimensions has made the diagnosis of diseases such as glaucoma and diabetic macular edema far more straightforward. The current gold standard for achieving this is optical coherence tomography (OCT), however, this piece of equipment is often too expensive so is not widely available throughout the world. Plenoptic imaging offers a new method for capturing three dimensional data about the

world from a single image, with the only additional cost from a standard digital camera being the addition of a microlens array. This technology therefore has the potential to achieve the diagnostic results of an OCT at a greatly reduced cost.

The development of photographic cameras have come a long way in terms of both the digitisation and improved resolution of the images over the years. However the basic concept has not changed; a two dimensional representation of the three dimensional physical world. This representation fails to distinguish between two rays of light originating from different depths if they are deposited onto the same pixel. This concept has advanced recently with the development of computational imaging. The field has focused on encoding the rays of light entering a camera before they are incident on the sensor. Then by decoding the raw image more information about the scene can be gained.

One example of a computational imaging technique is a plenoptic camera, first developed by Ng *et al* [3] to capture the full 4D light field of a scene (more details on the light field can be found in Section 1.1). A plenoptic camera is constructed by placing a microlens array between the main lens and the sensor, splitting rays which would usually hit a single pixel onto different ones, depending on the depth of the source. The information recorded by a plenoptic camera allows a number of different computations to be performed on a single image, including digital refocusing, changing the perspective and extending the depth of field. However it is the ability to quantitatively measure the depth of an object and provide a surface map of a scene which is most exciting when applying this technology to the field of retinal imaging, or indeed, any form on quantitative imaging.

Imaging of the retina has developed over the years from two dimensional fundus images to the current gold standard of optical coherence tomography (OCT). OCT has the ability to gain high resolution three dimensional images of the retina, which help with the diagnosis of diseases such as age-related macular degeneration and glaucoma [4]. However, the cost of OCT can be large and too expensive for many small medical centres and many hospitals and surgeries in the developing world. Without the ability to image the retina in 3D many

retinal diseases may either not be diagnosed, or diagnosed much later on increasing the severity of the disease.

The aim of this thesis is to investigate whether it is possible to extract the relevant 3D information using a plenoptic camera, and whether this technology may bring advantages to retinal imaging.

1.1 The Light field function

The term light field dates back to the work from Gershun [5] into expressions for light sources projecting illumination patterns, with Ashdown [6] continuing this area of research. In terms of imaging, light fields can be seen as a set of two-dimensional images of the same scene but from viewpoints with slight variations [7, 8]. The light field function is derived from the plenoptic function, which can be described as the radiance, L which is a function of position and direction in a static scene with a fixed illumination, where radiance is measured in watts / m^2 steradians [7]. This radiance is said to be a function of 5-dimensions, $L(x, y, z, \theta, \phi)$ where x, y and z represent the position and θ and ϕ describe direction (see Figure 1.1) [9]. However, in the absence of occluders (opaque objects which block all light), this function becomes 4-dimensional as the radiance of a single ray of light becomes constant between two points along the length of the ray, i.e. we lose a dimension because we can collect no information about the origin of that ray of light in one dimension as there are no occluders to block the light [10]. This new 4D function is generally called the light field and is what is used to capture light field images [5].

The concept of the 4D light field is conceptionally very similar to that of epipolar volumes from the field of computer vision [11]. It has been shown that these epipolar volumes, along with holographic stereograms, can be captured by translating a camera [12]. This concept has been adopted to capture the light field, along with using arrays of cameras [13]. A more comprehensive look at how the 4D light field can be captured, other than with a plenoptic camera, is given in Section 1.4.

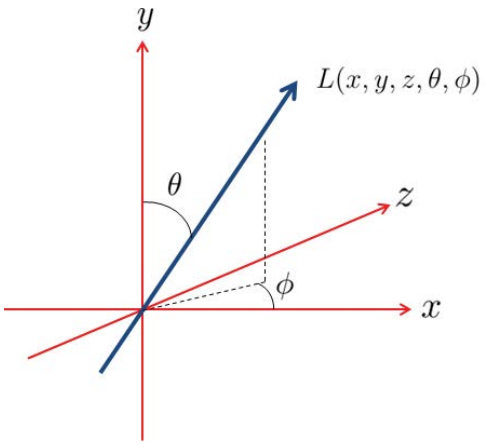


Figure 1.1: The plenoptic function indicating the 5-dimensions of light. From this the 4D light field is derived.

1.1.1 Parametrising the light field

Unlike the 5D plenoptic function, where the parameters are easily visualised, the 4D light field this is not so obvious. A number of ways have been suggested to parametrise the 4D light field. The first candidate for this is by placing a point on a plane, then assigning a direction with θ and ϕ. Therefore, the new function is $L(x, y, \theta, \phi)$. The second candidate is by placing two points onto a sphere then drawing a ray between the latitude and longitude of the first point with that of the second point. This method has one nice characteristic that if the points on the surface of the sphere are chosen at random then there will be an equal chance of measuring the light in all directions and positions, so the entire light field will be uniformly parametrised. This new function is then represented by $L(\theta_1, \phi_1, \theta_2, \phi_2)$. A third method to parametrise is by a rays intersection with two planes [7]. One of the planes is represented by (u, v), the other by (s, t), choose a (u, v, s, t) quadruple and this denotes a single ray of light travelling through the space, hence the new function is $L(u, v, s, t)$ (see Figure 1.2). From these four spatial coordinates, a unique ray of light can be drawn between them so the light field is parametrised. Although this parameterisation cannot represent all rays, for example a ray that is parallel to the two planes providing the planes are parallel to each other, it is advantageous due to its close proximity to the analytic geometry of perspective imaging [10].

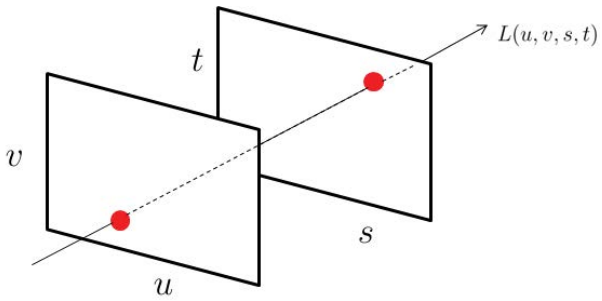


Figure 1.2: With no occluders, the plenoptic function becomes a 4D system. This parametrisation is of two planes, (u, v) and (s, t) with a ray of light between them.

1.2 History of plenoptic imaging

Integral photography was first introduced in the early 20$^{\text{th}}$ Century by Lippmann [14] who proposed the use of a microlens array to capture the full radiance of a scene not just a two-dimensional image [15]. This design was then improved by Ives in 1928 [16] who added an objective lens to improve the sharpness of the images. With the invention of digital photography, many new opportunities became available to investigate integral photography, such as Adelson and Wang [9], who in 1992 placed a lenticular array on the focal plane of the cameras main lens, with the information being used to estimate the depth of the scene from a single image. They proposed that using a plenoptic camera improved the reliability of the depth estimation compared to a binocular stereo system as the correspondence problem is minimised. This is because information about both vertical and horizontal parallax are available and a greater number of views are recorded. This design was then improved in 2005 by Ng [3], who introduced new digital processing techniques such as digital refocusing and extending the depth of field. The design of this new camera has since been dubbed the plenoptic 1.0 or traditional plenoptic camera (see Section 1.7) [17]. Although the basis of most current plenoptic cameras, this design was flawed as the spatial resolution was dependent on the number of microimages, so is typically too small for most applications. Developments by Georgiev and Lumsdaine [18] improved the spatial resolution by the introduction of the focused plenoptic camera (or plenoptic 2.0 camera). The spatial resolution has been improved by creating a relay

system between the microlens array and the objective lens so decoupling the dependency of resolution to the number of microimages (see Section 1.8). The focused plenoptic camera has been developed further to include microlenses with varying focal lengths interwoven into each other to improve the depth resolution [19, 1]. Raytrix GmbH have adapted this technique which has been used to gain up to 200 focal planes from a plenoptic camera.

Plenoptic image formation and manipulation requires substantial computation, and many improvements have been made through algorithm development. Levoy *et al.* first developed the algorithms for light field rendering [7]. Isaksen further developed the light field image rendering method to extend their utility by advancing the image based rendering algorithms through the use of a new parameterisation method [20]. Ng improved digital refocusing by deriving algorithms to process in the Fourier space [21]. Improvements in the spatial resolution by implementing new super-resolution algorithms have been developed [22, 23, 24]. Work has also been performed into image blending to create a more natural blur of out of focus planes in a rendered image [25]. Efforts to reduce the artifacts in a plenoptic camera were also performed [26]. Spatial resolution has also been improved by manipulating the colour demosaicing process [27]. This work has then been developed to the point of full sensor resolution in a rendered image [28]. Along with improvements into the spatial resolution of rendered images, Favaro and Bishop have developed full resolution depth maps by using multiple aliased views [29, 30, 31].

1.3 Applications of plenoptic imaging

When Ng developed the first hand-held plenoptic camera [3], his work only focused on using the technology to enhance the field of photography [32]. This research then lead on to the formation of Lytro, Inc., the worlds first commercial plenoptic camera company [33]. The first generation Lytro, however, only boasts nine different focal planes, so using this camera for scientific research is not feasible. Raytrix is another plenoptic company

specialising in creating cameras suitable for scientific and manufacturing applications as the number of focal planes is ≈ 200 [34]. These cameras can then be used for scientific research, such as the quantification of depth of an angled plane [35], the development of a single snapshot 3D scanner [36] or to image medical tissue in three dimensions [37].

As well as using commercial plenoptic cameras, research for specific applications has involved the development of custom cameras. Levoy *et al.* [38] developed the first light field microscope. Applications in microscopy can benefit hugely from imaging in 3D, but most methods of achieving such data require multiple images such as confocal microscopy. The invention of a single image acquisition to gain 3D information, which Levoy achieved, opens up the ability to image 3D samples in real time *in vivo*. The main drawback though is the sacrifice of spatial resolution to gain angular. This does, currently at least, limit the size of the samples which can be imaged with a light field microscope. The method was then utilised by Prevedel *et al.* [39] who studied the real-time neuronal activity in 3D of zebrafish. The development of the CAFADIS plenoptic camera by Rodríguez-Ramos *et al.* [40] has investigated the use of a plenoptic camera as a wavefront sensor for extremely large telescopes [41]. The field of particle image velocimetry has also benefited from the development of plenoptic cameras, with many publications in the field [42, 43, 44, 45]. This field has seen a lot of activity due to the capability of plenoptic imaging to gain 3D images from a single acquisition, so real time imaging in 3D is possible. As well as practical experiments, computational simulations of light propagating through plenoptic cameras have been developed [46, 47, 48]. A computational study into plenoptic retinal imaging has also been performed to assess the potential benefits of manufacturing a physical system [49, 50].

The available resolution in plenoptic cameras varies dramatically depending on the sensor, microlens array and the placement of the microlens array in the optical system. As will be discussed in detail later in this chapter, a plenoptic system can be placed into either a traditional or focused plenoptic configuration. In the traditional set-up, the

microlens array is placed on the focal plane of the main lens. The lateral resolution of a processed image is then limited to the number of microlenses in the array, and the depth resolution is then a function of the number of pixels under each microlens. This was the first type of plenoptic camera produced [3], and was used throughout the Stamford group for projects such as the light field microscope described earlier [38]. The focused plenoptic configuration places the microlens array into a relay system with the main lens, decoupling the resolution of the computed images from the properties of the microlens array. This provides greater flexibility in the lateral-angular resolution trade-off, so most recent plenoptic systems have adapted this approach. Raytrix, a market leader in commercial plenoptic cameras, claim to have approximately 200 focal planes whilst only reducing the lateral resolution of computer images compared to the image sensor by a factor of around 5. By knowing the capabilities and limitations of both configurations, it is much easier to select the set-up required for a particular application.

1.4 Alternative methods of capturing the light field

As well as plenoptic cameras, there have been many different approaches developed to capture the light field. These different methods can be grouped into a few categories.

Multi-image

Photogrammetry is one alternative to using a plenoptic camera, which works by using a single traditional sensor to take multiple images [51, 52]. This can be achieved by mechanically translating the sensor around the object being imaged [7], or by rotating the object with a static camera whilst imaging at known orientations. In situations where neither of the two previous methods are possible, a mirror can be rotated effectively creating a series of images from a set of virtual cameras [53, 54]. These methods have the advantage over plenoptic imaging in that no sacrifice in sensor resolution has to be made. However, the objects which can be imaged are limited to ones which are static as

movement of the object between images would cause artefacts in the processed images. Sequential imaging using a static camera and scene can also be achieved instead of any of the above. Liang *et al.* [55] proposed a method in which a sequence of conventional images are taken, but with a dynamic aperture mask placed close to the sensor and changing mask between each image acquisition. The speed of the process was improved with the use of Hadamard aperture patterns [56]. These methods do suffer from the same limitations as the earlier examples, as movements in the objects being imaged will add artefacts in the processed images.

Single image multiplexing

Although a plenoptic camera does fall under this category, there are other methods in which the 4D light field can be captured from a single image acquisition. One of these methods was to create an array of lenses external to the main lens in a conventional system [57]. This method was then extended further to incorporate variable focal length liquid lenses, allowing for optimisation of the system depending on the scene [58, 59]. Another method for acquiring light field data is with the use of a heterodyned camera [60]. In this method a coded aperture is placed between the main lens and the sensor, with the technique advanced further by the use of an adaptive mask [61]. Although this method is much cheaper and simpler than a plenoptic camera, as a large majority of the light is blocked, it requires a longer acquisition time to achieve an acceptable signal to noise. This is not seen as a viable option for retinal imaging, where exposure times are always kept to a minimum to reduce the risk of retinal damage.

Multiple camera

The 4D light field can also be captured with the use of a camera array. One of these was developed at Stanford [13] [62], in which hundreds of webcams were placed into a rectangular array, then algorithms to change the focal plane, perspective, and effectively see around or through semi-transparent objects can be implemented. A similar system

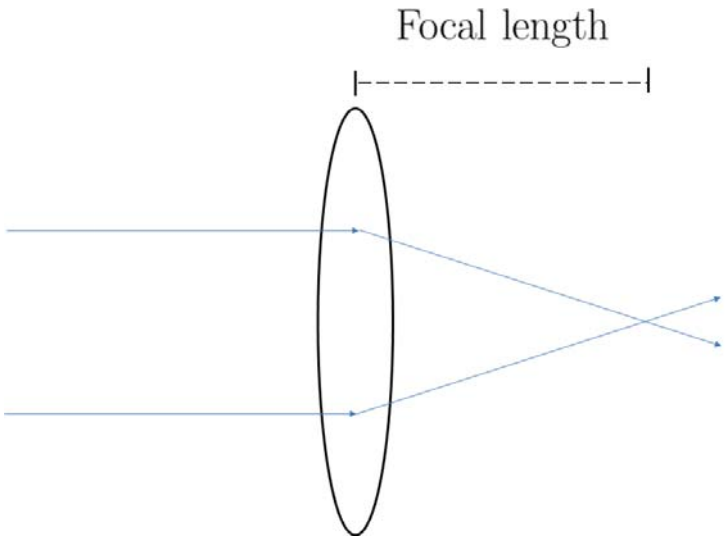


Figure 1.3: A diagram highlighting what the focal length of a lens is

but with fewer cameras was proposed by Yang *et al.* [63]. A custom arrangement using cameras attached to a flexible sheet has also been developed [64]. These systems are the most similar to a plenoptic camera, as the multiple cameras are analogous to the microimages formed under each microlens.

1.5 Basic optics

Throughout this thesis, particular terms will be used which need to be understood for both the theory and practical results to have merit. The first of these is focal length. Basically, the focal length is the distance between the centre of a convex lens and the point at which parallel rays passing through the lens converge. Figure 1.3 helps to illustrate this point.

Thin lens approximation

Now that the focal length of a lens has been defined, it is important to understand how an image is formed of an object by a lens. The equations used throughout this thesis, unless otherwise stated, use the thin lens approximation. This states that the thickness of the lens is negligible compared to the radius of curvature of the lens in question.

This approximation is used as it simplifies the ray tracing and ray transfer matrices calculations. The object distance, u, can then be related to the image distance, v, by the focal length of the lens, f, by the thin lens equation;

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \tag{1.1}$$

F-number

The f-number of an optical system is a dimensionless value that represents the ratio of the focal length of the system to the diameter of the entrance pupil [65]. It can be calculated from the following equation, where F is the f-number, f is the focal length of the lens and D is the diameter of the lens;

$$F = \frac{f}{D} \tag{1.2}$$

The f-number becomes especially important in this thesis in Chapter 7 when novel methods are used to create custom microlens arrays. As will be discussed in more detail in that chapter, it is essential for the f-number of the main lens to match that of the microlens array. For more information, see Section 7.1.2.

1.6 Image formation

To understand how light field imaging can be useful, it is important to understand how an image is formed inside a conventional camera. Image formation can be seen as determining the irradiance, I, from the aperture of the main lens, on the (u, v) plane, to a particular position on the sensor plane (s, t) and can be expressed as

$$I_F = \frac{1}{F^2} \iint L_F(u, v, s, t) \cos^4\theta \, \mathrm{d}s \, \mathrm{d}t, \tag{1.3}$$

where F is the distance between the aperture of the main lens and the image plane, L_F is the light field specified on the sensor plane, and θ is the angle between the sensor plane normal and the ray (u, v, s, t). For simplicity the $\cos^4\theta$ term is absorbed into the definition of the light field, L_F, so the irradiance becomes

$$I_F = \frac{1}{F^2} \iint L_F(u, v, s, t) \, \mathrm{d}s \, \mathrm{d}t, \tag{1.4}$$

For simplification, the light field function $L_F(u, v, s, t)$ can be seen as a radiance with a spatial component q and an angular component p, such that

$$L_F(u, v, s, t) = r(q, p) \tag{1.5}$$

Each pixel on the sensor receives lights from all directions. For an image to be formed, all of the directions for a given spatial coordinate are integrated to get a single brightness value. This is equivalent to performing a line integral parallel to the p axis, as shown by Figure 1.4 on the following page and represented by

$$I_F(q) = \int r(q, p) \, \mathrm{d}p \tag{1.6}$$

To refocus a conventional camera, the main lens is simply shifted by a distance d. Mathematically this can be seen as the main lens being moved along the optical axis by a distance of s as

$$s = F(\alpha - 1)$$

where $\alpha = \frac{F'}{F}$, F' is the new distance between the aperture on the main lens and the new image plane, and F is the original distance between the aperture and image plane before refocusing as previously defined. When applying this to the ray space representation, this is equivalent to a shearing of the radiance. A positive shear represents moving the

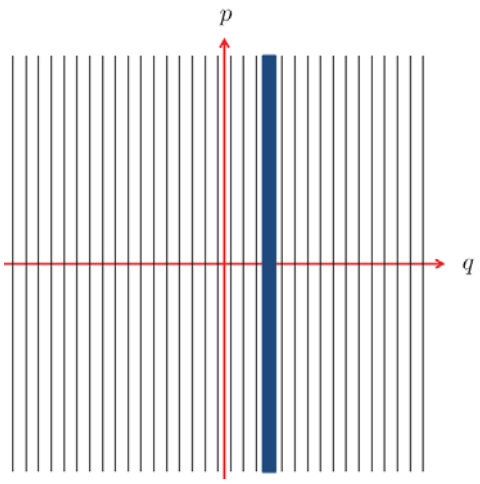


Figure 1.4: A line integral taken parallel to the p axis will result in a pixel outputted onto the q axis

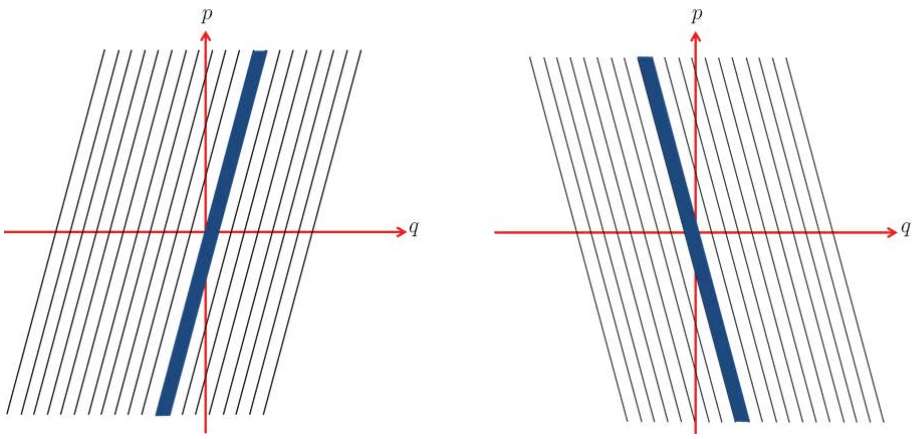


Figure 1.5: Line integrals showing how the shearing of the radiance corresponds to refocusing. Left: Lens far away from sensor, positive slope. Right: Lens close to sensor, negative slope

lens further away from the sensor, imposing a positive slope in ray space. A negative shear is the opposite, with the main lens being moved closer to the sensor and creating a negative slope in ray space. A visual representation of this can be seen in Figure 1.5.

The light field for refocusing is also sheared and the irradiance will then become

$$I_{\alpha \cdot F}(q) = \int r(q + (\alpha - 1)Fp, p) \, \mathrm{d}p \tag{1.7}$$

$$= \frac{1}{\alpha F} \int L_F \left(\frac{u}{\alpha} + s \left(1 - \frac{1}{\alpha} \right), s \right) \, \mathrm{d}s \tag{1.8}$$

Applying this transformation to the light field of a system in which the distance between the main lens and the sensor is changed from F to F' the light field becomes

$$L_{F'} = L_F \left(s \left(1 - \frac{1}{\alpha} \right) + \frac{u}{\alpha}, t \left(1 - \frac{1}{\alpha} \right) + \frac{y}{\alpha}, s, t \right) \tag{1.9}$$

so the irradiance at a separation of F' and with spatial co-ordinates of (u, v) follows [21]

$$I_{\alpha \cdot F}(u, v) = \frac{1}{(F')^2} \iint L_{F'}(u, v, s, t) \, \mathrm{d}s \, \mathrm{d}t \tag{1.10}$$

$$= \frac{1}{\alpha^2 F^2} \iint \left(s \left(1 - \frac{1}{\alpha} \right) + \frac{u}{\alpha}, t \left(1 - \frac{1}{\alpha} \right) + \frac{y}{\alpha}, s, t \right) \, \mathrm{d}s \, \mathrm{d}t \tag{1.11}$$

From evaluating equation 1.6, it can be shown that refocusing the light field is a four-dimensional shearing of the original function. Although this equation is in relation to traditional photography, it will be shown later that it is also applicable to plenoptic images but instead of a physical shift of the main lens, the distance F' is changed computationally.

1.6.1 Image formation in a plenoptic camera

In this section the formation of images inside a plenoptic camera is described, as first put forward by Bishop *et al.* [24]. For this work a 3D point in space, P $= [x, y, z]^T \in \mathbb{R}^3$ is transferred through the main lens by using the thin lens law. This states that a point in space at a distance from the lens, z, produces an in-focus point at a distance of $\frac{Fz}{z-F}$

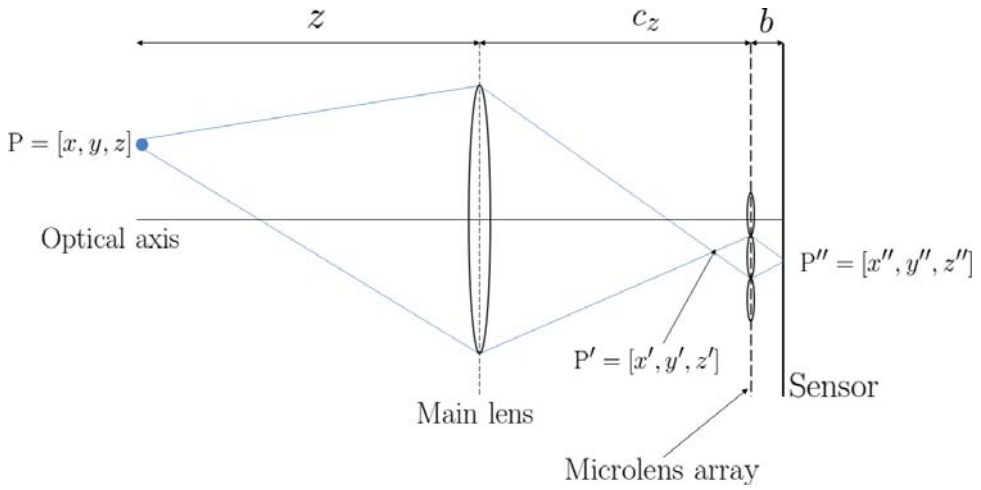


Figure 1.6: A diagram showing the transfer of a point through a plenoptic camera.

from the lens [66] where F is the focal length of the lens. The position at which point P comes into focus is $\mathrm{P}' = [x', y', z']^T \in \mathbb{R}^3$, where

$$\mathrm{P}' = \frac{F}{z - F} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathrm{P} \tag{1.12}$$

and z is the distance between the point, P, and the main lens as per Figure 1.6. The final projection of the point, P, is onto the sensor, through a microlens centred at point C= $[C_x, C_y, C_z]^T \in \mathbb{R}^3$ by

$$\mathrm{P}'' = \mathrm{C} + \frac{b}{C_z - z'} \left(\mathrm{P}\frac{z'}{z} + \mathrm{C} \right) \tag{1.13}$$

where b is the distance between the microlens array and the sensor.

1.6.2 Optical transfer matrices

The transport of a ray of light in free space and the refraction of light due to a lens can be described by the two matrix transforms, T and R_L respectively [17], where d is an arbitrary distance and f is the focal length of the lens.

$$T = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix}, R_L = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix} \tag{1.14}$$

Using the notation proposed by Lumsdaine and Georgiev [22] the radiance at a plane perpendicular to the optical axis is represented by $r(q, p)$, where q is the position and p is the slope of the ray relative to the optical axis [25]. From this any coordinate in ray space can be represented by $x = (q, p)^T$. Using this notation, the light field [7] is the radiance as a function in ray space, $r(x)$. Taking an arbitrary ray transfer matrix, A, each ray is transformed by

$$x' = Ax \tag{1.15}$$

If A is an optical transform of equation 1.15, then considering the transformation of $r(x)$ to $r'(x)$ since all optical tranfer matrices follow det $A = 1$ and assuming all energy is conserved, $r'(x') = r(x)$. Substituting in equation 1.15 then $r'(Ax) = r(x)$. By considering a ray $y = Ax$, it can be seen that $r'(y) = r(A^{-1}y)$ [17], but since y is an arbitrary ray, the radiance transformation equation can be derived as

$$r'(x) = r(A^{-1}x) \tag{1.16}$$

1.7 Traditional plenoptic (1.0) camera

A traditional plenoptic camera, or plenoptic 1.0 camera is created by placing a microlens array on the focal plane of the main objective lens and the sensor one microlens focal length from the microlens array [3] as shown in Figure 1.7. In this set up, parallel rays coming from a given direction are focused by a microlens onto a single point, effectively recording angular information as a spatial distribution.

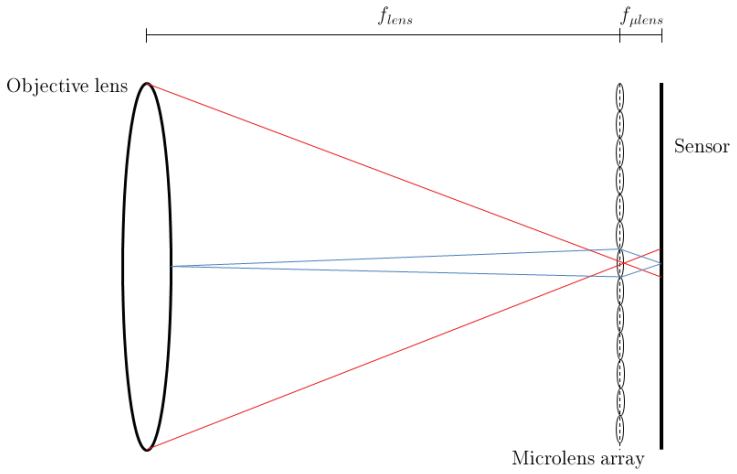


Figure 1.7: A diagram showing the set up of a plenoptic 1.0 camera, with the microlens array on the focal plane of the objective lens, and the sensor one microlens focal length away from the microlens array.

Considering a single microlens, it is possible to show that each pixel measures the intensity of light incident onto the main lens from a specific angle. To show this, the transformation matrices A and A^{-1} need to be calculated for rays that are incident onto a plane one microlens focal length, f in front of the microlens. From using the matrix transforms in equation 1.14, it can be seen that

$$A = \begin{bmatrix} 1 & f \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix} \begin{bmatrix} 1 & f \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & f \\ -\frac{1}{f} & 1 \end{bmatrix}, A^{-1} = \begin{bmatrix} 0 & -f \\ \frac{1}{f} & 0 \end{bmatrix} \tag{1.17}$$

Considering equation 1.17, a single pixel on the sensor has a uniform response to rays from all angles, so its sampling kernel in ray space is represented by a vertical line, as shown in Figure 1.8 [25]. This figure also shows the rays after transformation by A^{-1}, which transforms vertical lines into horizontal lines. This is due to the slope of the input ray relative to the optical axis, p, having no influence on the output. Mathematically, this occurs because of the zero in the bottom right element of the matrices in equation 1.17. To render an image in the plenoptic 1.0 configuration, all angular samples from a specific spatial coordinate need to be integrated. However, since a single spatial coordinate is sampled by a single microlens, all pixels under a microlens need to be integrated. This

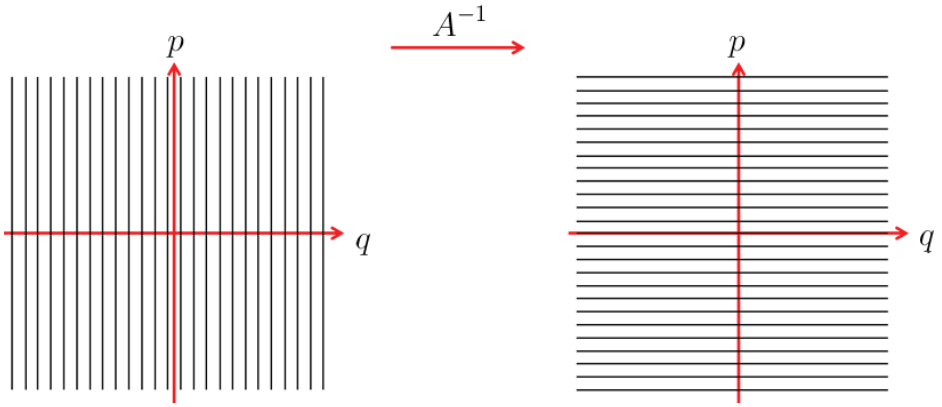


Figure 1.8: The sampling kernel for rays incident onto a single microlens in the plenoptic 1.0 set up.

puts a limit on the spatial resolution of the image rendered to that of the number of microlenses visible on the raw image. Therefore, even if the sensor has many megapixels, if the microlens array only has 500 × 500 microlenses, then the final image is limited to a maximum spatial resolution of 0.25MP. Many applications, especially photography for which this methodology was first devised, require a greater spatial resolution. The way to overcome this is to use a larger number of microlenses, which may require the use of a higher resolution sensor in order to preserve the angular information, or to change the configuration of the components slightly. The change in configuration is what was proposed by Lumsdaine and Georgiev who created the focused plenoptic camera, or plenoptic 2.0 [18].

1.8 Focused plenoptic (2.0) camera

In the plenoptic 2.0 set up, the microlenses are focused onto the image plane of the objective lens instead of being placed in that plane. This effectively creates a relay system with the main lens with focal length f, satisfying the lens equation $1/a + 1/b = 1/f$. The principal difference between this and the plenoptic 1.0 is that in the traditional configuration the microimages created by each microlens are completely defocused with

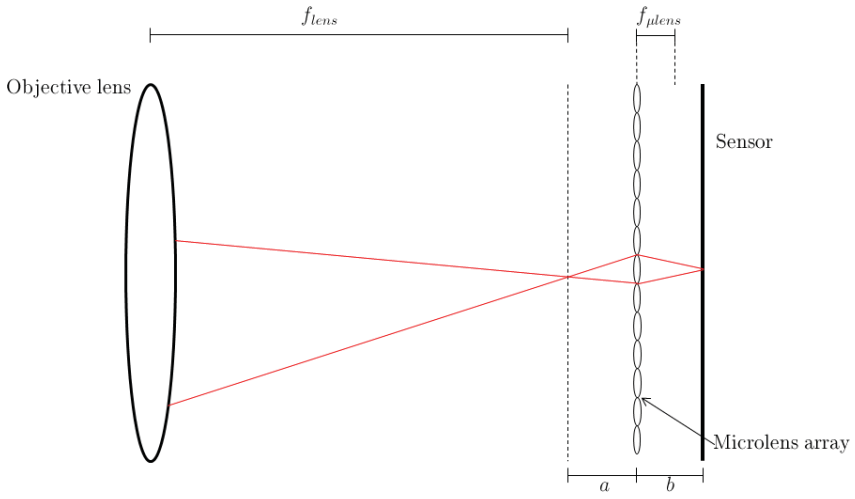


Figure 1.9: A diagram showing the set up of a plenoptic 2.0 camera, with the microlenses focused onto the image plane of the objective lens. The microlenses then form a relay system with the objective lens.

respect to the scene being imaged, whereas in the plenoptic 2.0 they are focused due to the rays entering into a relay system [17].

Using the optical transformation matrices derived earlier, the total transfer matrix from the image plane of the objective lens inside the camera is

$$A = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix} \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{b}{a} & 0 \\ -\frac{1}{f} & -\frac{a}{b} \end{bmatrix}, A^{-1} = \begin{bmatrix} -\frac{a}{b} & 0 \\ \frac{1}{f} & -\frac{b}{a} \end{bmatrix} \tag{1.18}$$

The sampling kernel for the plenoptic 2.0 is also very different from the traditional arrangement, as each microlens image captures both the angular and positional distributions of the radiance of the image plane of the objective lens. From this, sampling is no longer related to microlens size and the spatial resolution of the computed image is b/a multiplied by the resolution of the sensor. A result of this is that the spatio-angular trade-off for the plenoptic 2.0 is no longer determined by the number of microlenses, but instead by the relationship between the distances a and b. To reduce edge effects in the images, larger microlenses can be used. As with the plenoptic 1.0, the images are rendered by integrating the angular samples of the radiance at all spatial coordinates. However, in the focused plenoptic camera the angular samples from a specific spatial point are recorded

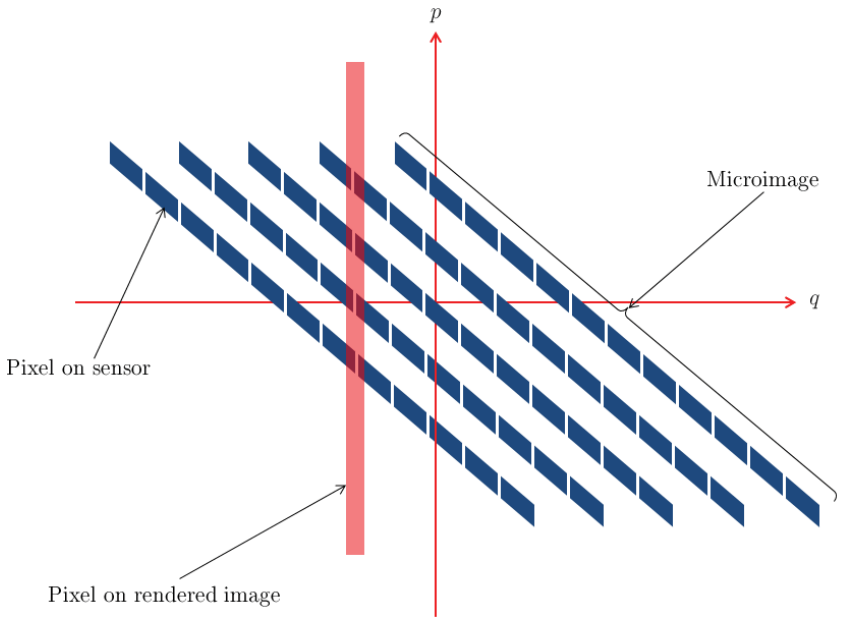


Figure 1.10: A ray space diagram indicating how an image is rendered when in the plenoptic 2.0 configuration. Note that integration of a pixel takes place across different microimages, and that a single microimage will contribute to more than one pixel in a rendered image.

by different microlenses, so rendering the plenoptic 2.0 involves integrating across many different microimages, rather than just from one (see Figure 1.10).

1.9 Plenoptic rendering

The previous sections explained how a plenoptic camera is constructed in both the traditional and focused plenoptic configurations. However, a raw image from a plenoptic camera needs to be processed computationally before even a traditional 2D image can be rendered. This section will focus on the algorithms used to render images, digitally refocus, change perspective, extend the depth of field and create depth maps from the raw image in both the traditional and focused configurations. Although the raw images have been taken from various sources, all rendering code used in this section was written by the author of this thesis.

1.10 Traditional plenoptic (1.0) rendering

As the angular samples for a given spatial point are sampled by a single microlens array, basic rendering in the traditional plenoptic camera is relatively simple. To recreate a traditional image, all the rays from a microimage need to be integrated together, as shown in equation 1.6. This can also be seen as averaging over the directions at each position, so follows equation 1.19

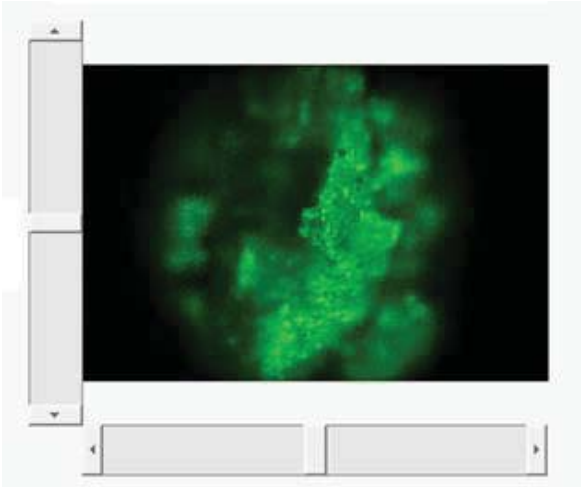
$$I[u,v] = \frac{1}{N^2} \sum_{s,t} r[u,v,s,t] \tag{1.19}$$

1.10.1 Synthetic aperture

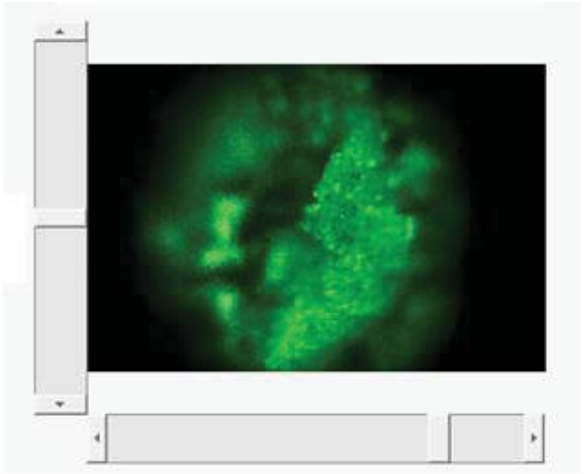
The aperture of the lens in the rendered image can be changed synthetically. This is achieved by instead of averaging the entire microimage, a smaller section is summed. To render an image equivalent to that of a pinhole camera, only the central pixel from each microimage should be taken into the final image.

1.10.2 Multiple viewpoints

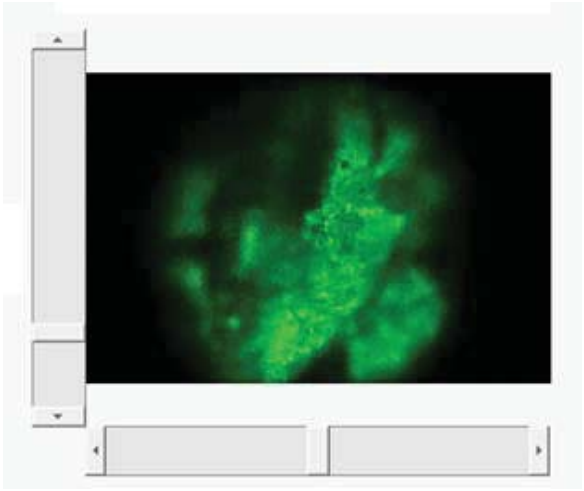
To change the perspective view of the scene when in the plenoptic 1.0 configuration, only a minor change to the basic rendering is needed. Take for example the case where a synthetic pinhole aperture has been used. Instead of extracting the central pixel from each microimage, depending on what perspective is desired, pick the corresponding pixel from under that microimage, i.e the bottom left pixel under a microimage if the desired viewpoint is from the bottom left. As long as the same pixel is taken from each microimage, a different perspective view will be generated. In order to visualise this, as graphical user interface (GUI) was created in Matlab. Examples of a few different viewpoints of some fluorescent wax taken under Stanford's light field microscope [67] is shown in Figure 1.11.



(a) Image taken from the central points of the object

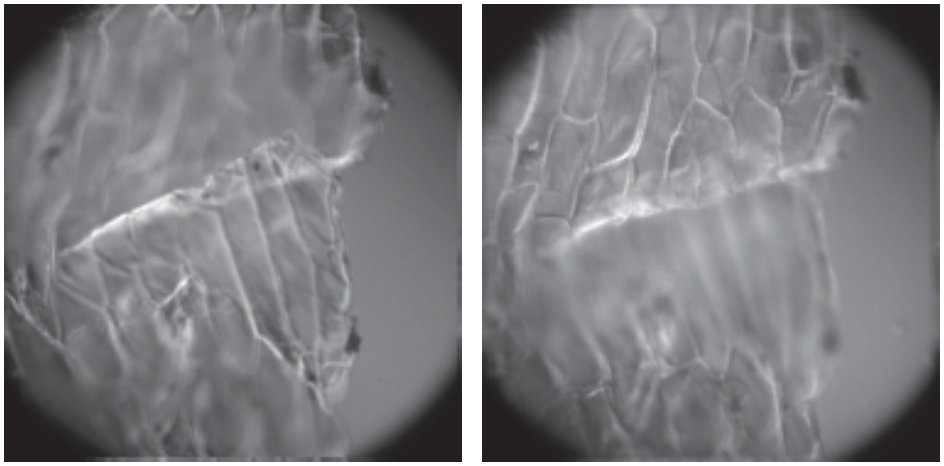


(b) Viewpoint taken from the right of the object



(c) Viewpoint taken from below the object

Figure 1.11: Images of fluorescent wax taken by Stanford light field microscope [67], all at different perspective views.



(a) Image rendered with the focal plane in the background

(b) Image rendered with the focal plane in the foreground

Figure 1.12: Images of onion cells taken by Stanford light field microscope [68], taken at two different focal planes highlighting the ability to refocus digitally.

1.10.3 Digital refocusing

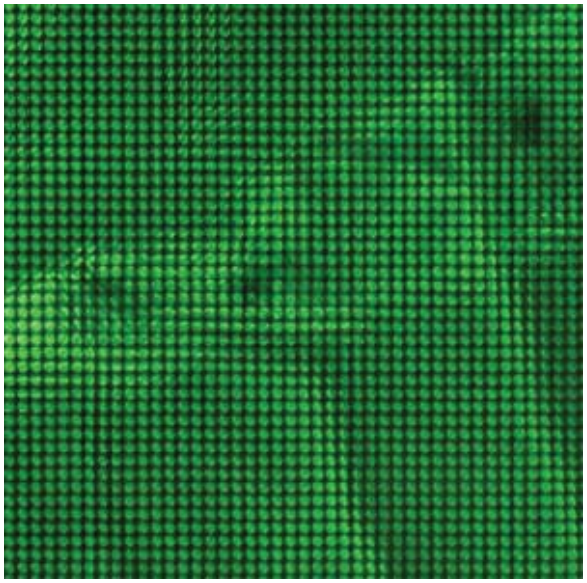
To digitally refocus an image taken with the traditional plenoptic camera, a shearing transform needs to be applied to the radiance $r(q, p)$ as shown by equation 1.20

$$r'(q, p) = r(q - tp, p) \tag{1.20}$$

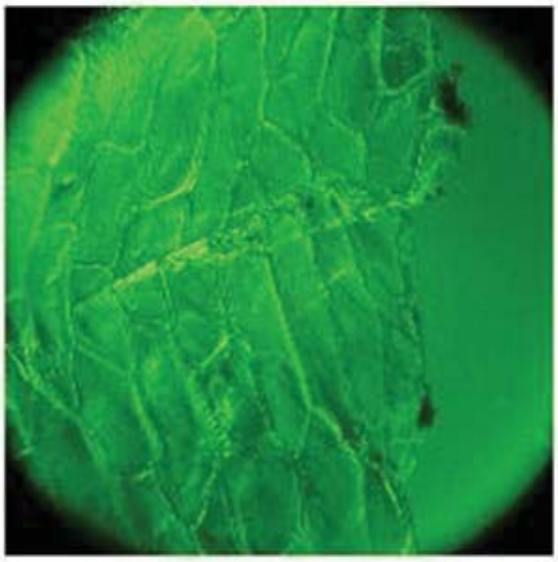
The new image is then rendered in the same way as previously (see Section1.10),

$$I'(q) = \int r'(q, p)dp \tag{1.21}$$

This shearing transform is also identical to the refocusing of a traditional camera, except that the movement of the main lens is no longer a physical movement, but is done computationally (see Section 1.6 and equation 1.6). An example of refocusing for plenoptic 1.0 data is shown in Figure 1.12.



(a) A crop of a raw image of some onion cells taken under a light field microscope developed at Stanford University [68]



(b) The rendered image from Figure 1.13a by taking the central pixel from all of the microimages.

Figure 1.13: A crop of the raw image then the rendered image by taking the central pixel from under each microlens. Raw image taken from website of Stanford Light Field Microscopy [68].

1.10.4 Extended depth of field

An all in focus, or extended depth of field image can be rendered by simply extracting the central pixel behind each microlens array, then placing these values into a new image as this replicates a pinhole aperture. The resolution of this image as previously described is limited by the number of microlenses. An example of a raw image and the corresponding rendered image is shown in Figure 1.13.

1.11 Focused plenoptic (2.0) rendering

Rendering in the plenoptic 2.0 configuration is different to that of the traditional camera for many reasons. In the conventional plenoptic 1.0 camera, all of the directional information is contained within a single microimage, and the spatial information in different microimages. In the plenoptic 2.0 camera, different views are spread across many microimages. Also, as the microlenses now act as a relay system rather than splitting

rays incident on the focal plane of the main lens, the microimages under each microlens are inverted. The resolution of the system is dependent on b/a multiplied by the sensor resolution as described previously in Section 1.2, so it is clear that more than a single pixel must be extracted from each microimage unlike in the plenoptic 1.0 configuration when a single pixel is taken from a microimage. The number of pixels extracted, or patch size, is also dependent on the depth of the object in the scene. As objects move closer to the main lens, the relationship between a and b become closer to equality, hence the resolution increases. In relation to patch size, the patch size increases the closer to the main lens the object is. It is also this attribute of the plenoptic 2.0 configuration that leads to the ability to digitally refocus the images.

1.11.1 Multiple viewpoints

To change the viewpoint of a focused plenoptic image, it follows the same algorithm as for the traditional configuration. Once a particular patch size has been chosen to be taken from a microimage, to change the perspective view the patch should be moved around the microimage depending on the desired view.

1.11.2 Digital refocusing

Objects in the scene will be shifted by different amounts on adjacent microimages dependent on the depth of the object. Objects far away from the main lens will be translated a small amount between adjacent microimages, so a small patch size is needed in order to reconstruct the image correctly. This consequently causes a lower resolution for objects far away from the camera. For objects close to the camera, a larger translation between adjacent microimages means a larger patch size is required to reconstruct the object correctly. As more pixels are taken from the microimages, objects close to the camera will have a higher resolution than those further away.

(a) Patch size = 8

(b) Patch size = 9

(c) Patch size = 10

(d) Patch size = 11

Figure 1.14: Rendered images taken with varying patch sizes. Raw image taken from website of Todor Georgiev [69].

As objects at different depths require different sized patches taken from the microimages, providing the objects are not all in the same plane, some objects will always be reconstructed incorrectly if a single patch size is chosen. By changing the patch size incrementally, this can be seen as changing focus of the image. Figure 1.14 shows examples of images rendered with varying the patch size.

Examining the trees in the background of Figure 1.14a, it can be seen that patch size 8 is the correct size for that depth of the scene. This is because the trees appear in focus in this image. In contrast, the trees in Figure 1.14d appear to be out of focus. This is because a patch size of 11 is too large for that particular depth, so the same part of the tree is repeated. If the camera and watch in the foreground is taken into consideration,

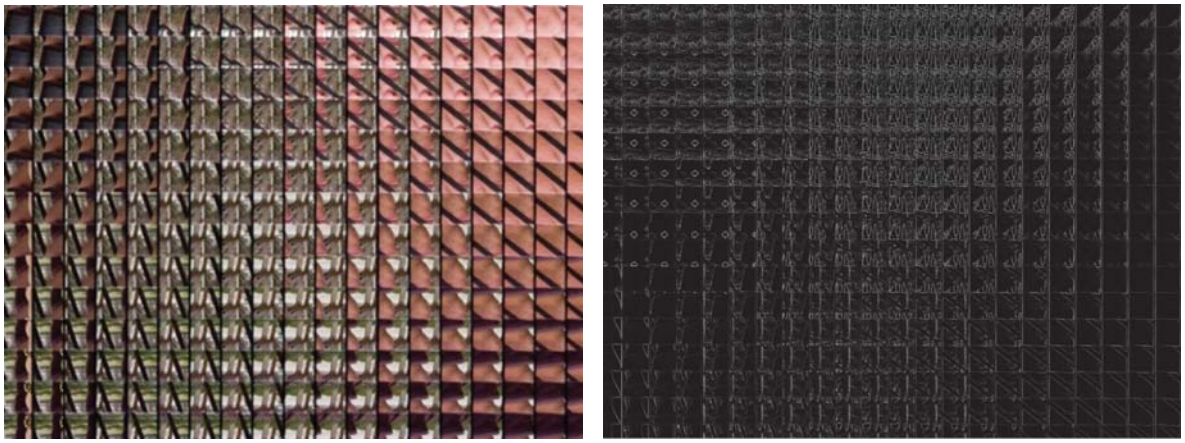
the opposite effect can be seen, with it being in focus in Figure 1.14d and out of focus in Figure 1.14a. This is due to this object being closer to the main lens of the plenoptic camera, so these objects on the raw image are shifted more between microimages, so a larger patch is needed to render the image correctly. Figure 1.14 shows that by changing the size of the patch taken inside the microimage, the plane of focus is effectively changed, so this method can be seen as a way of digitally refocusing the image after acquisition.

1.11.3 Depth map generation

Depth maps can be produced from a plenoptic raw image. They have been investigated and solved in a number of different ways, such as through the use of aliased light fields by Bishop *et al.* [70] or through registration of adjacent microimages by Georgiev *et al.* [25]. Below is a method which has been developed based upon the ideas presented by Georgiev.

In a focused plenoptic raw image, the size of the patch taken represents the depth of that region. For the image used in Figure 1.14, this would mean using patches of different sizes for different parts of the image. The problem then arises of what patch size to use for each microimage. In order to determine the correct patch size, the repetition of the same regions on adjacent microimages is used, along with the linear relationship between depth and the size of the translation. The problem then results in a basic image registration problem, with the best match between adjacent microimages being the optimum patch size for that microimage. In order to register the images together, an edge detection algorithm was implemented to make quantitative analysis on the microimages easier, as shown in Figure 1.15b.

To determine the optimum patch size, take a box from the centre of a microimage on the edge enhanced image (see Figure 1.15b), then taking a box of the same size place it at the same y-coordinate, but at the furthest left point on the box to the right. Find the difference between the two images, then translate the box on the edge of the box



(a) A crop of the raw image used to render the images in Figure 1.14

(b) A crop from the same raw image but with the novel edge detection algorithm applied

Figure 1.15: Crops of the raw image and the raw image with the edge detection algorithm applied to aid the patch size estimation. Raw image taken from website of Todor Georgiev [69].

by a single pixel to the right and perform the same calculation as before. Repeat this procedure until the translated box reaches the end of the adjacent microimage. The distance from the left edge at which the difference is a minimum is the optimum patch size from that particular microimage. The same procedure is repeated but keeping the x-coordinate constant and taking the box below instead of to the right and the optimum patch size from both these operations are averaged to give the final patch size. In order to test the accuracy of the patch sizes calculated, a map of the depths can be generated, as shown in Figure 1.16.

In this image, the lighter regions indicate the foreground of the image, or larger patch size, and darker regions the background, or smaller patch size. As the outline of the person in the image can be seen as being in the foreground, this provides an indication that the depth estimation was implemented correctly. To visualise this further, a surface plot of the depth map with the originally rendered image is shown in Figure 1.17.

Figure 1.16: An image showing the different patch sizes used for depth based rendering. The lighter regions indicate a larger patch size (the foreground of the image) and darker regions the smaller patch sizes (background of the image)

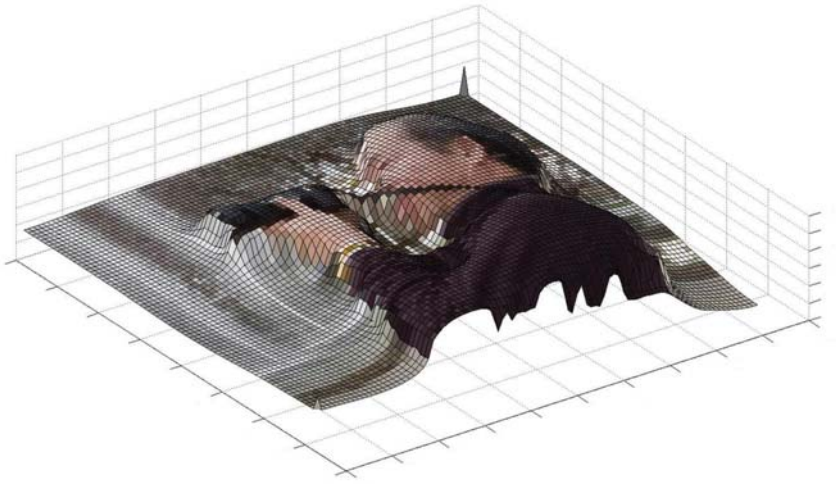


Figure 1.17: A surface plot using the values in the depth map in Figure 1.16

Figure 1.18: A rendered image using the values for patch size generated by the depth map. The image has all planes in focus resulting in an extended depth of field image

1.11.4 Extended depth of field

An extended depth of field image can be generated with all parts of the image being in focus, however this is non-trivial unlike the plenoptic 1.0 rendering. To extend the depth of field from a focused plenoptic raw image, the depth map from the previous section needs to be used. Using the ideal patch sizes generated for each microimage, a final image can be generated using these patch sizes, and is shown in Figure 1.18. The patches from the microimages may need to be resized to all be the same size so they all fit together to create one single all in focus image.

1.12 Commercial plenoptic cameras

There are currently two main commercial producers of plenoptic cameras; Lytro and Raytrix. Although any results from a commercial plenoptic camera throughout this thesis have been captured with a Raytrix, both the Lytro and Raytrix will be explained and compared. Through this comparison, it should become clear as to why only a Raytrix was used.

1.12.1 Lytro

Lytro is a spin-out company by Ren Ng after the successful completion of his thesis on digital light field photography [32]. As the name suggests, the Lytro camera was developed for the use of commercial photography. The first generation camera produced 9 focal planes with an output image size of approximately 1 mega-pixel [33]. The camera is placed into the traditional plenoptic configuration as described in Section 1.7 as this is the technology Ng developed throughout his research.

The small number of focal planes outputted from the camera is for a few reasons. Firstly, as it is in the traditional configuration, the output rendered image resolution is equal to the number of microlenses in the system. Therefore a microlens array with small diameter lenses was used. This in turn created a microimage with only 23 pixels beneath it. If gaining an accurate depth of an object in a scene is what you desire, these parameters are not ideal. However, Lytro was designed as a consumer camera for photography, with the main function being the ability to refocus the image after it has been taken. The first generation Lytro was first sold at a price of around £500 in the UK retail market.

1.12.2 Raytrix

Raytrix was founded in 2009 and began selling its first generation cameras in 2010. Unlike Lytro which focusses on the consumer market with a main focus being on digitally refocussing the images, Raytrix cameras are aimed at gathering full 3D models of a scene for research and industrial applications. This does, however, require significantly more depth resolution whilst maintaining a similar or greater lateral resolution as the Lytro. The high rendered image resolution is achieved by placing the system into the focused plenoptic configuration as described in Section 1.8, as in this set up the image resolution is decoupled from the number of microlenses. The greater number of focal planes is achieved by using three different focal length microlenses in the same array, as described by Georgiev [19]. By using this method around 200 focal planes can be reconstructed,

whilst maintaining a rendered image over a 1 MP and a sensor with a resolution of no higher than 10 MP. However this does come at a much greater price, with the cheapest models costing approximately £10,000 and going up to around £100,000.

Raytrix depth maps

As a majority of Chapters 4, 5 and 6 will focus on the analysis of depth maps generated using the Raytrix software (RxLive 2.10), it may be useful to first describe their format. They are presented in the form of an RGB image, with the red end of the spectrum representing larger values than the blue end. The depth values presented are only relative depths between different areas on the image, with the large values (closer to the red end of the spectrum) representing objects close to the camera and smaller values (blue) representing further away. If part of the depth map appears black, this represents that no depth value could be calculated for this point. If an area of the depth map appears grey, this indicates the calculated depth for this area is outside the maximum or minimum depth values set in the software.

1.13 Conclusion

This chapter has given an overview of plenoptic imaging, from when it was first hypothesised by Lippmann in the early 20^{th} Century [14, 15], through to Ng who developed a hand-held plenoptic camera [3] from which the first commercial plenoptic camera company, Lytro, was founded. Different applications and alternatives to plenoptic imaging were then explored before going into details on both the traditional and focused plenoptic configurations. An in depth analysis of plenoptic rendering in both configurations was shown, including methods for obtaining depth maps. The depth maps from the focused plenoptic set-up will be explained and derived further in Chapter 3 as they are used extensively in Chapters 4, 5 and 6.

Now that the theory of plenoptic imaging has been explained and derived, it is worth looking at how this can be applied to imaging of the retina. To achieve this, different capabilities are explored computationally and the results can be found in the following chapter. On the back of the simulation results, it is determined that recovering the surface topography of the retina is the path which is most likely to yield positive results. However, it was then also found that plenoptic depth recovers struggled when imaging scenes with minimal features. Work-flows were then developed to investigate how projecting texture onto featureless scenes improved depth recovery from a commercial plenoptic camera. This then led to the final part of the thesis, which is on the development on novel, cost-effective microlens arrays for use in the manufacturing of cost-efficient plenoptic cameras.

Chapter 2

Plenoptic retinal imaging

The previous chapters have explained the background theory of what the light field is and how it can be recorded by the means of a plenoptic camera. A thorough investigation into the algorithms used to digitally refocus, shift perspective and calculate the depth of objects in a scene have been explored. An understanding of these algorithms is important when investigating how plenoptic imaging may be beneficial to the field of retinal imaging. In this chapter, we develop a comprehensive understanding of which features of plenoptic imaging could be applied to retinal imaging. This is achieved by performing simulations to model the transport of light through a computationally modelled human retina and looking at the transmittance through the different layers of the retina, as well as the angular reflectance from the surface.

2.1 Eye and common diseases

The human eye is a sensory organ which allows vision, and a diagram highlighting the main areas is shown in Figure 2.1. The visible parts of the eye include the white sclera, the pupil in which light enters the eye, the coloured pupil iris and the transparent cornea. When light enters the eye, it is partially focused by the cornea, before passing through the pupil, and being further focused by the lens. The light then finally passes through

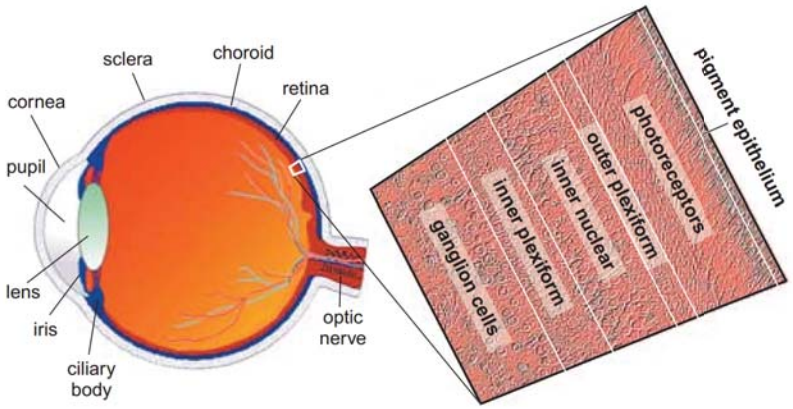


Figure 2.1: A diagram indicating the main structures of the eye, including the different layers of the retina. The photoreceptors are located against the pigment epithelium in the top layer of the retina. Image from: [72].

the vitreous before being focused onto the retina [71]. The retina, a thin layer of tissue approximately half a millimeter thick [72], can be divided into different layers as shown in Figure 2.1. When the photons reach the photoreceptors (rods and cones) in the retina, this triggers a series of neural connections where the information is transported to the brain via the optic nerve.

2.1.1 Diabetes and diabetic retinopathy

Imaging of the eye can provide information about many different diseases, many not directly associated with the eye itself. One of the most common diseases that can be studied via eye imaging is diabetes. In 2007, the number of patients in the U.S. with diabetes reached 23.5 million [71], and with a complication of diabetes, diabetes retinopathy, being the second most common cause of blindness in the U.S. the early diagnosis of these diseases is important to prevent blindness and visual loss [73]. For patients with diabetes, the main cause of visual loss is diabetic macular edema (DME) which is a breakdown of the blood-retinal barrier [74] and damages the macular structure of the retinal tissue both in the short and long term. The management of diabetes primarily involves the lowering of blood sugar levels through a change in diet and anti-diabetic drugs, however if diabetic retinopathy is present, management of clinically significant macular edema

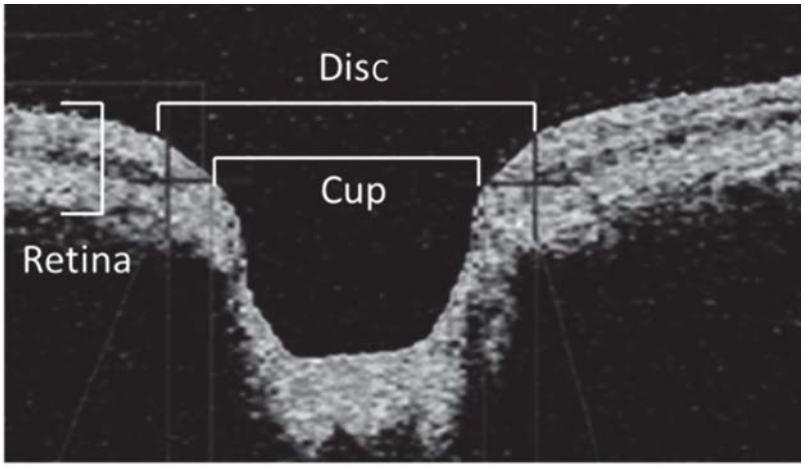


Figure 2.2: An OCT image of an optic nerve, with the cup and disc illustrated. The cup-to-disc ratio is used for the diagnosis of glaucoma.

(CSME) needs to be implemented by the administration of anti-vascular growth factors and steroids, which have been shown to prevent blindness and further visual loss [74][75].

2.1.2 Glaucoma

Glaucoma is the third largest cause of blindness in the U.S. and is caused by damage to the ganglion cells and their axons in the retina [76]. It can be managed by the administration of ocular pressure lowering drops and visual loss can be reduced by early diagnosis. Diagnosis is generally achieved by assessing the cup-to-disc ratio of the optic nerve head, which is easier to visualise in three dimensions. This can be visualised in Figure 2.2 which shows an optical coherence tomography image of retina highlighting the cup-to-disc ratio.

Figure 2.2 is a profile view of the optic nerve head, meaning moving from the bottom of the image towards the top is the same direction as moving from the retina to the pupil in the eye. In this view, which can be achieved by OCT, it is much easier to visualise the ratio between the cup and disc than from a fundus image. To diagnose glaucoma effectively, a 3D imaging technique is essential. This is thought to be a major area in which plenoptic imaging can benefit the field in the absence of OCT.

2.1.3 Age-related macular degeneration

Age-related macular degeneration (AMD) is the leading cause of blindness in the U.S. with it being responsible for 54% of all legally blind Americans [77]. As well as diminishing quality of life, AMD is estimated to cost the U.S. $30 billion annually by reducing the likelihood of employment [78]. The two main types of AMD are dry, which leads to the gradual loss of vision, and wet, which is the most visually threatening and is characterised by increased vascular permeability [71].

2.2 Current retinal imaging methods

To help understand the benefits that adding a plenoptic camera to a retinal imaging system, it is first important to understand the current state of the field. This section will give an overview of fundus imaging, a method which could potentially be improved by the addition of a plenoptic camera, and the current gold standard of retinal imaging, optical coherence tomography.

2.2.1 Fundus imaging

Fundus imaging is the collective name for the capturing of a two dimensional image of reflected intensities representing the three dimensional semi-transparent retinal tissues. Advances in fundus imaging have helped reveal new information about the structure of the eye such as hyperspectral imaging, and improved spatial resolution with adaptive optics attachments. Adaptive optics increases spatial resolution by measuring the distortion of the imaging wavefront, then with the addition of a deformable mirror, changes the illumination wavefront by modeling the aberrations created by the eye and produce a higher resolution output image. Cell structure cannot be seen through standard fundus imaging as the point spread function width from a normal eye is approximately the size of a photoreceptor. With the aberration corrections from adaptive optics, individual photoreceptors have been imaged *in vivo* [79]. Imaging of the retina poses many technical

challenges. The first is that the retina is not naturally illuminated, so external illumination has to be used. The second problem is because of the small size of the pupil, if the illumination and imaging paths overlap corneal and lenticular reflections will diminish image contrast.

2.2.2 Optical coherence tomography

Optical coherence tomography is a method by which depth is estimated by recording the time of flight of backscattered photons. It was first developed by Huang *et al.* [80] as a non-invasive method for gaining cross-sectional images of biological systems. Contrast between different layers can be seen as varying amounts of photons are backscattered depending on the refractive index of the layer, and depth is estimated by the distances further away taking longer to backscatter than objects closer. As the thickness of the retina is approximately 500μm, the time differences are so small the layers in the retina are measured via interferometry [80]. An image is created by light, usually near infrared or even infrared for retinal imaging due to the deeper penetration through tissue, passing through a beam splitter. One beam is then incident onto the retina and reflected back, the other off a reference mirror. The interference energy between these beams is then recorded by a sensor. A depth scan, know as an A-scan, is created from the interferogram intensities from each point. An image is then created by performing A-scans at all x and y co-ordinates. The first *in vivo* OCT images of a human retina were taken by Swanson *et al.* in 1993 [81], which achieved the highest depth resolution retinal images at the time.

2.3 Simulating retinal plenoptic imaging

The aim of this study is to see whether plenoptic imaging could to some extent replicate OCT and yield depth information about retinal tissues. As previously discussed, plenoptic imaging has a wealth of post processing techniques that can be performed on the raw images. These techniques include digital refocusing, change of viewpoint, depth map gen-

eration and three dimensional surface topography images. However, these have only been studied in free space models, and not by light that has transported through scattering media such as the eye. The aim of this study is to simulate the transport of light through the retina, gaining quantitative data on the angular distribution and intensity of the transmission and reflectance of photons through the different layers. These distributions will help us to understand the potential for plenoptic retinal imaging. One important aspect is the ability to determine, either quantitatively or relatively, the thickness of the neural retina for help with diagnosis of diseases such as DME, which is located between the retinal pigment epithelium (RPE) and the surface of the retina which is bound by ocular media. The neural retina is a moderately scattering layer which is 240μm thick and has a scattering co-efficient, $\mu_s = 5.735 \times 10^3 \text{cm}^{-1}$ at 800nm [82]. The thickening of this layer or changes to the shape of the retinal surface are important factors in the diagnosis of diseases such as glaucoma and diabetes. To study these phenomena, transmission of light through to the RPE and information about the reflectance from the neural retina has been explored. In order to perform the required simulations, a multi-layered Monte Carlo software package (MCML) was used [83]. This has been chosen due to the ease of simulating light transport through multiple layers, which is perfect for modelling the retina which itself contains many layers. This model is explained in more detail in Section 2.3.2. The model of the retina has been taken from Styles *et al* [82], and is explained further in the following section.

2.3.1 Computational eye model

In order for this experiment to provide useful results, it is imperative that an accurate model of the reflectance and scattering of the occular fundus is used. Over time, many different attempts have been made to create this model, with Berendshot *et al.* [84] providing a thorough summary of this work. In 1986, the first model was developed which incorporated the entirety of the ocular fundus by van Norren and Tiemeijer [85]. This model comprised of neutral reflectors and absorbing compounds, but it wasn't until it

was modified by Delori and Pflibsen [86] that scattering effects in the choroid were added through the use of Kubelka-Munk transport theory [87]. Using work done by Hammer *et al.* [88] on the optical properties of the choroid and the retina, Hammer and Schweitzer [89] included scattering in these tissues through the use of the adding-doubling method of radiation transport. Through the use of Monte Carlo simulations [90], Preece and Claridge [91] used the data from Hammer *et al.* [88] and properties about haemoglobin [92] and melanin [93] in the fundus to develop their model. This is the basis of the model used by Styles *et al.* [82] which is the model used throughout this chapter.

2.3.2 Light propagation model

The simulations performed throughout this chapter were done using MCML, a Monte Carlo method for modelling light transport through multi-layered tissues [83]. There are certain features of the program which make it highly applicable to the experiments devised throughout this chapter. The layered approach to the software relates well to the anatomy of the retina, as can be visualised well in Figure 2.4. Also, the ability to input the absorption and scattering coefficients, the refractive indices of different tissues and that these can be all be a function of the wavelength chosen, which is utilised throughout the results presented in this chapter, is highly beneficial. Figure 2.3 illustrates the geometry of the incident photon into the neural retina. This photon will then either transport through the tissue, scatter or be absorbed. The outcome of each individual photon is determined by the parameters of the tissue, and by the random sampling of variables from probability distributions [83]. This is the principle of Monte Carlo modelling, with [94, 95, 96] providing a comprehensive review of these methods. For a full description of the maths behind the MCML model, it is recommended to read the paper by Wang [83], however, there are some key points to highlight. Firstly, should there be a mismatch in the refractive indices at boundaries between tissues (as there is between the neural retina and the ocular fluid, and the neural retina and the RPE), then specular reflectance, R_{sp}, follows:

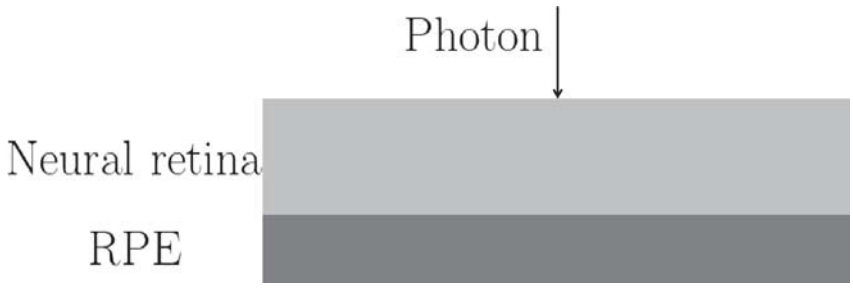


Figure 2.3: A diagram highlighting the input of a photon using the MCML Monte Carlo simulations

$$R_{sp} = \frac{(n_1 - n_2)^2}{(n_1 + n_2)^2} \tag{2.1}$$

Where n_1 and n_2 are the refractive indices of the two media. The photon weight entering the new medium is then computed as:

$$W = 1 - R_{sp} \tag{2.2}$$

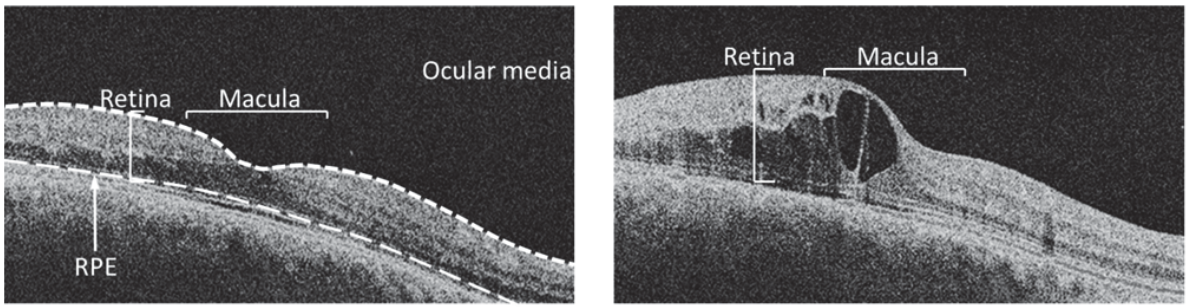
Upon scattering once inside a medium, the probability distribution for the cosine of the deflected angle, $\cos\theta$, is defined as [97]:

$$p(\cos\theta) = \frac{1 - g^2}{2(1 + g^2 - 2g\cos\theta)^{3/2}} \tag{2.3}$$

Where g is the anisotropy factor and has a value between -1 and 1.

2.3.3 Propagation of light to the RPE

Diabetic macular edema (DME) is the leading cause of blindness for patients with diabetes, and causes thickening of the retinal tissue. It is commonly visualised through OCT, by imaging the depth of the RPE with respect to the surface of the retina. Figure 2.4 shows an OCT of a patient with DME, and one with a healthy retina for reference. The increased distance between the retinal surface and the RPE can be seen clearly. As



(a) (b)

Figure 2.4: (a) OCT of a healthy retina. (b) OCT of a retina with thickening due to diabetic macular edema (DME)

plenoptic imaging can only recover depth in free space models, this can be thought of as recovering the depth of the last point of scattering when applied to turbid media. Therefore, if plenoptic imaging is going to be a useful tool in quantitatively measuring the distance from retinal surface to the RPE, then a proportion of photons must travel to the RPE ballistically.

The hypothesis for this experiment is that if a large enough proportion of the light which is transported through the retinal tissue to the RPE is ballistic, then algorithmically these particular rays of light can be selected and used to determine the thickness of the RPE using a plenoptic camera.

To investigate the feasibility of diagnosing DME via plenoptic retinal imaging, simulations were executed looking at the transmission of light to the RPE. Although MCML does not allow direct selection of only ballistic photons, photons that transmit to the RPE at the same angle and position as the inputted photon can be selected. It is possible that a photon that is scattered multiple times could emerge from the tissue with the same angle and position as a ballistic photon, however if small enough positional and angular bins are used this becomes less likely. As DME causes thickening of the retinal tissue, simulations were run for a variety of distances between the retinal surface and the RPE, with the results being shown in Figure 2.5.

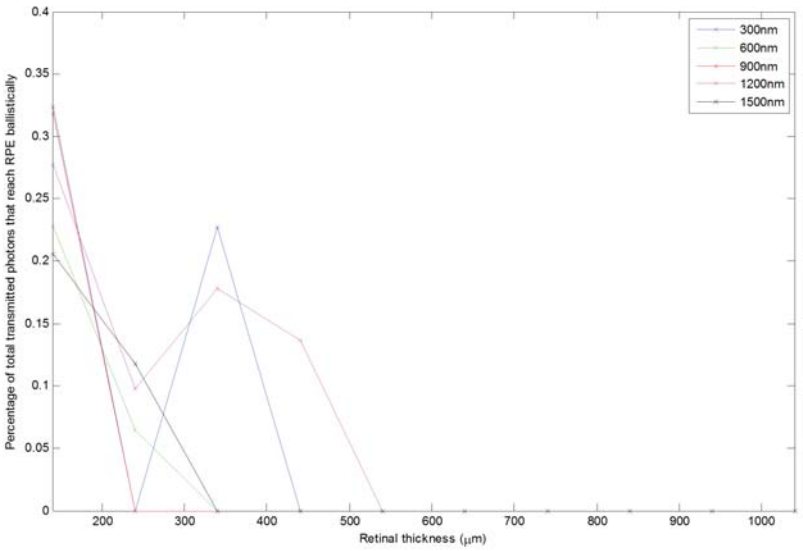


Figure 2.5: A graph indicating the percentage of photons that have propagated through retinal tissue to the RPE without scattering. The experiment was repeated for a variety of wavelengths and thickness's between retinal surface and RPE, and with 10^7 photons.

The results in Figure 2.5 show that at 300nm, 600nm, 900nm, 1200nm and 1500nm the maximum percentage of photons which reach the RPE with the same position and angular orientation of the ray is less than 0.35% of the initial incident intensity. This percentage would then also be expected to decrease dramatically as the photons then have to be reflected back off the surface of the RPE, and propagate without scattering through the retinal tissue again if they are to be ballistic. To get any contrast large enough for the camera to detect high levels of illumination would be required which may be in excess of safe levels. Wolffe [98] discusses the extent of damage to the retina from different imaging modalities. A quantitative analysis on the amount of light exposure which is safe is very difficult as it is different depending on the person being imaged. However, it is indicated that current methods could be close to the level in which damage could occur. Therefore, if we take 0.35% each way, this would equate to an increase in intensity of more than 800 times than currently used. This then provides a strong indication that plenoptic retinal imaging can not be used for determining quantitative data on retinal thickening via the use of ballistic photons. The original hypothesis is therefore indicated to be disproved.

2.3.4 Angular reflectance from retinal surface

By looking at the angular distribution from the retinal surface, a lot of information about the potential for plenoptic retinal imaging can be deduced. Although it has already been found that retinal thickening can not be visualised via the use of ballistic photons, if a relationship between the angular distribution and distance between the retinal surface and the RPE can be made, then plenoptic retinal imaging may be able to diagnose DME. If no relationship is found, this could provide strong evidence that surface topography could be measured by plenoptic imaging.

Determining retinal thickness by angular reflectance

The hypothesis for this experiment is that if the angular reflectance off the surface of the retina changes as a function of the thickness of the retinal tissue, then this distribution could be used to determine the retinal thickness.

To see if any major changes to the angular distribution occurs during retinal thickening, Monte Carlo simulations were performed with the thickness of the neural retinal layer selected as 60μm, 140μm, 240μm, 440μm and 1040μm. These were chosen to see a full range of thickness's from thinner than a healthy neural retinal layer to much thicker (for reference, healthy neural retinal layer has a thickness $\approx$ 240μm). The angular distributions can be seen in Figure 2.6.

From looking at this graph, the angular distributions do not appear to have any identifiable differences from each other, apart from at a thickness of 60μm, where an increased intensity is apparent at very small angles. However, in relation to DME a thickness of 60μm is not a very useful thickness as this particular disease causes retinal thickening, so it is thickness's larger than 240μm that are of particular interest. In this region there are no identifiable changes, which indicates that it will not be possible to deduce the thickness of the neural retina simply by knowing the reflected angular distribution. Therefore, this indicates that original hypothesis for this experiment would be false.

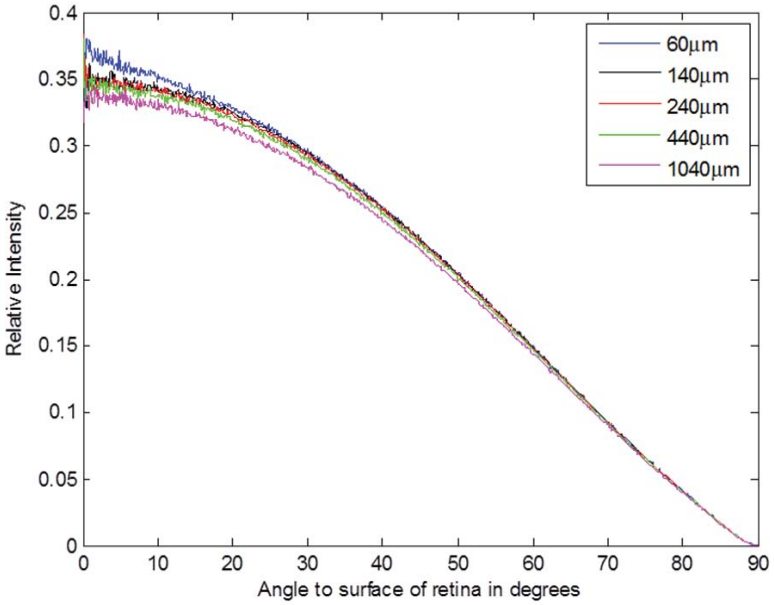


Figure 2.6: The angular distributions of light reflected from the surface of the retina at different thicknesses ranging from 60mm to 1040mm. This experiment was performed at 800nm and with 50,000,000 photons.

Surface topography

Topographic surface images have been proven to be possible in plenoptic imaging in free space systems. However, their response in scattering media systems such as the eye have not been studied. By looking at the angular distribution of photons reflecting from the surface of the retina, more information about a plenoptic cameras response to this system can be identified. One way in which topographic images of the retina would be beneficial is in the diagnosis of glaucoma. This is done by assessing the cup-to-disc ratio. However, this is difficult to quantify from standard fundus photography, and benefits greatly from imaging modalities that provide three dimensional data, as shown in Figure 2.2 from an OCT image of the optic nerve with the cup and disc illustrated.

In order to gain accurate topographic images of the retinal surface, the reflectance from the surface needs to be independent from any abnormalities or changes in the retinal layers below. By looking at Figure 2.6, in which the angular reflectance is presented with varying retinal thickness, the shape of each graph does not appear to change as

the thickness varies. This therefore shows that retinal thickening does not affect the angular distribution which will be recorded by the plenoptic camera and is independent of abnormalities under the retinal surface. This is a good indicator that surface topography images of the retina could be taken.

2.4 Conclusion

This chapter has focused on computational simulations of light propagation through the retina in an attempt to understand how a plenoptic camera could aid diagnosis of retinal diseases. It was first thought that if enough photons were transmitted through the neural retina to the retinal pigment epithelium without scattering, then these ballistic photons could be selected from the raw plenoptic image and be used to focus on the RPE. However, it was found that less than 0.35% of light did reach this layer without scattering, with an even lower percentage reflected and then propagate unscattered to the surface of the retina. It was therefore deemed that refocussing via ballistic photons would not be possible in the retina to view the RPE.

The reflected angular distribution was then studied. The shape of this distribution with the thickness of the neural retinal layer varying was examined, and did not appear to have identifiable differences. This then ruled out its use to determine the retinal thickness. However, as the angular reflectance is independent of abnormalities or changes below the surface, an accurate topographic image of the retina could be possible. It is important to note however that the simulations only included the retina and not the rest of the eye. Although surface topography of the retina would be possible if isolated, it is still unclear from these simulations whether angular data would be preserved after being transported through the ocular media, the lens of the eye and the cornea. Further work in the following chapters however will focus on practical experiments gathering accurate, quantitative depth data from a plenoptic camera with the goal of developing a retinal topography system.

Chapter 3

Plenoptic depth maps and pattern projection

Throughout Chapters 4, 5 and 6, analysis on improving the depth estimation from a Raytrix R11 focused plenoptic camera is conducted. For all of these experiments, the Raytrix camera is treated as a black-box system. This is due to the limited information which is provided about the internal configuration of the camera itself. Access to the RAW image data is also not available, and the RAW images are processed using the Raytrix software (RxLive 2.10). Therefore, predictions about the algorithms used to produce the virtual depth have to be made. It is, however, known that the Raytrix uses a hexagonal grid microlens array with three different focal length lenses interlaced into the array, which can achieve a much greater depth of field (dependent on the focal lengths in the microlens used) whilst only reducing the effective resolution ratio in each dimension to 1/2 [1].

Raytrix depth algorithm prediction

The standard algorithms used to render focused plenoptic images, along with those to change the perspective, digitally refocus and generate depth maps are described throughout Section 1.11. Later chapters focus mainly on the use of depth estimation from a

focused plenoptic camera, which can only be estimated in areas which have sufficient local contrast. The depth algorithms are thought to be calculated via triangulation [1]. For this to work, pixels of the same image point need to be incident in at least two microimages. Therefore, triangulation is only possible in areas with sufficient local contrast. This point is the premise of future chapters, which looks at adding contrast to plain scenes to allow for the triangulation algorithms to be successful.

3.1 Plenoptic depth maps

As future chapters focus on the use of a Raytrix plenoptic camera, it is first important to develop a photogrammetric review of plenoptic imaging to give rigour to future analysis. This section will focus on the formation of depth maps inside a plenoptic camera. It is also worth referring back to Section 1.6.1 to look at the image formation inside a plenoptic camera.

3.1.1 Reconstruction from two views

Depth from plenoptic imaging is a reconstruction of a scene from multiple images. In this section, how this reconstruction is formed will be explored mathematically, by using only two images and a scene made purely of points for mathematical simplicity. In this case, a set of correspondences $\mathrm{x}_i \leftrightarrow \mathrm{x}'_i$ in two images is considered. Taking two camera matrices, P and P′, and a set of three dimensional points, X_i, then there is the image correspondence such that $\mathrm{PX}_i = \mathrm{x}_i$ and $\mathrm{P}'\mathrm{X}_i = \mathrm{x}'_i$. It then follows that the two given data points are projected to by the point X_i. However, neither the point X_i nor the projection matrices P and P′, which represent cameras, are known and have to be determined.

In plenoptic imaging, it is often the case that depth is outputted in the form of virtual depth. Part of Chapters 4, 5 and 6 focus on converting this virtual depth into actual depth in real space. Virtual depth is output as it is impossible to determine positions of points uniquely without knowing anything about the calibration of the two cameras [99]. This is

a general ambiguity that is independent of the number of images given, as it is impossible to know the absolute orientation, scale or position of the imaged object. This ambiguity, in the absence of camera calibration, is expressed as a projective transformation. There is this ambiguity as a projective transformation, H, can be applied to each point X_i and to the right of the camera matrix, P_j, without the projected image points being changed, such that:

$$\mathrm{P}_j\mathrm{X}_i = (\mathrm{P}_j\mathrm{H}^{-1})(\mathrm{H}\mathrm{X}_i) \tag{3.1}$$

In these circumstances, the choice of H is arbitrary as there is no reason to pick one set of camera matrices and points over another. This reconstruction, therefore, has projective ambiguity, or is sometimes known as projective reconstruction. However, for this reconstruction, there must be at least seven points which must not lie in one of the various critical configurations [99].

The tool which is used to compute the points from the two images is through the use of the fundamental matrix, which can be thought of as the constraint on image points x and x′ if they are images of the same point in 3D space. This is true because of the coplanarity of the centres of the cameras of both views, the image points themselves and the space point. Given a fundamental matrix, F, and pair of matching points $\mathrm{x}_i \leftrightarrow \mathrm{x}'_i$ it must follow the below where F is a 3×3 matrix of rank 2 [99]:

$$\mathrm{x}'^{\mathrm{T}}_i \mathrm{F}\mathrm{x}_i = 0 \tag{3.2}$$

If F is unknown, it can be computed from a set of point correspondences as the equations are linear in the entries into matrix F. This fundamental matrix F can be uniquely determined from a pair of camera matrices, P and P′. Alternatively, the fundamental matrix, F, can determine the pair of camera matrices, P and P′, up to a 3D projective ambiguity. Therefore, the fundamental matrix, F, holds the full projective geometry of the

camera matrices, P and P′, and critically is not changed by 3D projective transformation. One method for reconstructing the imaged scene using the fundamental matrix, F, is as follows:

1. Using several correspondences between points such that $\mathrm{x}_i \leftrightarrow \mathrm{x}'_i$ from two views, then linear equations are formed in the entries of F from the coplanarity equation ${\mathrm{x}'_i}^{\mathrm{T}}\mathrm{F}\mathrm{x}_i = 0$
2. Compute the linear equations to find a solution to F
3. Compute a pair of camera matrices, P and P′, from F
4. Using the pair of camera matrices, P and P′, and the corresponding image point pairs, $\mathrm{x}_i \leftrightarrow \mathrm{x}'_i$, find the point X_i in three dimensional space.

The method above for solving a point in 3D space, X, is known as triangulation. As a plenoptic raw image can be thought of as many two view representation, depth can be computed by applying this method to neighbouring microimages.

3.1.2 Theoretically achievable depth resolution

As previously discussed, depth can be calculated from a plenoptic camera. Work performed by Perwass *et al.* [1] calculated the theoretical depth resolution a plenoptic camera can achieve. For an understanding of the derivation of these limits, Figure 3.1 shows the geometric parameters used. Take an image point at distance x from the centre of a microlens, this intersects the line between two identical microlenses at point *z(x)*. If Δx is taken as the width of a single pixel, the line at $x - \Delta x$ intersects the line as before at the point $z(x - \Delta x)$. The depth resolution for a pixel at a distance of x from the centre of a microlens is given by $z(x) - z(x - \Delta x)$. The relation between z and x is described by Perwass as [1]:

$$\frac{z}{\kappa D/2} = \frac{B}{x} \tag{3.3}$$

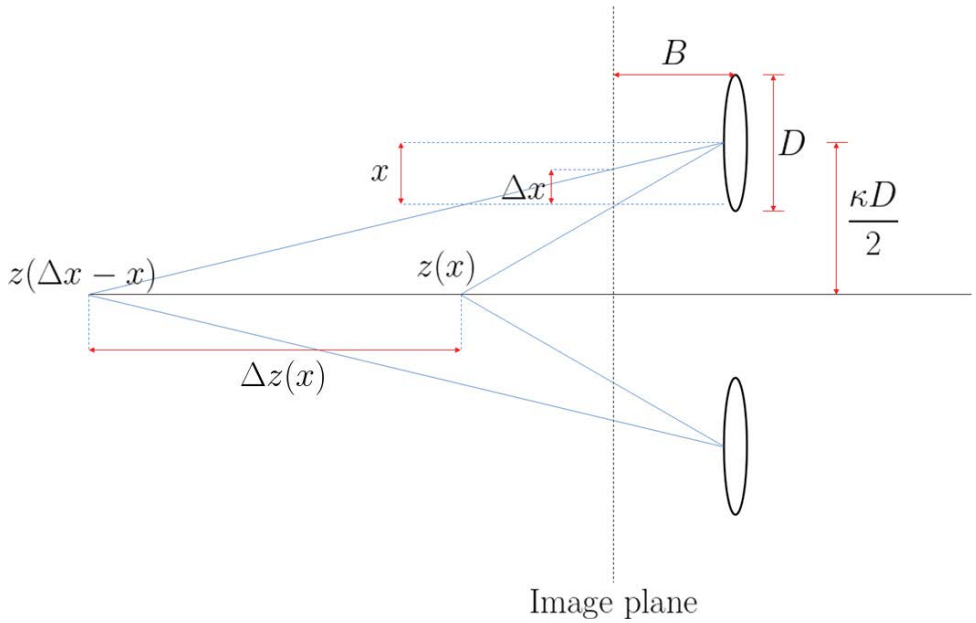


Figure 3.1: Diagram visualising the relationship between depth resolution and pixel resolution in a microimage. Image adapted from [1]

As the virtual depth of a point increases, in a plenoptic camera that point becomes visible in multiple microimages. It therefore follows that in a plenoptic system, any of the neighbouring microimages upon which the point is repeated can be used as a basis for triangulation, and this does not necessarily have to be the closest neighbouring microimage.

In Figure 3.1, κ denotes the distance between the centres of the microlenses used for the triangulation. In this geometry, z is at a minimum if the pixel at distance x from the microlens centre is at the border of that microlens. Under these circumstances, $\kappa = z/B = v$, where v is the virtual depth, i.e. κ denotes the minimum virtual depth that a plenoptic system can perform a triangulation to calculate the depth of an object.

3.1.3 Calculation of a depth map

This section will focus on the theoretical understanding of how a depth map is calculated in a focused plenoptic camera, and the accuracy of this depth calculation from work first put forward by Zeller *et al.* [100]. By looking at Figure 3.2 it is possible to show how the

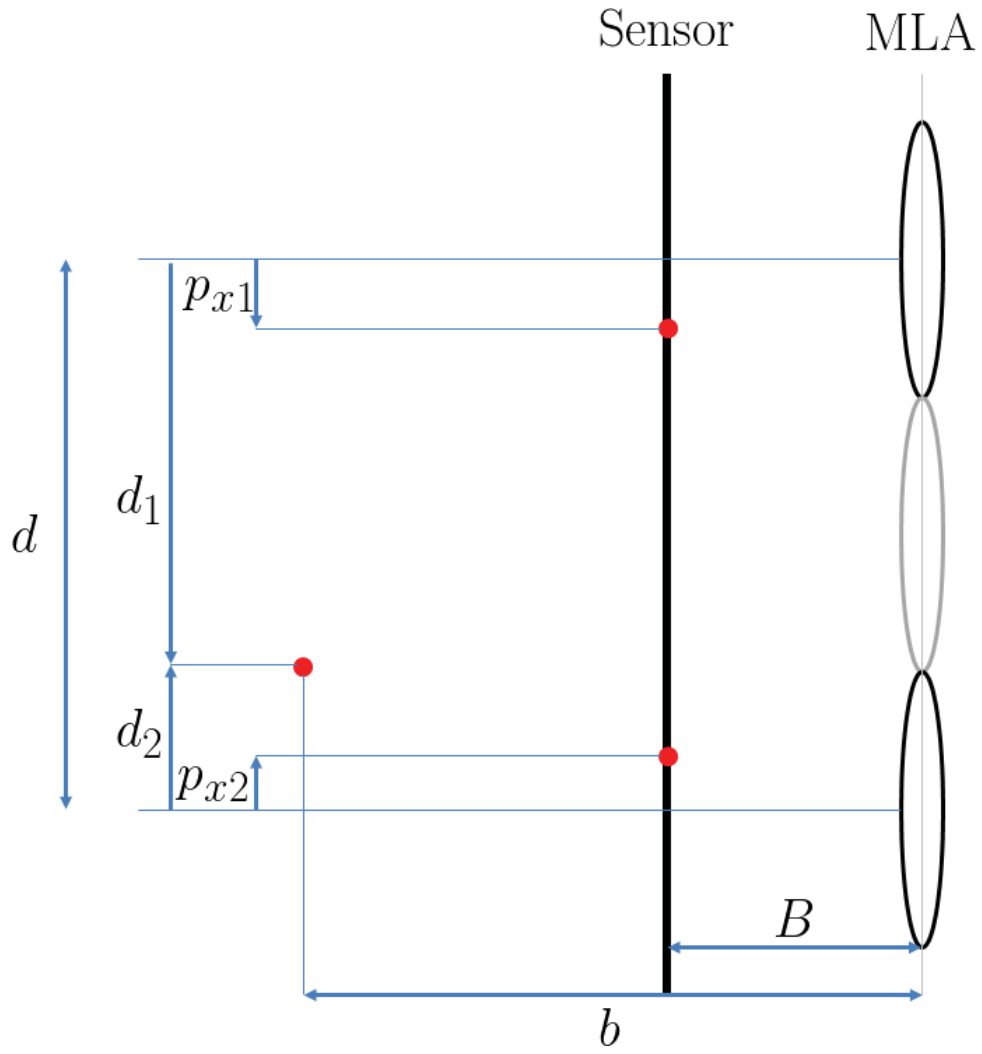


Figure 3.2: Diagram showing the principle of depth estimation in a focused plenoptic camera

distance between the sensor and a virtual image point (b in Figure 3.2) can be calculated from the points projected onto the sensor from two microimages. From Figure 3.2 it is also possible to show that the following holds true.

$$\frac{p_{xi}}{B} = \frac{d_i}{b}, \quad p_{xi} = \frac{d_i \cdot B}{b} \quad \text{for} \quad i \in 1, 2 \tag{3.4}$$

If the parallax of the virtual image point is defined as p_x, using Equation 3.4 it can be shown that.

$$p_x = p_{x2} - p_{x1} = \frac{(d_2 - d_1) \cdot B}{b} = \frac{d \cdot B}{b} \tag{3.5}$$

which can then be rearranged to be in terms of b, the distance between the virtual image point and the microlens array, and is a function of the distance between similar microlenses in the array, d, the distance between the sensor and the microlens array, B, and the previously defined parallax, p_x with the following equation.

$$b = \frac{d \cdot B}{p_x} \tag{3.6}$$

If the triangulation used in the focused plenoptic camera is performed using two neighbouring microlenses, then the distance d from Figure 3.2 is equivalent to the aperture of a single microlens. However, in a Raytrix it is known that a microlens array with varying focal lengths is used. Therefore, no two neighbouring microlenses can ever be focused on the same virtual image point, so d will always be greater than the microlens aperture.

In a Raytrix camera, which is used in later chapters, the actual distance between the sensor and microlens array (B in Figure 3.2) is an unknown value. Therefore, the depth which is outputted from the focused plenoptic camera is not a depth in real space, but instead a relative depth value, v, known as the virtual depth. This can be derived as:

$$v = \frac{b}{B} = \frac{d}{p_x} \tag{3.7}$$

This virtual depth, as previously discussed, can only be calculated if the virtual image point is in focus on at least two microimages. In a Raytrix camera, the microlens array which is used is multi-focal with three different lenses, arranged in a hexagonal array. As discussed by Perwass *et al.* this produces a minimum virtual depth value, $v_{min} = 2$ [1].

Plenoptic image synthesis

Every virtual image point, $p_n = (x_n, y_n)$, which is inside the depth of field of a plenoptic camera must be focussed in a minimum of one microimage. Therefore, a fully focussed image, $I(p_n)$, can be calculated from the virtual depth, $v(p_n)$. This is synthesised by calculating the average of each corresponding focused points in the microimage [101]. For a more in depth analysis on this please refer to the paper by Perwass *et al.* [1].

Accuracy of virtual depth

From the theory of propagation of uncertainty, it is clear that there will be some error in the depth accuracy due to the estimation of the parallax described previously. This section of theory helps to quantify this error which is inherent in a focused plenoptic imaging system such as the Raytrix. By using the following approximation, the standard deviation for the virtual depth σ_v can be approximated as [100]:

$$\sigma_v \approx \left| \frac{\partial v}{\partial p_x} \right| \cdot \sigma_{p_x} = \frac{d}{p_x^2} \cdot \sigma_{p_x} = \frac{v^2}{d} \cdot \sigma_{p_x} \tag{3.8}$$

By looking at Equation 3.8 it clearly shows that the accuracy of the virtual depth, σ_v decays proportionally to v^2. The distance d from Figure 3.2 isn't a continuous function with respect to v, as the different microlenses which see the image point changes depending on depth. This then follows that the dependency of the accuracy of calculated depth as a function of the object distance, a_L, is also discontinuous. However, d is proportional to v on average, so overall the accuracy of the virtual depth decays approximately proportional to the virtual depth [100].

A linear function, as shown in Equation 3.9, can be used to show the relationship between the image distance, b_L, with virtual depth, v, where b_{L0} is the unknown distance between the microlens array and the main lens of the camera. The object distance, a_L, can then be expressed as a function of virtual depth, v, by using the thin lens equation $(1/f = 1/a + 1/b)$.

$$b_L = b + b_{L0} = v \cdot B + b_{L0} \tag{3.9}$$

The standard deviation of the object distance, σ_{a_L}, can then be approximated by taking the derivative of the object distance with respect to the image distance [100] as follows, where f_L is the focal length of the main lens of the camera.

$$\sigma_{a_L} \approx \left|\frac{\partial a_L}{\partial b_L}\right| \cdot \sigma_{b_L} = \frac{f_L^2}{(b_L - f_L)^2} \cdot \sigma_{b_L} \tag{3.10}$$

$$= \frac{(a_L - f_L)^2}{f_L^2} \cdot \sigma_{b_L} = \frac{(a_L - f_L)^2}{f_L^2} \cdot B \cdot \sigma_v \tag{3.11}$$

When the object distance is much greater than the main lens focal length, Equation 3.11 can be further simplified to Equation 3.12, which shows how in these circumstances and for a constant object distance, a_L, the accuracy of the depth is proportional to f_L^2.

$$\sigma_{a_L} = \frac{(a_L - f_L)^2}{f_L^2} \cdot B \cdot \sigma_v \approx \frac{a_L^2}{f_L^2} \cdot B \cdot \sigma_v \quad \text{for} \quad a_L \gg f_L \tag{3.12}$$

3.1.4 Calibration of a depth map

Throughout Chapters 4, 5 and 6 depth maps from a Raytrix R11 camera are used for investigations into how texture can be added to featureless scenes to artificially add reference points which as previously described are necessary for the calculation of a depth map. Part of this process involves converting between the previously described virtual depth, v, that is output from a Raytrix camera to a physical object distance. This topic has been researched extensively by Zeller *et al.* [100] using three different methods to convert virtual depth to actual object distances. The focus of the following chapters is not on improving this calibration but on gaining a more accurate depth value in scenes with minimal features as this is the case with fundus imaging. Therefore, it is not claimed that the method put forward in this thesis produces results to a higher level of accuracy

than those by Zeller [100], but merely that a variation on the curve fitting model is used to enable the investigation to progress further. For completeness, below is a derivation of the problem trying to be solved in converting virtual depth, v, into a physical object distance.

As shown in Equation 3.9, the image distance, b_L, has a linear dependency on the virtual depth, v. Also, as the position of the main lens cannot be measured explicitly, the object distance, a_L, has to be estimated from the measurable object distance, o, and an unknown offset, a_{L0}, by the following equation.

$$a_L = o + a_{L0} \tag{3.13}$$

By then placing Equations 3.9 and 3.13 into the thin lens equation, this can then be rearranged to show that the object distance, o, can be described as a function of virtual depth, v, by the following equation.

$$o = \left(\frac{1}{f_L} - \frac{1}{v \cdot B + b_{L0}} \right)^{-1} - a_{L0} \tag{3.14}$$

Equation 3.14 shows that the object distance, o, is related to the virtual depth, v, by four unknown but constant variables. These are the main lens focal length, f_L, the distance between the sensor and the microlens array, B, the distance between the main lens and the microlens array, b_{L0}, and the unknown offset, a_{L0}. Due to the number of unknown variables, the function in Equation 3.14 has to be estimated. Zeller *et al.* [100] present three methods to solve this function, but these will not be discussed in more detail here.

3.2 Pattern projection

The physical world in which we live is in three dimensions, however, traditional cameras record a two dimensional representation of this world. Different methods have been

devised to overcome this limitation, such as plenoptic imaging which has been described in depth in Chapter 1. Another method for gaining 3D surface imaging is through the use of structured illumination. The principal of this method is that an image sensor is used to collect a 2D image of a surface illuminated by the pattern. If the surface being imaged does not have a 3D variation, then the pattern in the image will be the same as the projected pattern. However, should the scene have some 3D variations, this will distort the projected pattern. Three dimensional information about the surface of the scene can then be computed using various algorithms and principles about structured light [102].

There are currently numerous techniques available to achieve surface imaging by using structured light illumination. These techniques can be categorised into either sequential, also known as multi-shot, or single-shot methodologies. In general, sequential techniques will yield more accurate and reliable results if the object being imaged is static. However, should the object be moving, single-shot techniques have distinct advantages over sequential techniques and will tend to bring about the more accurate results. It is these single-shot techniques which bare the most resemblance to plenoptic imaging, which achieves 3D information from a single image acquisition. The single-shot techniques can be categorised into three different groups: continuously varying structured light patterns, one dimensional and two dimensional encoding schemes [102]. It is possible to combine these different techniques together depending on the application and benefits which are required.

3.2.1 Sequential techniques

As previously mentioned, sequential or multi-shot techniques bring about the best results when imaging a static object. The different types of sequential projections can be categorised into binary code, grey code and phase shifted. It is also possible to combine both the grey code and the phase shift.

Binary code

This method uses black and white stripes to create a sequence of projection patterns [103, 104]. The advantages of this method are that it is very reliable and as only binary values are used, the technique is less sensitive than other methods to surface characteristics. The main disadvantage is that many patterns need to be projected sequentially, increasing the acquisition time. This means that the scene must be static throughout imaging, and the acquisition time may be too long for some 3D imaging applications.

Grey code

One way to reduce the number of patterns required to compute a 3D image is to change from a binary code to a grey pattern. For example, if 64 stripes are needed with a binary code, only 6 grey patterns will be needed [102].

Phase shift

This method is achieved by a set of three sinusoidal patterns being projected onto an object. The intensities of every pixel in (x, y) of the projected sinusoidal patterns are

$$
\begin{aligned}
I_1(x,y) &= I_0(x,y) + I_{mod}(x,y)\cos(\phi(x,y) - \theta),\\
I_2(x,y) &= I_0(x,y) + I_{mod}(x,y)\cos(\phi(x,y)),\\
I_3(x,y) &= I_0(x,y) + I_{mod}(x,y)\cos(\phi(x,y) + \theta),
\end{aligned}
$$

where $I_1(x,y)$, $I_2(x,y)$ and $I_3(x,y)$ are the intensities of the three fringes, $I_0(x,y)$ is the background, $I_{mod}(x,y)$ is the modulation signal amplitude, $\phi(x,y)$ is the phase and θ is the phase shifted angle [102].

To retrieve the phase information, phase unwrapping needs to be performed as per the equation below.

$$\phi' = \arctan\left[\sqrt{3}\frac{I_1(x,y) - I_3(x,y)}{2 \cdot I_2(x,y) - I_1(x,y) - I_3(x,y)}\right] \tag{3.15}$$

There is an inherent discontinuity to this function, which can be removed by the following equation.

$$\phi(x,y) = \phi'(x,y) + 2k\pi \tag{3.16}$$

where k is an integer. The 3D surface information about the object can then be calculated from the difference between the phase measured at each point compared to the phase from a reference plane [105]. It is worth noting that the unwrapping does not solve the absolute phase, but merely a relative unwrapping.

Combining grey code and phase shift

As previously discussed, the phase shift method has a major problem in that it only calculates a relative unwrapping and cannot calculate the absolute phase. If two surfaces are separated by more than 2π, then these two surfaces will not be unwrapped correctly, causing a problem which is known as ambiguity. This can be solved by combining the phase shift method with the previously described grey code. This is because the grey code determines the absolute range of the phase and the phase shift can then achieve sub-pixel resolution [106, 107].

3.2.2 Single-shot techniques

As previously discussed, the drawbacks of sequential imaging are when the object being imaged is not static. This is particularly an issue when applied to the biomedical sciences, as it can be difficult to keep living samples still. It may then be more applicable to use a single-shot illumination technique, for which some examples are described.

Continuously varying pattern

Examples of a continuously varying pattern to gain single-shot 3D information is through the use of a Rainbow 3D camera [108]. This technique illuminates a scene with a spatially varying wavelength projection. By knowing a baseline value and a viewing angle of the camera, a 3D surface can be computed through simple triangulation algorithms from a single image. Another example is through the use of a continuously varying colour coded illumination [109]. In this method, spatial location information is encoded by the varying illumination. By then knowing the ratio of contributions from each colour channel, the decoding is easily derived and 3D information from a single shot can be achieved.

Stripe indexing

To achieve robust results with striped illumination, indexing is required. This is because the ordering in which the stripes are projected is not necessarily going to be the same as the observed pattern, due to parallax in the triangulation method, or stripes being missed because the surface being imaged occludes particular parts of the projection. There are various different ways to index the pattern, one of which is by using colours [110, 111]. Another method is called segment pattern, which uses a unique pattern in each stripe to distinguish one from another [112]. This method though is only useful for surfaces which are smooth and continuous, and does not provide robust results from discontinuous surfaces. Another way to index is by using grey patterns, where by using different intensities of grey in unique patterns each area can be distinguished from another [113]. A fourth way to index the pattern would be through a De Bruijn sequence [114], which can be used to construct a striped pattern that has enough local variations to make it unique [115, 116, 117].

Grid indexing

The concept of grid indexing is that any subwindow of the projected 2D pattern should be unique and identifiable to its 2D position in the overall pattern. This method has

the greatest resemblance to plenoptic imaging, as the subwindows can be though of as the microimages under each microlens in a raw plenoptic image. For this reason, in subsequent chapters where projected patterns are used, a grid pattern is chosen over a striped one.

Just as there was with stripe indexing, there are many different methods to index in a grid. The first is through a pseudo-random binary array, where a patterns of dots, or any other shape, are placed onto a grid such that the pattern in any subwindow is unique [118, 119, 120]. Another method for indexing is by colour coding vertical and horizontal stripes so that a colour coded 2D grid is created [121, 122, 123]. A 2D grid of colour coded dots which were not randomly generated has also been developed by an algorithm which still preserves the uniqueness of each subwindow [124].

Accuracy limitations of using a projected pattern

Although it will not be reviewed in depth in this thesis, it is worth noting that there are some accuracy limitation involved in digital projection for optical metrology systems. Baker *et al.* review the theoretical limitations of projections, which is then backed up by practical experimentation [125]. The model describes how the accuracy is a function of the distance between the projector and object, the gradient of the object and both the angle and resolution of the projector.

3.2.3 Pattern projection in plenoptic imaging

Projecting patterns in the field of plenoptic imaging is a relatively unresearched area. This is largely due to the fact that the design of a plenoptic camera enables 3D surface imaging from a single image acquisition. Cai *et al.* [126] have however combined the fields of structured illumination and plenoptic imaging for the application of gaining 3D high dynamic range images. Their work used a Lytro plenoptic camera with a sinusoidal striped pattern projected onto the object. The phase was then unwrapped as described

in Section 3.2.1 to help gain higher resolution 3D images of objects with differing reflectiveness.

In Chapters 4, 5 and 6 patterns are projected onto scenes which are imaged with a Raytrix R11 plenoptic camera. The patterns are not used to calculate the depth, but they are used add texture to featureless scenes so that the depth map algorithms described earlier in the chapter can function. This is because they require identifiable texture as the algorithms are based on triangulation.

3.3 Metrology review

Metrology, as defined by the Bureau International des Poids et Mesures (BIPM), is the science of measurement [127]. Looking back at the results from Chapter 2, for the application of retinal imaging the results indicate that efforts should be focused on surface metrology as opposed to recovering layered information through the use of ballistic photons when imaging with a plenoptic camera. This section will therefore present current methods of surface metrology, which relate to future chapters and then refer back to place in the context of plenoptic imaging.

Surface metrology can be split into two main categories; contact and optical. Contact methods include the use of a stylus profilometer, in which a calliper is used to contact both a test surface and a reference plane, with the stylus on the test plane having sufficient dimensions so that the details of the surface can be recovered [128, 129, 130]. Another stylus method which can be used to measure the surface is through the use of a scanning tunnelling microscope or atomic force microscope [131, 132]. A great advantage of using these methods is that the resolution is limited to the geometric size of the stylus rather than the much larger wavelength of light through optical methods. However, these surface metrology instruments are not applicable to retinal imaging, so will not be explained further.

Although optical measurement instruments typically cannot gain the resolution found in atomic force microscopy, there are some advantages which imaging with light has over the contact methods previously described. As optical imaging is non-contact, samples which would be destroyed by stylus methods can be imaged such as for *in vivo* analysis of the retina. The speed in which measurements can be recorded is typically much faster than contact methods, and can be specifically designed for 3D imaging as is the case for plenoptic imaging. To measure the 3D surface profile of an object, patterns can be projected onto scenes as described in Section 3.2. These methods highlight one of the major advantages of plenoptic imaging in that 3D data can be collected from a single image acquisition. Many different methods have been devised for gaining 3D information, such as confocal microscopy [133] or the gold standard in retinal imaging, optical coherence tomography [80].

Another method to gain surface profile is through the use of optical profilers. This in essence uses a beam of light as a stylus tip similar to the profilometer described previously, with the most common methods using either the focus detection method or triangulation [134]. Some of the first work in this field was done by Simon [135] who used a normal microscope, but split the illumination at the eyepiece and recorded with two photo detectors. By separating these two detectors by the same distance, but one before and one after the image plane, sub-nanometre vertical resolution was gained. A different type of optical profiler uses light sectioning [136]. This method uses the movement of a laser spot on a surface which reflects onto a sensor which can be used to measure the topography of a surface. There are many other optical imaging techniques, from microscopes to interferometers, which depending on the object being measured. However, as the methods described are for surface topography, these relate to the future chapters which focus on gaining the surface topography of planar and spherical models.

3.4 Experimental set-up

Throughout Chapters 4, 5 and 6 images are taken with a Raytrix R11 plenoptic camera of a flat plane and a spherical object. The scene is illuminated with a Optoma PK201 Pico LED Projector only. For the images in which a projector is not used, a diffuse white light source located behind the Raytrix R11 camera is used to illuminate the scene.

3.4.1 Limitations

The results presented in Chapters 4, 5 and 6 were captured in limiting circumstances, which should be taken into account when evaluating the results. The major limiting factor to the results is the amount of time available with the Raytrix R11 camera. All the results presented which involve Raytrix images and depth maps were captured in a 2 week period, as this was the only time with which the author had access to the equipment. During this time, many problems occurred including the melting of the motherboard on the only machine with the Raytrix software installed due to a malfunctioning fan and long waiting times for Raytrix to reply when difficulties arose in the initial camera calibration (see below). This severely limited the amount of data which could be captured, so many of the results presented throughout later chapters were not able to be repeated as was desired when designing the experiments. The results presented should therefore only be taken as a proof of concept and method development, which can be used as a base for future research to build upon.

Due to these limitations, the repeatability, reproducibility and stability of the results are largely unknown. In attempts to reduce the errors caused from the initial calibration described in the following section, all sets of data presented were done with the same initial calibration. The camera and projector were turned on first then left to stabilise, with the temperature of the room kept constant to minimise the effect of environmental noise.

Inherent noise analysis

Although they have not been investigated in detail throughout this thesis, it is worth noting the different forms of noise which could be present in the imaging described throughout. Firstly, fabrication errors in the sensor could cause small variations in quantum efficiency, so if uniformly illuminated would cause local non-uniformity in the charge collected on the sensor [137], which is often called the fixed pattern noise [138]. Shot noise is a product of the quantum state of light, and causes an uncertainty in the number of electrons stored in a collection site and is a fundamental limitation which cannot be removed [138]. Dark current is generated by free electrons due to thermal energy in the sensor. These can be reduced by cooling the sensor, or calibrated against by doing a process of dark current noise subtraction. In this method, a black image is taken along with the image of the scene. The black image will show some noise due to the dark current, which can then be subtracted from the image of a scene.

The noise can be analysed by looking at how scene radiance relates to measured pixel values. Only a brief outline will be presented here, for further reading on the subject please refer to Hamey or Forsyth [139, 140]. The process can be split into two main parts; before the light hits the image plane and after. The first process is a linear mapping from the scene radiance through the lens to the image irradiance. The second part has non-linear mapping from the image irradiance, through the camera electronics to the measured pixel values. The image irradiance and the measured pixel values are related by a general response function, with many shown in the database presented by Grossberg and Nayar [141]. Although not analysed in detail in this thesis, these sources of noise will be present in the Raytrix and hence the results through the following three chapters.

Raytrix camera calibration

In order to use a Raytrix R11 camera, it first has to be calibrated. This is performed using a predefined method inside the cameras information booklet, and involves illuminating the camera with a diffuse white light and aligning the images of the microlenses so that it matches those in the software. This often involves warping the shape of the microlens array on the software, especially towards the outer edges of the array. As this calibration relies on human judgement of alignment, it does incur an unknown amount of error in the results outputted in the depth maps created from the camera. Through experience, slight alterations to this initial calibration can make substantial differences to the calculated depth. This often causes large, unquantifiable errors in the depth calculations from the Raytrix.

Chapter 4

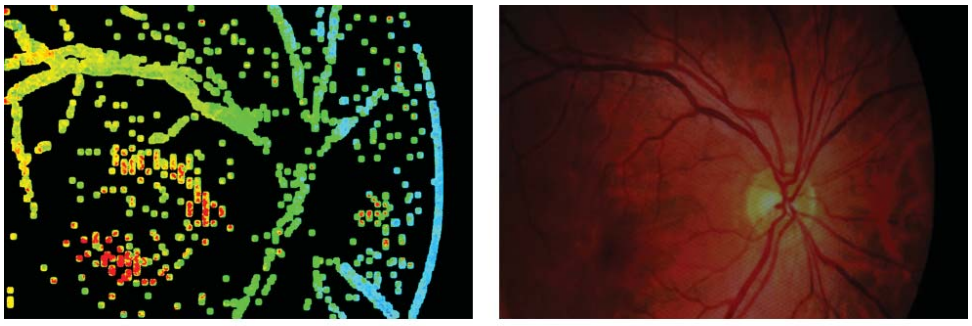
Improving depth estimation

Although a plenoptic image can provide many new features that are not possible in a conventional camera, it is the recovery of quantitative depth information from a scene that is most exciting for the field of biomedical imaging. In retinal imaging, information on the depth of various layers and the surface of the retina can be crucial in the early diagnosis of diseases such as glaucoma and diabetic macular edema (see Chapter 2). An accurate surface measurement is important in many optical tomography methods, such as bioluminescence tomography, in order to accurately reconstruct the light source [142]. Light field microscopy also utilises the depth measurements from the plenoptic system in order to create three dimensional volume renderings, such as Levoy *et al* [38] who created a 3D volume of a silkworm mouth.

As discussed previously in Section 1.12.2, Raytrix produce commercial plenoptic cameras. Throughout this chapter, images were taken by a Raytrix R11 camera. The R11 is a high-end Raytrix camera with a 43.3 mm large format colour interline 10.7 MP CCD sensor. Attached is an f/8 lens with variable zoom. Rendered images can reach up to 3 MP with up to 200 different layers from a single image. All of the depth maps generated in this chapter are produced using the Raytrix software (RxLive 2.10).

Whilst first imaging different scenes with a Raytrix R11 plenoptic camera, the qualitative accuracy of the depth maps appeared under simple visual inspection to change dramat-

ically dependent on the objects being imaged. To try to understand why the results were so inaccurate for some scenes, it was important to look into the depth estimation algorithms (see Section 1.11.3). Some of the features of plenoptic imaging, such as digital refocusing and changing perspective, require selecting predetermined pixels from the raw image. However, depth is generated by looking at local variations of features in neighbouring microimages (see Section 1.11.3). It was therefore hypothesised that accurate reconstruction of depth could not be produced from scenes with limited or no features, such as a retina. To test this, a picture of a retina was projected onto a flat sheet at an angle to the sensor. Should the depth be recorded accurately, the depth map should show a gradual change due to reflect the changes in depth. It was thought, however, that only features in the image would be registered and show a valid depth. Figure 4.1 shows the original image and depth map taken with a Raytrix R11 camera of a projected retina onto an angled plane. It is worth noting that this imaging, and all subsequent images, are performed in the visible spectrum. This was chosen by default because of the Raytrix camera which was available. The retina is often imaged in the near-infrared region as this will get greater depth penetration.



(a) (b)

Figure 4.1: a) The depth map produced when an image of a retina is projected onto an angled sheet and imaged with a Raytrix R11 camera. b) The original image of retina.

From looking at Figure 4.1, as expected, only the features of the retina register a depth. As the primary aim of this project is to gain depth of the entire retina form a plenoptic image, this is clearly not sufficient for purpose. Thoughts then went into ways this could

be rectified, and a preliminary set of experiments were devised to see if depth could be recovered from a scene when both a featureless object and the same object but with an artificial pattern placed in front of both a plain and patterned background.

Recovering depth from a featureless scene with a plain background

In order to test if features are needed in a scene to accurately reconstruct depth, a table tennis ball was placed in front of a black background, and imaged by a Raytrix R11 camera, as shown in Figure 4.2. The scene was illuminated with a light source placed next to the Raytrix R11 camera. It is assumed for this experiment that the light source is a diffuse white light source. A table tennis ball was chosen as it has a plain, featureless texture and has three dimensional shape that should be qualitatively recognisable from the depth maps and line plots generated from the depth data. The rendered image and the corresponding depth map can then be seen in Figures 4.3 and 4.4 respectively.

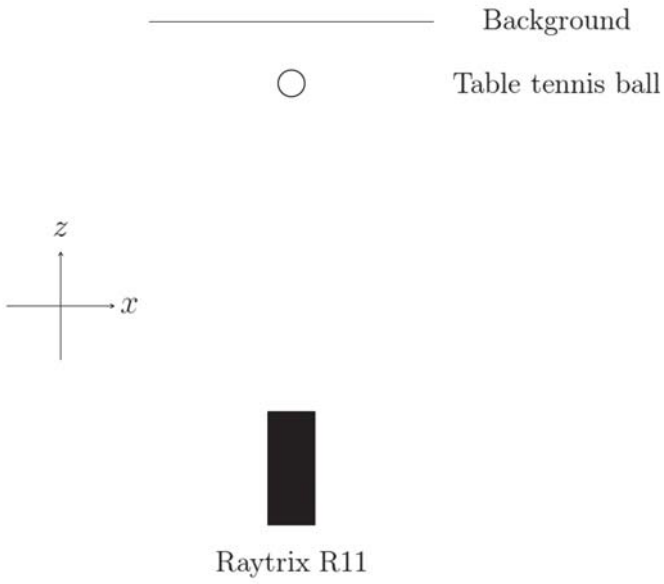


Figure 4.2: An overhead view of the experimental set-up for the preliminary experiments. The apparatus consists of a background, a table tennis ball and a Raytrix R11 plenoptic camera.

As predicted from the initial hypothesis, qualitatively no depth information is present in the shape of the ball as it appears to all be the same colour on the depth map. To more

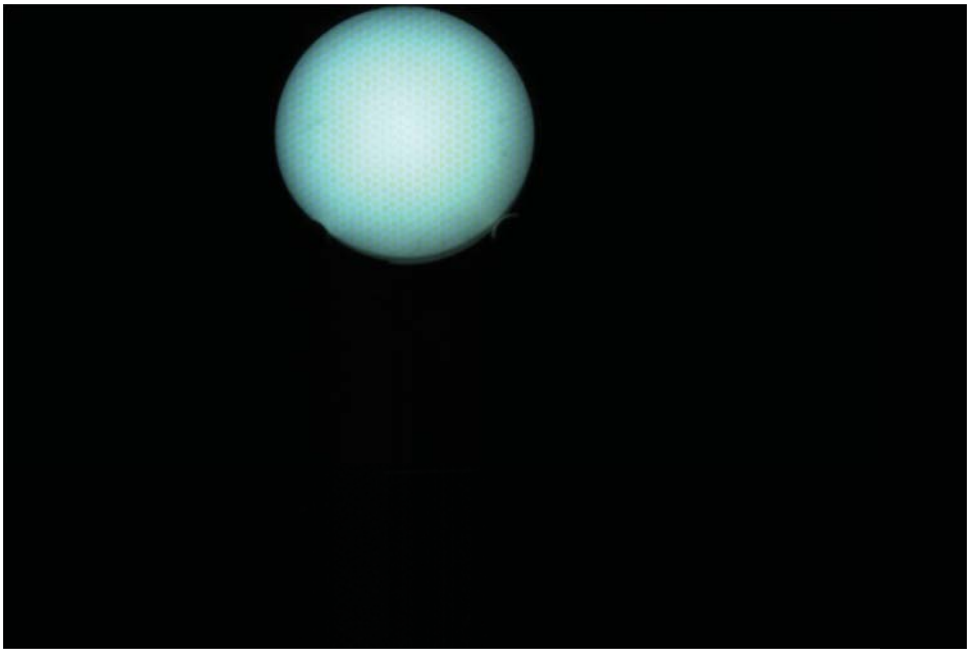


Figure 4.3: Rendered image of a table tennis ball when captured by a Raytrix R11 plenoptic camera in front of a featureless black background.

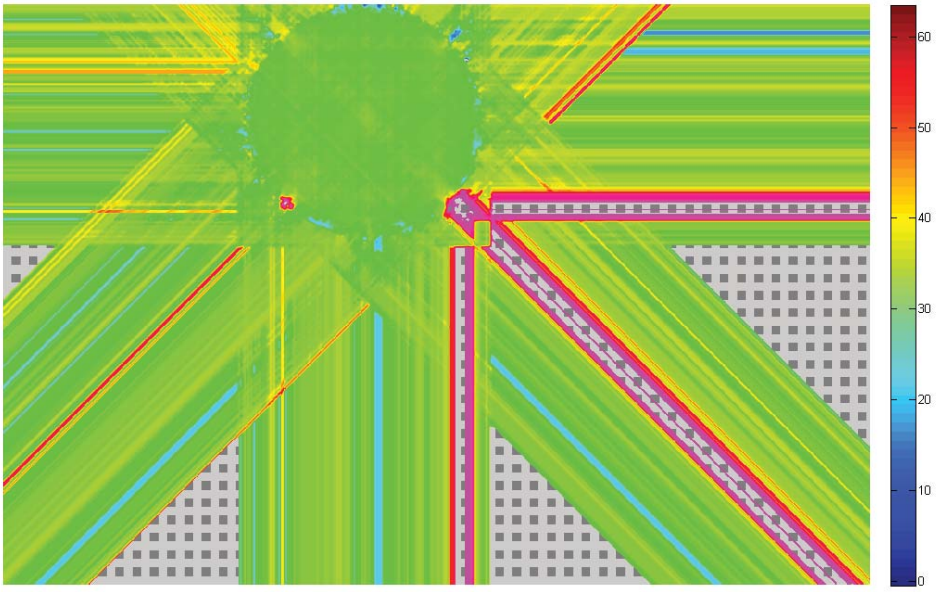


Figure 4.4: Depth map generated from the same raw image used to render Figure 4.3. The grey areas in the depth map represent areas in which no depth was recorded. This is further explained in Section 1.12.2

.

accurately asses if this is the case, a line plot has been taken horizontally through the centre of the ball across the width of the depth map in Figure 4.4, with the results being shown in Figure 4.5.

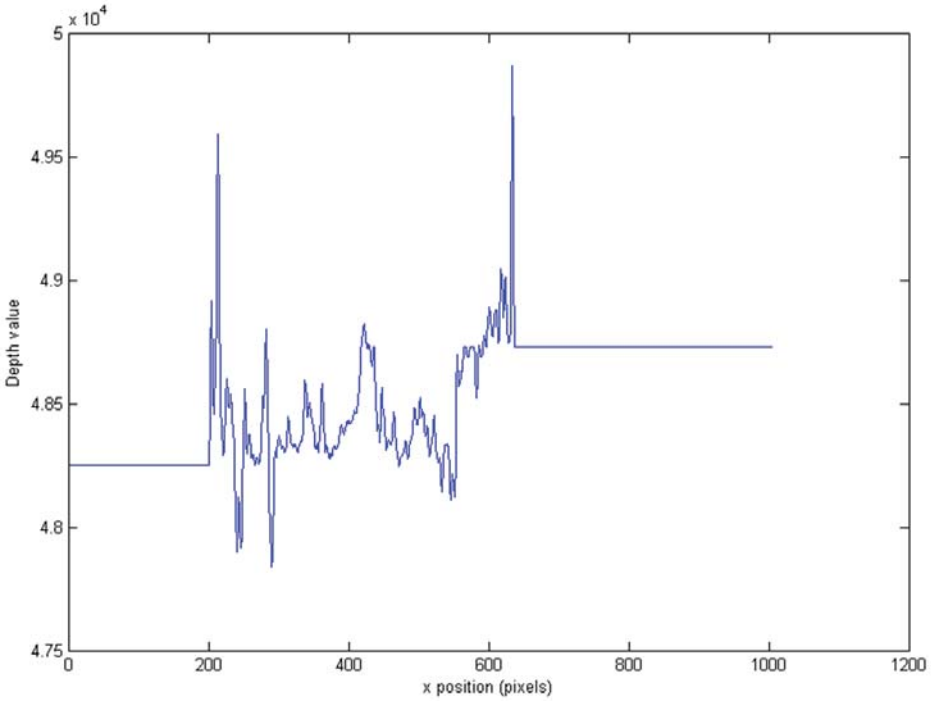


Figure 4.5: Line plot taken from the depth map through the centre of the table tennis ball shown in Figure 4.4. The scene has a featureless black background. The depth value has arbitrary units.

The graph in Figure 4.5 does agree with the initial analysis that the shape of the ball has not been recovered. This is evident from the shape of the line plot between x position of 200 to 600 (approximate values for the area on the depth map where the ball is present) which does not follow the spherical shape of the ball, but instead appears to have no topographic information. Analysing the areas of the line plot which correlate to the background of the image, it can also be stated that these are incorrect values. When analysing the background of the image, the shape of the line plot indicates it has been calculated correctly, as it shows a straight line below 200 and above 650 on the x-axis, and the background was a flat plane parallel to the camera sensor as shown in Figure 4.2. However, there are two features which indicate that the depth of the background has not

been calculated accurately. The first is the depth of background in relation to the table tennis ball. In the experimental set-up, the ball was placed closer to the camera than the background. Therefore, the areas on the line plot which represent the background should have a lower depth value than that of the ball. When assessing the graph, it is clear that the background has not in fact been calculated as being further away than the object, but at the same depth on one side of the ball and even closer to the camera at the other. As previously stated, the shape of the ball has not been recovered accurately, so the depth of this area of the image cannot be trusted as being accurate. Consequently, this shows conclusively that the background is outputted incorrectly, and the line plot can therefore be used to prove that the depth of the background and the object cannot be accurately measured when neither have significant detail on them.

Recovering depth from a scene with features and a plain background

To investigate if the depth and shape of the object could be recovered more accurately when it has features present on it, the same experimental set-up as in Figure 4.2 was used, however instead of a plain table tennis ball being the object, black dots were drawn on the ball. It is predicted that this will provide enough texture for the depth algorithms to reconstruct the shape of the ball. A rendered image from the experiment and the corresponding depth map can be seen in Figures 4.6 and 4.7 respectively.

A qualitative evaluation of the depth map in Figure 4.7 shows promising results, as the area in the centre of the ball has the largest depth value, with the calculated depth gradually decreasing towards the edges of the ball. This is what we would expect, as the centre of the ball is closest to the camera, with the distance gradually increasing towards the edges of the ball. To assess the shape more accurately, a line plot has been taken horizontally through the centre of the ball as previously. The line plot can be seen in Figure 4.8.

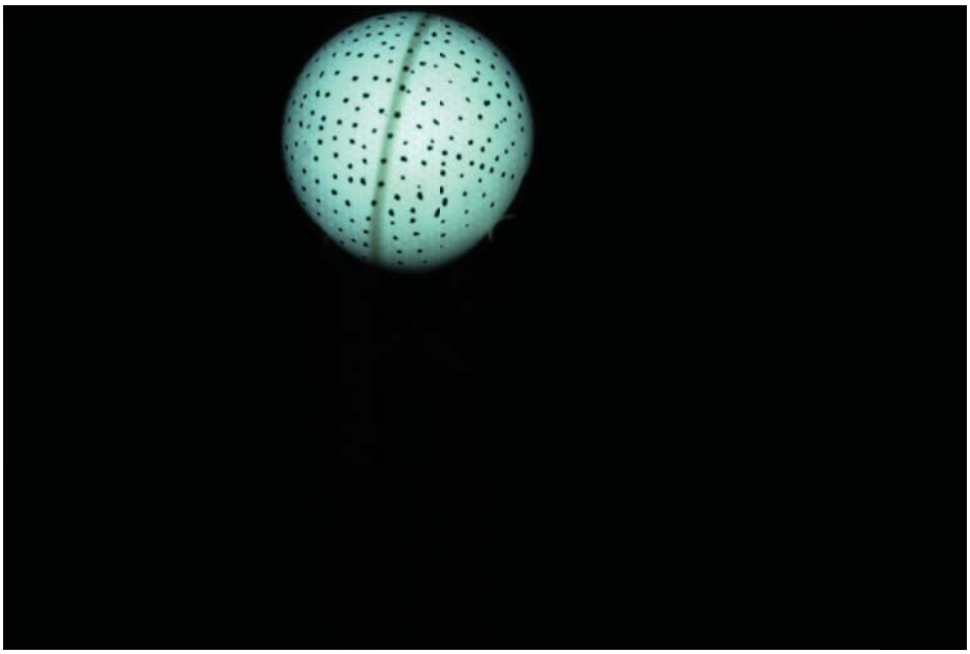


Figure 4.6: Rendered image of a table tennis ball with black dots drawn on when captured by a Raytrix R11 plenoptic camera in front of a featureless black background.

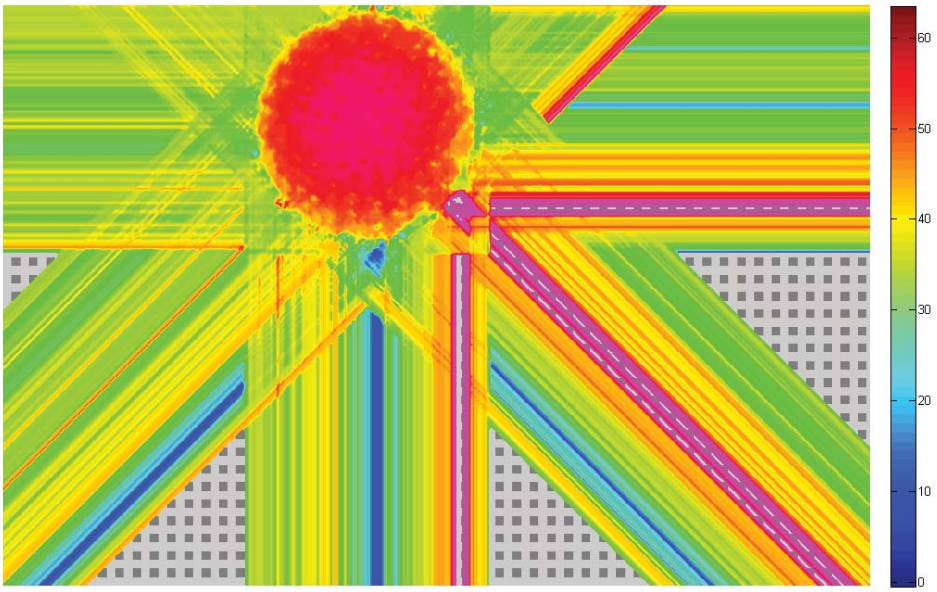


Figure 4.7: Depth map generated from the same raw image used to render Figure 4.6.

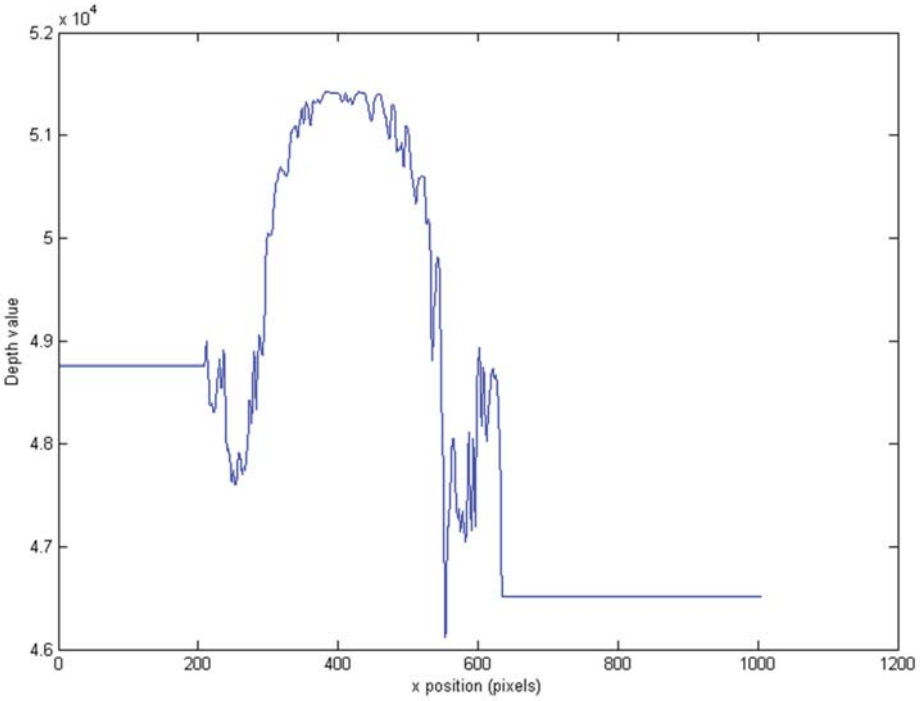


Figure 4.8: Line plot taken from the depth map through the centre of the table tennis ball with black dots drawn on it as shown in Figure 4.7. The scene has a featureless black background. The depth value has arbitrary units.

The graph in Figure 4.8 confirms initial analysis from the depth map that applying dots to the table tennis ball improves the recovered depth. Although the shape of the curve is not a smooth hemisphere as it should be, it is a clear improvement from a plain ball in Figure 4.5, as the software now correctly calculates the ball to be closer to the camera than the background. However, one area that has not been improved from the previous experiment is the background values. Either side of the ball again show different depth values from each other, and different from the previous experiment in Figure 4.5. This gives a stronger validation that the calculated depth of the background is not accurate. The next step then involves changing the plain background to one with texture to see if this issue can be rectified.

Recovering depth from a featureless scene with a textured background

The next preliminary experiment devised involved imaging the plain table tennis ball again but this time with a background with some texture. As the black dots on the table tennis ball did improve the depth calculation significantly, a similar design for the background was chosen. A plain white sheet with a regular pattern of small black circles was used as the background in front of which a plain white table tennis ball was placed. The scene was once again imaged with the Raytrix R11 plenoptic camera, with a rendered image visible in Figure 4.9 and the corresponding depth map in Figure 4.10.

The background of the depth map in Figure 4.10 shows considerable improvement to that from the previous two experiments. As the background is a flat sheet parallel to the camera sensor, if the software calculates the depth accurately it should appear as a single colour on the depth map, which is what occurs but with a certain amount of noise error. Other important attributes to note are the depth values in the areas representing the location of the ball. Although the whole ball appears to be calculated to be a flat circle at the same depth, this result is not surprising as the ball did not have any detail, so the algorithms used to calculate depth will fail to see differences in neighbouring microimages. Unlike the first experiment the depth of the ball has been calculated as correctly being

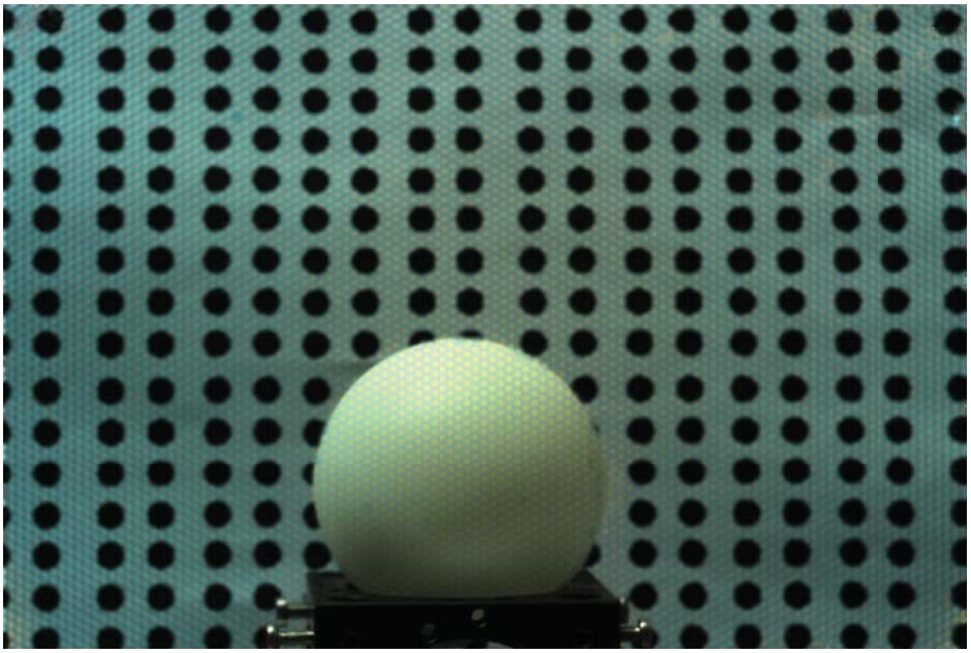


Figure 4.9: Rendered image of a table tennis ball when captured by a Raytrix R11 plenoptic camera. The background is a white sheet with a regular pattern of black circles printed onto it.

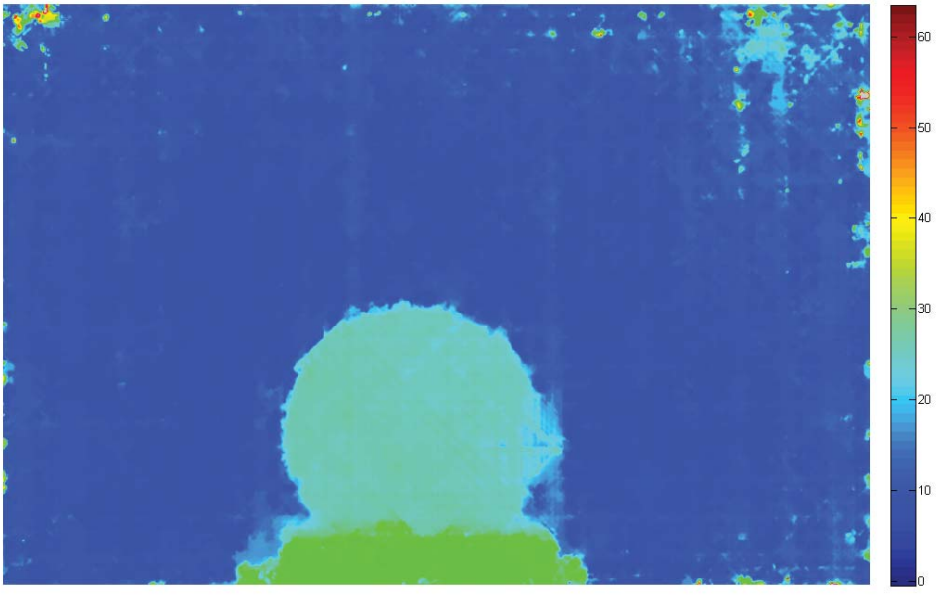


Figure 4.10: Depth map generated from the same raw image used to render Figure 4.9. The area at the bottom of the ball shown in green is the stand on which the ball was placed for this image.

in front of the background. This shows that even with a featureless object, if it is placed in front of a textured background then a difference in depth between the two objects can be observed. To examine this further a line plot through the centre of the ball has been taken as before, and is shown in Figure 4.11.

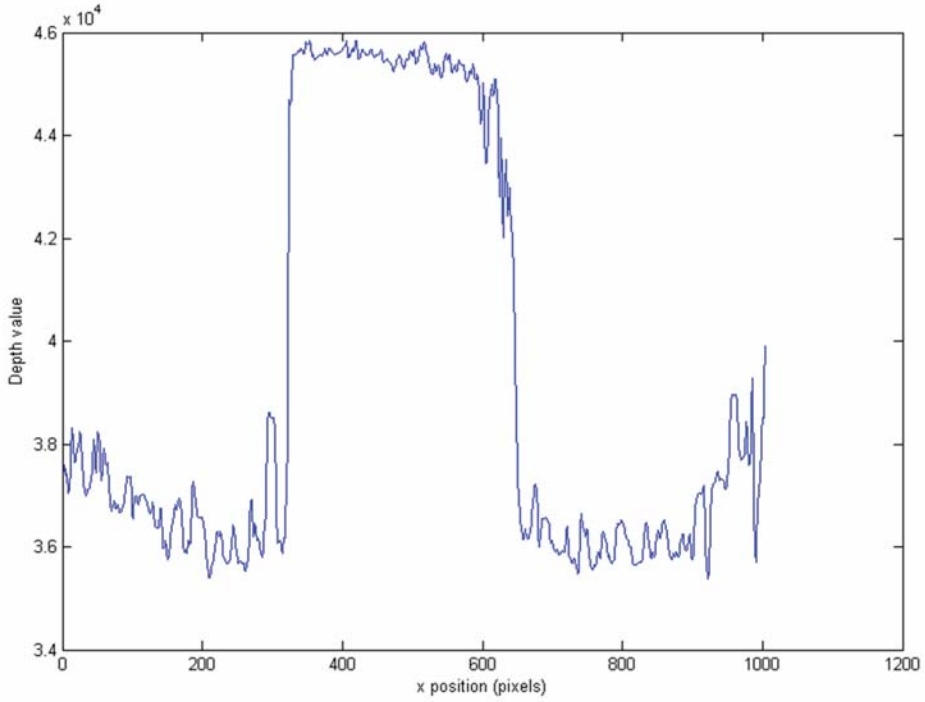


Figure 4.11: Line plot taken from the depth map through the centre of the table tennis ball as shown in Figure 4.10. The background is a white sheet with a regular pattern of black circles printed onto it.

The graph in Figure 4.11 does highlight very well the difference between the ball and the background. It also confirms that no depth information about the shape of the table tennis ball has been recovered, but it does recover the distance between the background and the ball accurately. This is thought to be due to the edge of the ball providing enough detail that the depth of the edges can be calculated. In general, the background now appears flat, however it does exhibit a periodic noise pattern. It is thought that this is due to the regular pattern of black circles on the background, but future experiments in Section 4.6.2 will investigate the origins and effect this noise has on the recovered depth.

Recovering depth from a scene with features and a textured background

The last of the preliminary experiments involved imaging the table tennis ball with texture in front of a textured background. It is predicted that following on from previous experiments, this one should yield results in which the spherical shape of the table tennis ball, the flat shape of the background and the distance between the background and the ball are accurately measured, subject to a certain degree of error and noise. A rendered image and the corresponding depth map can be seen in Figures 4.12 and 4.13 respectively.

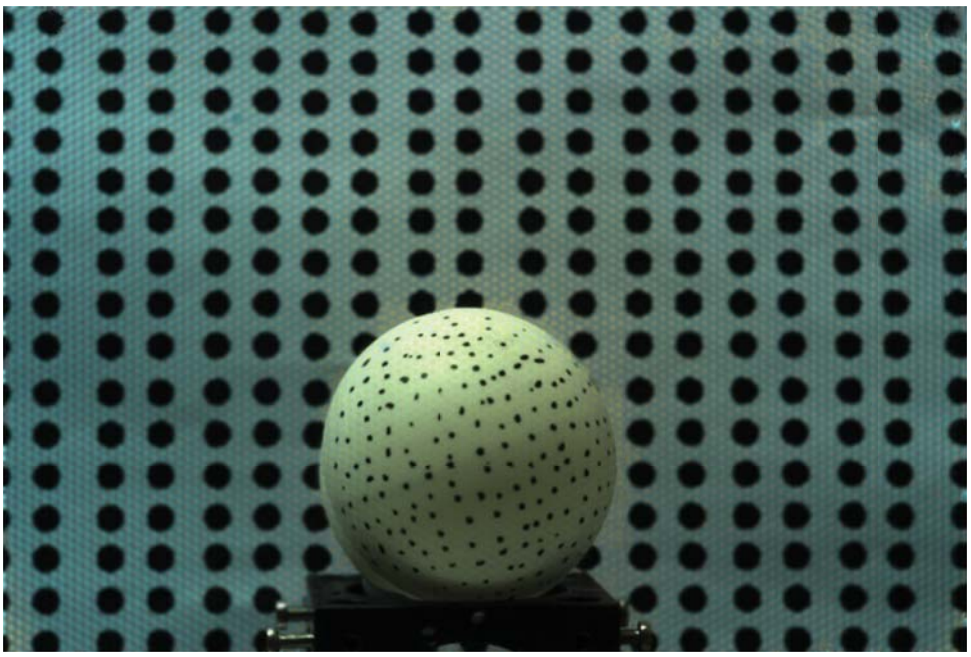


Figure 4.12: Rendered image of a table tennis ball with black dots drawn on when captured by a Raytrix R11 plenoptic camera. The background is a white sheet with a regular pattern of black circles printed onto it.

A qualitative analysis of the depth map in Figure 4.13 does suggest that the depth has been calculated as expected. The background of the image is very similar to that of the previous experiment in Figure 4.10. The table tennis ball is calculated to be at a depth significantly closer to the camera than the background as it was in physical space. The ball itself does appear to be closer to the camera in the centre of the ball, which would be correct due to its spherical shape. As before, the associated line plot has been taken and can be seen in Figure 4.14.

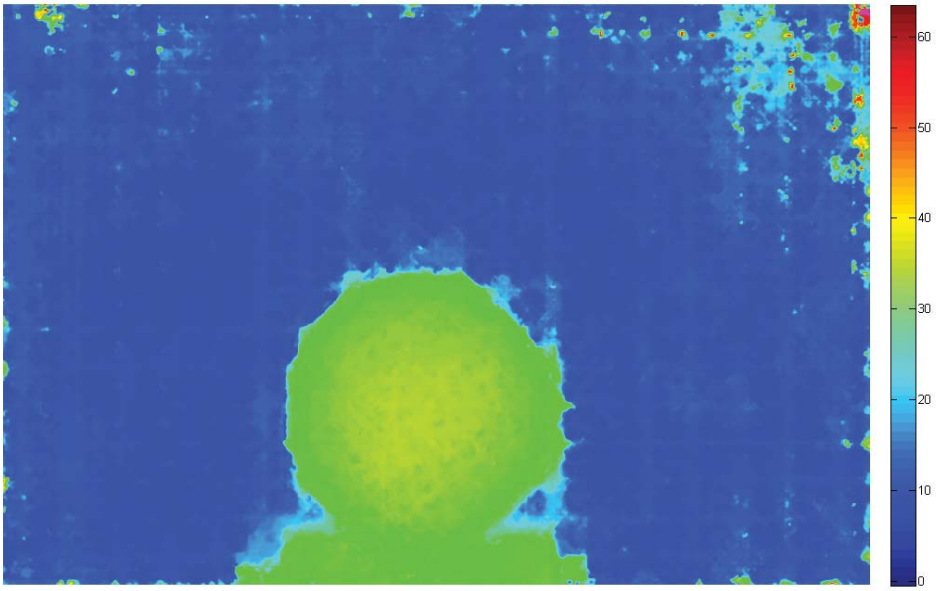


Figure 4.13: Depth map generated from the same raw image used to render Figure 4.12.

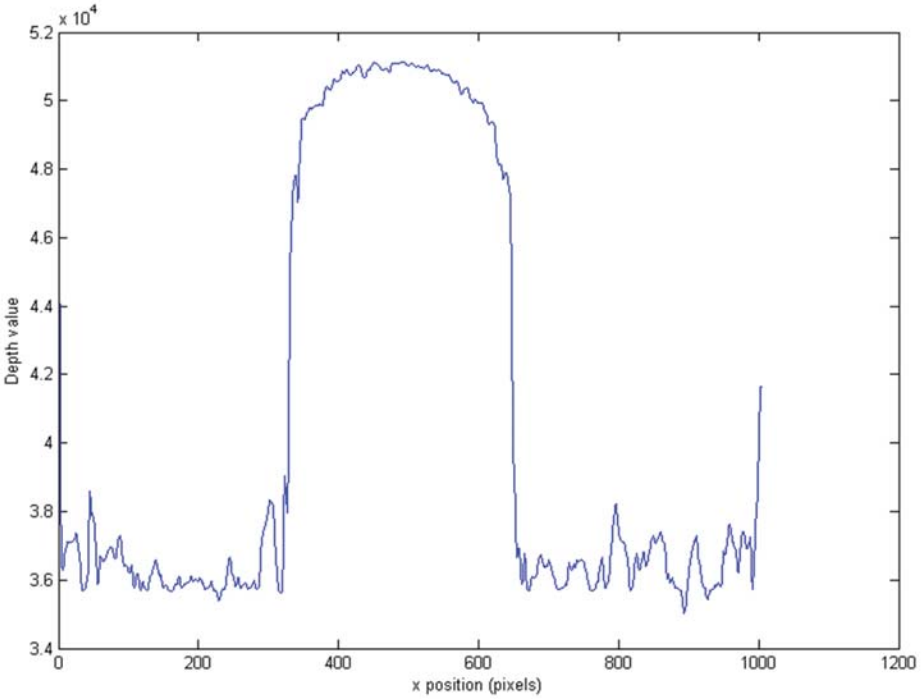


Figure 4.14: Line plot taken from the depth map through the centre of the table tennis ball with black dots drawn on it as shown in Figure 4.13. The background is a white sheet with a regular pattern of black circles printed onto it.

The profile of the line in Figure 4.14 confirms the previous analysis, as the shape is exactly as expected. The background is roughly constant, with the periodic noise as before which is thought to be due to the regular pattern. The ball is at a depth closer to the camera than the background, and the ball itself does have a spherical shape. This study therefore provides evidence that when a scene has objects with no features or significant detail, the depth or shape of that object cannot be accurately measured from a plenoptic camera. However, if detail is added to the objects in the scene, then the shape and depth of the objects can be recovered more accurately. The method of drawing on patterns to featureless objects may not always be possible and is certainly not ideal. An alternative method of projecting detail onto a scene was then explored.

Noise analysis

By looking at the line plots in Figures 4.5, 4.8, 4.11 and 4.14 it shows planar noise throughout all the plots. It is believed that this is due to the periodic nature of the microlens array. Section 4.6.2 removes this noise to see if it can improve the accuracy of the calculated angle of a slope. As this experiment indicated that it did not improve the accuracy of the results, no further work into investigating its origins was undertaken.

4.1 Experimental limitations

The limitations have been discussed in Section 3.4.1, but it is worth reiterating these points due to the importance when analysing the results. The major limiting factors in the data presented throughout Chapters 4, 5 and 6 was the time the author had with the Raytrix R11 camera, and the number of technical issues which occurred during this short time. In only a 2 week period, all data was collected for the mentioned chapters. Much of this time was spent trying to fix the computer with the Raytrix software installed, or in communication with Raytrix as the initial calibration would not function correctly. Due to a malfunctioning fan on the computer, after each image was taken the machine would

fail and have to be restarted, with many images lost as this happened before they could be saved. This significantly reduced the number of useful images which were recorded, so throughout these chapter many of the experiments were not able to be repeated as was first desired in the designing of the study. The results presented should then be taken as a proof of concept and method and algorithm development, providing a base for future experiments.

It is also worth noting that the protocol for recovering metric depth values is reliant on the software methods encoded in the Raytrix. Therefore, this protocol will change when Raytrix updates their software depth decoders, thus limiting the fundamental utility of this work. There are also some limitations due to mechanical alignment of the system. The methods presented throughout this chapter rely on the physical alignment of the imaged plane and the sensor of the Raytrix plenoptic camera. As a proof of concept that projected texture onto a featureless plane improves the recovered depth, the errors incurred from the mechanical alignment was seen as an acceptable error.

4.2 Projecting patterns

As shown previously, at least qualitatively, the depth estimation from a plenoptic camera can be improved when imaging scenes with limited features by drawing dots onto the objects. However, this method may not always be as drawing on objects may damage or destroy the object, or in some cases may not be physically possible. Therefore the idea to project a pattern onto a featureless scene has been proposed in order to create the same texture as drawing on the objects but without making any permanent changes. This section of work will include a preliminary study into whether texture can be projected onto a featureless scene to improve depth recovery. A study into the differences between using a laser or an LED projector when imaging a featureless slope is explored, then once the optimum type of projector to use, a study into the differences between texture being

inherently present in a scene to the same texture being projected onto the objects has been investigated.

4.2.1 Improving depth by patterned projections on a featureless scene

An experiment was devised to test if projecting patterns onto a featureless plane improves the calculated depth, with a Microvision SHOWWX+HDMI Laser pocket projector being used to illuminate a plain sheet at angles of 0° to 45° to the plane of the sensor at intervals of 15° with both white light and the pattern as shown in Figure 4.15(b). Depth maps calculated from the Raytrix are then analysed and compared. A schematic of the set-up is shown in Figure 4.15(a).

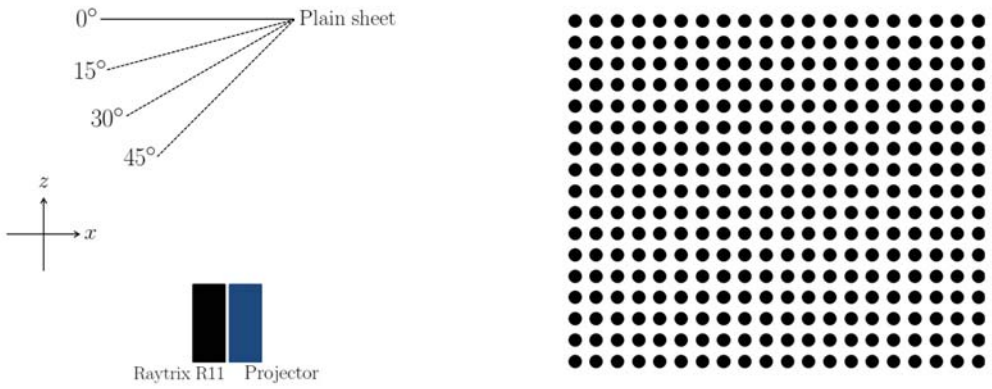


Figure 4.15: (a) An overhead view of the experimental set-up. The blue box represents a laser projector and the black box the Raytrix R11 plenoptic camera. The black lines represents the plain sheet being imaged at varying angles from parallel to the camera sensor to 45° in steps of 15°. When imaging the plane at 0° the camera is 700mm from the plane. Image is not to scale. (b) The pattern projected onto the plain sheet.

Due to the orientation of the slope to the camera, when the plane is parallel to the sensor the depth map should be a consistent colour indicating constant depth, but otherwise should produce a gradual depth change from left to right, with pixels in a column having

the same colour (or depth). Figure 4.16 shows the depth maps for parallel to and 45° to the sensor when illuminated with white light and the pattern shown in Figure 4.15(b).

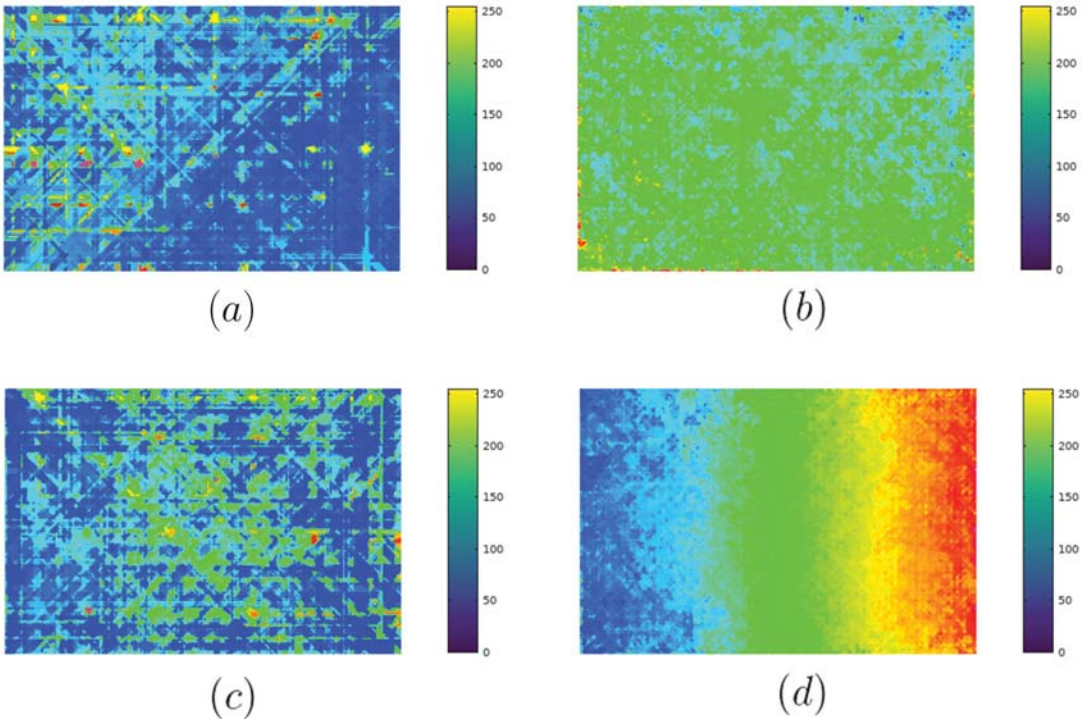


Figure 4.16: Depth maps produced when imaging a flat sheet (a) parallel to the sensor illuminated with white light, (b) parallel to the sensor illuminated with a pattern (see Figure 4.15(b)), (c) at 45° to the sensor illuminated with white light, and (d) at 45° to the sensor when illuminated with a pattern. All illuminations are performed with a laser projector.

As predicted, when the plane is parallel to the sensor, both the depth map illuminated with white light and a pattern produced a relatively uniform colour representing a flat depth. When at 45° to the sensor, the depth map produced from the plane illuminated with white light appears similar to that when the plane is parallel to the sensor. This shows that depth information is not present in the absence of texture, and this is consistent with previous experiments. When the 45° slope is illuminated with a pattern, the depth map behaves as expected. This shows that projecting texture onto a featureless scene does improve the depth calculated from a plenoptic camera.

4.2.2 Laser v LED projector

Although they demonstrate that projected texture improves depth estimate, the results in Figure 4.18 show a large amount of noise. This is thought to be speckle noise from the laser projector, so in an attempt to reduce the noise an LED projector was also used. To compare the effect of using an LED or a laser projector, the same experimental set-up as in Figure 4.18 was used. Images at all angles between parallel to the sensor to 45° at intervals of 15° were taken, with the same laser projector as previously and with a Optoma PK201 Pico Projector as the LED projector. The experiment has be performed somewhat differently for the two projectors. Since the laser projector projects a scanning collimated beam, it is "in focus" at all distances. However, the LED projector can only be focussed on one plane and must therefore be rotated along with the angles plane in order to maintain the focus of the pattern on the plane.

To compare the difference in noise from the two systems, depth maps can be analysed qualitatively in Figure 4.17 and quantitatively from Table 4.1. The image quality is measured as the standard deviation of the entire depth map when the plane is parallel to the sensor, as each pixel is at the same physical depth, the Raytrix should compute the same virtual depth value for every pixel. For angles greater than this, the standard deviation of each column is taken, and then a mean of the standard deviations is used. This is because the pixels in every column on the depth map are at the same distance away from the sensor, so should be computed as the same value in the depth map.

	0° slope	15° slope	30° slope	45° slope
Laser projector	0.0786	0.0536	0.0283	0.0294
LED projector	**0.0368**	**0.0328**	**0.0237**	**0.0273**

Table 4.1: A comparison of the standard deviation of depth values produced from a Raytrix R11 camera when imaging a featureless slope at 0°, 15°, 30° and 45° to the plane of the sensor with a pattern projected using both a Laser and an LED projector.

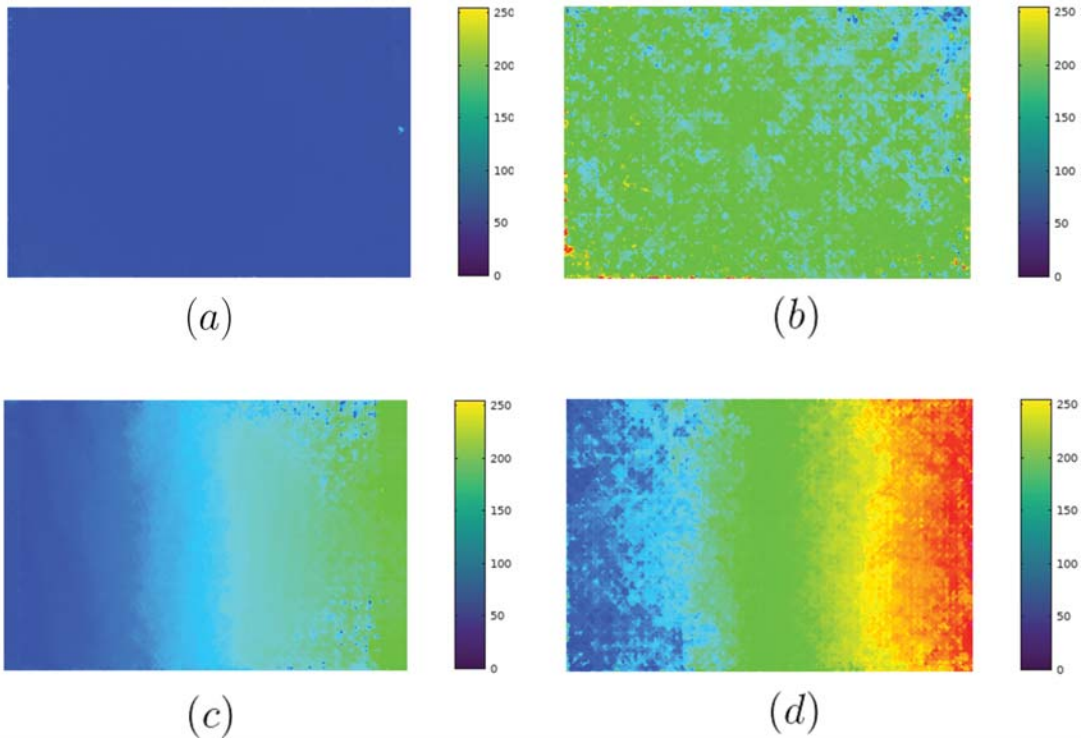


Figure 4.17: Depth maps produced when imaging a flat sheet illuminated with a pattern (see Figure 4.15(b)), (a) parallel to the sensor illuminated with the LED projector, (b) parallel to the sensor illuminated with the laser projector, (c) at 45° to the sensor illuminated with the LED projector, and (d) at 45° to the sensor when illuminated with the laser projector. *Note*: the colours in these depth maps do not correlate to the same exact depth values. They are only relative to other pixels in that particular depth map.

Looking at the depth maps in Figure 4.17 it appears qualitatively that the depth maps from the LED projector are less noisy. This correlates with the data in Table 4.1, which shows for every angle a lower average standard deviation compared to the laser projector. Therefore, for future experiments in this chapter the LED projector shall be used.

Another trend which appears from Table 4.1 is that as the angle increases, the standard deviation decreases. It is not clear as to the reason for these discrepancies, however, as the aim of this experiment was to determine whether to use an LED or laser projector, it is not seen as necessary to investigate this further.

4.2.3 Validation of using a projector

As previously shown, the errors from using an LED projector are smaller than those of a laser projector. However, it is important to investigate the difference between having texture inherent in the scene and adding the detail with a projector. In order to investigate this difference, an experiment was devised with the schematic shown in Figure 4.18. For this investigation, a Raytrix R11 camera was used to image a flat plane at angles of parallel, 10° and 20° to the sensor. To see the difference between projected vs intrinsic texture, identical patterns as shown in Figure 4.15(b) were printed and projected onto a flat plane and then imaged. Multiple images were taken of both the printed and projected pattern, and then compared, with the results shown in Table 4.2. Examples of the depth maps that have been compared can be seen in Figure 4.19.

By taking multiple images of the sheet with both the printed and projected pattern applied, a full comparison can be made. The maximum and mean differences from Table 4.2 are calculated by comparing identical pixels on the depth maps between two images where the pattern was printed onto the sheet, two images where the pattern was projected onto the sheet, and between a printed and a projected image. By looking at the difference between the two images taken when the pattern was printed onto the sheet, the maximum difference for all three angles is approximately 3% and the mean difference less than 0.5%.

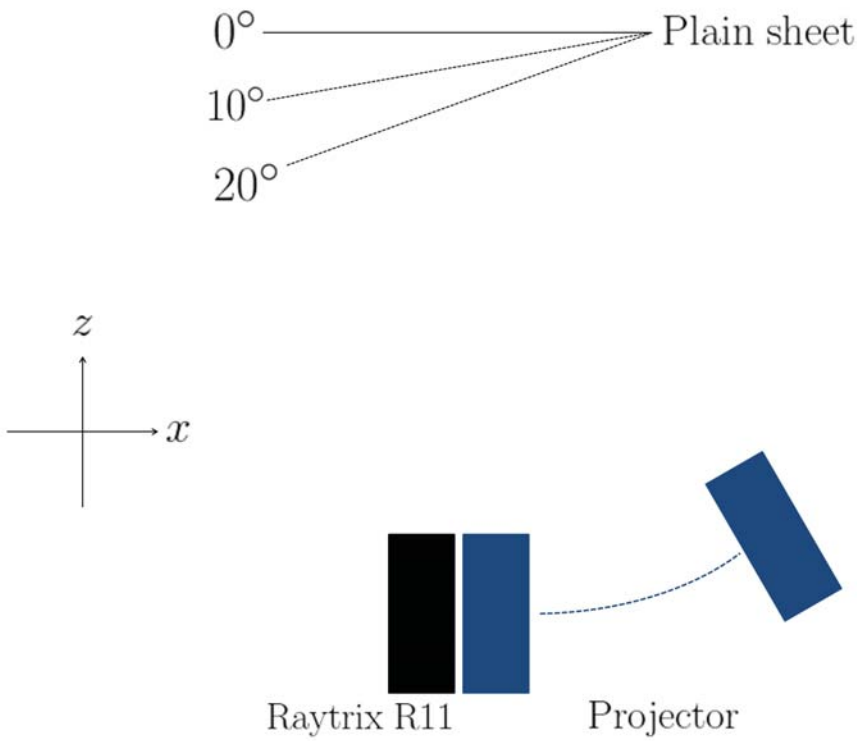


Figure 4.18: (a) An overhead view of the experimental set-up. The blue box represents the LED projector and the black box the Raytrix R11 plenoptic camera. The black lines represents the plane sheet being imaged at varying angles from parallel to the camera sensor to 20°. The projector is also rotated to ensure that the projected image stayed in focus on the plane. The projector and camera are approximately 700mm from the plane being imaged. Diagram is not to scale.

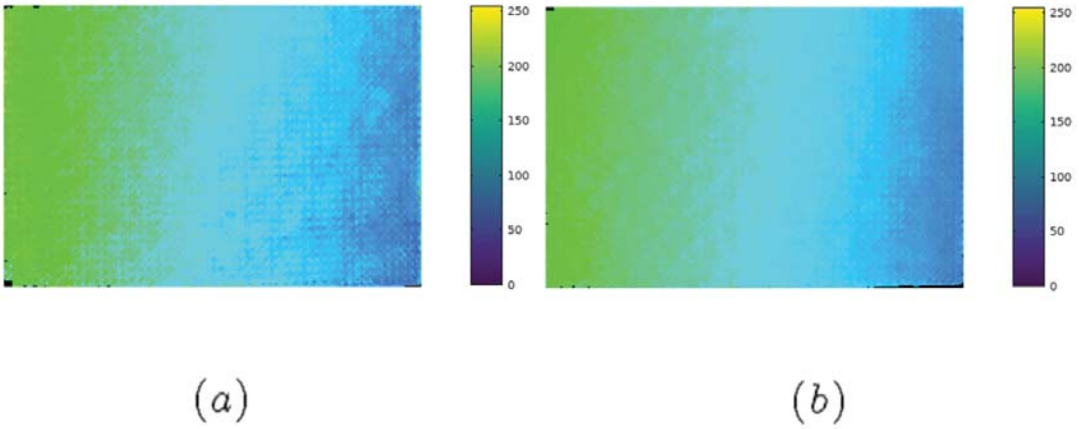


Figure 4.19: (a) Depth map produced when the pattern in Figure 4.15(b) is projected onto a slope at 20° to the plane of the sensor, (b) same as (a) but with the pattern printed onto the slope not projected. The scale is the virtual depth outputted from the Raytrix software.

Comparison	Plane angle (degrees)	Maximum percentage difference (%)	Mean percentage difference (%)
Printed/Printed	0	3.19	0.39
	10	2.72	0.26
	20	3.11	0.20
Projected/Projected	0	4.16	0.44
	10	6.86	0.65
	20	4.81	0.44
Printed/Projected	0	5.34	0.78
	10	5.37	0.67
	20	5.71	1.19

Table 4.2: Mean and maximum percentage differences between depth values for two images with a pattern printed, two images with the same pattern projected, and one printed and one projected pattern. The plane angle is in reference to the x-axis in Figure 4.18.

For the purpose of this experiment, this level of accuracy is taken as the gold standard and assumed to be the error inherent in the plenoptic imaging system. By then comparing the two images with the pattern projected, the maximum and mean differences have increased, however only up to around 5% and less than 0.7% respectively. Although there is an increase in the difference, it is still very small and on the same order of magnitude as for that from the two images with a printed pattern. The final comparison is between an image with a printed and one with a projected pattern. In this, the maximum difference does not exceed 6% and the mean is below 1.2% at its largest. This again is on the same order of magnitude as the differences between the two images with a printed pattern, so would seem that projecting texture onto a scene is very similar to that of the texture already being present. To confirm this analysis more conclusively, a Paired t-test was performed on all comparisons at all angles, yielding a p-value of < 0.001 for every test. This does therefore validate the use of a projector to apply texture, as it has very little variation from the same texture being part of the scene.

Although the differences between a projected pattern and a pattern inherent in the scene has been accepted as being in an acceptable range for this study, there are still some variations due to the projector. Section 3.2.2 states from the literature the accuracy limitations of projecting a pattern on a scene for three dimensional reconstruction. There will also be variations due to the thermal and mechanical stability of the projector. Efforts were made to minimise these errors, by allowing the projector to reach a stable temperature before the experiments were conducted, keeping the room, as much as possible, to a constant temperature and attaching the projector to a vibration proof optical bench.

4.2.4 Pattern projection geometry

Throughout the remainder of this chapter and in the following one, the results presented use a projector to impose texture onto a featureless plane. In all experiments, the projector is always at the same angle as the plane, and at a distance of 700 mm from the plane. The constant angle was done to minimise errors from part of the projection going out of focus, as the LED projector was chosen as per the results from Section 4.2.3. Also, by keeping the distance from the projector to the plane, the projected pattern will remain constant on the plane. This becomes more important in the following chapter when different patterns are analysed.

With regards to the spectral properties of the projection pattern, this was chosen not to be investigated. This is because the premise of these experiments was a proof of concept that adding texture to a scene can improve depth reconstruction from a plenoptic camera. The spectral analysis of the projections and how it is sampled onto the cameras imaging plane could be done as an extension to the work presented in this thesis.

4.3 Converting virtual to real depth

When using a Raytrix plenoptic camera, the depth maps outputted only contain virtual depth values. Although this can be useful for comparing relative depth changes between

objects in a scene, it would be more beneficial if these can be converted into a real depth in physical space. One reason why this is not inherent in the Raytrix is because the camera has no knowledge of the main lens. Although the f-number of the lens needs to be matched to that of the microlens array (see Section 7.1.2), the focal length of the lens can be altered depending on the desired working distance and field of view. If the lens focal length is changed and an object at the same distance from the sensor is imaged, the Raytrix computes a different virtual depth value. It therefore is impossible to gain real depth values directly from the depth maps. However, a calibration procedure has been devised in order to convert the virtual depth outputted from the Raytrix into real distance values.

The calibration procedure involves imaging a flat plane parallel to the sensor from the closest to furthest points in space of the depth of field at small intervals whilst recording the distance between the sensor and the plane. The ends of the depth of field can be found relatively easily, by increasing the distance between the camera and the plane until no depth is registered from the camera. The end of the depth of field closest to the camera can then be calculated by simply translating the plane closer to the camera until it no longer registers depth. Figure 4.20 shows an overhead schematic of the experimental configuration.

As each point on the plane is at the same distance from the sensor along the optical axis (z-axis in Figure 4.20), the same virtual depth should be calculated for every pixel in the depth map. In reality this is not true due to imperfections in the camera system and errors in the calibration of the camera. Therefore an average value for the entire depth map is taken. A physical measurement from the plane to the sensor of the camera is then taken, and this physical depth can then be directly related to a virtual depth value. This does incur some errors due to inaccuracies of human measurement and the estimation of where the sensor is inside the camera module. However, this estimation can be done accurately by looking up the flange focal distance for the given lens mount providing an

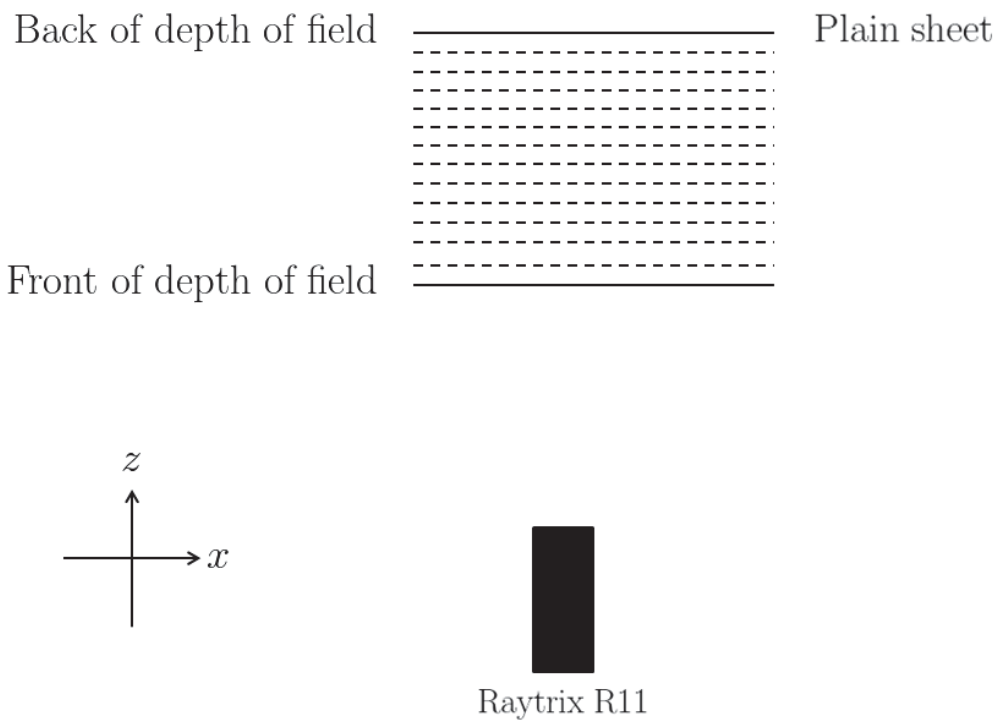


Figure 4.20: An overhead schematic of the novel calibration experiment devised in order to convert virtual depth from a Raytrix plenoptic camera into a real depth in physical space. The back of depth of field is placed at 710mm from the camera, and the front of depth of field at 460mm.

accuracy in the order of a few millimetres. Since the working distance for the Raytrix R11 is in the order of 0.4m - 0.8m depending on the lens settings, this error becomes very small and insignificant. If this process is repeated for many depth maps taken at different depths, a graph relating virtual to real depth can then be created. By fitting a function to this graph it creates a conversion from virtual to real depth. An example of this plot can be seen in Figure 4.21.

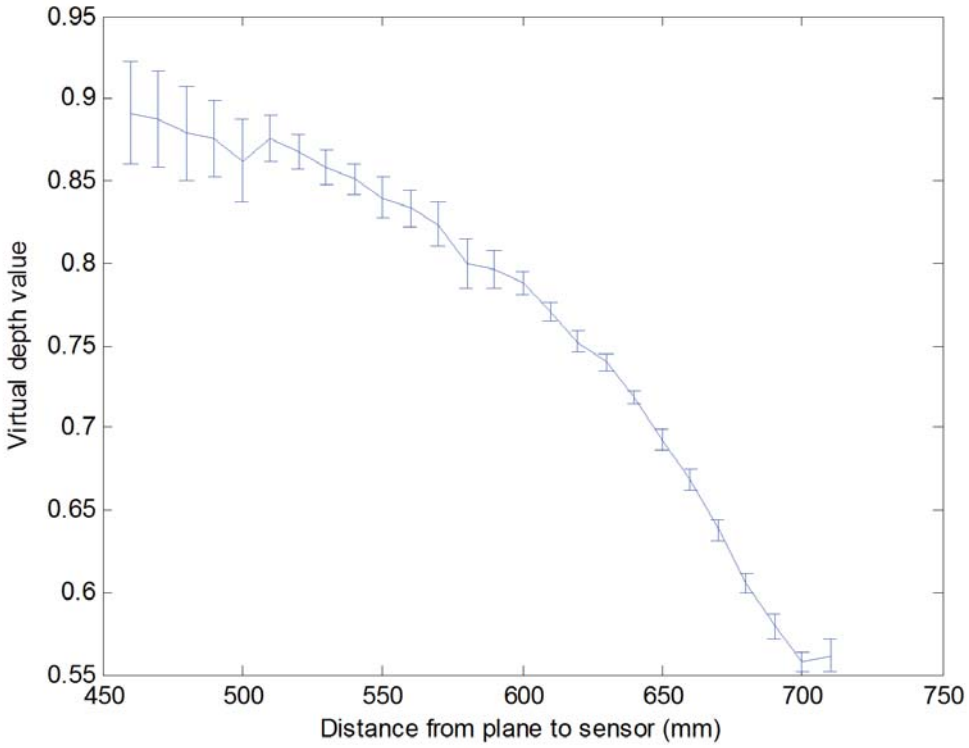


Figure 4.21: A graph showing the relationship between the distance between a flat plane and the sensor with the virtual depth value outputted from the Raytrix depth map. The values for virtual depth are calculated as the mean value from an entire depth map, and the error bars are ± the standard deviation. The lens in this experiment had a focal length of 700mm and an f-number of *f/8*.

A few interesting points can be concluded from analysing Figure 4.21. Firstly, the shape of the graph is not as expected. In an ideal situation, the graph would fit a straight line as the virtual depth should linearly correlate to depth in real space. As can be seen this is not the case. This could either be due to an inherent property of the relationship between virtual and real depth in the Raytrix, or could be due to small inaccuracies in

the initial calibration. To try to identify the cause, the experiment was repeated with a different initial camera calibration required in the set-up of the Raytrix with the results shown in Figure 4.22

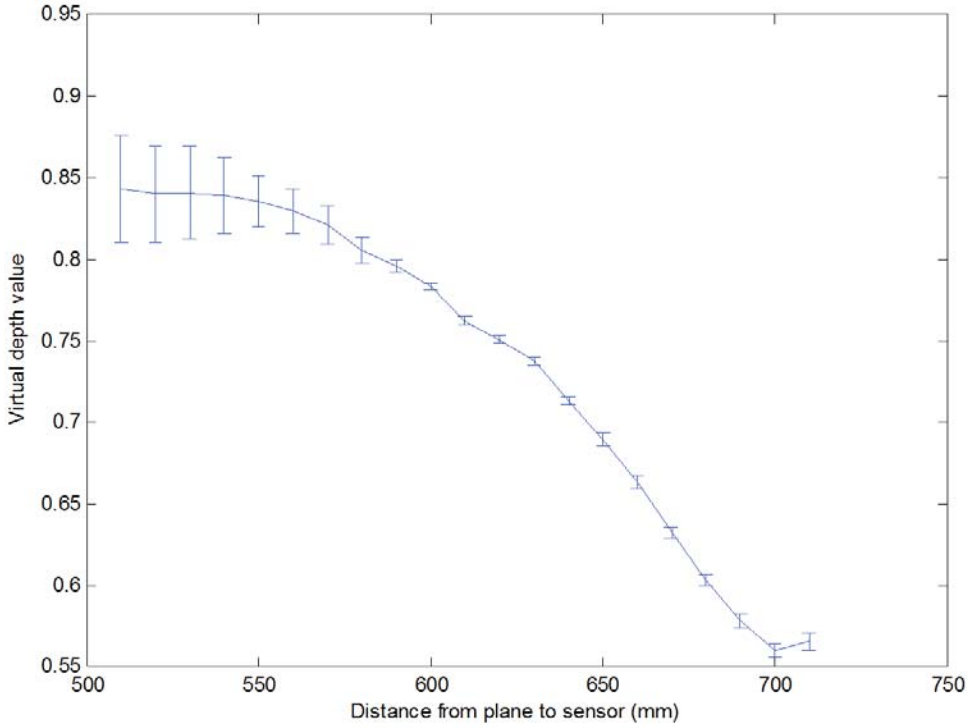


Figure 4.22: A repeat of the experiment shown in Figure 4.21 except with a different initial camera calibration.

From qualitatively analysing the shape of the graphs in Figures 4.21 and 4.22, they both appear to have a similar curved shape. This indicates that the Raytrix does not have a linear relationship between virtual and real depths as the shape is consistent independent of the initial camera calibration applied. This could be due to the Raytrix having a microlens array that is interlaced with lenses of different focal lengths. The image formation could then be linear for each individual lens type, but produces the shape shown when using them interlaced. A way to test this would be to repeat the experiment with a different plenoptic camera which used a microlens array with a single focal length. Due to practical restrictions this could not be tested. For more information

on the response of the interlaced microlens array, it is worth reading the paper by Perwass *et al.* [1].

Other similarities include the increase in virtual depth value when the distance is greater than 700mm. This is likely to be due to the object distance of the main lens being set to 700mm, so at distances greater than this the plenoptic camera no longer registers depth correctly. This will be investigated further in future experiments by repeating the experiment but changing the main lens focal length. Another interesting feature to note is the increase in the standard deviation and flattening of the curve at distances closer to the camera. The relationship between the virtual and real depth is currently calculated from the mean value of the depth map outputted from the Raytrix. However, this may only be the correct average to use if the spread of values in the depth map corresponds to a Gaussian shape centred around the mean. To see if this is the case, the range of values has been divided into percentiles and a histogram of the spread of points has been plotted. Figure 4.23 shows a selection of histograms for a variety of distances from the sensor to the plane from the results presented in Figure 4.21, and Figure 4.24 shows the same but for the results presented in Figure 4.22.

Both Figures 4.23 and 4.24 follow similar patterns. At the distances further away from the sensor, the spread of data appears to be close to a Gaussian distribution, with the mean value being on or very close to the most common value. However, when the distance between the sensor and the plane becomes small this distribution changes. The peak value no longer aligns with the mean value, as a large spread of data to smaller virtual depth values skews the mean to a lower value. This could help to explain the flattening of the curves in Figures 4.21 and 4.22. As the mean is not representative of the data at these smaller distances, it may be more statistically appropriate to use the modal value of each histogram. Figures 4.25 and 4.26 show the data present in Figures 4.21 and 4.22 respectively but with the modal values included as well as the means.

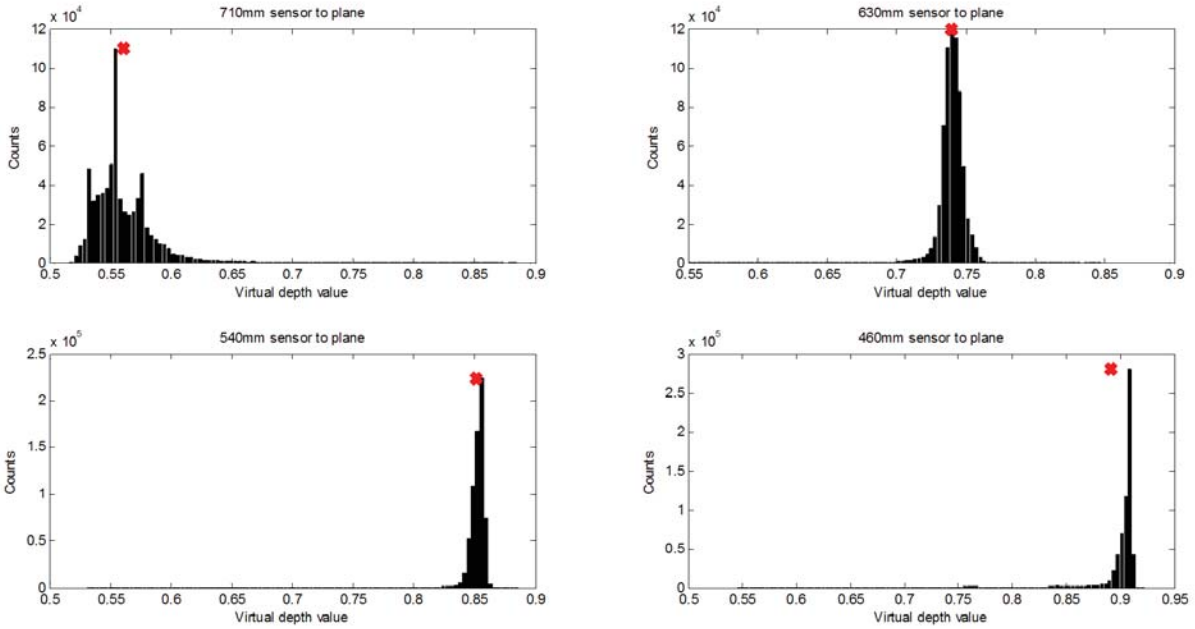


Figure 4.23: Histograms showing the distribution of virtual depth values for the data shown in Figure 4.21 at *Top-left:* 710mm, *Top-right:* 630mm, *Bottom-left:* 540mm and *Bottom-right:* 460mm from the plane to the sensor. The red cross indicates the mean virtual depth value for the data.

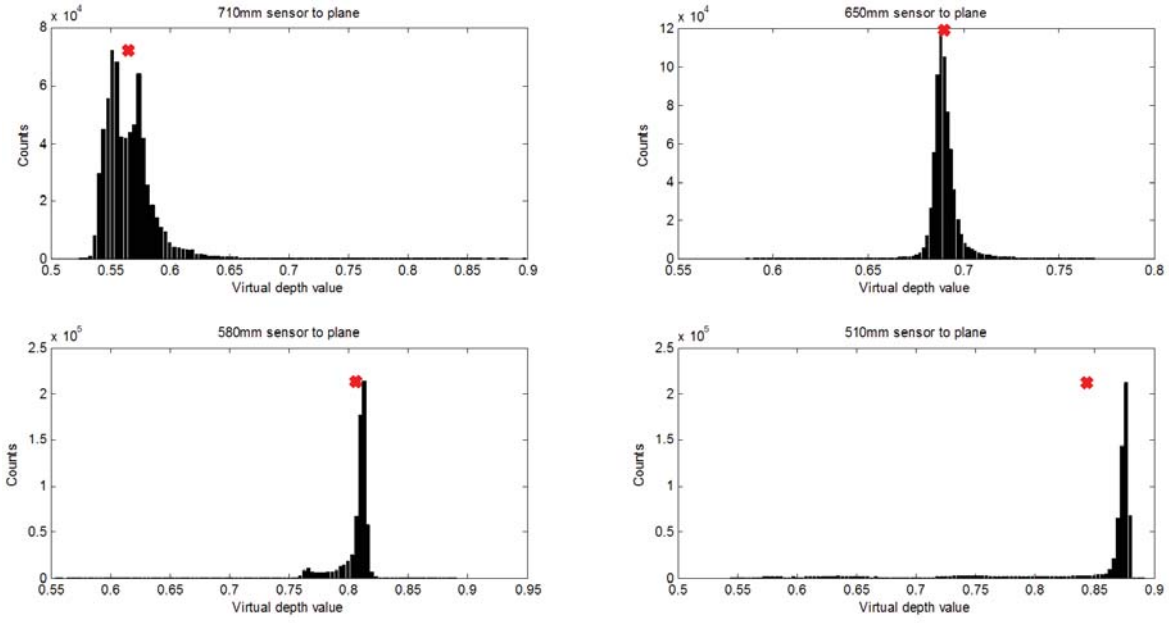


Figure 4.24: Histograms showing the distribution of virtual depth values for the data shown in Figure 4.22 at *Top-left:* 710mm, *Top-right:* 650mm, *Bottom-left:* 580mm and *Bottom-right:* 510mm from the plane to the sensor. The red cross indicates the mean virtual depth value for the data.

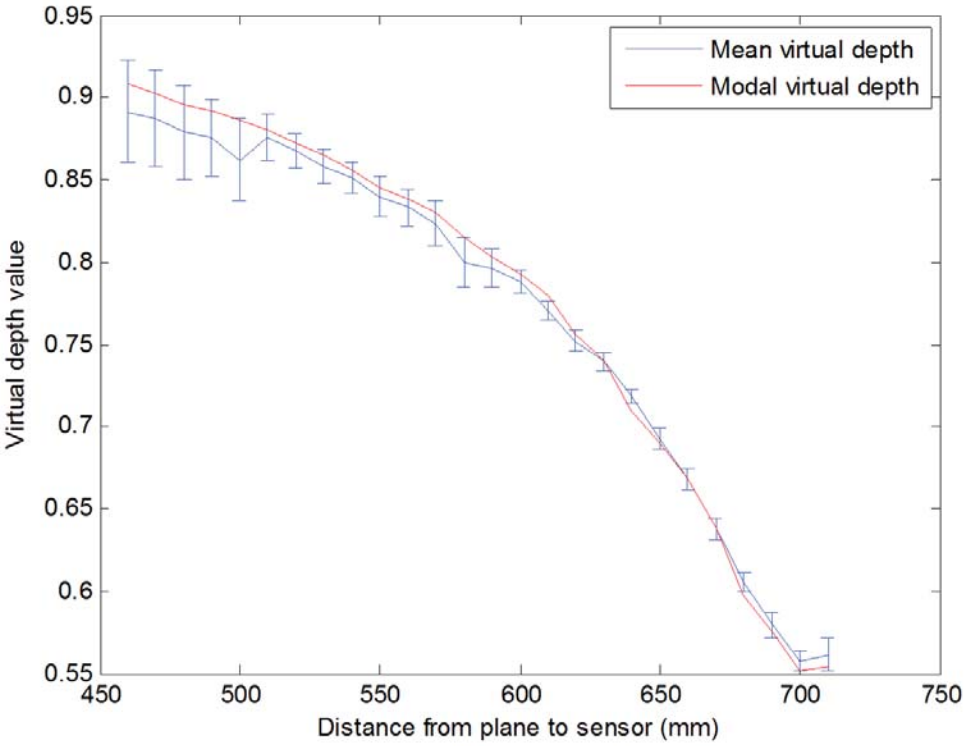


Figure 4.25: The same data as presented in Figure 4.21 but with the modal as well as mean values for the virtual depth presented

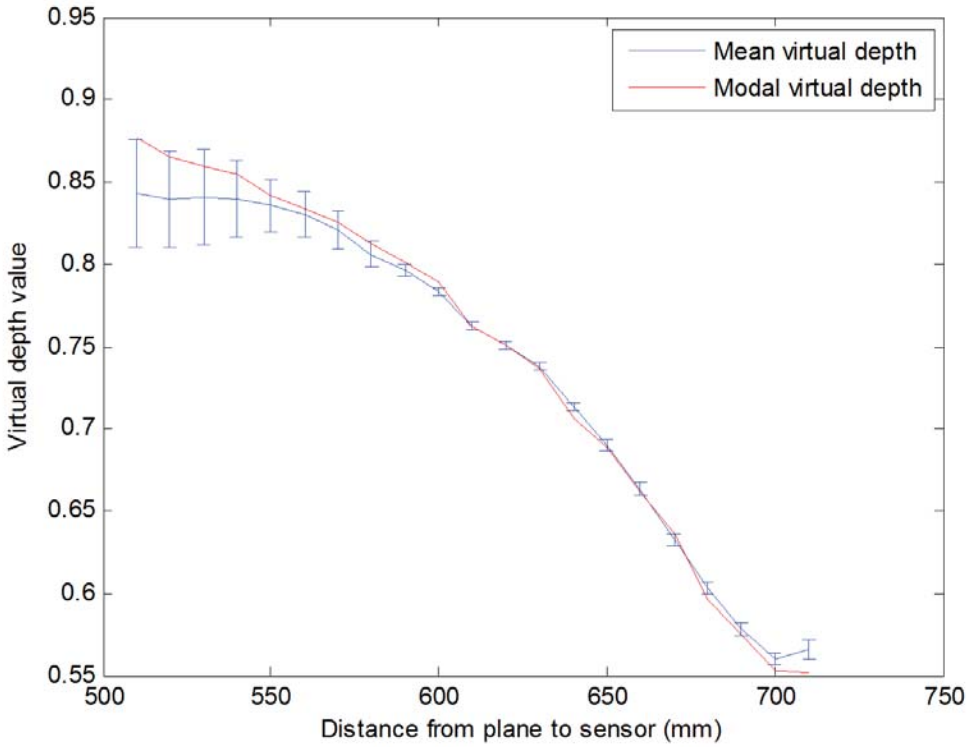


Figure 4.26: The same data as presented in Figure 4.22 but with the modal as well as mean values for the virtual depth presented

As predicted, by using the modal values instead of the means, the flattening of the calibration curves when at close distances is reduced. The modal values also appear to be more representative of data, as the means are highly skewed by erroneous values. Therefore, the modal average from each depth map will be used as the calibration value for that depth. The reason why this phenomena occurs is unknown. It is thought this could be due to the dynamic range of the camera rather than the distribution. This could be investigated further by trying to determine the dynamic range of the camera, however, as this is not seen as essential for driving the project forward this was not investigated further.

The next step is to fit a function to the curves. This is an essential step so any virtual depth value in a depth map can then be converted into a real physical depth. In order to achieve this, the POLYFIT function in MATLAB is used. Figure 4.27 shows the modal average virtual depth and the fitted function for the data in Figure 4.21. As can be seen, the fitted function follows the actual data to a high degree of accuracy. It is then going to be this function that is used to convert the virtual depth values in the Ratrix depth map into real depths.

From looking at the data in Figure 4.27 it appears that the data points above 700mm are erroneous. The fitted function to include this point had to go to the tenth polynomial to include this point, which will then also mean the function is likely to be fitting to noise in the other points. Therefore, it would be worthwhile removing the data points above 700mm and fitting a function with a lower polynomial value. Figure 4.28 shows this with a cubic function fitted to the data. As can be seen, this function fits relatively well to the data, so no order higher than cubic is necessary to fit to the data. The fitting function was also performed with a quadratic, which qualitatively did not fit the data as well as a cubic.

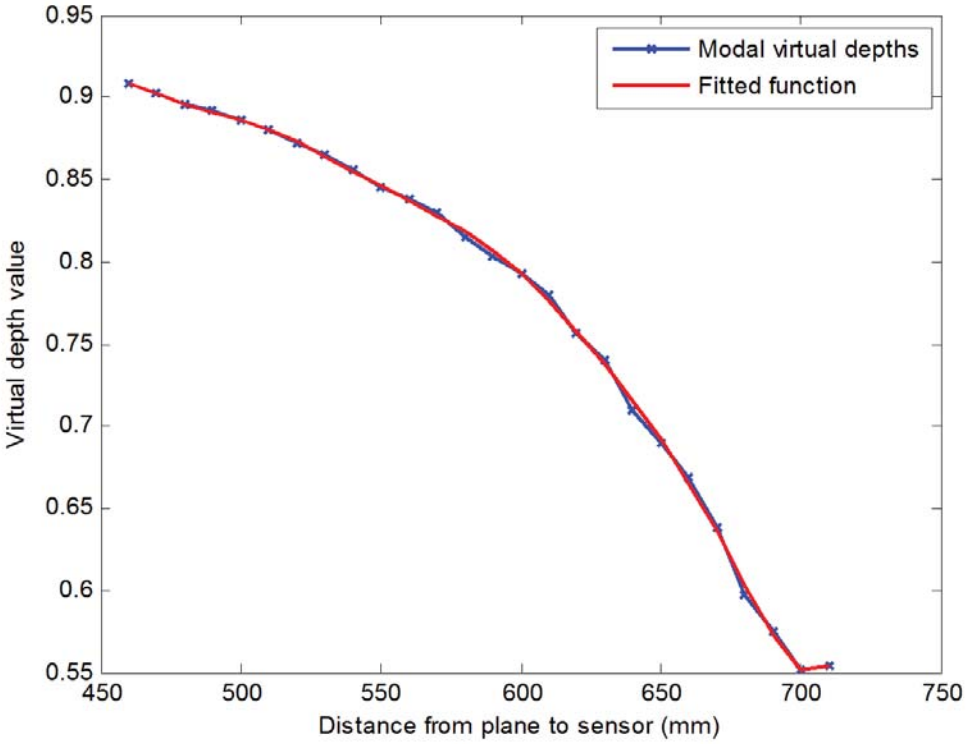


Figure 4.27: The same data as presented in Figure 4.21 but with the modal virtual depth values in blue and the fitted function to this data in red.

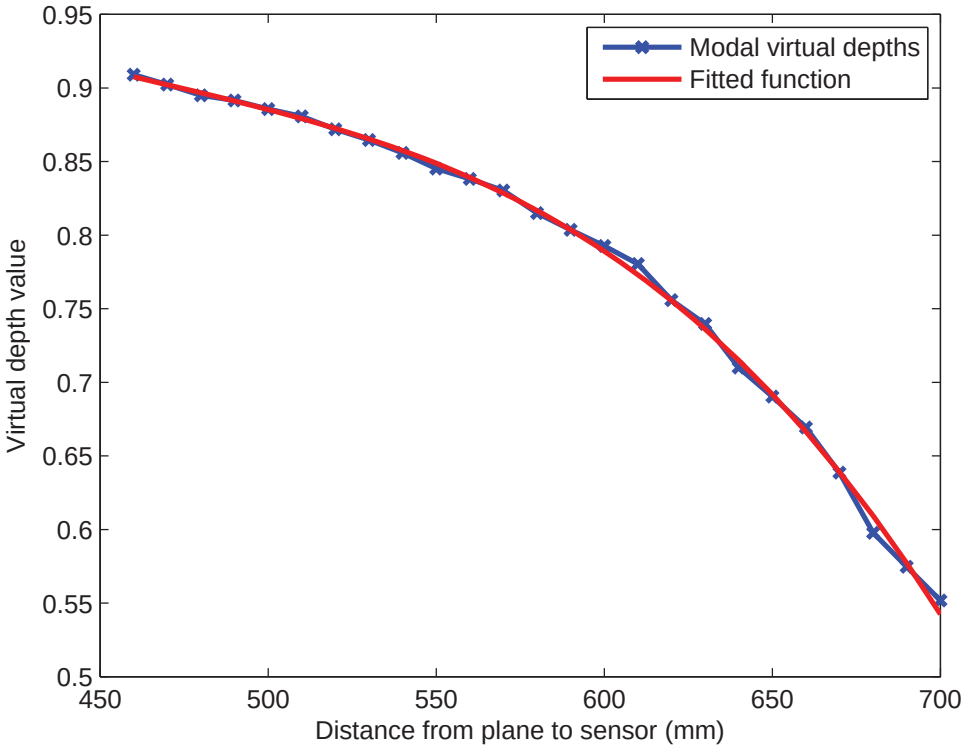


Figure 4.28: The same data as presented in Figure 4.21 but with the modal virtual depth values in blue and the fitted function to this data in red. The data above 700mm has been removed, and a cubic function fitted to the data.

4.4 Ground truth accuracy for sloped scene

Throughout the remainder of this chapter and the next, images are taken of flat planes at varying angles to the sensor of a Raytrix R11 camera. As this study is designed as a proof of concept that adding texture to a featureless scene improves the recovered depth rather than examining the achievable accuracy of a plenoptic camera at recovering depth, some estimations were taken to speed up the data gathering due to the limitations set out in Section 4.1. One of these, which is investigated further by Zeller *et al.* [100], is the unknown distances between the main lens and the microlens array, and the microlens array and the sensor. Raytrix do not provide information on these distances, but an estimation of the location of the placement of the sensor in the camera housing can be made to within 5mm through physical inspection. The plane was then placed at distances of around 700mm away from the camera, so the error relates to 0.7% of the total distance which was deemed acceptable to achieve the aims set out for the experiment. With respect to the angle of the sloped scene, this was measured with respect to a plane parallel to when the imaged plane was at 0° to an accuracy of 0.25°. The plane of the sensor in the Raytrix camera was measured to within the same degree of accuracy, but it is possible that these will become compound errors so the accuracy is to within 0.5°.

It is also assumed throughout the study that the plane is both flat and is a Lambertian surface. However, in reality the flatness of the plane is accurate on a millimetre scale. Relating this to imaging of the retina, which is neither of these things, it does not provide a good representation for what imaging of the retina may be like. However, these assumptions have been made to simplify the experiments and to allow for more results to be taken within the time restraints set out in Section 4.1. The retina does also have features, but as shown in Figure 4.1 there are large, relatively featureless, regions of the retinal surface which do not register depth on a plenoptic depth map. It is for this reason why imaging a featureless plane can be a first step towards plenoptic imaging of the retina.

It is also assumed that pose-related intensity changes do not affect the image formation. As the plane is flat, and the illumination from the project maintained at the same angle as the plane, it can be assumed that the plane is illuminated uniformly throughout the experiments. The intensity of the light reaching the camera, as the plane is assumed to be Lambertian, will then decrease as the angle is increased but it is assumed that the intensity of light reaching the camera when at different angles is negligible.

4.5 Definition of field of view

The following section refers to the field of view of the camera. For clarity, it is worth noting how the field of view is defined in these circumstances. The field of view is the physical size of the world which is observable from the camera for a given distance from the sensor. The field of view in these circumstances increases as a plane parallel to the sensor moves further away from the camera. This can be visualised in Figure 4.34.

4.6 Calculating the angle of a sloped scene

The aim of this next investigation is to gain quantitative depth of a scene. In the study a simple flat sloped scene was chosen. This is because the entire scene geometry is known, so a quantitative comparison between calculated and actual depth can be performed. The experimental set up is the same as shown in Figure 4.18 except at angles from parallel to the sensor up to 45° at intervals of 5°. The pattern in Figure 4.15(b) is projected with an Optoma PK201 LED Pico Projector. The lens on the Raytrix has an aperture of $f/8$ to match the f-number of the microlens array, and a focal distance of 0.75m. The plane when parallel to the sensor is placed at a distance of 0.71m from the sensor.

In order to calculate a quantitative angle of a flat sloped scene from the depth maps, the following steps need to be performed:

1. Perform the virtual to real depth conversion as described in Section 4.3 for the chosen lens parameters and Raytrix calibration file.

2. Measure the field of view (FOV) of the camera when parallel to the sensor.(*Note*: For simplicity, it is assumed that the change in FOV is insignificant.)

3. Image the slope.

4. Take the mean of each column in the depth map. Due to the geometry of the experiment, each column represents a single depth in real space.

5. Use the conversion function from point 1 to convert the average virtual depths into real depth values.

6. Plot the mean column values from the depth map on a graph.

7. Fit a straight line to the data. This has been performed using the POLYFIT function in MATLAB, but any straight line fit method would suffice.

8. The angle can then be calculated from the fitted straight line by using trigonometric functions by the change in depth with the measured field of view.

A selection of the graphs from which the angles were calculated can be seen in Figure 4.29 to Figure 4.32 with the complete set of the graphs in Appendix A1. A table with the results can be seen in Table 4.3.

From analysing the data in Figures 4.29 to 4.32 and the data in Table 4.3, the depth maps show an increase in slope angle when the slope increases from 0° to 45°. However, the difference between the real and calculated angle can reach as high as 6.9°. This is a significant error if attempting to perform quantitative studies on the angle of objects. An important note to make is the shape of the graph in Figure 4.29. The method used to calculate the slope involves fitting a linear line to the data, however for the 0° data it clearly does not follow a linear shape. As the graph should produce a flat line, it is a

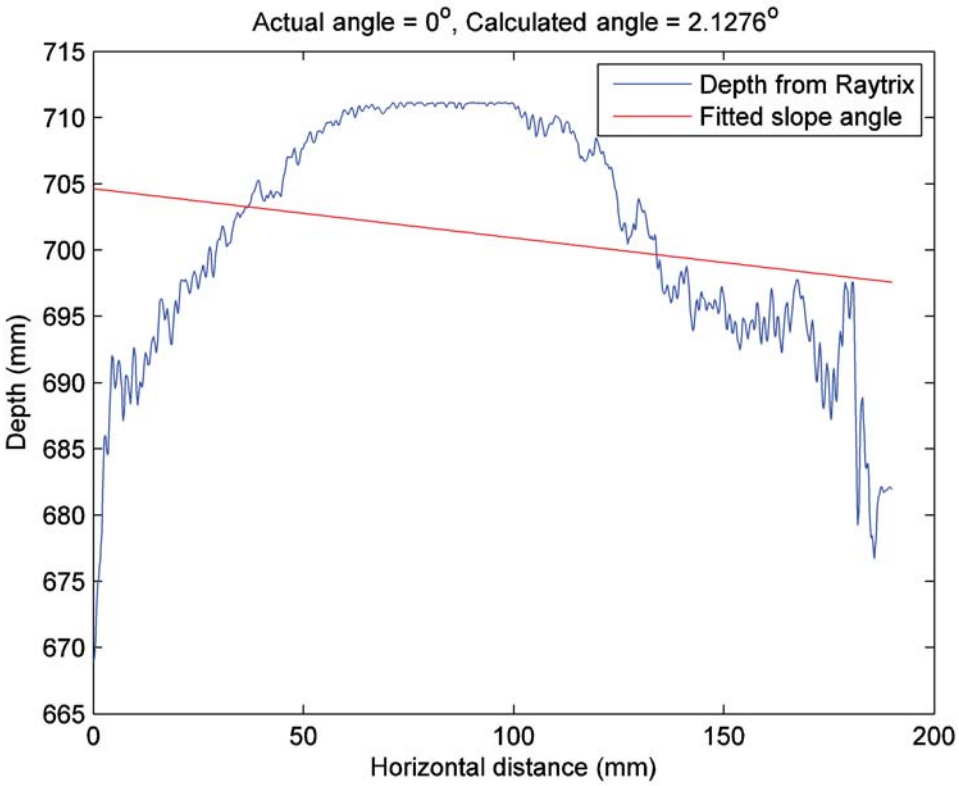


Figure 4.29: A graph showing how the angle of the flat sloped scene was calculated when at 0° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths. The red line shows the linear fit to the data.

Real angle	Calculated angle	Error	Percentage difference (%)
0°	2.13°	-2.13°	-
5°	6.02°	-1.02°	20.46
10°	6.39°	3.61°	36.08
15°	12.49°	2.51°	16.72
20°	16.75°	3.25°	16.23
25°	19.48°	5.52°	22.09
30°	25.57°	4.43°	14.76
35°	29.51°	5.49°	15.69
40°	34.89°	5.11°	12.77
45°	38.10°	6.90°	15.33

Table 4.3: The results from imaging the flat sloped scene with the Raytrix R11. The real angle is the actual angle of the slope, the calculated angle is what is outputted from the method described in this section, the Real - Calculated is the difference between the two values and the percentage difference is the difference as a percentage of the real angle.

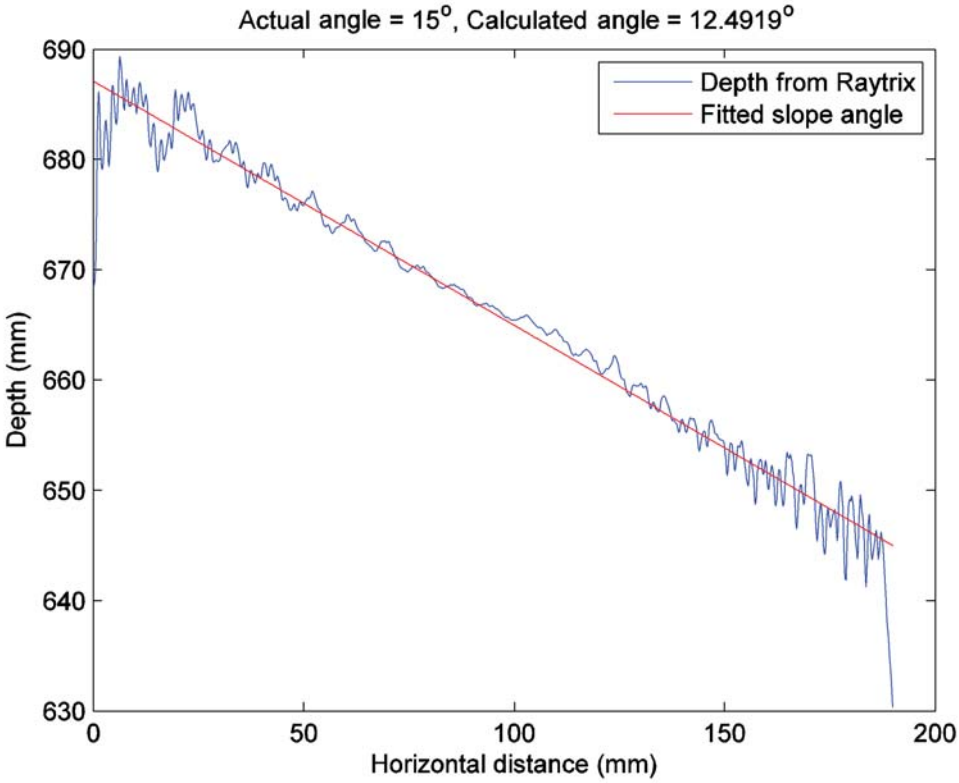


Figure 4.30: A graph showing how the angle of the flat sloped scene was calculated when at 15° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths. The red line shows the linear fit to the data.

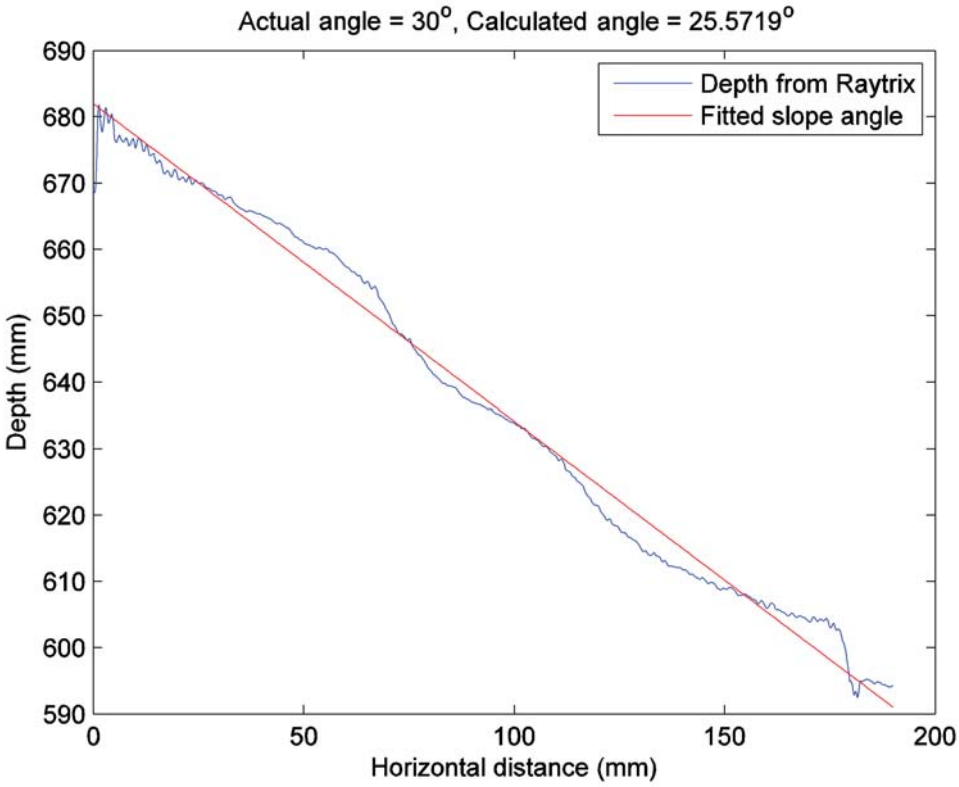


Figure 4.31: A graph showing how the angle of the flat sloped scene was calculated when at 30° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths. The red line shows the linear fit to the data.

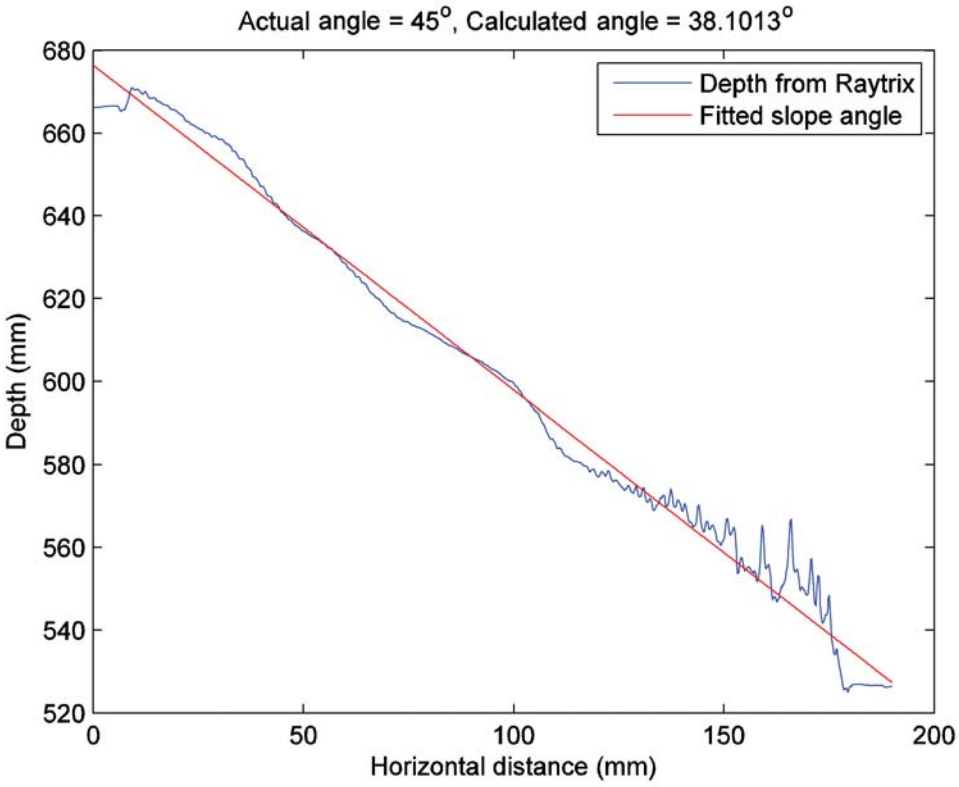


Figure 4.32: A graph showing how the angle of the flat sloped scene was calculated when at 45° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths. The red line shows the linear fit to the data.

reasonable assumption that the only accurate data is between the horizontal position of 60 - 100mm. This also correlates with the set-up of the experiment, as when the plane was parallel to the sensor it was at a distance of 0.71m, which the graph accurately measures in this small section. It could therefore be proposed that only the central section of the depth map accurately records the depth of the scene. This does seem logical, as when initially calibrating the Raytrix it is usually the outer microlenses that are misaligned. Also aberrations will be larger in the peripheries of the main lens. To see if this is actually the case, the previous experiment was repeated but with the depth maps cropped to only include the area which correctly recovered the depth from Figure 4.29. The resulting data can be seen in Table 4.4.

Real angle	Calculated angle full image	Calculated angle cropped image
0°	2.13°	0.82°
5°	6.02°	4.19°
10°	6.39°	17.30°
15°	12.49°	12.44°
20°	16.75°	13.50°
25°	19.48°	18.87°
30°	25.57°	32.57°
35°	29.51°	31.64°
40°	34.89°	37.54°
45°	38.10°	33.54°

Table 4.4: The results from imaging the flat sloped scene with the Raytrix R11. The real angle is the actual angle of the slope, the calculated angle is what is outputted from the method described in this section and the calculated angle cropped image is the outputted angle when the depth map is cropped to only include the pixels that correctly calculated the depth in Figure 4.29.

Looking at the data in Table 4.4, by cropping the image the calculated angles do not appear to have improved in accuracy, indicating that this is not the most significant error in the system. As stated in the method previously, the field of view of the camera was assumed to be consistent for all angles to simplify the calculations. However, this could

be causing errors as at the larger angles the field of view will decrease. Therefore, a more thorough investigation into how much this affects the results is required.

4.6.1 Correcting the field of view

To investigate whether the change in field of view (FOV) does make a significant difference to the results, the FOV must first be known for all depths with the selected lens parameters. To do this, the field of view is measured from a distance at the back depth of field to the front at regular intervals. The data is then plotted, and a function is fitted to the data as before using the MATLAB function POLYFIT. This will then provide a relationship between depth and the field of view. As only the field of view in the plane that changes (x-axis in Figure 4.18), it is only this field of view that needs to be fitted. The measured data and the fitted line can be seen in Figure 4.33 for when the main lens has a focal distance of 0.75m.

From initially looking at the data, it appeared to follow a straight line. Therefore a linear function was fitted to the curve, which can be used to relate any depth to the field of view. The new complete field of view can then be calculated by using the new function that relates depth to FOV to get the FOV at the maximum and minimum depths. The FOV for any angle can then be calculated by taking the average as shown in Equation 4.1. Figure 4.34 can be used to help visualise why the FOV of a sloped scene can be calculated by this equation.

$$\mathrm{FOV} = \frac{\mathrm{FOV}_{\mathrm{zmin}}}{2} + \frac{\mathrm{FOV}_{\mathrm{zmax}}}{2} \tag{4.1}$$

The method described earlier in Section 4.6 has now been modified to take into account the field of view correction, and now reads as:

1. Perform the virtual to real depth conversion as described in Section 4.3 for the chosen lens parameters and Raytrix calibration file.

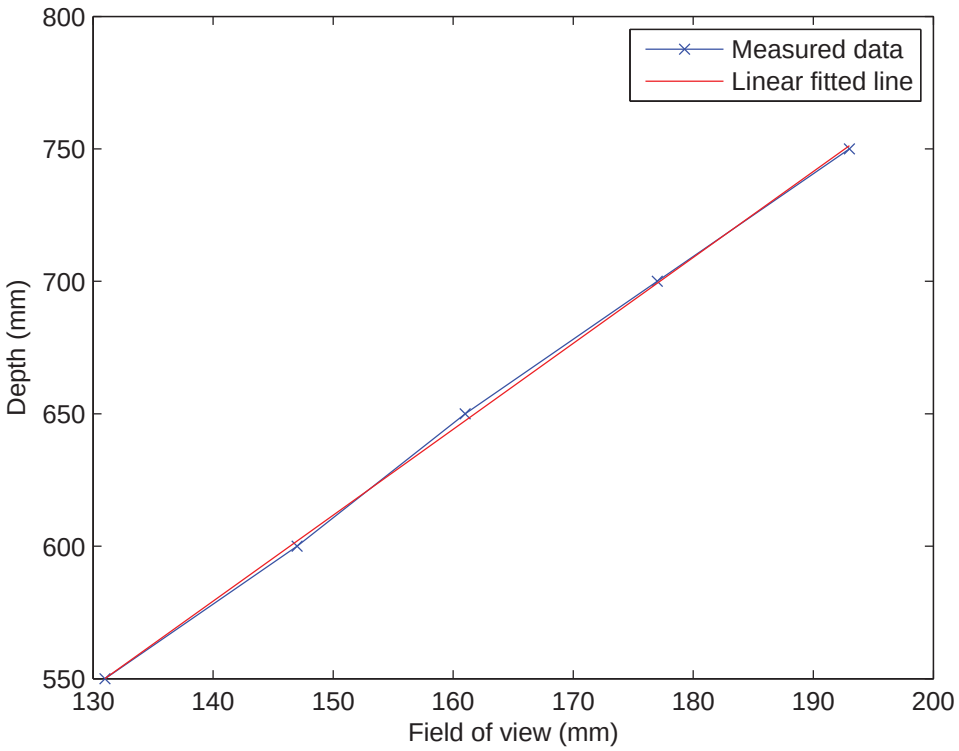


Figure 4.33: A graph showing the relationship between distance from sensor to the plane and field of view. The blue line indicates the measured data, and the red line the linearly fitted function.

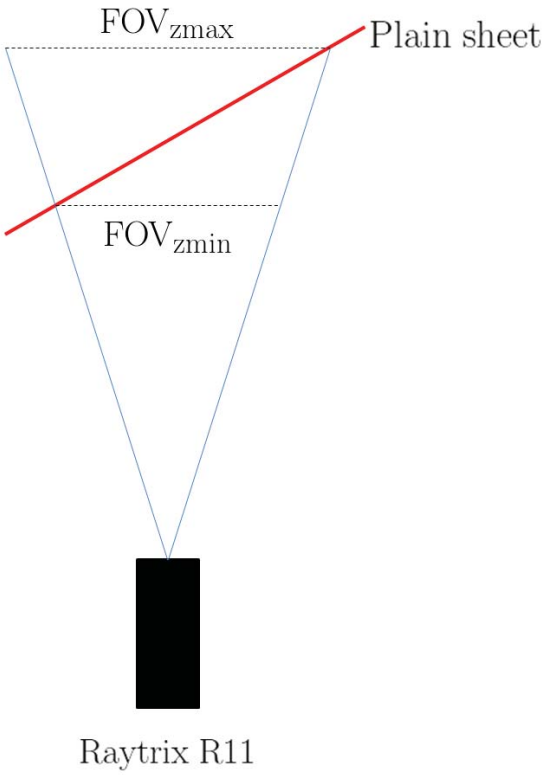


Figure 4.34: An image to help visualise the field of view of a sloped scene. It can be calculated as the mean of the field of views at maximum and minimum depth. (See Equation 4.1)

2. Measure the field of view (FOV) of the camera at different depths in the depth of field of the camera.

3. Fit a straight line to the data to gain a relationship between depth and FOV for any depth.

4. Image the slope.

5. Take the mean of each column in the depth map. Due to the geometry of the experiment, each column represents a single depth in real space.

6. Use the conversion function from point 1 to convert the average virtual depths into real depth values.

7. Plot the mean column values from the depth map on a graph.

8. Fit a straight line to the data.

9. Using the FOV and depth relationship, find the FOV for the largest and smallest depths.

10. Calculate the new FOV by following Equation 4.1.

11. The angle can then be calculated from the fitted straight line by using trigonometric functions by the change in depth with the measured field of view.

By following the new method, some of the inaccuracies from the calculated angle should be corrected. The results of using the FOV correction compared to without can be seen in Table 4.5, with the graphs for 0°, 15°, 30° and 45° in Figures 4.35, 4.36, 4.37 and 4.38 respectively, with the complete set of graphs in Appendix A2.

From analysing the results, significant improvements can be seen when using the corrected field of view method. These improvements are greater at larger angles, as predicted theoretically because the change in field of view is greater with a larger angle. The

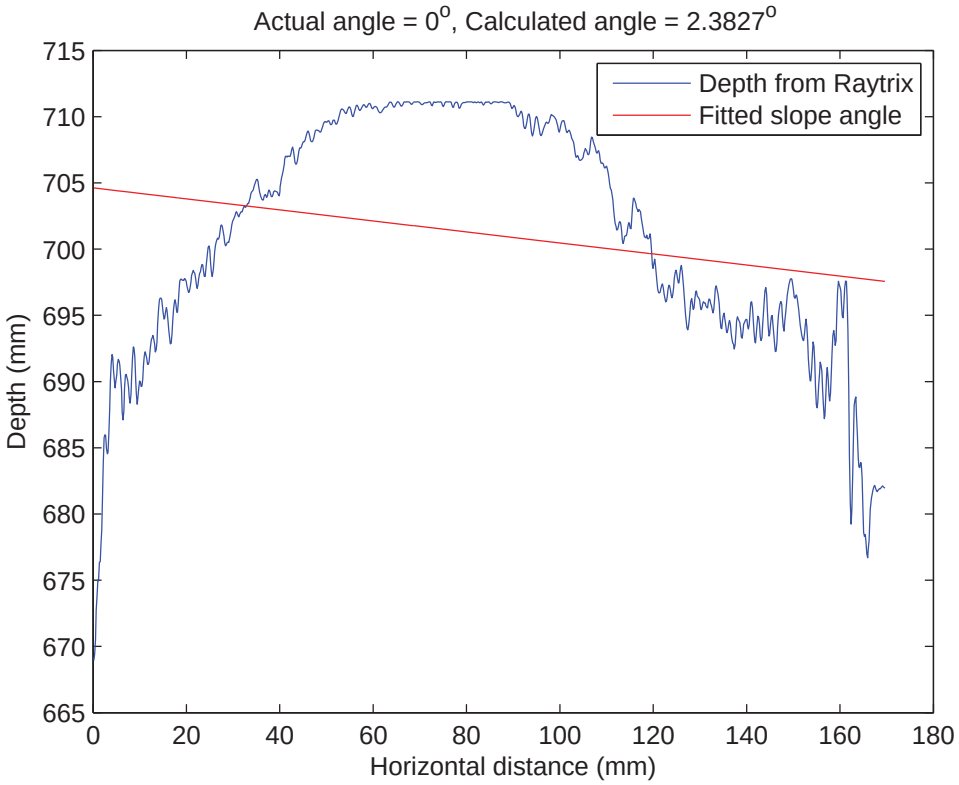


Figure 4.35: A graph showing how the angle of the flat sloped scene was calculated when at 0° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths. The red line shows the linear fit to the data.

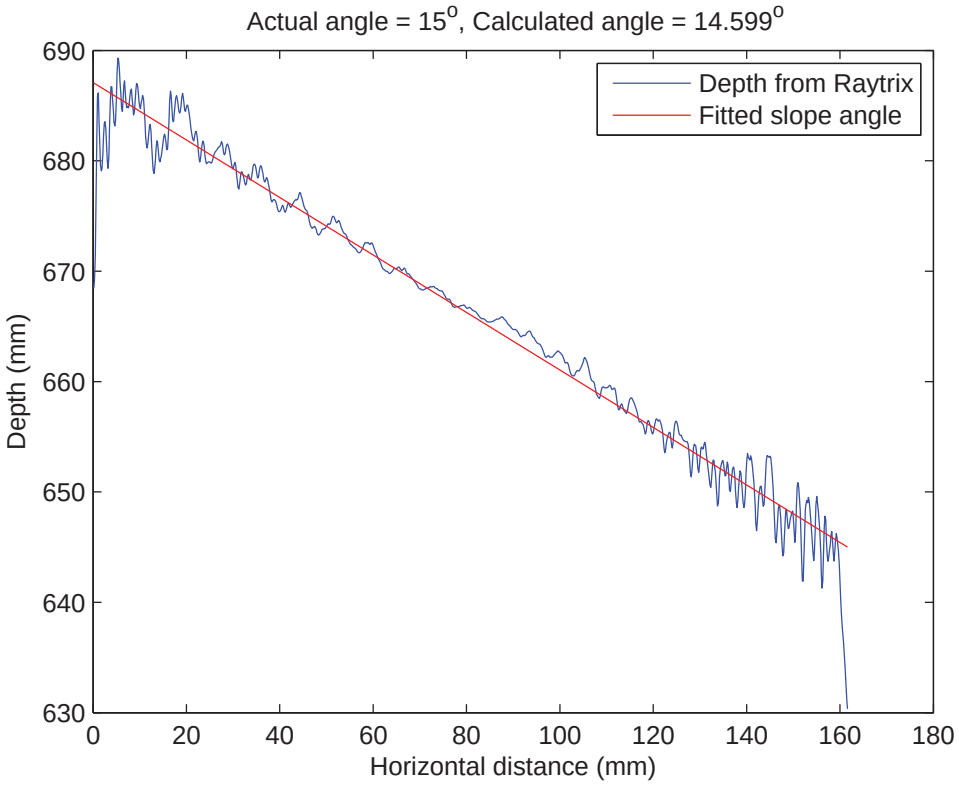


Figure 4.36: A graph showing how the angle of the flat sloped scene was calculated when at 15° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths. The red line shows the linear fit to the data.

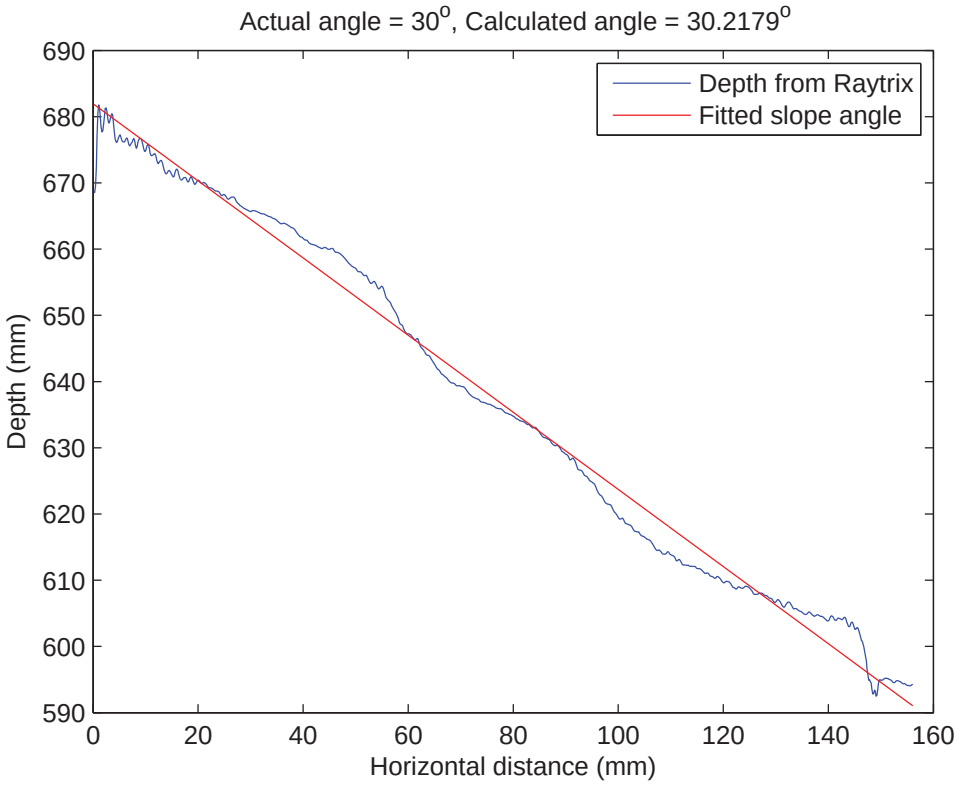


Figure 4.37: A graph showing how the angle of the flat sloped scene was calculated when at 30° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths. The red line shows the linear fit to the data.

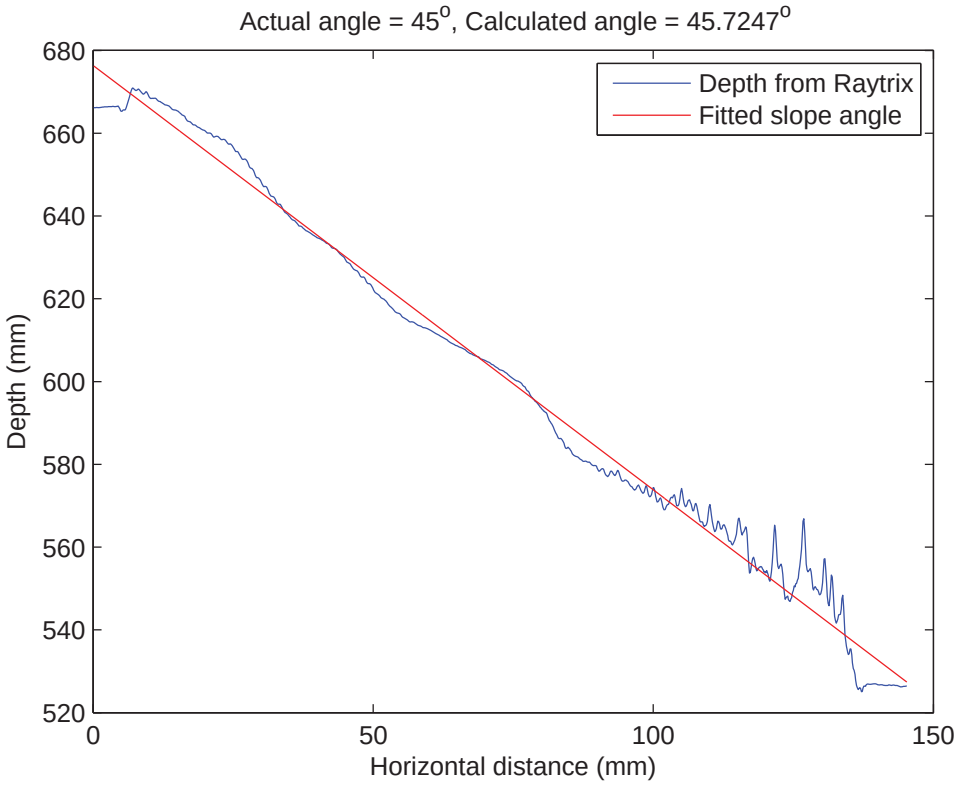


Figure 4.38: A graph showing how the angle of the flat sloped scene was calculated when at 45° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths. The red line shows the linear fit to the data.

Real angle	Calculated angle before FOV correction	Calculated angle after FOV correction	Real - Calculated angle after FOV correction
0°	2.13°	2.38°	-2.38°
5°	6.02°	6.96°	-1.96°
10°	6.39°	7.43°	2.57°
15°	12.49°	14.60°	0.40°
20°	16.75°	19.73°	0.27°
25°	19.48°	22.89°	2.11°
30°	25.57°	30.22°	-0.22°
35°	29.51°	35.13°	-0.13°
40°	34.89°	41.60°	-1.60°
45°	38.10°	45.72°	-0.72°

Table 4.5: The results from imaging the flat sloped scene with the Raytrix R11. The real angle is the actual angle of the slope, the calculated before FOV correction is the angle when the field of view of the camera is assumed to be constant at all depths, and the calculated angle after FOV correction is the angle when the changing FOV at different depths is taken into consideration.

difference between the real and calculated angles at 45° has decreased from 6.90° to 0.72°, almost a ten fold improvement in accuracy. It can therefore be concluded that including the field of view correction to the methodology for calculating quantitative angles from a plenoptic camera can produce accurate results with errors in the order of a degree.

4.6.2 Removing the high frequency noise

In the preliminary studies for this chapter in Section 4, it was seen that qualitatively high frequency noise was introduced to a flat plane when a regular pattern of black dots was used to create texture. An investigation into the origins of this noise and into whether it is caused by the regularity of the pattern will be investigated in Section 5.2.1, however it is also important to calculate whether this noise has any effect on the calculated angle of a slope. This has been investigated by taking the graphs from Figures 4.35, 4.36, 4.37 and 4.38, then applying a low pass filter to the Fourier transform of these plots. By then

applying an inverse Fourier transform, the sharp high frequency noise is then removed. The filter used can be seen in Figure 4.39. Following on from step 8 in the method described in Section 4.6.1, the angle of the new fitted line can be calculated then compared to both the real angle of the slope and the calculated angle before the high frequency noise has been removed. The new plots with the aforementioned angles can be seen in Figures 4.40, 4.41, 4.42 and 4.43 for angles of 0°, 15°, 30° and 45° respectively.

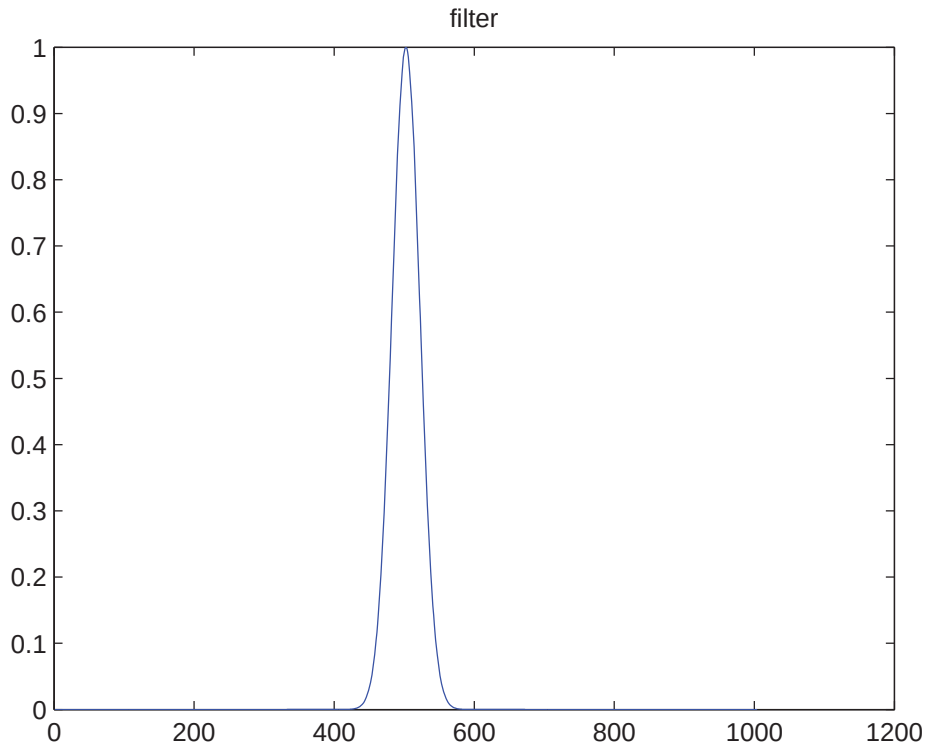


Figure 4.39: A graph showing the filter used to remove the high frequency noise. The y-axis of the graph is the value to multiply the inverse Fourier transform of the angle graphs by, and the x-axis is the pixel number in the x-axis of the angle graph.

By assessing these graphs, some strong conclusions about the effect of the high frequency noise can be made. In all the graphs, the calculated angle does not change by any significant amount, with the largest being for 15° at 1.1°. On the graphs for 15°, 30° and 45° the calculated angle once the high frequency noise had been removed are all less than before. This is mainly due to the downward peak at low horizontal distances and an upward peak at high horizontal distance values which have been created by the Gaussian

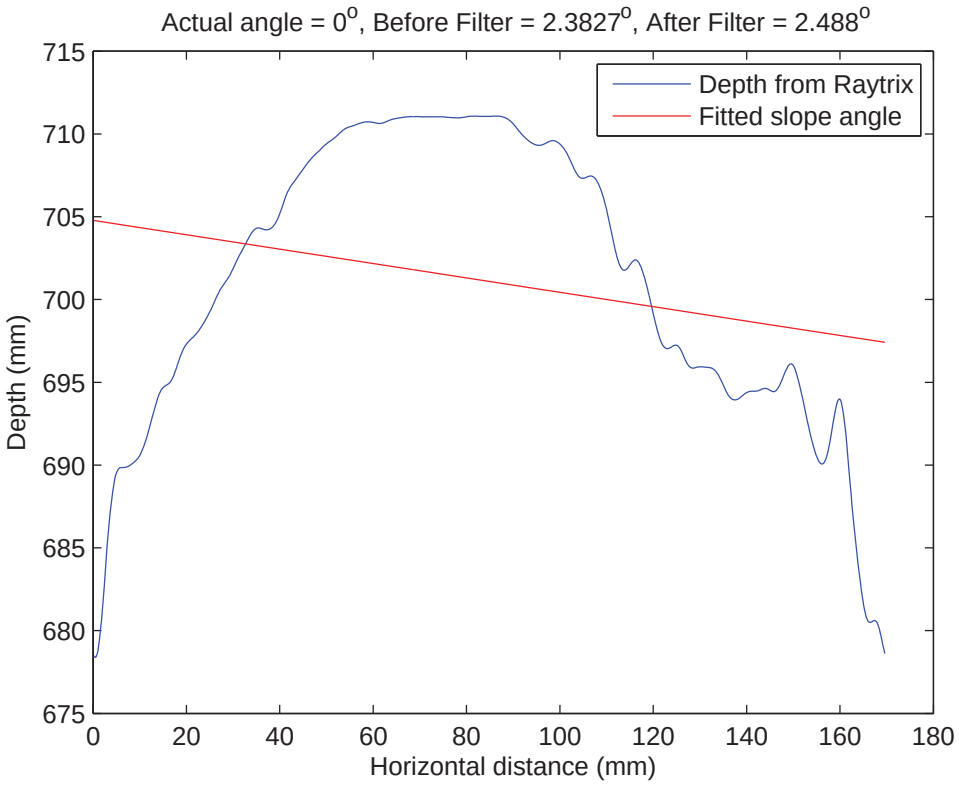


Figure 4.40: A graph showing how the angle of the flat sloped scene was calculated when at 0° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths with the high frequency noise removed. The red line shows the linear fit to the data.

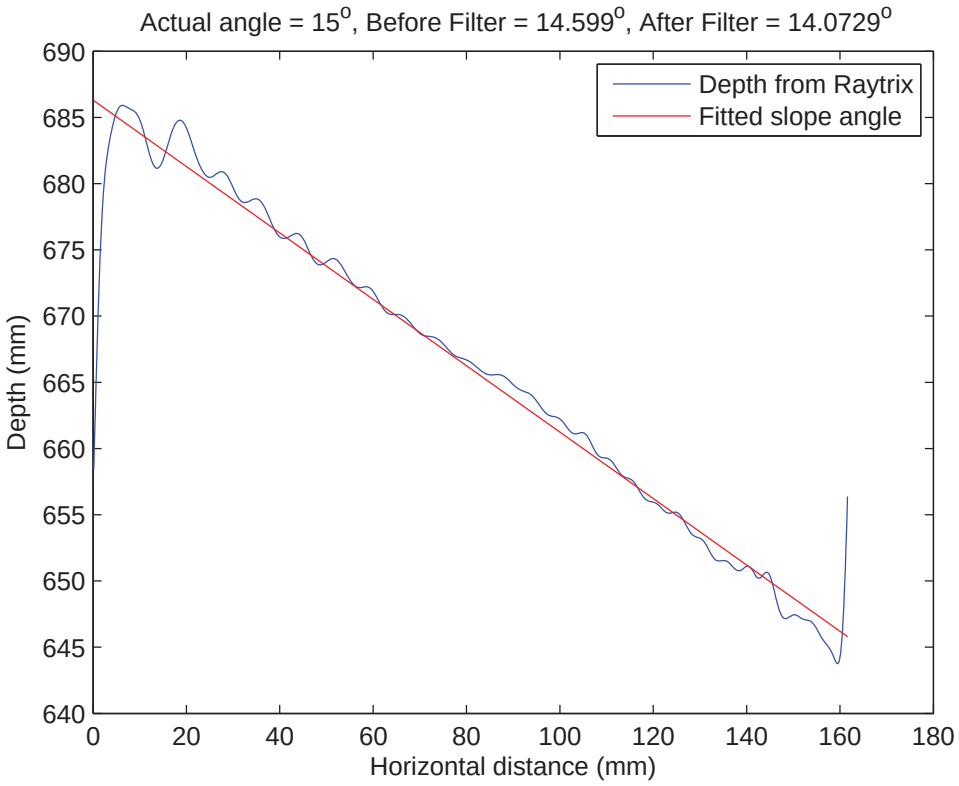


Figure 4.41: A graph showing how the angle of the flat sloped scene was calculated when at 15° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths with the high frequency noise removed. The red line shows the linear fit to the data.

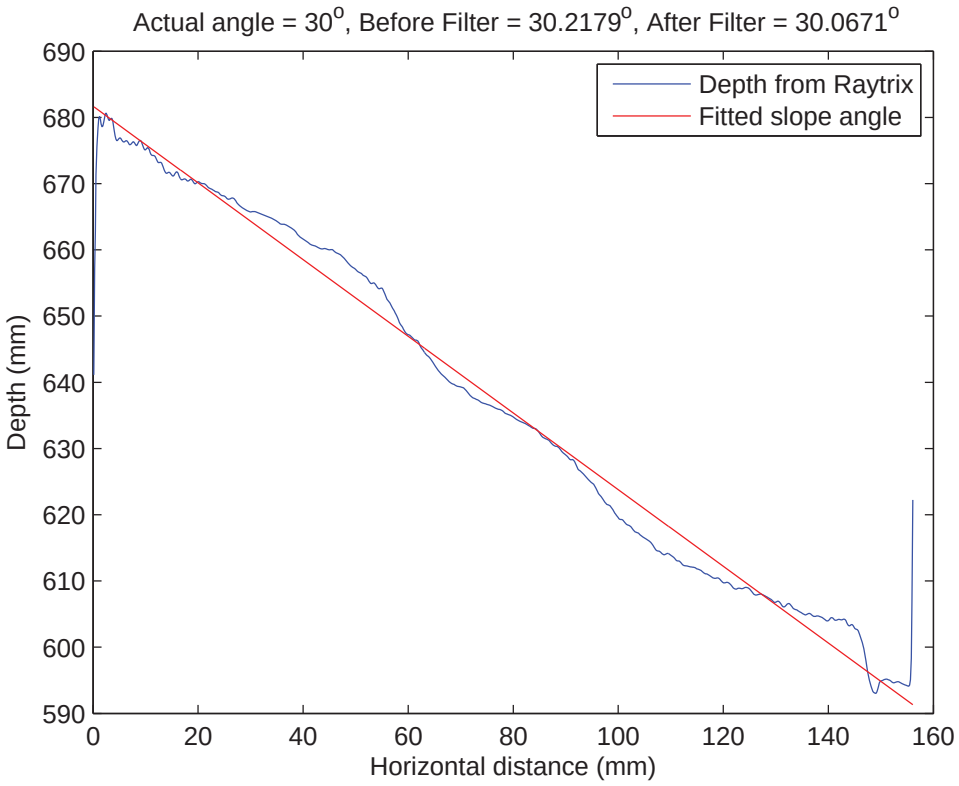


Figure 4.42: A graph showing how the angle of the flat sloped scene was calculated when at 30° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths with the high frequency noise removed. The red line shows the linear fit to the data.

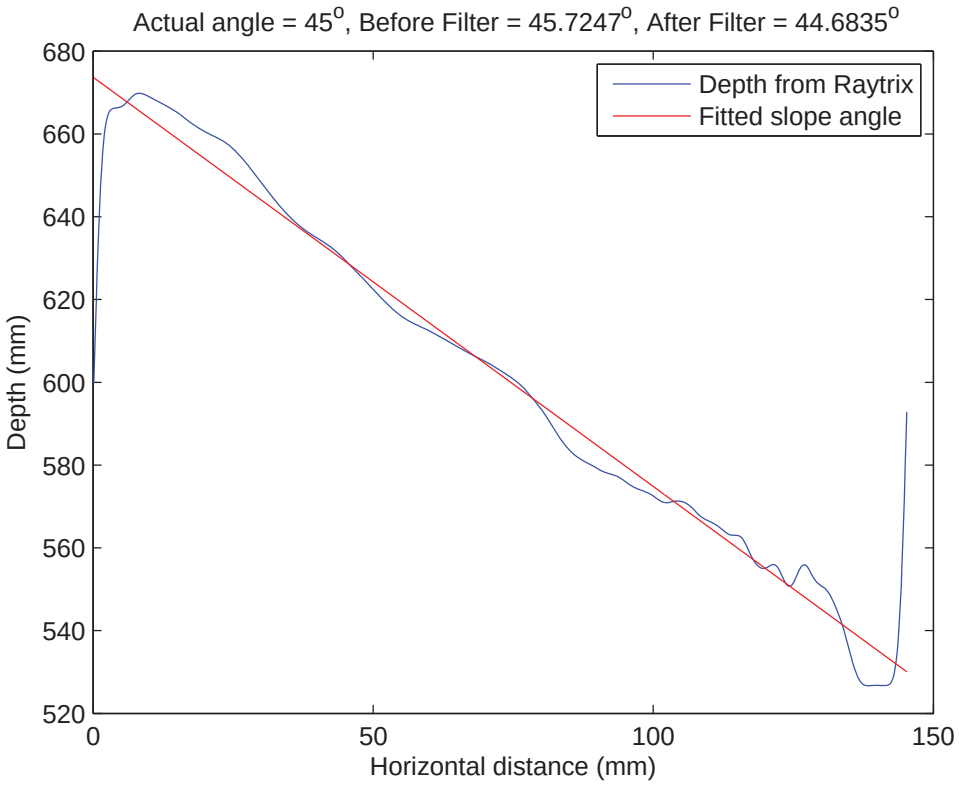


Figure 4.43: A graph showing how the angle of the flat sloped scene was calculated when at 45° to the plane of the sensor. The blue line shows the real depth values calculated from the Raytrix virtual depths with the high frequency noise removed. The red line shows the linear fit to the data.

filter used to remove the high frequency noise. Even with this however, the calculated angles are the same as with the high frequency noise within a small error. It can therefore be concluded that although the source of this noise is still unknown, it does not affect the accuracy of the calculated angle.

4.7 Fitting a plane to the data

As mentioned in Section 4.1 the methods presented throughout this chapter incur some errors due to the reliance on the mechanical alignment of the camera and plane. Methods to overcome this problem have been presented by Zeller *et al.* [100] where the orientation of the imaged plane was calculated with respect to the camera. This was achieved by using many reference points on a chessboard with known square sizes. This method is not applicable to the experiments presented throughout this chapter, as the features are not inherent in the scene but are projected onto it. This projection in itself creates features of unknown size, as the size of the pattern depends on the distance between the projected and the scene. The orientation of the pattern also depends on the mechanical alignment of the projector, adding to the error which is trying to be eradicated. It is, however, possible to fit a plane to the data rather than simply assume that the alignment is correct and take the average value of each column which if perfectly aligned would all be at the same distance from the camera.

4.7.1 Method of fitting a plane

To determine the plane which best fits the data, a least-squares method was chosen. For a plane, there are the general equations:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

$$z = \frac{1}{c}(ax_0 + by_0 + cz_0) - \left(\frac{a}{c}x + \frac{b}{c}y\right)$$

Let,

$$C = \frac{1}{c}(ax_0 + by_0 + cz_0), A = -\frac{a}{c}, B = -\frac{b}{c}$$

Then for $M = [C, A, B]$ the plane can be fitted by solving the overdetermined equation $Z = X \times M$ where,

$$X = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}$$

$$Z = \begin{bmatrix} z_1 & z_2 & z_3 \end{bmatrix}$$

The equation for the plane to fit the data can then be solved by performing the backslash function in Matlab, which follows a procedure of linear regression as,

$$M = X \backslash Z$$

To show an example of a plane fitted to the data, Figure 4.44 shows a surface plot of the data at 15° in blue with the fitted plane in red. It is worth noting, the depth values are virtual depths outputted from the Raytrix, and the scale for the x and y axes are in pixels. The following section will look into the misalignment of the sensor and the plane when imaging, which should help shed some light on some errors incurred through the method for calculating the angle of a plane described earlier in this chapter.

4.7.2 Using a plane to identify misalignment

The methods described through Section 4.6 could potentially incur some errors if the angled slope and the plane of the sensor are not aligned, as it is assumed that each column of pixels in the depth map relate to a single depth. In reality this may not be

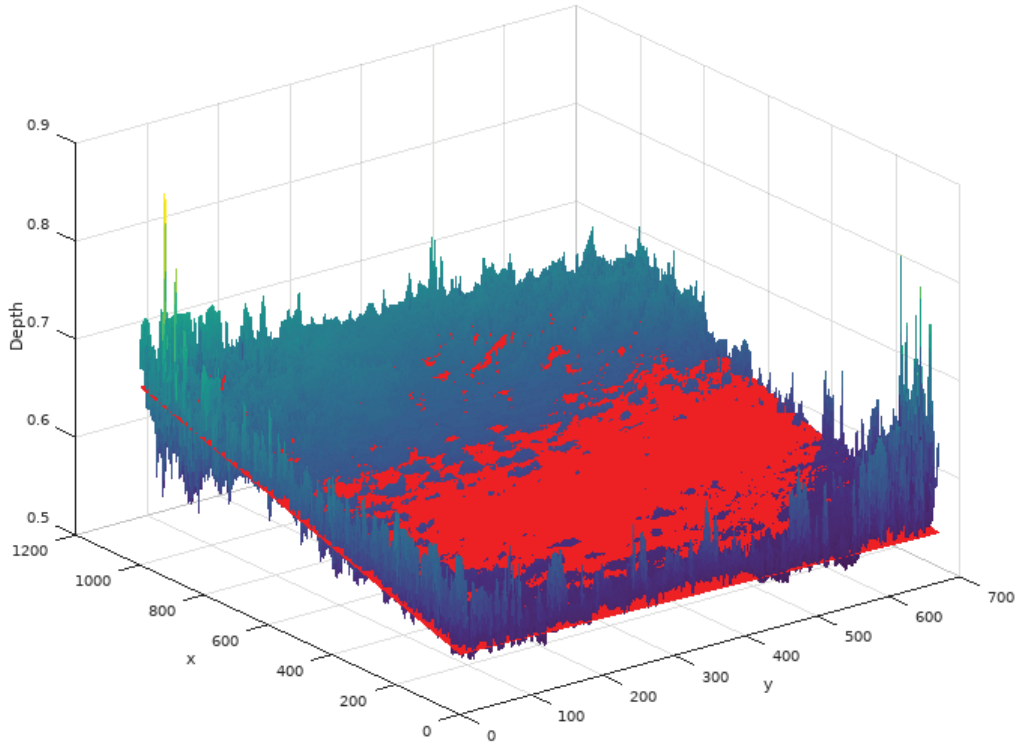


Figure 4.44: An image showing a plane fitted to data collected at 15°. The x and y axes are in pixels, and the depth is the virtual depth from the Raytrix R11 camera.

correct, as this requires a high level of mechanical alignment. To visualise this, a heat-map image of the fitted plane for data at different angles is taken, with varying values depending on the calculated depth for that pixel. Figures 4.45, 4.46, 4.47 and 4.48 show these heat-map images for the planes fitted to depth maps when at angles of 0°, 15°, 30° and 45° respectively. The x and y axes on the images represent pixels on the depth map and the values in the image are the real depth of that point in space in mm from the sensor to the plane.

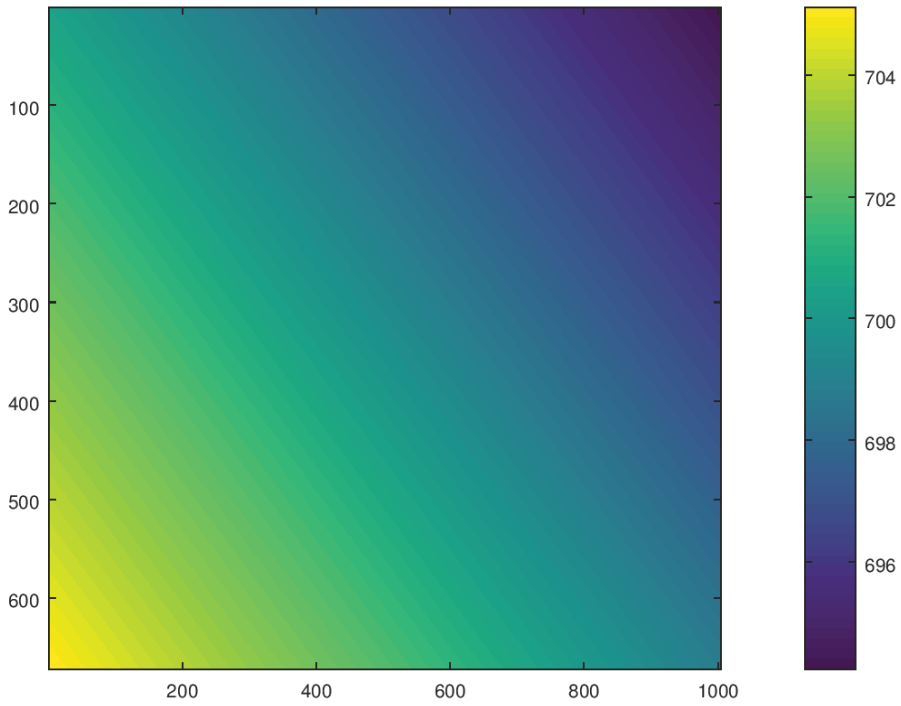


Figure 4.45: A heat-map image highlighting the mechanical misalignment of the planes of the slope and the sensor when at 0° from each other. The x and y axes are in pixels, and the values in the image represent the real depth of the fitted plane in mm.

By analysing the images in Figures 4.45 to 4.48 it is clear that the planes of the slope and sensor were not aligned in exactly the same plane. This is evident by in all figures at any given value on x, the depth value decreases as y increases. It is also worth noting that the apparent misalignment decreases as the angle increases. This is not thought to be that the error is changing, just that it has less of an affect on the image as the range of

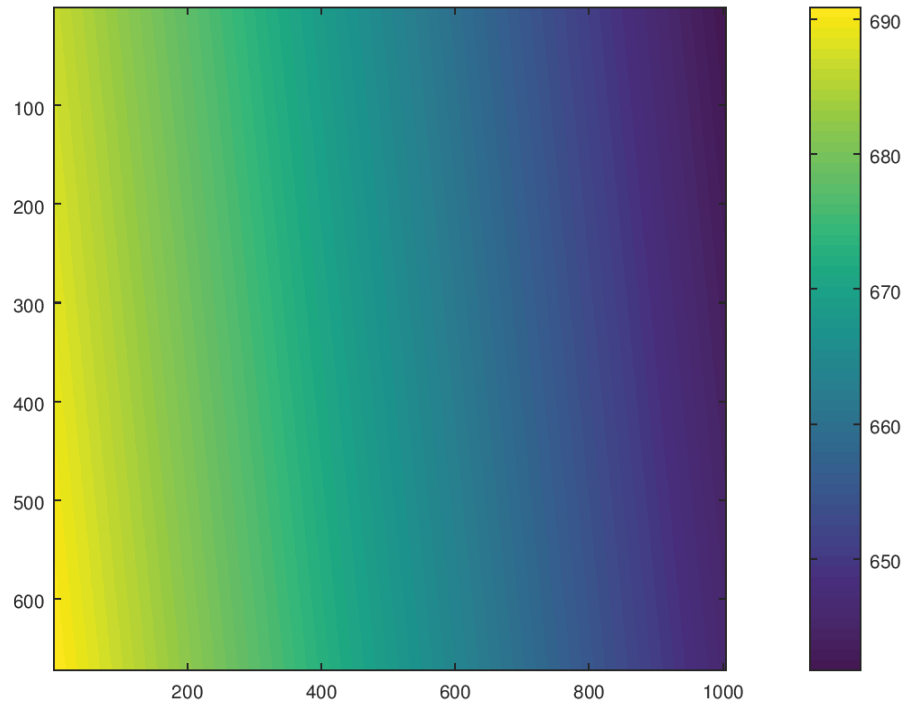


Figure 4.46: A heat-map image highlighting the mechanical misalignment of the planes of the slope and the sensor when at 15° from each other. The x and y axes are in pixels, and the values in the image represent the real depth of the fitted plane in mm.

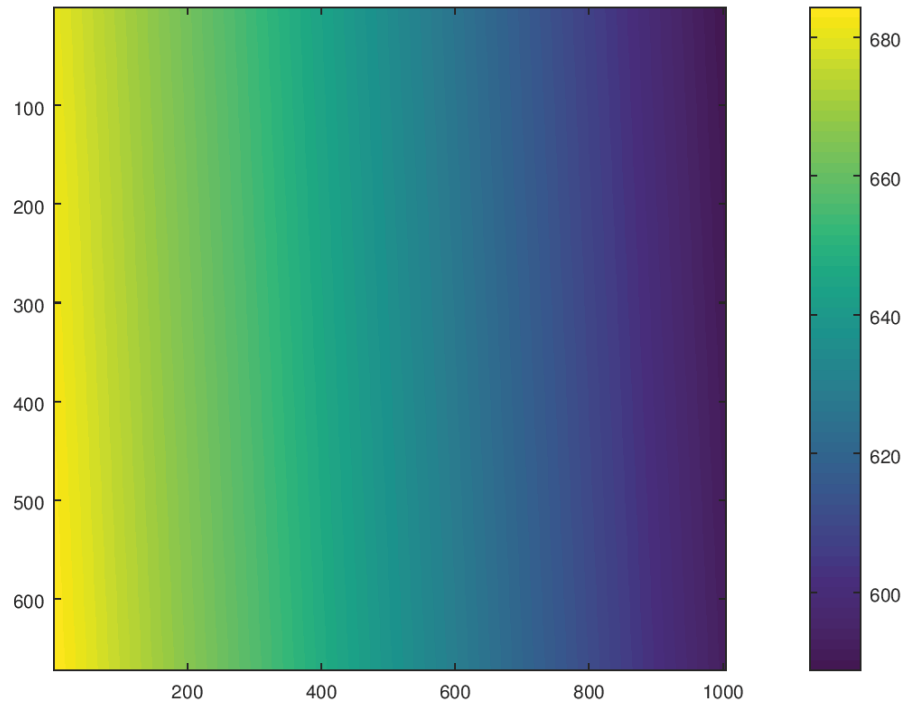


Figure 4.47: A heat-map image highlighting the mechanical misalignment of the planes of the slope and the sensor when at 30° from each other. The x and y axes are in pixels, and the values in the image represent the real depth of the fitted plane in mm.

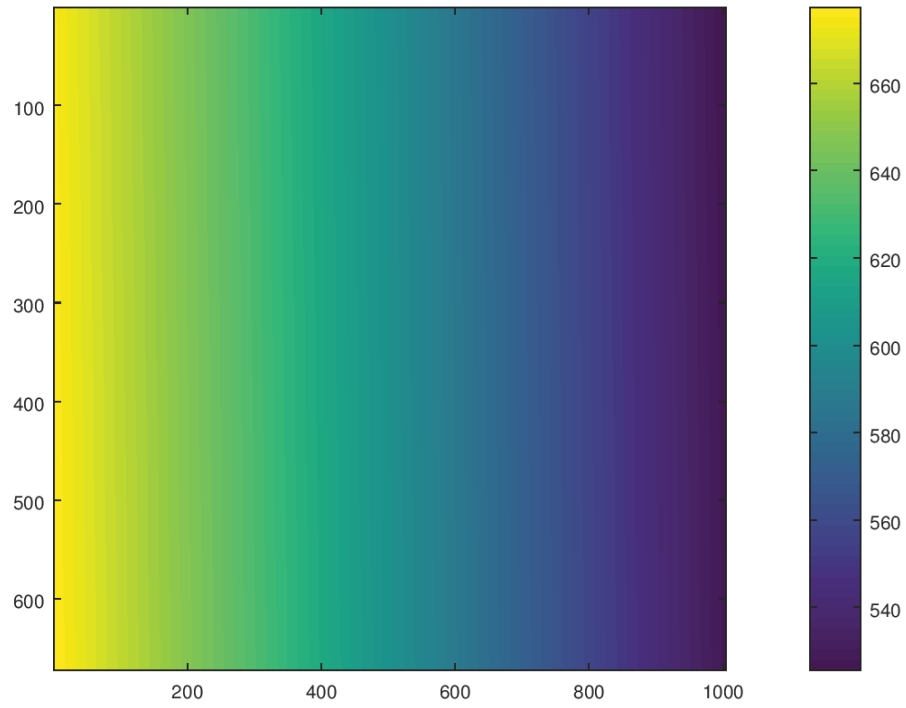


Figure 4.48: A heat-map image highlighting the mechanical misalignment of the planes of the slope and the sensor when at 45° from each other. The x and y axes are in pixels, and the values in the image represent the real depth of the fitted plane in mm.

depths present increases as the angle increases. However, this phenomena is not explored further.

4.7.3 Calculating the angle of a slope from a fitted plane

By following the method described in Section 4.7.1, a plane can be fitted to the data for a sloped surface. This has advantages over the previous method for calculating the angle of a sloped surface earlier in this chapter, as it takes away some errors incurred due to misalignment between the plane of the surface and the sensor, as highlighted in Section 4.7.2. The angle of this plane can then be used to see how accurately the Raytrix computes the angle of a sloped scene. This can be done by using the equation to fit the plane to the depth map data outlined in Section 4.7.1. As the plane is fit using a linear regression, the formula for the plane follows;

$$z = ax + by + c$$

By using the coefficient before the x axis, this can be used to calculate the angle of the plane by taking tan^{-1} of this coefficient. To ensure that the angle is correct, the pixel values which are currently on the x axis need to be converted into real distances, as per the method in Section 4.6.1. Figures 4.49, 4.50, 4.51 and 4.52 show surface plots of the fitted planes with the calculated angles when the plane is at 0°, 15°, 30° and 45° to the plane of the sensor respectively. Table 4.6 shows the results for all the angles from 0° up to 45° in intervals of 5°.

By looking at the results in Table 4.6 it can be shown that the calculated angle from the fitted plane is close to the real angle up to 25°. After this point, the calculated angles are significantly greater than the actual angle of the plane. This trend will not be examined here, but later in Section 5.3 in the following chapter.

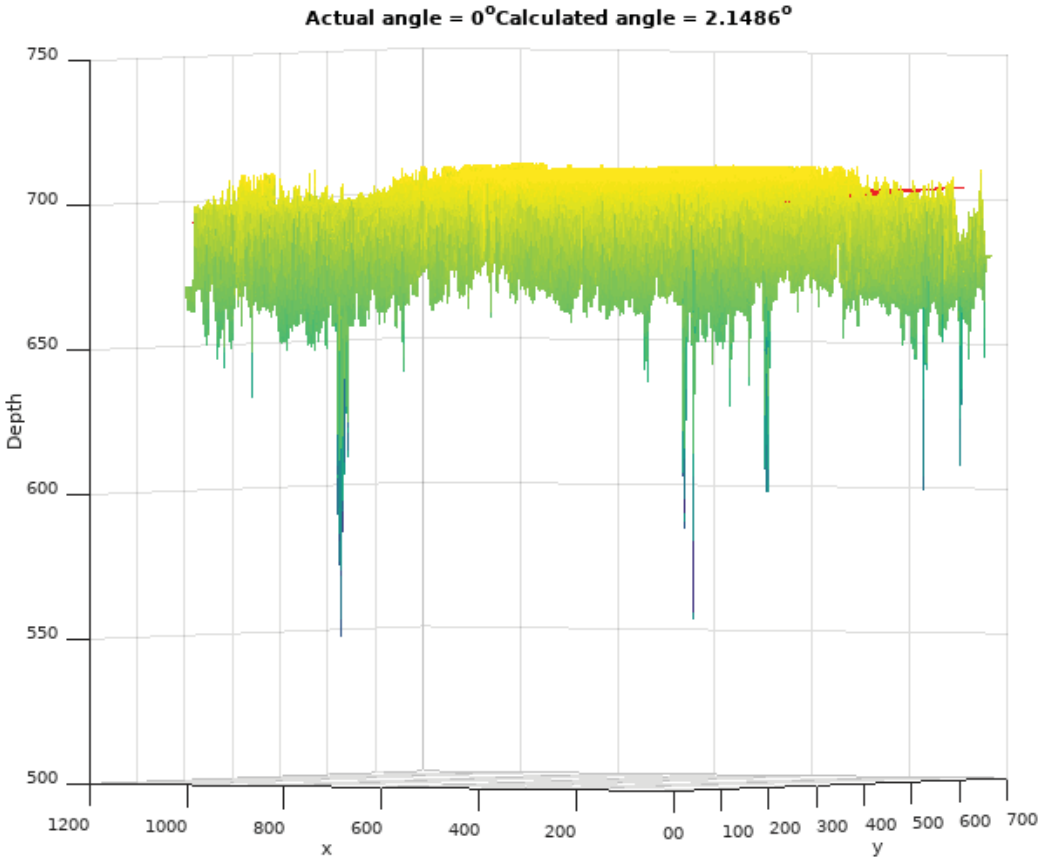


Figure 4.49: A surface plot showing the fitted plane and calculated angle from this plane when at 0° to the sensor. The x and y axes are in pixels, and the depth is in mm.

Real angle	Calculated angle	Difference
0°	2.15°	2.15°
5°	7.16°	2.16°
10°	8.11°	1.89°
15°	15.67°	0.67°
20°	21.28°	1.28°
25°	24.85°	0.15°
30°	35.45°	5.45°
35°	39.86°	4.86°
40°	50.33°	10.33°
45°	60.71°	15.71°

Table 4.6: The results from imaging the flat sloped scene with the Raytrix R11. The real angle is the actual angle of the slope, the calculated angle is the outputted angle from the fitted plane method described throughout this section.

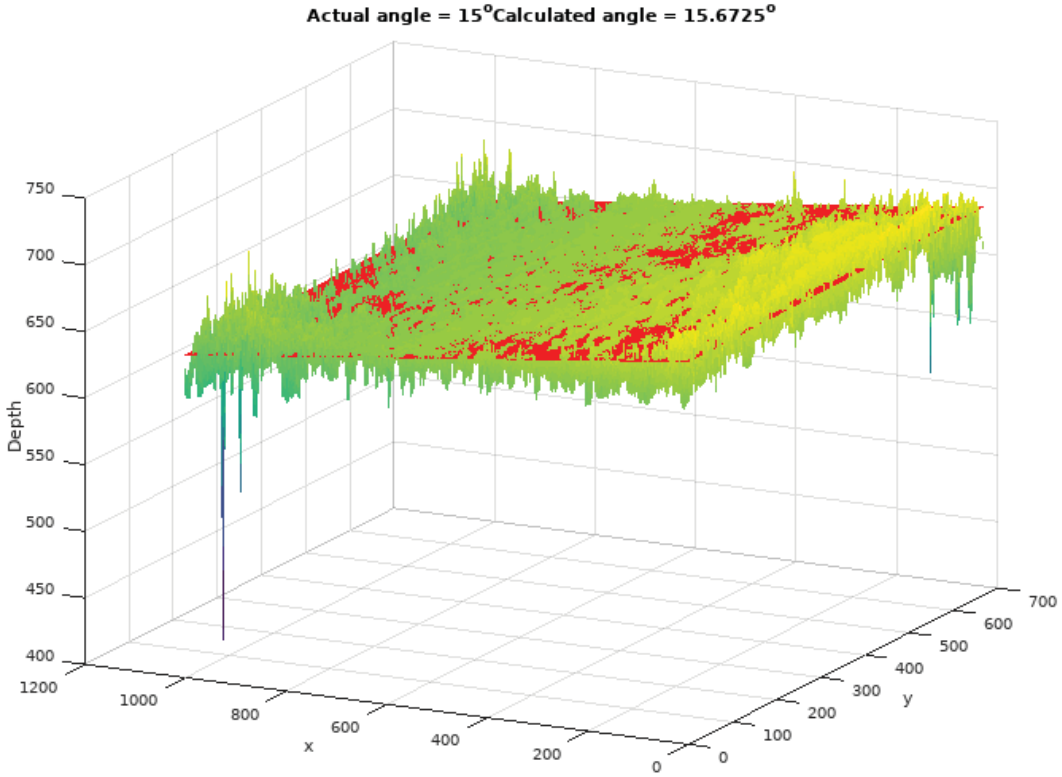


Figure 4.50: A surface plot showing the fitted plane and calculated angle from this plane when at 15° to the sensor. The x and y axes are in pixels, and the depth is in mm.

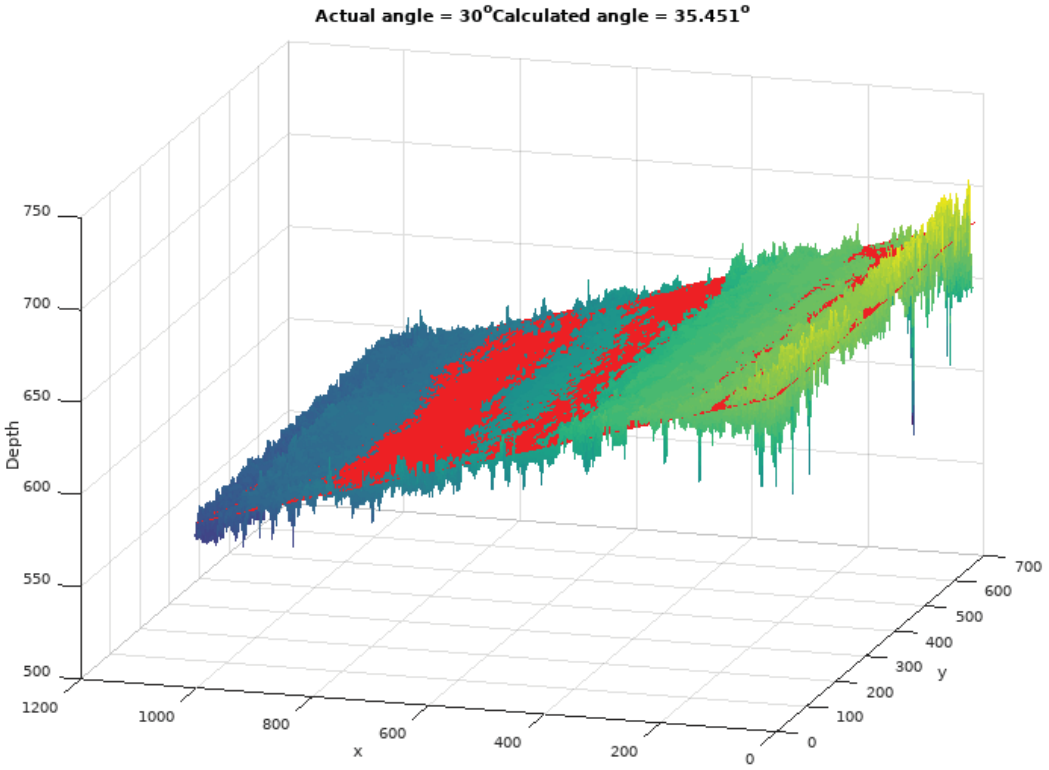


Figure 4.51: A surface plot showing the fitted plane and calculated angle from this plane when at 30° to the sensor. The x and y axes are in pixels, and the depth is in mm.

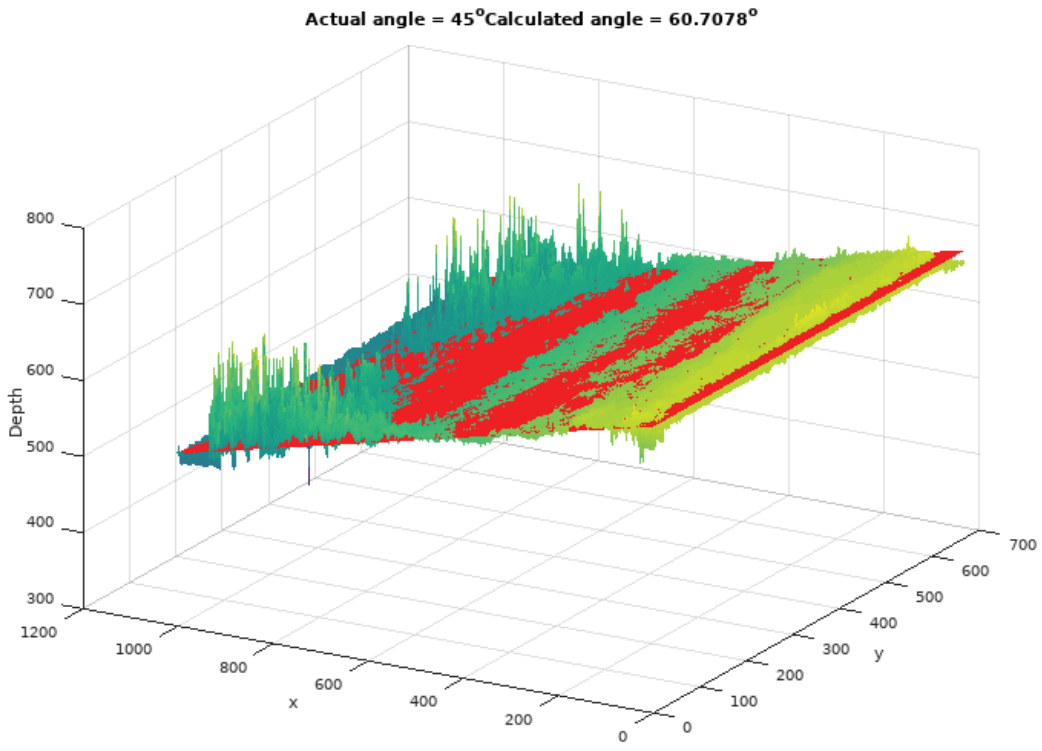


Figure 4.52: A surface plot showing the fitted plane and calculated angle from this plane when at 45° to the sensor. The x and y axes are in pixels, and the depth is in mm.

In this section, a linear plane was fitted to the depth maps produced from the Raytrix R11 camera. By doing so, it has shown that as predicted, since the scene and camera relied upon mechanical alignment, there were some errors inflicted because of this. The plane was then used to calculate the angle of the slope, with comparable results up to 25° with the methods previously described earlier in the chapter. Above 25° the results are much larger than the actual angle of the slope, but the reasons for this will be explored further in the following chapter.

4.8 Conclusion

In this chapter, the accuracy of recovered depth from a plenoptic camera in the absence of texture was explored. When there are few or no features on an object, incorrect information about its depth or shape is produced. However, it was proven that when a pattern was drawn on the objects the depth estimation was improved significantly.

Drawing on plain objects before being imaged could be destructive or not possible for certain objects, so the idea of projecting texture was proposed. Preliminary experiments showed promising results, as qualitatively the depth of a featureless sloped plane improved after being illuminated with a square grid of black dots by a laser projector. As some errors were apparent in the depth maps, a comparison between using a laser and an LED projector was explored. These results showed that the average error in a depth map is less when illuminated with an LED projector compared to a laser projector. It was therefore decided that from after this point the LED projector would be used to illuminate the scenes.

The difference between a scene already having texture present, and the features being projected onto the objects was then explored. It was found that the difference was insignificant as there was only a 1.2% average difference between the depth maps produced when the features were printed and projected onto the scene and a p-value of < 0.001 from a paired t-test. This provides strong statistical support for texture to be projected onto a scene. However, there are inherent limitations of using a projector. Firstly, by using an LED projector there is a range of depths which are in focus, and outside of this the pattern becomes blurred. There is also the thermal and mechanical stability of the projector to take into consideration, and a theoretical limit to the accuracy from projecting patterns (see Section 3.2.2).

A novel method to convert virtual to real depth was then devised. This involved creating a calibration function by imaging a flat plane at a series of known depths from the sensor. This needs to be done each time the lens on the plenoptic camera is changed as the virtual depth outputted from a Raytrix will correlate to a different depth in real space if the focus of the main lens is changed.

Attempts were then made to quantitatively assess the accuracy of a plenoptic camera to recover the angle of a known slope. Initially, results were promising but included errors of up to 6.9°. First attempts to improve the error included cropping the depth map, as it

was initially thought that the outer microlenses may increase the error due to aberrations or inaccuracies in the initial calibration of the Raytrix on the inbuilt software. However, this did not improve the results so a new idea was explored. A previous assumption that the field of view at all angles was examined, and decided that a more rigorous approach to calculating the field of view for all depths should be calculated. This was then applied to the same slope depth map data, and at large angles an improvement in accuracy by a factor of 10 was seen, with differences from the real angle as low as 0.13°. It is therefore demonstrated that a method to reconstruct the angle of a slope, with quantitative real depths, has been created. It is worth noting that as set out in Section 4.4 the accuracy of the angle of the slope can only be guaranteed to within 0.5°.

Some high frequency noise can be seen in most of the signals throughout the chapter. Although the origins of this noise will not be looked at until Chapter 5, the effect that this noise had on the calculated angle was investigated and shown to not have any significant effect on the angle.

Due to the methods for calculating the angle of the slope relying on the mechanical alignment of the camera and the plane, a virtual plane was fitted to the data to investigate if there was any error incurred because of this. The results showed that this was in fact the case. The fitted plane was then used to calculate the angle of the slope, with comparable results to the first method up to 25°. Above this, much larger errors were incurred which will be investigated further in the following chapter.

The main conclusions which can be drawn from this chapter are that texture does need to be present in a scene for a Raytrix R11 plenoptic camera to register depth. This texture can be inherent in the scene or projected onto it. Also throughout this chapter, the same grid of black dots was added to the scenes for consistency. However, this pattern was chosen at random and there is no evidence that this is the ideal pattern to help improve the depth accuracy from scenes with limited features. Therefore, the next chapter of

work will focus on examining if this pattern is the optimum for improving depth from a plenoptic camera, if a different pattern works better, or if it is insignificant.

Chapter 5

Testing the effect of different patterns on the accuracy of calculated depth

As shown in Chapter 4, the depth from a plenoptic camera can be improved when imaging a featureless scene by projecting a pattern. Throughout the chapter the same regular grid of black dots was projected onto the scene, however there has been no investigation of whether this is the optimum pattern to use. Throughout this chapter, different patterns will be assessed to try to determine whether there is an optimum pattern to use, and what features about the pattern affect the accuracy of the recovered depth.

The initial aim for this study was to identify theoretically the optimal pattern in simulation, then test this practically. This would have been performed using plenoptic simulation software which was being developed by Chris Meah at the University of Birmingham. Unfortunately, not enough progress had been made on the simulation software by the time this piece of work needed to be completed. It was therefore only completed practically, with the reasoning for the selection of patterns used explained below.

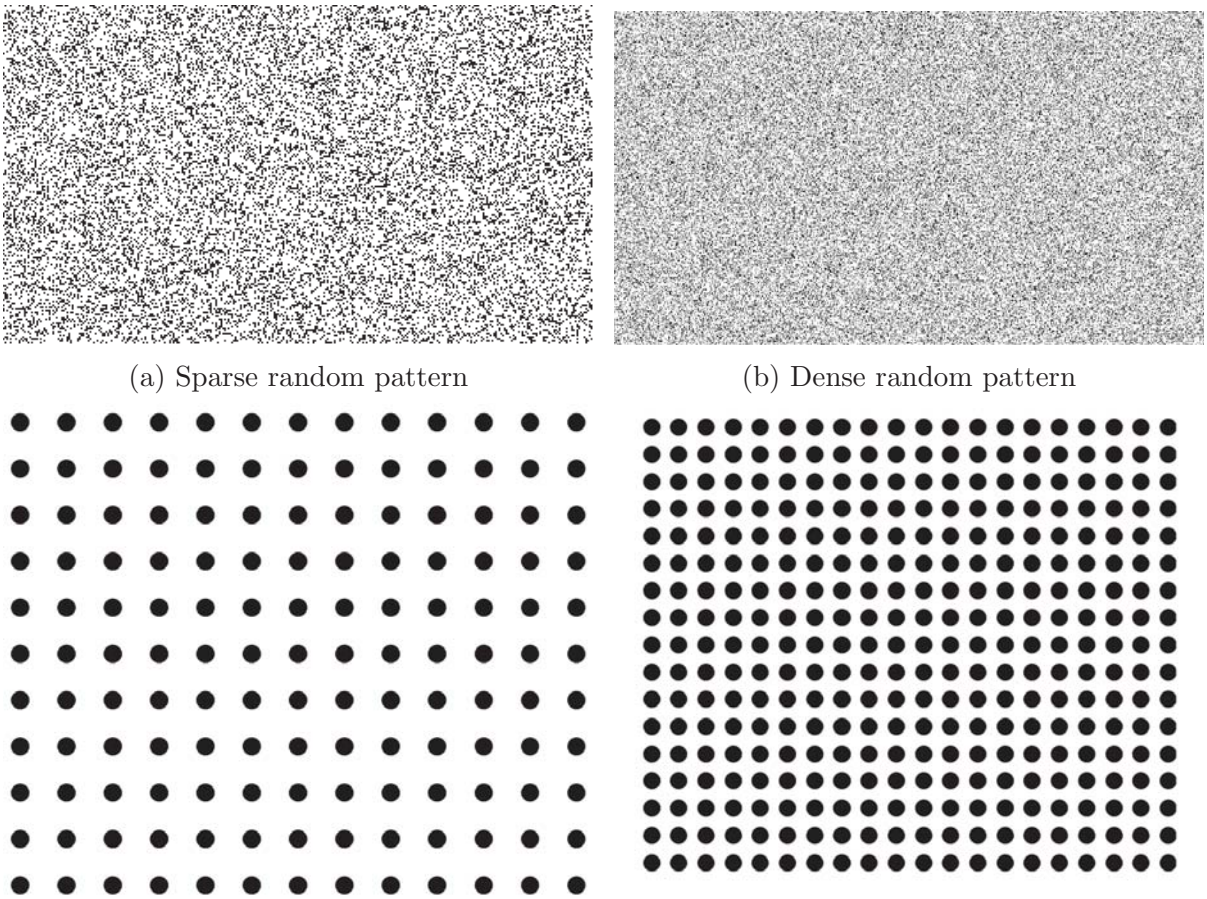
5.1 Selection of patterns

Following on from the previous chapter, a Raytrix R11 plenoptic camera will be used to produce the depth maps. This is advantageous as once the calibration is calculated, the camera will produce depth maps that can be compared to assess the accuracy of the recovered depth with different patterns. However, a drawback from using this apparatus is that direct communication between the Raytrix software and Matlab is not possible. In order to analyse a depth map, it first has to be saved manually from the Raytrix software and then uploaded into Matlab to be analysed. The initial plan for this piece of work was to have an automatic analysis of the accuracy of the depth map without any manual intervention. This would have then allowed the development of an evolutionary algorithm to adapt the pattern until it is optimised to produce the most accurate depth recovery. As this has not been possible, the patterns had to be selected manually.

To decide upon which patterns to select, it was important to hypothesise about what factors are most likely to affect the reconstructed depth. To do this, it is important to understand how the depth algorithms function. As previously described in Section 1.11.3, depth from a plenoptic system relies on the registration of features between neighbouring microimages. It therefore follows that to calculate depth accurately and reliably, there must be sufficiently many common features between microimages to enable them to be registered. This can then be seen to be dependent on two main factors; the regularity of the pattern and the sparsity of the features in the pattern. If a regular repeated pattern is used, different features of the pattern may be indistinguishable, leading to an incorrect correspondence between the images. If a very sparse pattern is used, there may be insufficient repeated features in neighbouring microimages to enable an accurate correspondence to be determined. Alternatively, if the pattern is not sparse enough, and the features are smaller than the size of the pixel on the sensor, then this may cause errors when trying to calculate the depth.

To test which of these factors affects the reliability of the depth calculation, 4 patterns have been chosen. The first pattern is the grid of regular dots that has been used throughout Chapter 4. This has been chosen as it has a repetitive pattern and is relatively dense. The second chosen pattern is the same as before but more sparse, *i.e.* the dots are the same size as previously but the white space between the dots is larger. The third and fourth patterns chosen were randomly generated images with 25% black pixels and 75% white space. The difference between the two is the density and size of the pixels, with the sparse image having 191×334 pixels, and the dense image having 821×1436 pixels for the same total image size. All patterns are shown in Figure 5.1. The four patterns have been chosen as they are dense and regular, sparse and regular, dense and unrepeated, and sparse and unrepeated. The two random patterns are equally sparse by definition (black/white ratio). By analysing the accuracy of the depth when imaging a featureless scene of known geometry when projected with each of these images, it should give an indication into how the regularity and sparsity of the pattern affect the recovered depth accuracy.

The choice of pattern can be related back to the literature in Section 3.2, which is split into either sequential or single-shot techniques. As plenoptic imaging is a modality which gains 3D information from a single acquisition, it seems intuitive to follow the single-shot techniques. Of these, there are many different options for indexing of the pattern with the overall aim being to have a uniquely identifiable subwindow, or subimage when referring to a plenoptic raw image. The chosen patterns throughout this chapter follow the literature by using a pseudo-random binary array and a 2D grid of dots. The grid patterns were chosen over the stripe indexing due to the similarity to the grid of the microlens array. As mentioned previously, should the plenoptic simulation software have been completed within the research group earlier, these patterns, and many others, would have been simulated first. However, due to practical limitations only the four mentioned have been used.



(a) Sparse random pattern

(b) Dense random pattern

(c) Sparse regular pattern

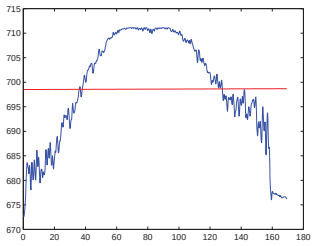
(d) Dense regular pattern

Figure 5.1: Images showing both the dense and sparse random and regular patterns used throughout the chapter.

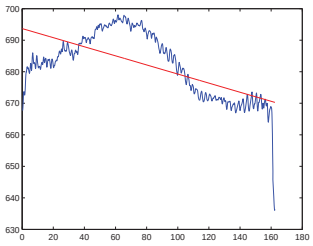
5.2 Imaging a flat plane with various patterns

To investigate the effect different patterns have on the accuracy of the depth, the experimental set-up is the same as in Figure 4.18 where a flat featureless sheet is imaged at angles between parallel to the sensor and 45° to that plane, except that the intervals are at 5° instead of 15°. This has been chosen as a protocol for quantitatively assessing the accuracy has already been designed (see Section 4.6.1). The results for the regular sparse, regular dense, random sparse and random dense at all angles can be seen in Figures 5.2, 5.3, 5.4 and 5.5 respectively. From these results, the error between the real physical angle of the plane and the angle computed from the Raytrix depth map has been collated and presented in Figure 5.6.

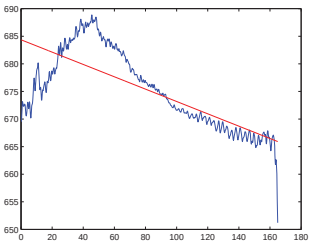
This figure shows how the error varies as the angles from the sensor plane increases. Generally, the error seems relatively small for all patterns, being below 5° for almost all data points except for the dense random pattern when at large angles. When at angles above 30° the error increases dramatically, growing to above 25° when at 45°. By looking at the Figures 5.5(g) to 5.5(j) it can be seen that the depth results from the Raytrix are no longer well modelled by a straight line as the angle increases. This does not follow the trend of the other patterns, as they are increasing in accuracy as the angle increases. It is expected that the accuracy of the calculated angle should improve as the angle increases, as errors in calculating the depth of a single pixel are less influential on the overall calculated angle due to a larger depth of field being utilised. This therefore correlates with the dense regular, sparse regular and sparse random, but not with the dense random pattern. It is thought this could be due to the fact that at larger angles, the features in the pattern will become denser with respect to the position of the camera. The features could then become too close and multiple features in the pattern be imaged by a single pixel. This would then make it difficult for the registration algorithms that are used to calculate depth to function accurately. To help identify if this trend is occurring, an evaluation of the standard deviations for every depth map has been conducted. To gain



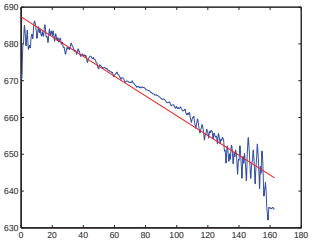
(a) True angle = 0°
Calculated angle = -0.06°



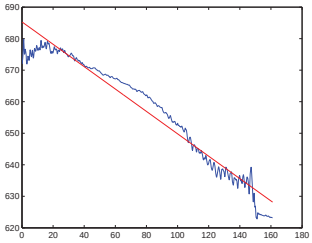
(b) True angle = 5°
Calculated angle = 8.20°



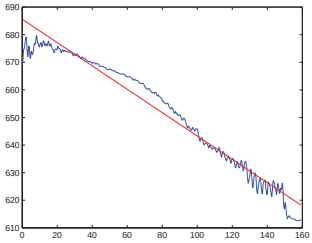
(c) True angle = 10°
Calculated angle = 6.38°



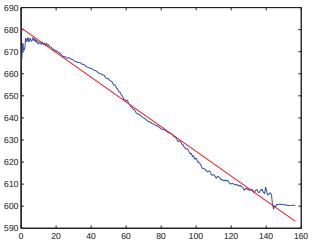
(d) True angle = 15°
Calculated angle = 15.04°



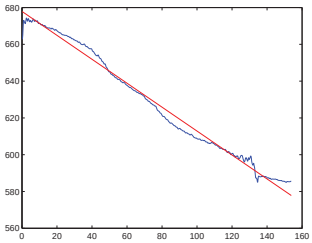
(e) True angle = 20°
Calculated angle = 19.54°



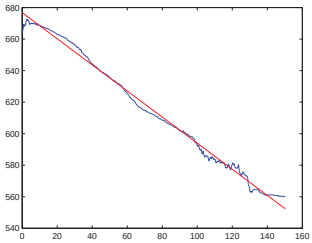
(f) True angle = 25°
Calculated angle = 22.94°



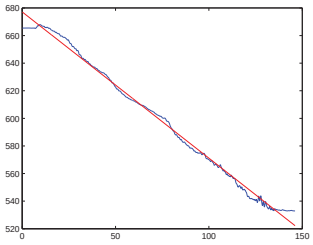
(g) True angle = 30°
Calculated angle = 29.23°



(h) True angle = 35°
Calculated angle = 33.05°

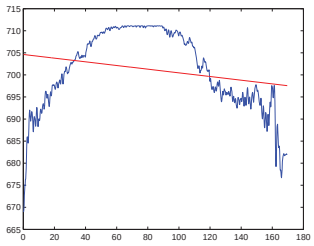


(i) True angle = 40°
Calculated angle = 39.62°

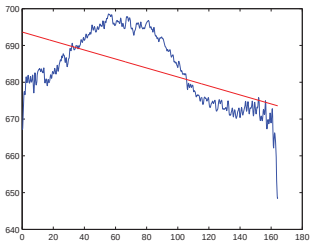


(j) True angle = 45°
Calculated angle = 46.68°

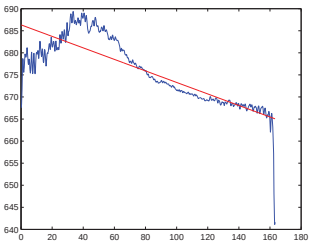
Figure 5.2: Graphs used to calculate the angle of a featureless scene when illuminated with a regular sparse pattern. The blue line is the actual data, and the red line is the fitted straight line.



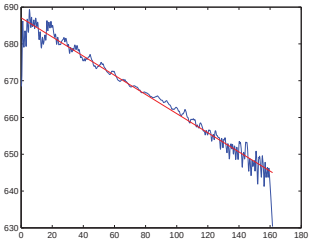
(a) True angle = 0°
Calculated angle = 2.38°



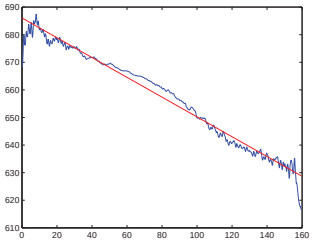
(b) True angle = 5°
Calculated angle = 6.96°



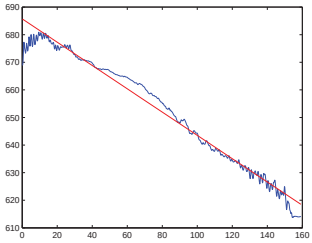
(c) True angle = 10°
Calculated angle = 7.43°



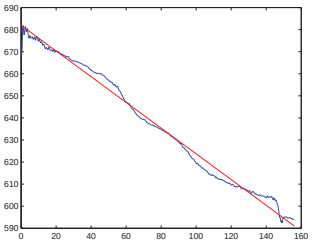
(d) True angle = 15°
Calculated angle = 14.60°



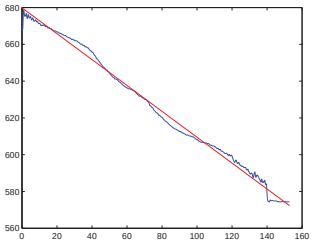
(e) True angle = 20°
Calculated angle = 19.73°



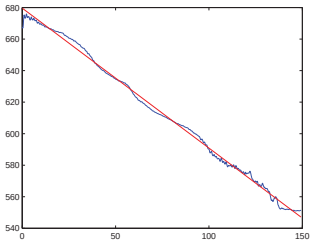
(f) True angle = 25°
Calculated angle = 22.89°



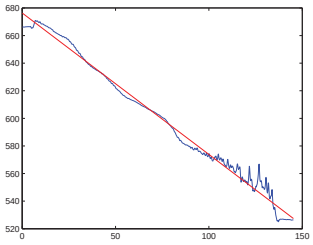
(g) True angle = 30°
Calculated angle = 30.22°



(h) True angle = 35°
Calculated angle = 35.13°

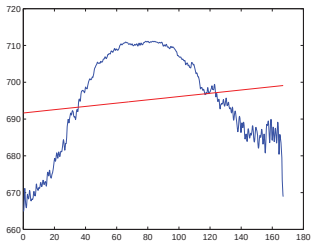


(i) True angle = 40°
Calculated angle = 41.60°

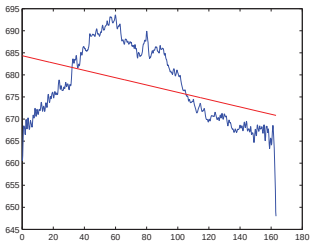


(j) True angle = 45°
Calculated angle = 45.72°

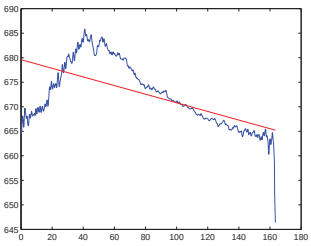
Figure 5.3: Graphs used to calculate the angle of a featureless scene when illuminated with a regular dense pattern. The blue line is the actual data, and the red line is the fitted straight line.



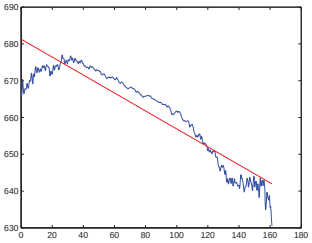
(a) True angle = 0°
Calculated angle = -2.26°



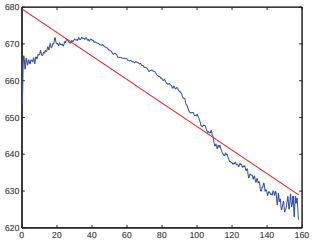
(b) True angle = 5°
Calculated angle = 4.74°



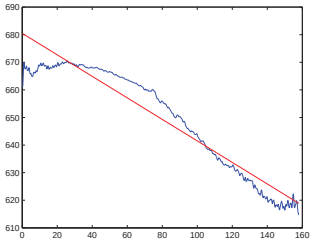
(c) True angle = 10°
Calculated angle = 5.02°



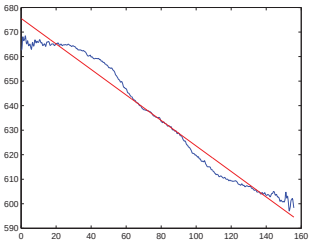
(d) True angle = 15°
Calculated angle = 13.71°



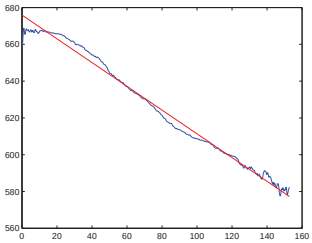
(e) True angle = 20°
Calculated angle = 17.73°



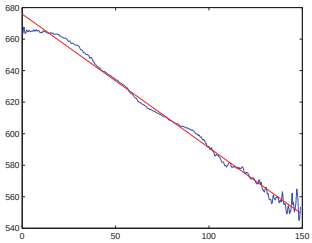
(f) True angle = 25°
Calculated angle = 21.30°



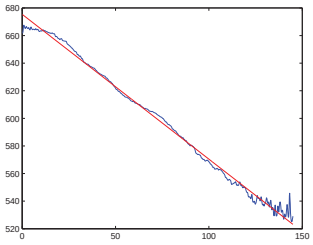
(g) True angle = 30°
Calculated angle = 27.47°



(h) True angle = 35°
Calculated angle = 32.85°

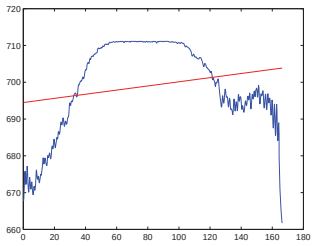


(i) True angle = 40°
Calculated angle = 40.27°

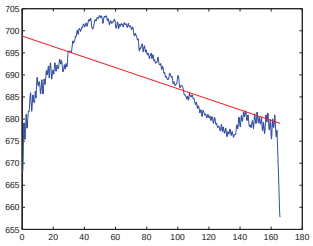


(j) True angle = 45°
Calculated angle = 46.39°

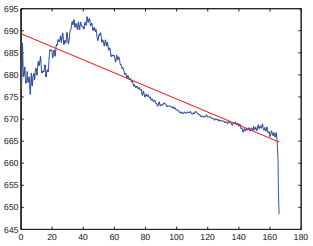
Figure 5.4: Graphs used to calculate the angle of a featureless scene when illuminated with a random sparse pattern. The blue line is the actual data, and the red line is the fitted straight line.



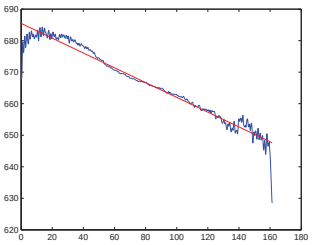
(a) True angle = 0°
Calculated angle = -3.22°



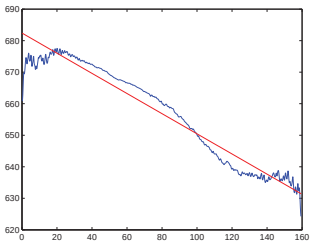
(b) True angle = 5°
Calculated angle = 6.81°



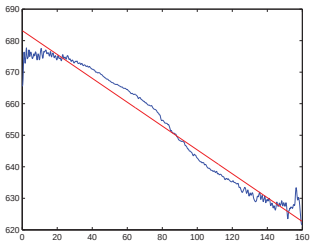
(c) True angle = 10°
Calculated angle = 8.41°



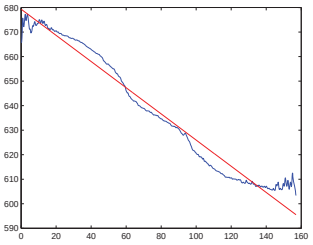
(d) True angle = 15°
Calculated angle = 13.20°



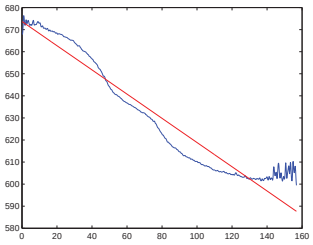
(e) True angle = 20°
Calculated angle = 17.72°



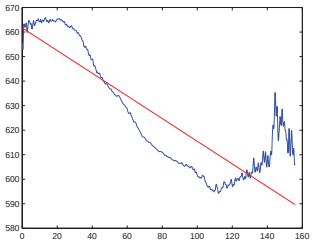
(f) True angle = 25°
Calculated angle = 20.72°



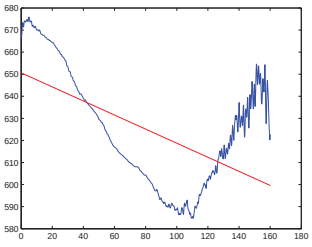
(g) True angle = 30°
Calculated angle = 28.09°



(h) True angle = 35°
Calculated angle = 28.74°



(i) True angle = 40°
Calculated angle = 24.81°



(j) True angle = 45°
Calculated angle = 17.68°

Figure 5.5: Graphs used to calculate the angle of a featureless scene when illuminated with a random dense pattern. The blue line is the actual data, and the red line is the fitted straight line.

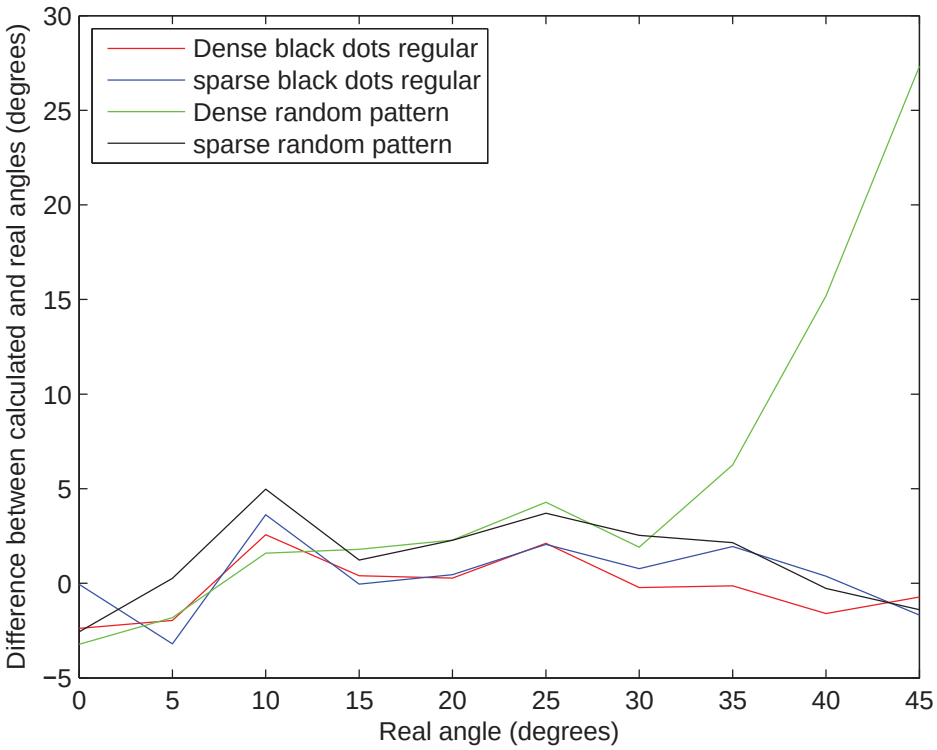


Figure 5.6: A graph showing the error in the calculated angle for the four tested patterns at a range of angles to the sensor plane from parallel to 45°

a single standard deviation value for each depth map, first the standard deviation of each column of the depth map has been calculated. Due to the geometry of the experiment, all values in a single column of the depth map are at the same depth, so should yield the same virtual depth value. This can be visualised in Figure 5.8. Once the array of standard deviations is calculated, a mean of these is then taken and used as the total image standard deviation. A plot of these values for all patterns at all angles can be seen in Figure 5.7. Tables 5.1 and 5.2 show all values for the error and standard deviations for the regular and random patterns respectively.

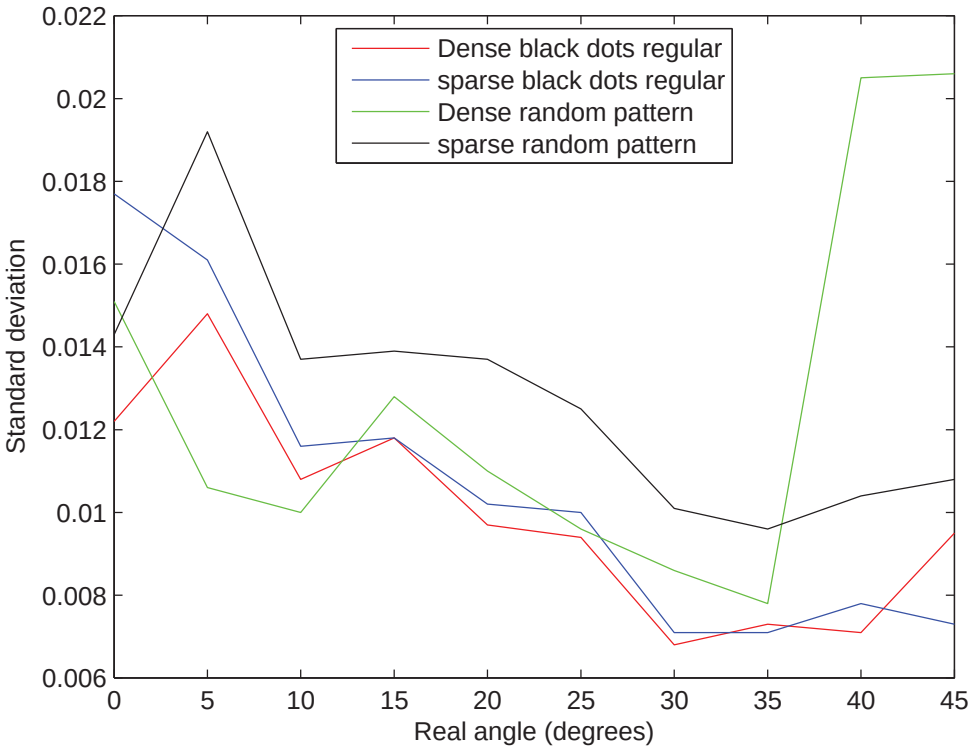


Figure 5.7: A graph showing the standard deviations for the depth maps used to calculate the angles represented in Figure 5.6

By studying Figure 5.7, it shows for the dense random pattern that at 40° and 45° the standard deviations show a considerable increase. This correlates with the previous graph in Figure 5.6 in which the difference between the calculated and real angle is much larger at these angles. If the standard deviation line for the regular dense pattern is studied in

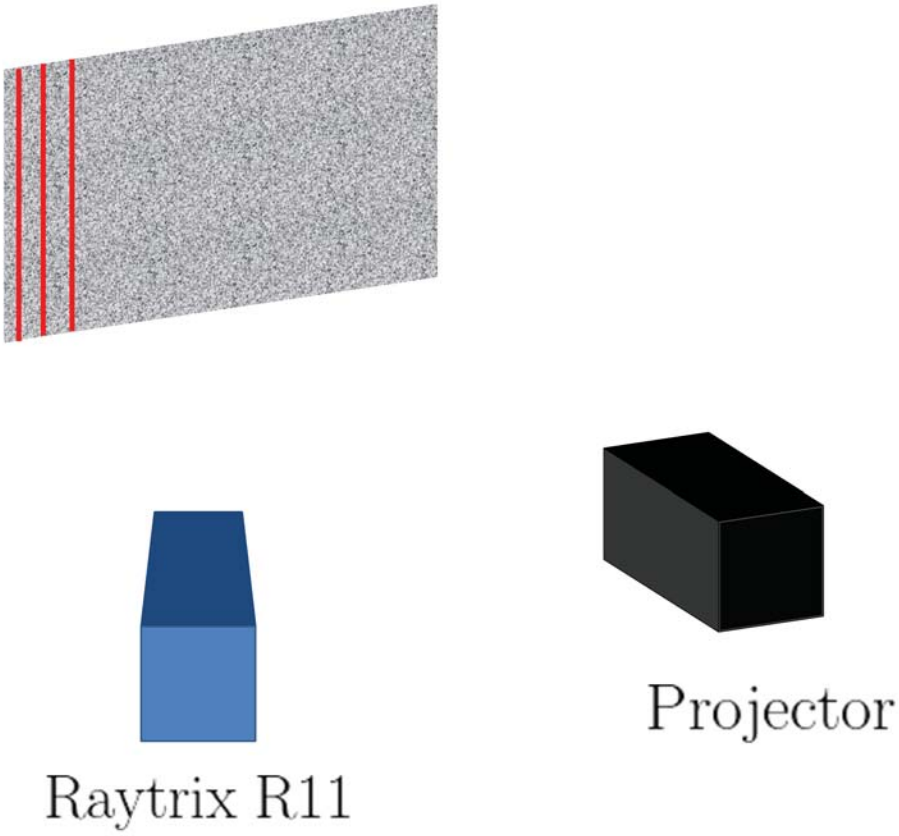


Figure 5.8: A three dimensional schematic of the experiment. The red lines represent the columns in the depth maps, with all points on this line at same physical distance from the Raytrix R11 plenoptic camera. In this system, the camera and projector remain approximately 70cm from the plane being imaged.

Pattern	Angle (degree)	Error	Standard deviation
Regular Dense	0°	2.38°	0.012
	5°	1.96°	0.015
	10°	2.57°	0.011
	15°	0.40°	0.012
	20°	0.27°	0.010
	25°	2.11°	0.009
	30°	0.22°	0.007
	35°	0.13°	0.007
	40°	1.60°	0.007
	45°	0.72°	0.010
	Mean	**1.24°**	**0.010**
Regular Sparse	0°	0.06°	0.018
	5°	3.20°	0.016
	10°	3.62°	0.012
	15°	0.04°	0.012
	20°	0.46°	0.010
	25°	2.06°	0.010
	30°	0.77°	0.007
	35°	1.95°	0.007
	40°	0.38°	0.008
	45°	1.68°	0.007
	Mean	**1.42°**	**0.011**

Table 5.1: The results from imaging the flat sloped scene with the Raytrix R11 with both a regular dense and sparse pattern. The errors and standard deviations are given.

Pattern	Angle (degree)	Error	Standard deviation
Random Dense	0°	3.22°	0.015
	5°	1.81°	0.011
	10°	1.59°	0.010
	15°	1.80°	0.013
	20°	2.28°	0.011
	25°	4.28°	0.010
	30°	1.91°	0.009
	35°	6.26°	0.008
	40°	15.19°	0.021
	45°	27.32°	0.021
	Mean	**6.57°**	**0.013**
Random Sparse	0°	2.56°	0.014
	5°	0.26°	0.019
	10°	4.98°	0.014
	15°	1.23°	0.014
	20°	2.27°	0.014
	25°	3.70°	0.013
	30°	2.53°	0.010
	35°	2.15°	0.010
	40°	0.27°	0.010
	45°	1.39°	0.011
	Mean	**2.13°**	**0.013**

Table 5.2: The results from imaging the flat sloped scene with the Raytrix R11 with both a random dense and sparse pattern. The errors and standard deviations are given.

Figure 5.7, it can be seen that the values do not increase at large angles to the same extent as for the random dense pattern. This is due to the features in the random dense pattern being closer together than in the regular dense, so only when at the greatest angles will the features be too close together for them to be indistinguishable. The regular dense line does however indicate at 45° that they are reaching this limit, as the standard deviation increases slightly relative to 40°.

It is interesting to note that as the angle increases from parallel to the sensor plane, in all patterns it is not until 15° that the graphs appear to follow a straight line (see (a), (b) and (c) in Figures 5.2 to 5.5). As this phenomena occurs for all patterns, this can be eliminated as the cause. Another possibility is that the original calibration of the Raytrix camera contained some inaccuracies. However, these should have been corrected for in the conversion from virtual to real depth procedure in Section 4.3. This is because although the relationship between virtual and real depths was found to not follow a straight line, the fitted function for the given Raytrix calibration will take into account any deviations from this line and correct for it. This therefore shows that the plane being imaged is not flat, or the error is occurring due to aberrations in the imaging system. As the angle of the plane was measured with an accuracy of less than a single degree, and the plane is known to be flat to a high degree of accuracy, it indicates that the error is caused by the imaging system. Therefore a further study into the errors on the graphs has been investigated in Sections 5.2.1 and 5.2.2.

To summarise the effect of the pattern on the accuracy of the recovered depth, it appears that the dense regular pattern is the best. However, the errors for this pattern compared to the others is minimal and usually within a degree or two except for the random dense pattern which is considerably different. The major errors occurred when the random dense pattern was projected onto a slope with a large angle. These errors are thought to occur because the pattern became too dense and merged together onto a single pixel. Therefore, in conclusion the regularity of the pattern plays a small part, but the major

factor is the density of the pattern. The depth is recovered more accurately the denser the pattern is, but if it becomes too dense then the depth is calculated with a very high degree of accuracy.

5.2.1 Investigating the origins of noise

As discussed in Sections 4.6.2, the high frequency noise is found to have an insignificant effect on the calculated angle. However, the origins of this noise was not investigated. It has been hypothesised that it is due to the regularity of the pattern. This can now be investigated as throughout this chapter the same experiment has been replicated but with a random pattern. If the regularity is the cause of the high frequency noise, then this same noise should not be present in the graphs for the random patterns in Figures 5.4 and 5.5, although the random pattern may cause a new type of noise in the data. However, under close inspection of these graphs the high frequency noise is still apparent. This shows that the regularity of the pattern is not the cause for the noise. It could be that this noise is due to the regularity of the microlens array, however as it has a minimal effect on the accuracy of the results, it is not the most important noise so a more thorough investigation into the low frequency noise will be made.

5.2.2 Investigation of the low frequency noise

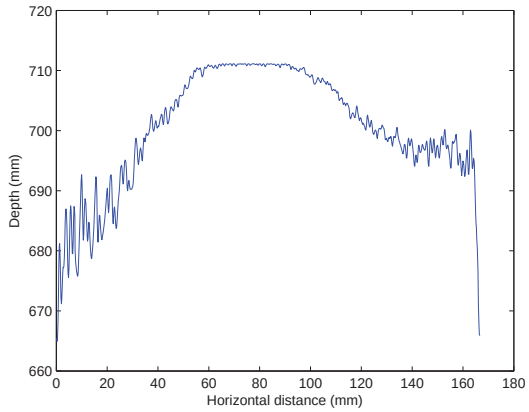
By analysing the graphs for when the plane is parallel to the sensor (Figures 5.2a, 5.3a, 5.4a and 5.5a), it is clear that some errors occur away from the predicted flat shape which the graphs should have. To help to understand if this error is systematic in the Raytrix or not, two experiments need to be conducted. The first is to image a flat plane parallel to the sensor but at a variety of distances from the sensor. As previously all these images were taken at the same distance from the sensor, it will be instructive to see if the shape of the graph changes dependent on distance. The second experiment will involve imaging the plane at 0° but with a different initial calibration of the Raytrix camera. Although it

is assumed that any errors from the calibration should be accounted for in the conversion from virtual to real depth, it is important to confirm this. The results from the first experiment can be found in Figure 5.9.

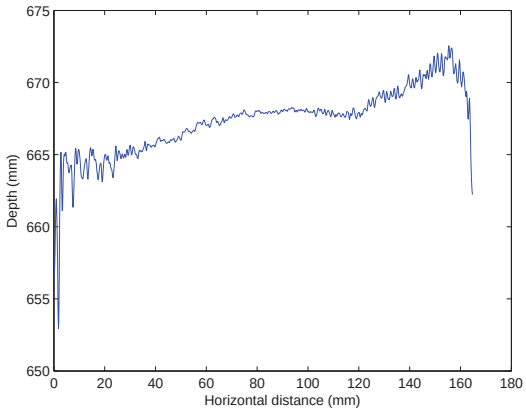
By studying these images, it is clear to see that the profile for a flat plane does change depending on distance from the sensor. This proves that the shape seen in the graphs for the plane at 0° to the sensor in Figures 5.2a, 5.3a, 5.4a and 5.5a are only consistent as they are taken at the same distance from the plane to the sensor, proving that the shape of the graph is a function of the distance. This could either be due to a systematic error in the Raytrix or from the initial calibration of the camera. To help with the identification of the error, the same experiment as presented in Figure 5.9 was conducted but with a different initial camera calibration. These results can be seen in Figure 5.10.

By comparing the graphs in Figure 5.9 to the equivalent graphs in Figure 5.10, qualitatively they appear to be very similar, especially in the top row of graphs. The graph in the top left has the characteristic shape of the graphs at 0° to the sensor throughout this and the previous chapter. This is due to the plane being at roughly the same distance from the sensor in each of these images. The interesting point to note is that the plot in Figure 5.10 is with a different initial camera calibration. This therefore provides strong evidence that this shape is not caused from the calibration but is a systematic error that is a feature of the camera itself. This is expected as the calibration method includes this feature.

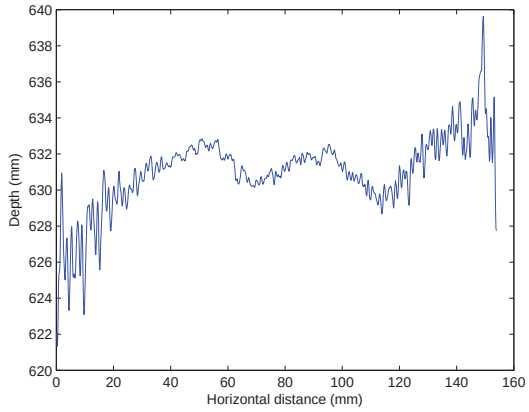
To try to understand why the graphs have the appearance they do, it is important to assess the shape of each graph individually as well as collectively. In general, the central sections of the graphs tend to have less variation and follow a flatter line with large variations towards the edges. (The exception is the middle and bottom right images in Figure 5.9. It is not fully understood why this effect is present, but it could be that the dominant microlens in the Raytrix is changing at those depths, causing different results. This has not been studied any further). The most common shape is most likely due to



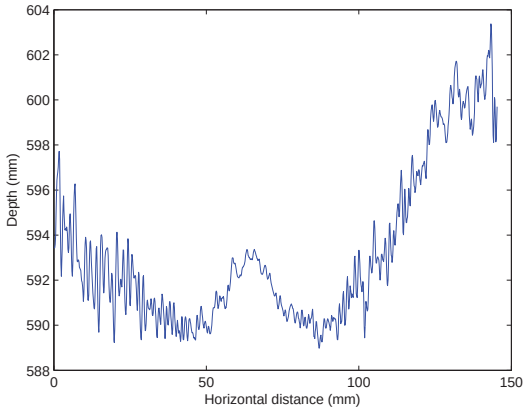
(a) Depth of plane = 710mm



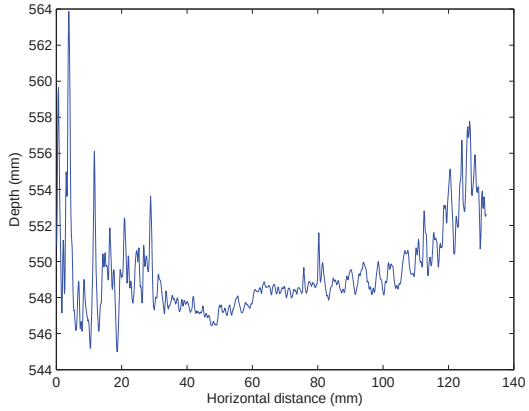
(b) Depth of plane = 670mm



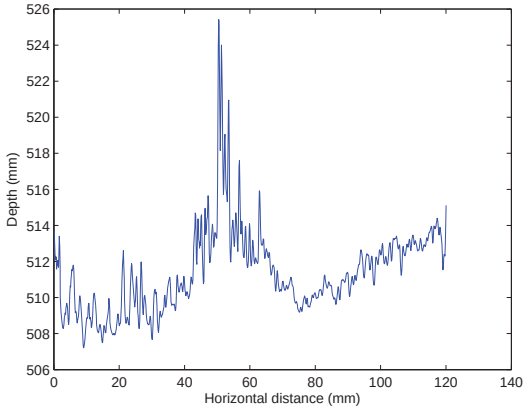
(c) Depth of plane = 630mm



(d) Depth of plane = 590mm

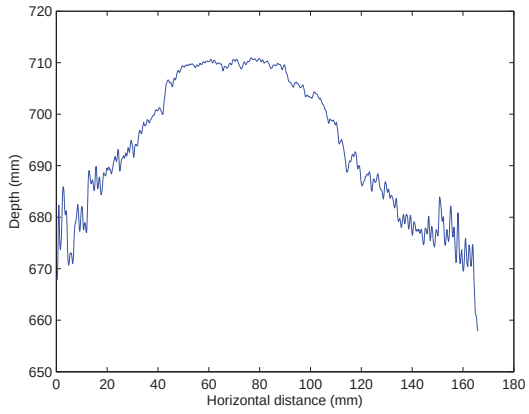


(e) Depth of plane = 550mm

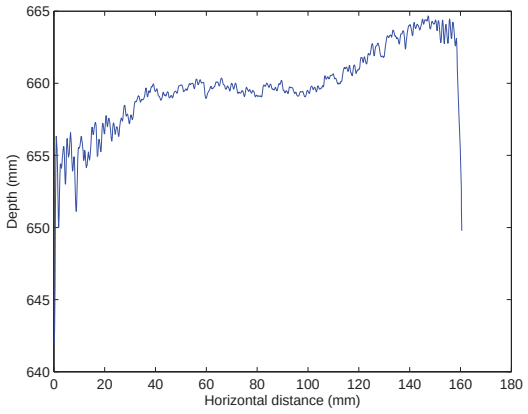


(f) Depth of plane = 510mm

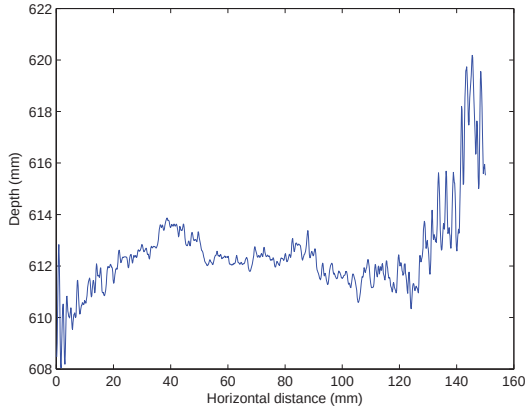
Figure 5.9: Graphs of a plane at 0° to the plane of the sensor when at varying depths.



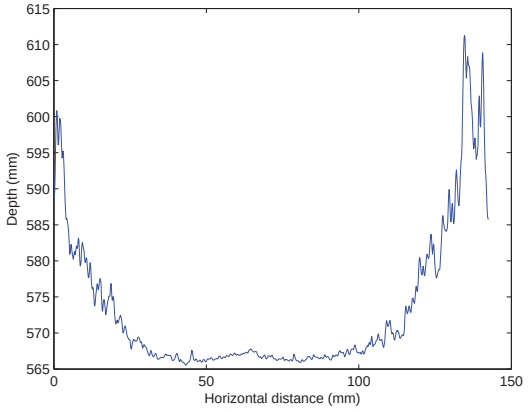
(a) Depth of plane = 710mm



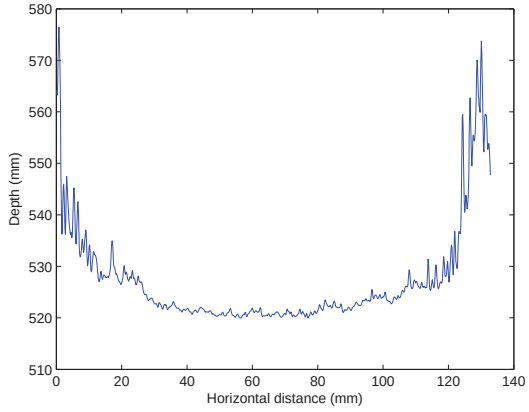
(b) Depth of plane = 660mm



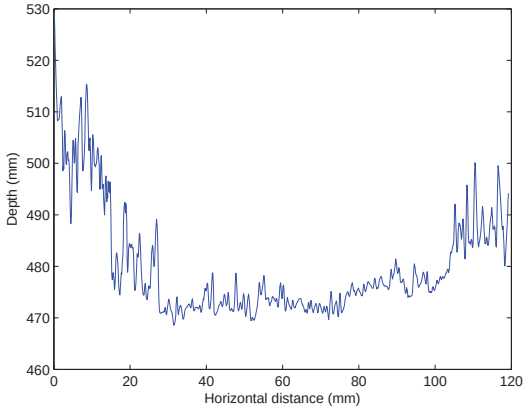
(c) Depth of plane = 610mm



(d) Depth of plane = 560mm



(e) Depth of plane = 520mm



(f) Depth of plane = 470mm

Figure 5.10: Graphs of a plane at 0° to the plane of the sensor when at varying depths with a different initial calibration to Figure 5.9.

aberrations being larger towards the outer part of the main lens of the camera, as well as in the outer lenses in the microlens array. An interesting point to note is that when at larger distances closer to the back of the depth of field (the top row in Figures 5.9 and 5.10), either ends of the graph drop to a smaller depth. This goes against intuition, as the edges of the field of view are further away from the sensor than the centre, so it would be expected if it is not calculated to be at the same depth as the central part of the image, then it should be further away. One explanation would be that the Raytrix is miscalculating the depth as towards the edges the distance is then greater than the depth of field so cannot produce an accurate depth. If this is in fact true, then when the plane is moved closer to the sensor, then the graphs should look like the mirror image of the previous images, with the edges of the plot increasing in distance as it is moved closer to the edge of the field of view. By assessing the bottom images in Figure 5.9 and especially in Figure 5.10 this characteristic does appear to be true.

Another explanation for the shape of the plots in Figures 5.9 and 5.10 could be because of the microlens array used in the camera. As previous explained in Section 1.12.2, the Raytrix plenoptic camera uses a microlens array containing microlenses with three different focal lengths. This is chosen to increase the depth resolution by creating more focal planes from the raw image. However, at depths at which the dominant microlens type is changing could cause some non linear effect, which is what could be causing the differences in profiles of flat planes when at different depths.

The discrepancies could be due to mechanical or thermal noise or from the environment. To help minimise these errors, the Raytrix camera and projector were allowed time to stabilise after being turned on, the plane was mounted on a static optical bench and the temperature of the room was kept constant as well as possible. Referring back to Table 4.2, the variations in Figures 5.9 and 5.10 are greater than those found to be due to projecting texture over the scene having inherent features. This indicates that the projector is not the main cause for the errors. To verify this the experiment could

be repeated but with texture inherent in the scene and not projected, but due to time restrictions this was not performed.

5.3 Fitting a plane

In Section 4.7 in the previous chapter a method was devised for calculating the angle of the slope by fitting a plane to the data, rather than the methods used throughout the earlier parts of this chapter. This was done to try to eliminate any potential errors through the reliance on the mechanical alignment of the scene and the camera throughout the experiments. The results from this method looked promising, except at angles above 25° where the difference between the calculated and real angles became much larger. To see if this trend occurs independent of the pattern projected, the method described in Section 4.7.1 has been used on the data for all the patterns, with Tables 5.3 and 5.4 showing the results for the regular and random patterns respectively, and Figure 5.11 showing the errors in the calculated angles with respect to the real angle of the plane.

By analysing the graph in Figure 5.11 a few interesting points can be noted. Firstly, the calculated angle is much less than the real angle above 30° for the dense random pattern. This is consistent with the results from Figure 5.6, and it is believed this is due to the details in the pattern becoming too densely packed so that they cannot be distinguished by the depth recovery algorithms in the Raytrix software, although this was never proven. The second interesting point to note is that above 30° for the other three patterns, the calculated angle is much larger than the real angle of the plane. These errors could either be due to errors in the plane fitting method, the angle of the slope recovery, or in the depth map data itself. To investigate this further, more images of slopes at different depths to the camera could be taken to help narrow down the cause of the errors. Due to the time limitations set out in Section 3.4.1 this was not possible to complete. However, as the lines for the dense regular, sparse regular and sparse random all tend to follow the same pattern, it could be inferred that to within a certain degree of accuracy the

Pattern	Real angle	Calculated angle	Difference
Regular Dense	0°	2.15°	2.15°
	5°	7.16°	2.16°
	10°	8.11°	1.89°
	15°	15.67°	0.67°
	20°	21.28°	1.28°
	25°	24.85°	0.15°
	30°	35.45°	5.45°
	35°	39.86°	4.86°
	40°	50.33°	10.33°
	45°	60.71°	15.71°
Regular Sparse	0°	0.06°	0.06°
	5°	7.80°	2.80°
	10°	6.96°	3.04°
	15°	16.27°	1.27°
	20°	20.43°	0.43°
	25°	24.13°	0.87°
	30°	32.70°	2.70°
	35°	37.96°	2.96°
	40°	47.06°	7.06°
	45°	64.97°	19.97°

Table 5.3: The results from imaging the flat sloped scene with the Raytrix R11 with both a regular dense and sparse pattern using the fitted plane method.

Pattern	Real angle	Calculated angle	Difference
Random Dense	0°	2.23°	2.23°
	5°	6.81°	1.81°
	10°	9.04°	0.96°
	15°	14.46°	0.54°
	20°	18.87°	1.13°
	25°	21.74°	3.26°
	30°	30.88°	0.88°
	35°	32.05°	2.95°
	40°	25.32°	14.68°
	45°	19.05°	25.95°
Random Sparse	0°	2.49°	2.49°
	5°	4.96°	0.04°
	10°	5.94°	4.06°
	15°	14.81°	0.19°
	20°	18.62°	1.38°
	25°	22.47°	2.53°
	30°	29.60°	0.40°
	35°	36.93°	1.93°
	40°	48.65°	8.65°
	45°	61.60°	16.60°

Table 5.4: The results from imaging the flat sloped scene with the Raytrix R11 with both a random dense and sparse pattern using the fitted plane method.

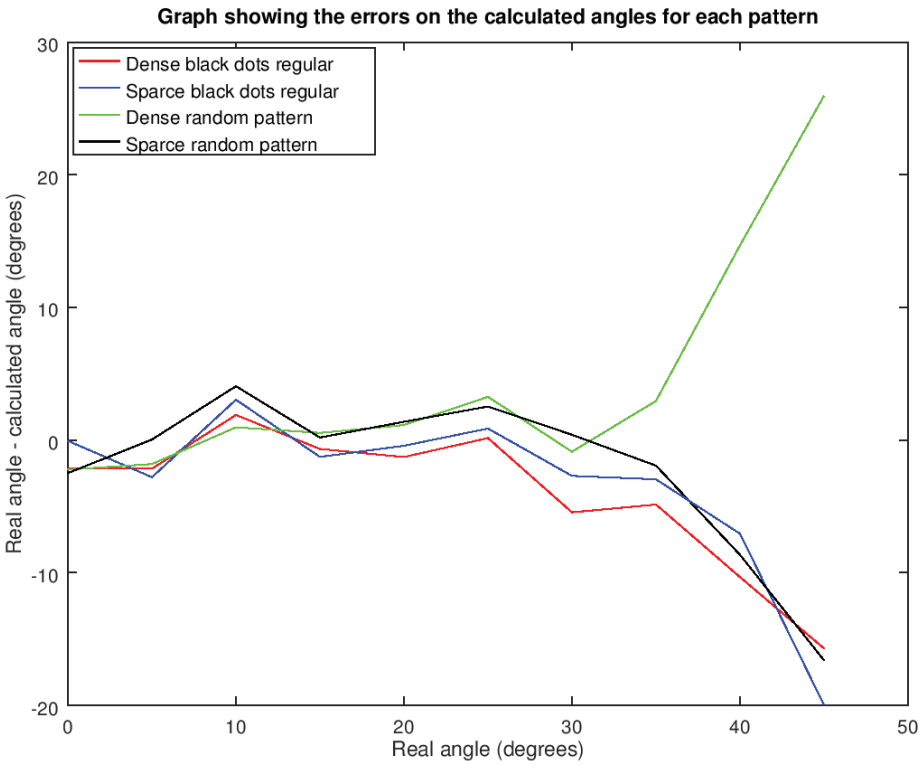


Figure 5.11: A graph showing the difference between the calculated angles from the fitted plane method to the real angle for all four patterns

pattern used does not make a material difference to the calculated depth. What does make a material difference is if the pattern becomes too dense, and the features become indistinguishable to the depth map algorithms.

5.4 Conclusion

This chapter has focused on the use of different patterns to improve the accuracy of the calculated angle from a sloped scene using the novel methods described in Chapter 4. Four patterns were chosen to be investigated, to see if the randomness or sparsity of the pattern affected the overall result the most. From imaging with these four patterns, both high and low frequency noise was observed. In Section 4.6.2 the high frequency noise had been removed, and found out to not be significant in the calculated angle of the slope. However the origins on this noise had only been hypothesised as being due to the regularity of the pattern. In this chapter it has been demonstrated that the high frequency noise is not due to the regular pattern, as the plots for the random pattern also yielded similar noise. This therefore points towards the cause being the regularity of the microlens array, however as the noise had been found to have an insignificant effect on the calculated angle, it was not investigated further.

The shape of the plane when parallel to the sensor did not follow a flat line as it should. It was investigated whether the same shape was apparent when the plane was at different depths, which it was found not to be. This shows that the perception of relative depth across the field of view is a function of distance from the sensor. To see if it was caused by the initial camera calibration, the experiment was repeated with a different calibration conducted. Qualitatively, this showed the same behaviour so indicates that the shape of the plots are not due to the initial calibration but instead a systematic error in the camera.

With regards to the accuracy of each pattern, the regular dense pattern produced the most accurate results. However, the error of all patterns at small angles is within a few

degrees so the type of pattern does not play a significant effect on the accuracy. When at large angles, the dense random pattern had very large errors. This is thought to be because the details became too dense at large angle due to parallax. This therefore shows that the pattern cannot be too dense if the angles of the objects in the scene are large with reference to the plane of the sensor.

A plane was then fitted to the data, and the angle calculated from this as per the previous chapter. Similar to the first method for calculating the angle of the slope, the dense random pattern produced very large errors when at large angles. The other three patterns then produced calculated angles much larger than the real angles when above 30°. They all followed the same pattern, indicating this was likely an error in the methodology, or a systematic error in the imaging system. Due to the time and equipment limitations set out in Section 3.4.1, this was not investigated further. However, it can be inferred that as the other three patterns followed the same trend, the computing of the depth does not rely upon the pattern used but rather just that one is.

Throughout this and the previous chapter, a simple sloped scene has been used to investigate the effects on the depth accuracy by projecting patterns onto a featureless scene. However, for retinal imaging a sloped scene is not very representative of any features in this area. Therefore, the following chapter will focus on the accuracy of the depth recovered from a Raytrix on a spherical object.

Chapter 6

Quantitatively assessing the accuracy of imaging a sphere with a plenoptic camera

In Chapter 4, a novel method for gaining quantitative measurements of a sloped scene was developed. An investigation into the effect of different patterns on this sloped scene was then investigated in Chapter 5. To advance this investigation further, the sloped scene is now replaced with a plain white sphere. This object has been chosen to be imaged to investigate how a plenoptic system reacts to a continuously changing angle. It has also been chosen as smooth curves and continuously changing angles relate more closely to the anatomy of the eye and retina [143]. To analyse the data, a novel method to quantitatively assess the accuracy of the depth maps produced by the Raytrix R11 plenoptic camera has been devised. This method will then be used to compare and contrast the accuracy of depth maps when illuminated with the same patterns as used in Chapter 5. For all experiments throughout this chapter, a Raytrix R11 plenoptic camera will be used to image the scene, and an Optoma PK201 LED Pico projector to illuminate the scene with patterns.

6.1 Imaging protocol

The results collected throughout this chapter all followed the same imaging protocol. As per the previous two chapters, a Raytrix R11 plenoptic camera had been used to capture the images. The object being imaged throughout the chapter is a spherical ball with a diameter of 70.0mm, measured using callipers. The diameter of the ball stayed constant at 70.0mm when measured across the ball, so is accurate to at least 10^{-4}m. When imaging, the temperature of the room was kept constant, and the ball's diameter was measured between experiments and afterwards, yielding the same diameter as before the first images were taken. Because of this, it is assumed that any variations due to the mechanical and thermal stability of the ball are negligible for the purpose of these experiments. The different patterns were projected onto the scene with an Optoma PK201 LED Pico projector.

6.2 Method for assessing accuracy when imaging a ball

In order to quantitatively assess and compare the accuracy of depth maps generated from the plenoptic camera when illuminated with different patterns, a specific comparison criterion had to be devised. For this investigation, the factors that will be used to assess the accuracy are the diameter of the ball and the surface profile of the ball. This can be achieved by using the following novel method.

1. Convert virtual to real depth

The first step in the process is to convert the virtual depth values given in the Raytrix depth maps into depths in real space. The method for this involved taking a set of calibration images of a flat plane at varying distances from the sensor, to which a function

can be fitted to map the virtual to real depths. For a full explanation of this method please see Section 4.3.

2. Segment the ball from the depth map

The next step of the process involves segmenting the ball from the background on the depth map. To aid in the explanation of this method, it will be performed step-by-step on the depth map produced when the sphere was illuminated with the dense regular pattern. For reference, the depth map of this can be seen in Figure 6.1.

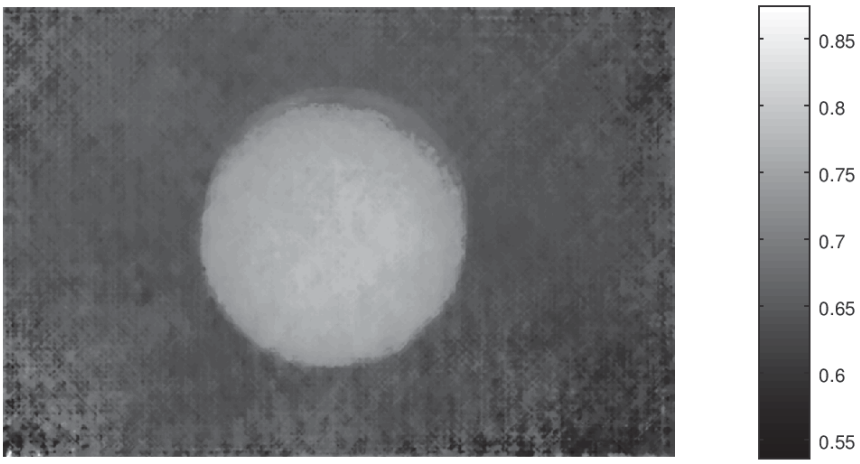


Figure 6.1: The depth map from the Raytrix of a ball when illuminated with the dense regular pattern. The depth values are in metres.

The first step to segmenting the image is to apply an appropriate threshold value. Algorithms such as Otsu gray-level threshold [144] can be applied to automatically select an appropriate value. However, by looking at the cross-section through the centre of the ball in Figure 6.2, the sphere contains depth values that vary considerably from the back-

ground so a threshold value can be selected manually relatively easily. The thresholding is performed on the real depth values rather than the gray values in Figure 6.1. Once an appropriate value has been chosen, the threshold should be used to create a binary mask with the area of the ball given the value of 1, and the background the value of 0. This mask can be seen in Figure 6.3.

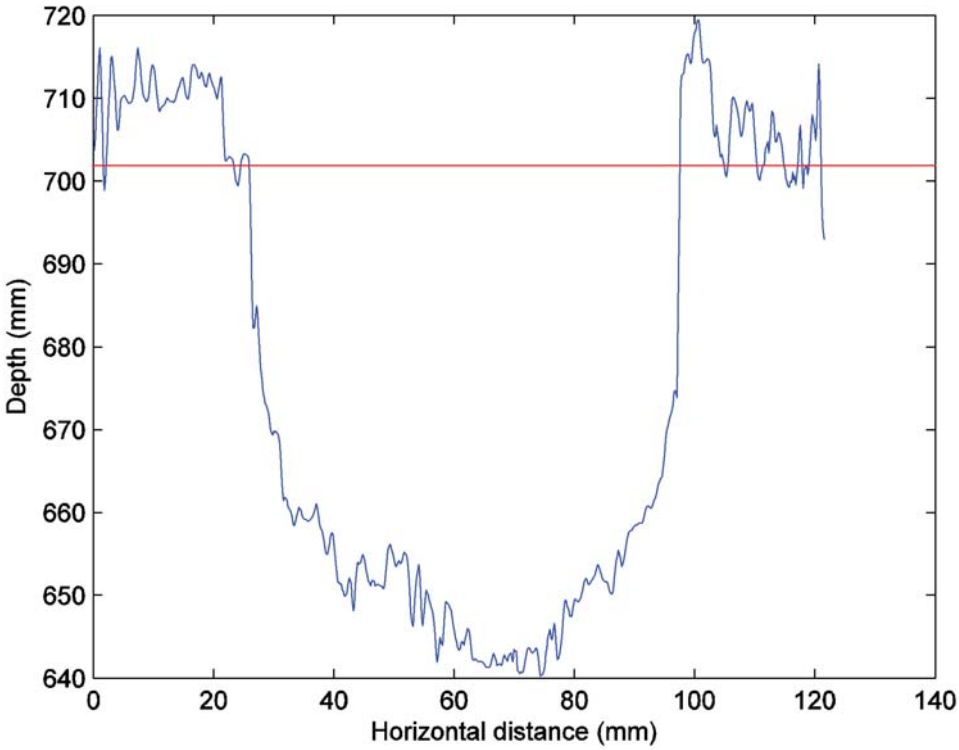


Figure 6.2: A cross-section through the centre of the ball in the depth map as shown in Figure 6.1. The red line represents the chosen threshold value

To check that the binary mask is representative of the original image, a boundary around the ball in the mask can be drawn onto the original depth map from Figure 6.1, with the results of this are presented in Figure 6.4. If by a quick visual inspection the boundaries align with the edges of the ball, then it is appropriate to move on to the next step of the method. If not, a new threshold value should be chosen. Under some circumstances, the thresholding method may yield more than one object. If the secondary object is small it may be removed though mathematical morphology protocols such as erosion or dilation

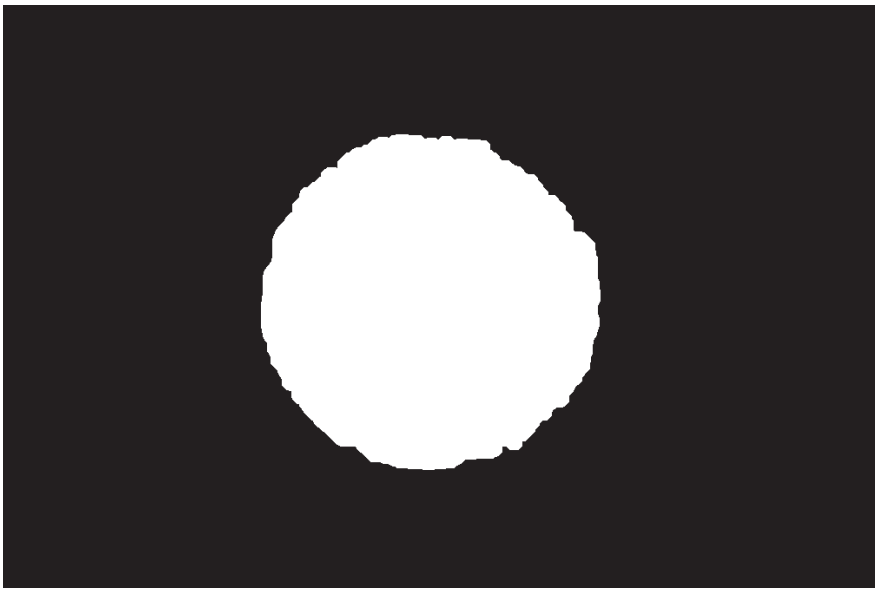


Figure 6.3: The binary mask created from using the selected threshold, in which the ball has a value of 1 and the background a value of 0.

[145]. If the unwanted objects are larger, the original depth map may need to be cropped to close to the edges of the ball.

3. Calculate the average depth of the background

For further analysis on the accuracy of the depth map, it is first important to calculate the mean value for the background. Without the ball present this is easy as the mean value of the entire depth map can simply be taken. However, when the ball is present in the depth map, this area of the depth map needs to be ignored so a more sophisticated method needs to be applied. This can be made easier by taking the inverse of the mask for the ball. This mask can then be applied to the image to remove the ball and isolate the background. This image can be seen in Figure 6.5. The mean value for the background can the be calculated from the segmented image.

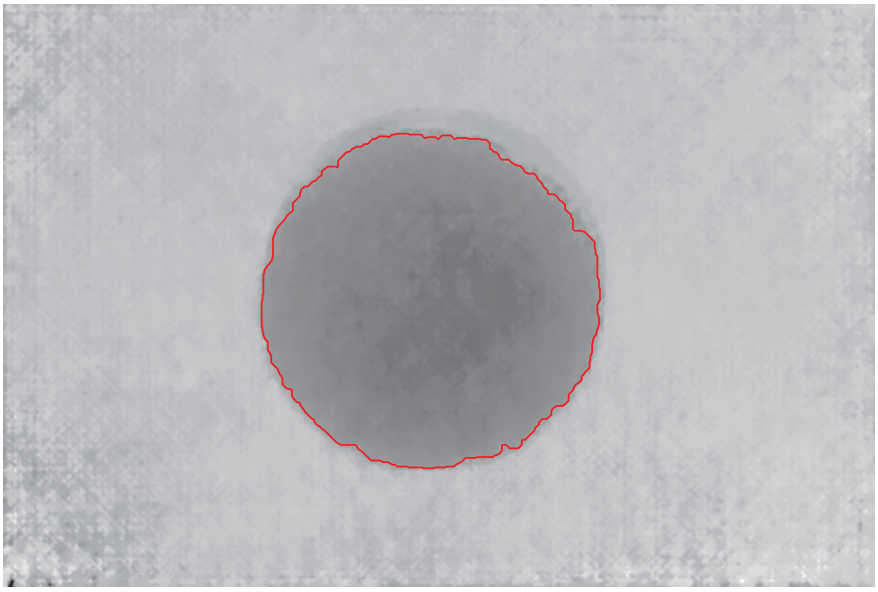


Figure 6.4: The depth map of the ball presented in Figure 6.1 with the boundaries from the binary mask overlaid.

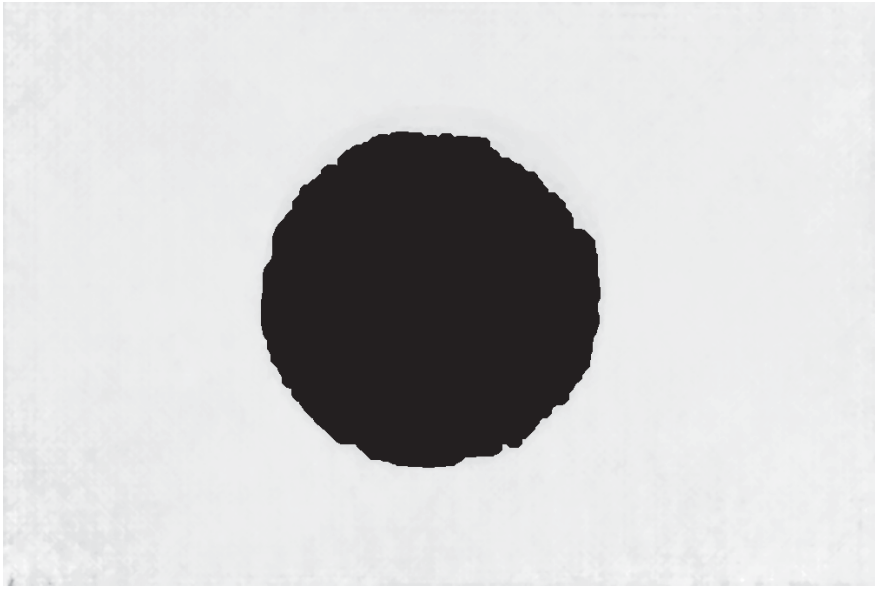


Figure 6.5: An image showing the binary mask for the background multiplied by the original depth map. The result shows the original values for the background with zero values in the location of the ball.

4. Calculate the real distance represented per pixel

The next step in the process involves calculating the physical distance which each pixel on the depth map represents. To achieve this, the mean value for the background from the previous step needs to be used and the method for relating real depth to field of view from Section 4.6.1 needs to be applied. In this method, a function relating the virtual depth values from the Raytrix under a particular camera calibration and lens settings to a field of view in real space is calculated. By using this method, and calculating what the field of view of the entire image is, by simply dividing the resolution of the image by the field of view, a value for real distance per pixel can be calculated.

5. Calculate the lateral diameter of the ball

One important measure for how accurately the ball has been calculated from the depth map is the diameter of the ball. The 'lateral diameter' is the value calculated from the lateral size of the ball on the depth map, rather than from the depth values of the ball itself. To gain this value, a segmentation of the ball and the real distance represented per pixel are vital to the calculation. By calculating the bounding box around the ball, this will give both the horizontal and vertical size of the ball in units of pixels. By simply taking the mean of these two values, and then multiplying this by the real distance per pixel, the diameter of the ball can be calculated. This can then be compared to the actual size of the ball if this has been measured prior to imaging.

6. Calculate the centre of the ball in 3 dimensions

The next step in the process is to calculate where the centre of the sphere is in 3 dimensions. Laterally this is relatively simple, as this can be achieved by taking the centre of the object. The red cross in Figure 6.6 shows its location on the original depth map.

To calculate the centre of the ball in the dimension of the depth (i.e. in relation to distance from the ball to the camera) it is not as trivial. First the mean depth value for

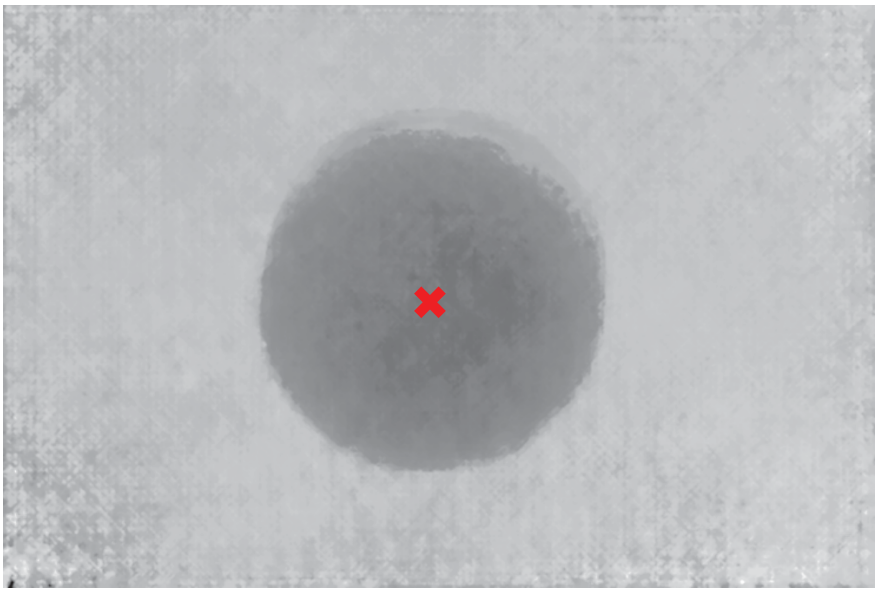


Figure 6.6: An image showing the original depth map with the calculated centre of the ball represented with a red cross.

the background needs to be taken, as calculated from step 3, and the diameter of the ball as calculated from step 5. By dividing the diameter by 2, then subtracting this from the mean depth of the background, the centre of the ball in this dimension can be found. This is shown by the red cross in Figure 6.7, where a cross section through the lateral centre of the ball is taken.

7. Calculate the diameter of the ball using depth

Another way to calculate the diameter of the ball is to use the depth values in the depth map. This can be calculated using the co-ordinates for the centre of the ball as presented in Figure 6.6 and taking the depth value at this position as this is the depth of the top of the ball. Subtracting this from the mean background depth will produce the diameter of the ball. This value can then be compared to the value for the ball diameter as calculated in step 5, along with the real diameter of the ball that can be measured before imaging.

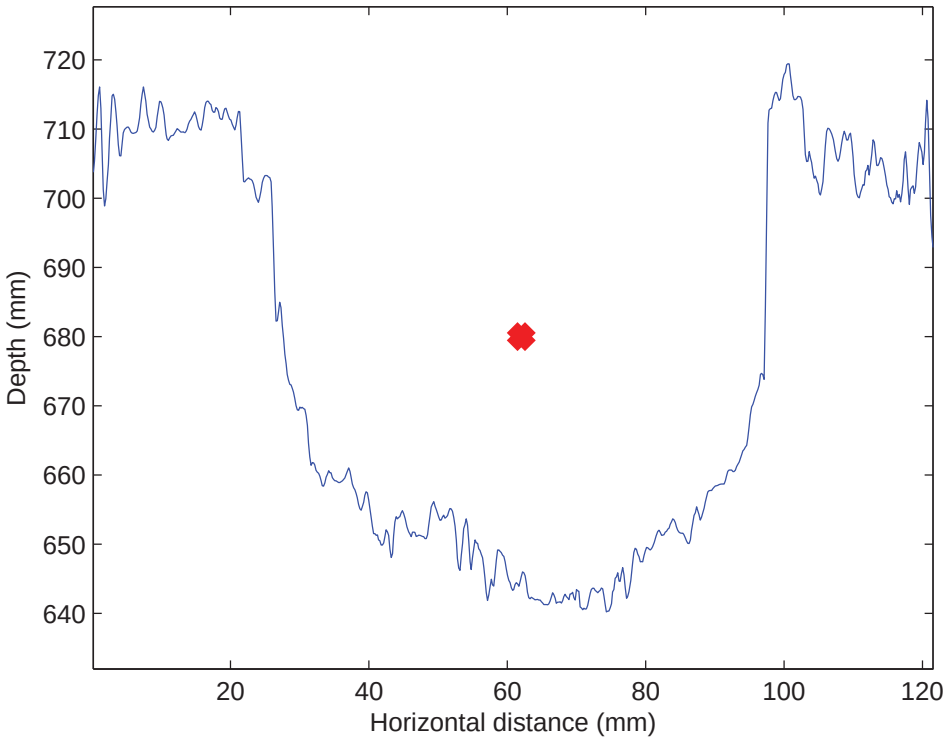


Figure 6.7: A graph of the cross-section through the centre of the ball, with the red cross representing the calculated centre of the ball.

8. Plot a circle centred around the calculated centre of the ball

The previous steps have allowed a comprehensive analysis on the ability to accurately calculate the diameter of the ball. These next few steps will help to study the accuracy of the curvature of the ball. To achieve this, first a circle needs to be plotted on the cross-section through the centre of the ball like in Figure 6.7, with the ball having the exact diameter of the sphere being imaged. This can be seen in Figure 6.8.

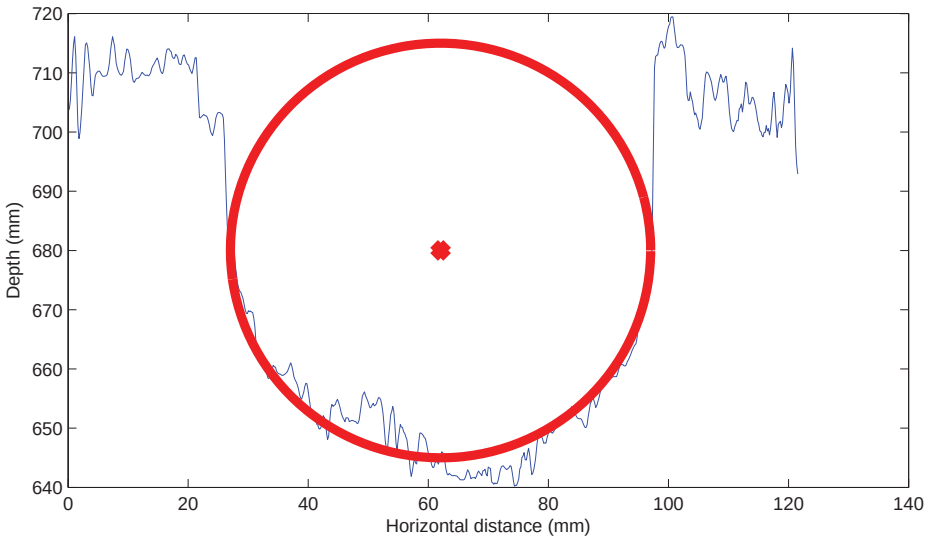


Figure 6.8: A plot showing a circle centred around the ball centre with a diameter equal the actual diameter of the ball.

9. Crop the cross-section and circle

The next step of the process is to crop the circle and the raw data plot down to only include the parts that represent the curve of the ball. This can be seen in Figure 6.9.

10. Calculate the absolute difference between the circle and the cross-section

The final part of the process is to calculate the absolute difference between the raw data (blue line in Figure 6.9) and the circle. By taking the mean of these values, the average error on the curvature of the ball can be seen. A plot of the absolute error can also be

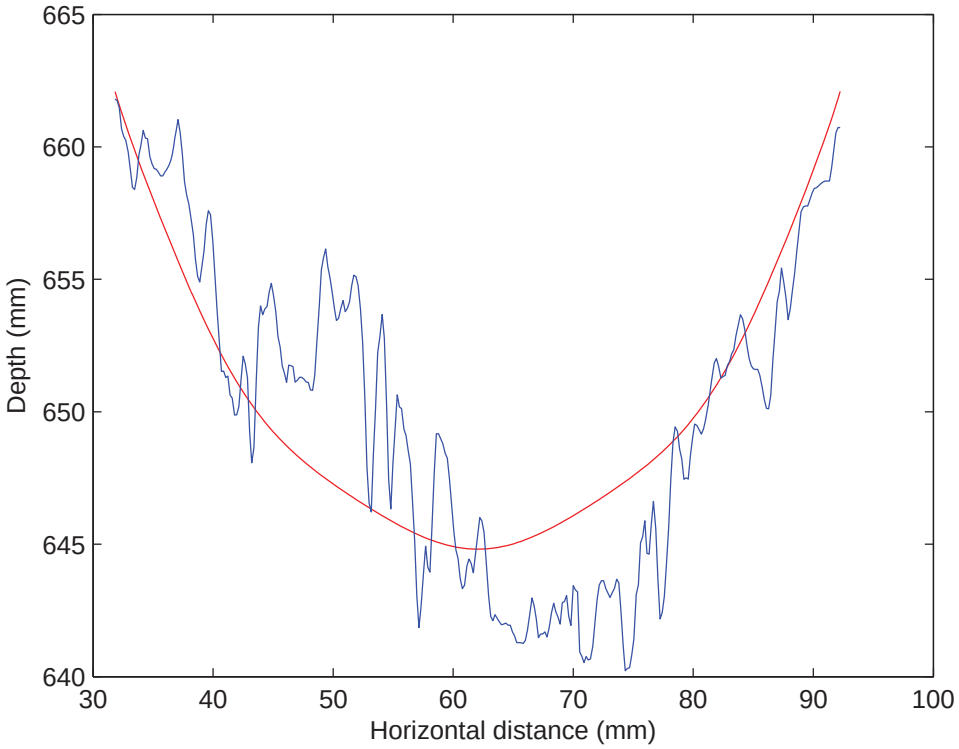


Figure 6.9: A plot showing a cropped circle centred around the ball centre with a diameter equal the the actual diameter of the ball along with the real data from the original depth map.

plotted to see whether the error accumulates more towards the centre of the ball or the outer edges. An example of this can be seen in Figure 6.10.

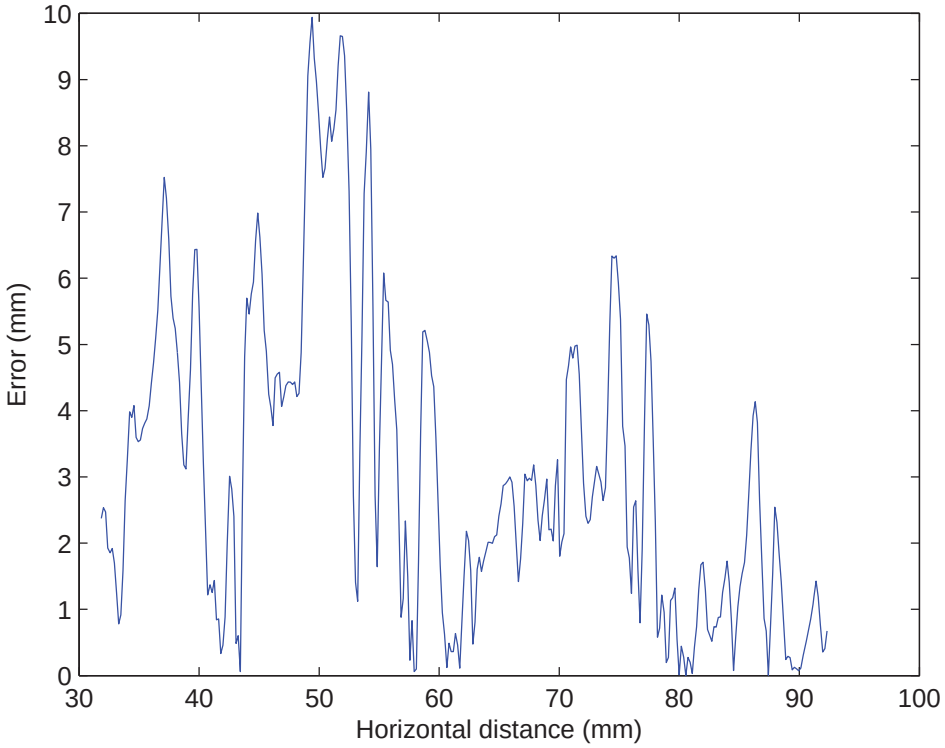


Figure 6.10: A plot of the absolute error between the circle and the raw data.

6.3 Experimental results from all patterns

In an attempt to see if the same conclusions can be drawn about different patterns being projected on a ball as it is for a sloped scene shown in Chapter 5, the method previously described in Section 6.2 has been followed for the regular dense, regular sparse, random dense and random sparse patterns from Section 5.1. The first value to be compared is the lateral diameter and diameter from depth as described in steps 5 and 7 of the method. These results are shown in Table 6.1.

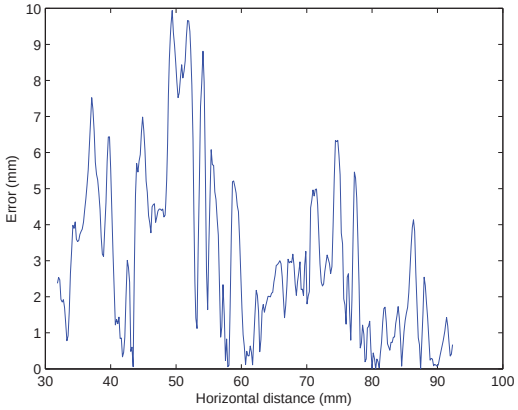
From looking at these results, one of the first conclusions that can be drawn is that the error on the lateral diameter is less than the depth diameter for all patterns. This is

Pattern	Actual Diameter (mm)	Lateral Diameter (mm)	Error (mm)	Depth Diameter (mm)	Error (mm)
Regular Dense	70.0	70.4	0.4	68.9	1.1
Regular Sparse	70.0	70.6	0.6	63.9	6.1
Random Dense	70.0	72.0	2.0	75.6	5.6
Random Sparse	70.0	70.8	0.8	66.9	3.1

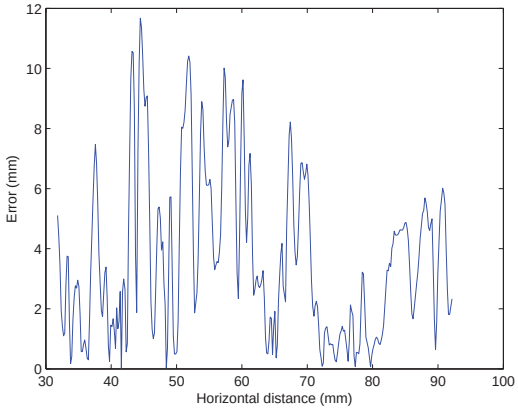
Table 6.1: A table showing the lateral and depth diameters when the ball is illuminated with different patterns

expected however, as the lateral resolution of the plenoptic camera is greater than the depth. It is also worth noting that the lateral diameter values are dependent on the initial threshold used when segmenting the ball, and can be changed to minimise the error on the lateral diameter. For this experiment, the threshold was not changed in order to minimise the error on the diameter, but instead was chosen to qualitatively fit the ball as accurately as possible. Due to this method being subjective by the user, and the errors on the lateral diameter being below 1mm for all patterns apart from the random dense, it is difficult to conclusively say if a particular pattern produces a more accurate measure for this value. When assessing the diameter which is calculated by taking the difference between the depth of the top of the ball and the mean of the background, the regular dense pattern has the lowest error compared to the other patterns. This does follow from the results in Chapter 5, in which the regular dense pattern produced the most accurate results when imaging a slope.

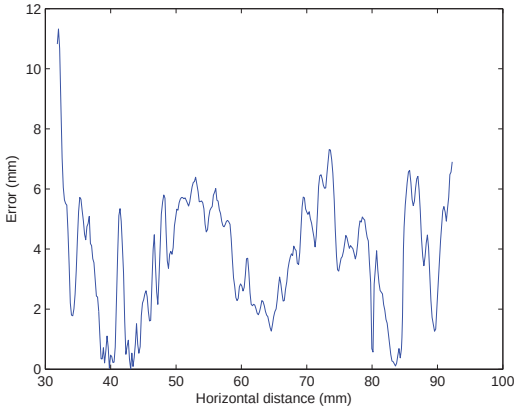
The second feature to examine the accuracy of the depth map is the curvature of the ball, as described in steps 8 - 10 from the method in Section 6.2. To assess the accuracy, both the mean absolute difference between the circle and the data and a plot of the difference across the ball are taken. These can both be seen in Figure 6.11.



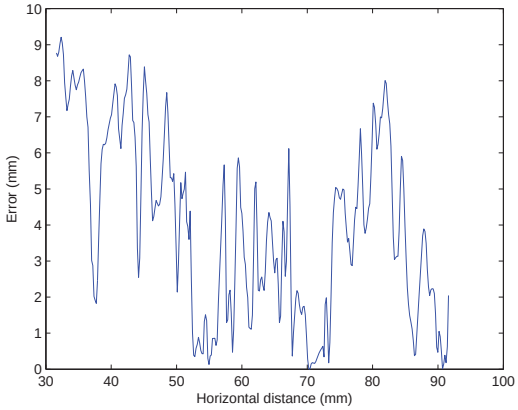
(a) Regular dense, mean error = 3.11mm



(b) Regular sparse, mean error = 3.71mm



(c) Random dense, mean error = 3.77mm



(d) Random Sparse, mean error = 4.01mm

Figure 6.11: Graphs showing the error in the curvature of the ball when illuminated with all four patterns.

By looking at the captions on the individual images in Figure 6.11, it is visible that the most accurate pattern is the regular dense, with a mean error of 3.11mm. This follows on from the diameter experiments in which this pattern yielded the most accurate results, and from Chapter 5 in which this pattern produced the most accurate reconstruction of a sloped scene. With regards to the shape of the plots in Figure 6.11, it is interesting to note that in general neither the edges or the middle of the plot appear to have a significantly greater or less error. There are some high frequency variations in the error, which is consistent with the high frequency errors present when imaging the slope in Chapter 4.

6.3.1 Fitting a sphere to the data

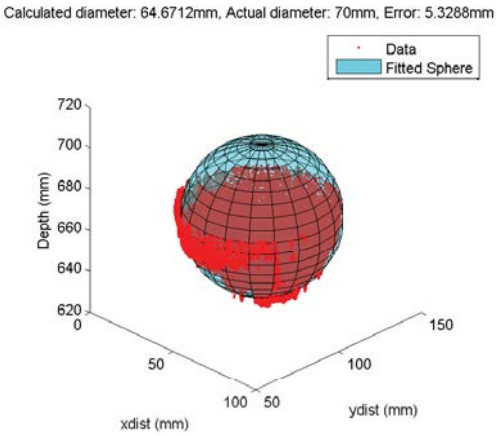
Further analysis on the data can be undertaken to help determine both the accuracy of the calculated depth, and about which pattern yields the most accurate measures of the ball by fitting a sphere to the data. The diameter of this fitted sphere can then be taken and compared to the known diameter of the imaged ball.

Method of fitting a sphere

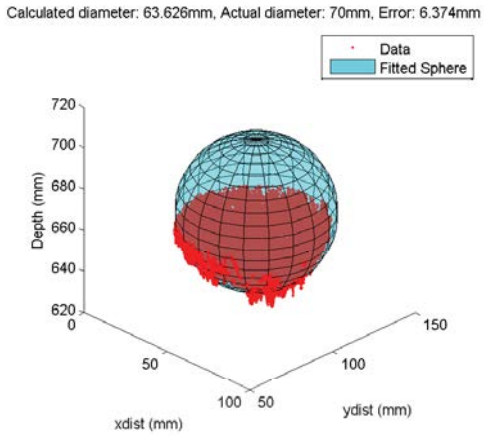
To calculate the diameter of the ball from a fitted sphere, a function written by Alan Jennings, University of Dayton [146] was used. The method works by using a closed form solution, minimising the following equation:

$$\Sigma \left((x - x_c)^2 + (y - y_c)^2 + (z - z_c)^2 - r^2\right)^2 \tag{6.1}$$

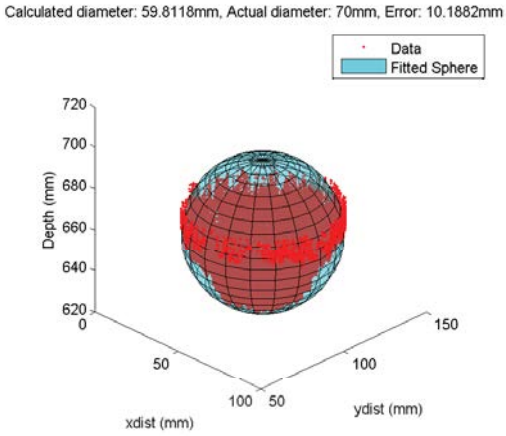
Where x, y and z are the data, x_c, y_c and z_c are the sphere's centre and r is the radius. The diameter of this fitted sphere can then be used to assess the accuracy of the depth data from the Raytrix.



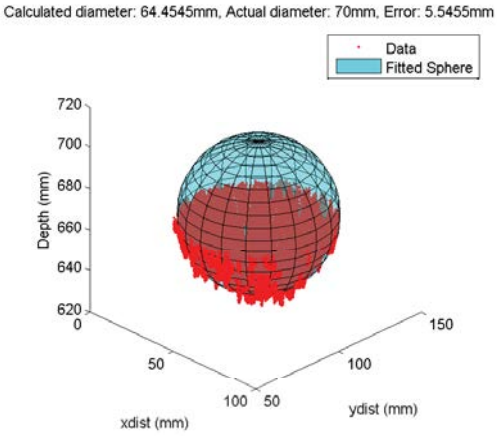
(a) Regular dense, diameter error = 5.33mm



(b) Regular sparse, diameter error = 6.37mm



(c) Random dense, diameter error = 10.19mm



(d) Random Sparse, diameter error = 5.55mm

Figure 6.12: Images showing the raw data fitted to a sphere with the calculated and actual diameters.

Results

The function written by Jennings [146] was used for depth maps of the same ball and patterns used throughout this chapter. Images showing the data points from the depth map along with the fitted sphere can be seen in Figure 6.12.

6.3.2 Comparing results of the flat plane to sphere

After the analysis just presented, it is now possible to compare the accuracy of results from the flat plane from Chapter 5 and of the sphere presented in this chapter. Table 6.2 highlight this comparison, taking the data from Tables 5.2 and 5.1 for the plane data

and from Figure 6.11 for the error on curvature of the sphere and 6.12 for the error in diameter of the sphere.

Pattern	Sphere curvature error (mm)	Sphere diameter error (mm)	Plane angle error (degrees)
Regular Dense	3.11	5.33	1.24
Regular Sparse	3.71	6.37	1.42
Random Dense	3.77	10.19	6.57
Random Sparse	4.01	5.55	2.13

Table 6.2: A table showing the comparison in error for the sloped plane and the sphere

Although the values in Table 6.2 cannot be directly compared as they are errors relating to different characteristics, certain conclusions can be drawn and inferred. Firstly is the larger errors on the random dense pattern. Looking into the errors for the sphere diameter, the smallest is from the regular dense pattern. If we take the error as a percentage of the actual diameter of the sphere (70mm), then this works out to be an error of 7.6%. Comparing this to the 1.24° error for the sloped plane, this intuitively seems to be much less than that of the sphere, although this can only be inferred.

6.4 Conclusion

Throughout this chapter, a new novel method for evaluating the accuracy of a plenoptic camera when imaging a sphere has been developed. This method was then applied to imaging a plain white ball when it was illuminated with a regular dense, regular sparse, random dense and random sparse patterns. The attributes used to assess the accuracy were the calculation of the diameter of the ball, which was physically measured before imaging, and the error in the curvature of a line plot taken through the centre of the ball. For both of these measures, the regular dense pattern produced the best results, with the diameter of the ball being accurate to 1.1mm and the mean error on the curvature of the

sphere being 3.11mm. It also provided the lowest error when fitting a sphere to the data, with the difference between the calculated and actual diameter being 5.33mm. This is consistent with the results from Chapter 5, when this same pattern was the one which produced the most accurate results when imaging a slope. It can therefore be concluded from these experiments that for a consistent or changing slope, a regular grid of densely packed dots produces the most accurate results.

It is important to relate these results to the overall aim of the project which is to use a plenoptic camera to image the retina. Firstly, the diameter of the sphere can be calculated accurately to 1.1mm. As the size of the ball used was 70mm, converting the accuracy to a percentage gives the diameter an accuracy of within 98.4%. If the error on the curvature of the sphere of 3.11mm is converted to a percentage, it yields an accuracy of 95.6%. If these accuracies can be replicated whilst imaging the retina, this would easily be adequate enough to diagnose many retina diseases such as glaucoma and diabetic macular edema.

The results throughout this chapter, along with in Chapters 4 and 5 a Raytrix R11 plenoptic camera has been used. This had been chosen due to the high number of focal planes present in the images as well as due to availability. However, the cost of this camera is in the order of tens of thousands of pounds sterling. For the application of retinal imaging, this causes many issues, as the simulations in Chapter 2 showed that the most plenoptic retinal imaging could hope for would be to gain a surface topography of the retina. This will provide less data than the current gold standard of retinal imaging, optical coherence tomography [81] which can gain both topographic images as well as information about deeper layers in the retina. Therefore, to make plenoptic retinal imaging worthwhile, it would need to be significantly cheaper than an OCT. It would then be very beneficial to be able to have a self built plenoptic retinal camera at the same cost as a fundus camera with the only added extra being a microlens array. To buy a microlens array, offcuts from previous orders can be bought for approximately £500, but often these do not have appropriate parameters for a plenoptic camera. Customised microlens

arrays can be purchased, however through face-to-face conversations with suppliers, to gain custom microlens arrays the cost would be around £10,000. This is increasing the overall cost of the system closer to the price of an OCT. It would then be beneficial to be able to build fully customisable microlens arrays cheaply.

As discussed briefly in Chapter 2, although it was found that surface topography images should in theory be able to be produced by a plenoptic camera when imaging the retina in isolation, the simulations did not include the transport of light through the cornea, lens and occular media. As these tissues are refractive they could produce errors into the data and gaining depth information of the retina by a plenoptic camera may be impossible. Another application for plenoptic imaging in the field of retinal imaging could be to image the outer surface of the cornea [147]. In this chapter it has been shown that with a Raytrix R11 camera, the accuracy of the curvature achievable is $\approx$ 3mm with a field of view of 180mm $\times$ 120mm. This is much larger than the size of a human eye, so much of the field of view of the camera would be wasted. A plenoptic camera optimised for that particular field of view and a depth of field comparable to the size of the corner would be necessary. A fully customised plenoptic camera would be the only solution to this problem. As previously discussed, a customised microlens array would be necessary to build a retinal plenoptic camera at a price comparable to a fundus camera. A method for developing cheap microlens arrays was devised and has been explored throughout the following chapter.

Chapter 7

Microlens arrays

7.1 Motivation

Commercial plenoptic cameras are available from companies such as Raytrix and Lytro. They can be very useful due to the built in rendering software, however in some situations they may not be applicable. They are often closed systems, cannot be self-customised and the cost of some of the cameras can also be in the order of tens of thousands of pounds, which could be too expensive. A viable option could then be to build a plenoptic camera, which would give more freedom dependent on the required use of the camera. As discussed in Sections 1.7 and 1.8, a plenoptic camera can be configured into either the traditional or focussed plenoptic system. If the traditional plenoptic configuration is chosen, the output resolution is dependent on the number of microlenses, so a large number of microlenses is required in order to obtain a high resolution rendered image. To increase the amount of viewpoints and virtual focal planes, the number of pixels under each microlens array needs to be high, requiring a high resolution sensor. Therefore to provide a plenoptic system with a high number of focal planes, viewpoints and rendered image resolution, a high resolution sensor and a microlens array with a high frequency of lenses is required. The system can be more flexible if it is placed in a focused plenoptic configuration, as in this case the rendered image resolution is decoupled from the number

of microlenses [17]. This can provide more flexibility in selecting appropriate parameters for the microlens array in the system, however there are a few important limitations independent of the configuration selected.

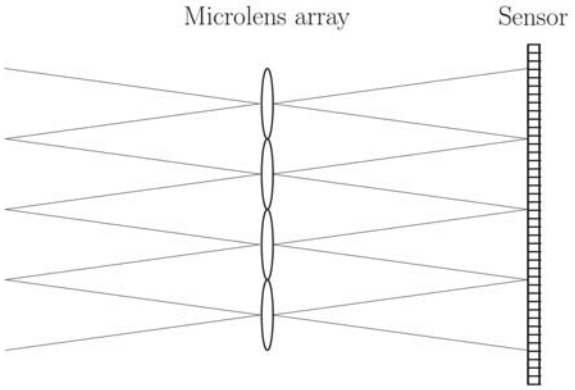
7.1.1 Microlens focal length

There is a minimum distance between the microlenses and the sensor due to the thickness of the array itself and the cover glass over the sensor. Typically this distance will be at least 1 - 1.5mm, so if the system is placed in the traditional plenoptic configuration the focal length of the microlens array will have to be exactly this distance. If the system is in the focused plenoptic configuration, the focal length of the microlens array can be shorter or larger than this distance but will need to be in the same order of magnitude.

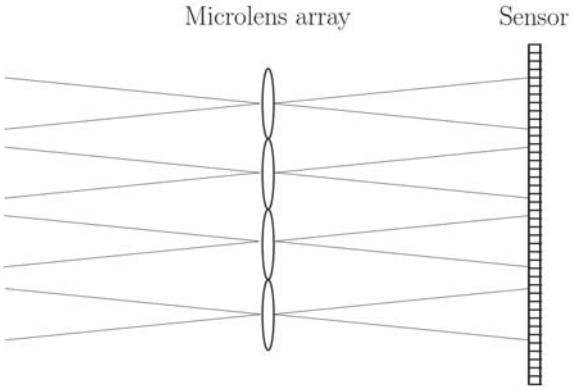
7.1.2 Matching f-numbers

Another important factor to consider is the matching of the f-numbers of the microlenses and the main lens as first described by Ng *et al* [3]. The aim of this is to maximise the number of useful pixels in the raw image by making the microimages created from each microlens as large as possible without overlapping. If the f-number of the main lens is less than the microlens array (or the aperture is larger in relation to the main lens focal length), then the microimages will overlap on the raw image. In contrast, if the f-number of the main lens is greater than the microlenses then the microimages will have spaces between then, creating wasted pixels on the raw image. Figure 7.1 shows this phenomena diagrammatically.

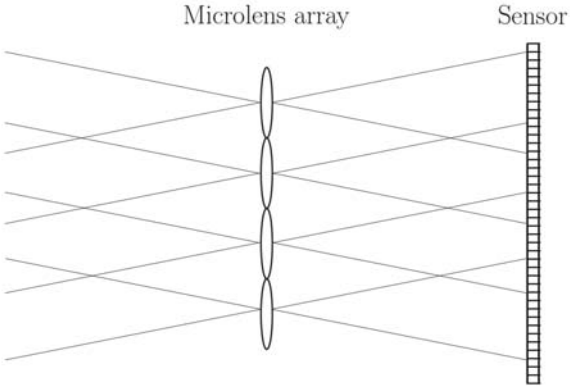
A major advantage of plenoptic imaging is the ability to gain an extended depth of field, similar to applying a small aperture on a conventional camera, whilst still gaining a high signal to noise ratio from having a large aperture, which is achieved by using a main lens with a small f-number. As the microlens f-number should match that of the main lens



(a) f-number main lens = f-number microlens



(b) f-number main lens > f-number microlens



(c) f-number main lens < f-number microlens

Figure 7.1: Images highlighting the importance of matching the f-number of the main lens to that of the microlens array. Image adapted from [3]

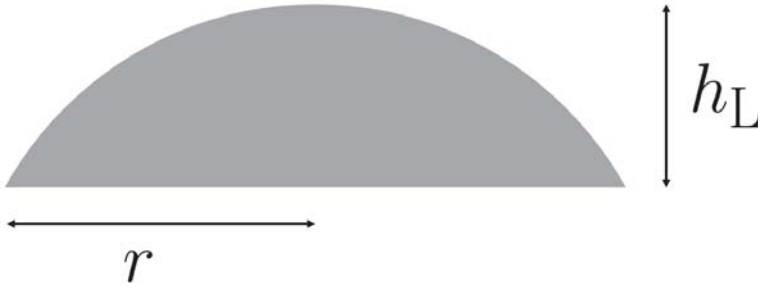


Figure 7.2: A diagram highlighting the parameters used in Equation 7.2.

to optimally fill the sensor with useful pixels. The f-number of a system, F is related to the focal length, f and the lens diameter, D as shown in Equation 7.1.

$$F = f/D \tag{7.1}$$

As previously discussed, the focal length of the microlens array needs to be in the order of 1 - 1.5mm. If we assume an f-number of $f/4$ and a focal length of 1.5mm, then using Equation 7.1 will produce a lens diameter of 0.375mm. If we then also assume a lens array of 10 $\times$ 10mm, the most common size of "off the shelf" commercial microlens arrays, a detector of at least this size and tightly packed hemispherical microlenses in a square array, then this would produce an array with 26 $\times$ 26 lenses. In a plenoptic system this is too few microlenses, especially if in the traditional configuration, where if no super-resolution algorithms are applied the rendered image would contain fewer than 700 pixels in total. Equation 7.2 shows the relationship between the focal length of the microlens, f, the lens height, h_L and the lens radius, $r = D/2$ for a spherical lens with wavelength dependent refractive index, $n(\lambda)$ [148]. A diagram showing these parameters can be seen in Figure 7.2.

$$f = \frac{h_{\mathrm{L}} + r^2/h_{\mathrm{L}}}{2(n(\lambda) - 1)} \tag{7.2}$$

With the use of Equation 7.2 it is possible to optimise the parameters of the microlens specific to the system required. New difficulties will then be faced when attempting to purchase a microlens array of the required parameters. Commercial companies, such as Thorlabs or SÜSS MicroOptics, will typically offer a catalogue of available microlens arrays, however these will usually be off-cuts from larger custom orders made, so there is little freedom is parameter selection, and typically these do not match those desirable for a plenoptic system. The second drawback of these "off the shelf" microlens arrays is the dimensions of them. Typically they come in 10 $\times$ 10mm arrays. To fully utilise all of the sensor in a plenoptic system, it would be optimal to cover the whole sensor with the microlens array. Due to the inevitable compromise between spatial and angular resolution, a sensor with a large number of pixels is required. Therefore sensors larger than 10 $\times$ 10mm are typically used, such as the 16MP medium format camera with a sensor size of 37.8 $\times$ 37.8mm used by Ng [32]. This creates significant limitations on using commercially bought off-cuts in a plenoptic system. It is possible to buy custom made arrays from industry, however, in order to make it worthwhile for the manufacturing company, they require a purchase of an entire 12 inch wafer of lenses which can cost in the order of tens of thousands of pounds. The aim of this chapter of work is to first assess the current methods of microlens fabrication, then to investigate new and innovative production methods for custom arrays at a fraction of the commercial cost.

7.2 Current methods of production

There are many suitable techniques for the production of refractive microlens arrays, but the most commonly used in industry is the resist-melting or reflow technique first achieved by Popovic *et al* [149]. In this technique, microlenses from tens to a few hundred microns in size were fabricated with an optical quality approaching the diffraction limit for high numerical apertures of N.A $\approx$ 0.5. The process used involved mask based lithography techniques, which is where a substrate is covered by a photoresist. A mask with a pattern

encoded on it via lithography is produced and aligned on top of the photoresist. The mask is then exposed with uniform illumination of either UV or X-rays from which the pattern is directly written onto the photosensitive layer [150]. This will then produce a rectangular pattern on the photoresist, but through heating, the columns melt into hemispherical microlenses [151]. There have also been advances in the fabrication of more exotic types of microlens arrays. Krupenkin *et al* [152] fabricated a tunable liquid microlens array which can vary the location and focal length of each individual microlens by electrowetting. Another approach involved liquid-filled lenses which are pneumatically activated for focal length tuning [153] [154]. This technique has been applied to a plenoptic camera by Lei *et al* [155] to electronically control the focal length of the microlens array in a plenoptic system.

7.3 Cost-effective production methods

After looking at how microlens arrays are currently fabricated in industry, it can be concluded that the methods are not suitable for production of custom lenses on a small scale due to the infrastructure required and the cost of this apparatus. Therefore, new methods need to be devised. As the overall aim of this investigation is to produce cost-effective and fully customisable lens arrays, it was important to fabricate them using inexpensive equipment. The methods which were developed for this investigation were done so with significant constraints in terms of available equipment. However, this was not seen as a disadvantage, because if successful this work could provide a method for fabricating custom microlens arrays which could be replicated with minimal costs.

The methods devised to attempt to solve this problem all follow the same basic production methodology. Firstly, an opposite mould to that of the microlens array needs to be fabricated. This has been attempted through a variety of methods that are shown throughout this section. Once the mould has been fabricated, polydimethylsiloxane (PDMS) has

been poured over the moulds and then left to solidify, producing the desired microlens arrays.

It is worth noting, all microlens arrays throughout this chapter have been fabricated into a rectangular grid pattern. Although a hexagonal grid pattern is known to maximise the number of lenses which can fit into a set space, the rectangular grid was chosen for ease of processing any images which would then be captured using the lens. This was not thought to be an issue for this study, as the aim is not to create the most space efficient microlens arrays, but instead to fabricate cost-effective lens arrays from new methodologies.

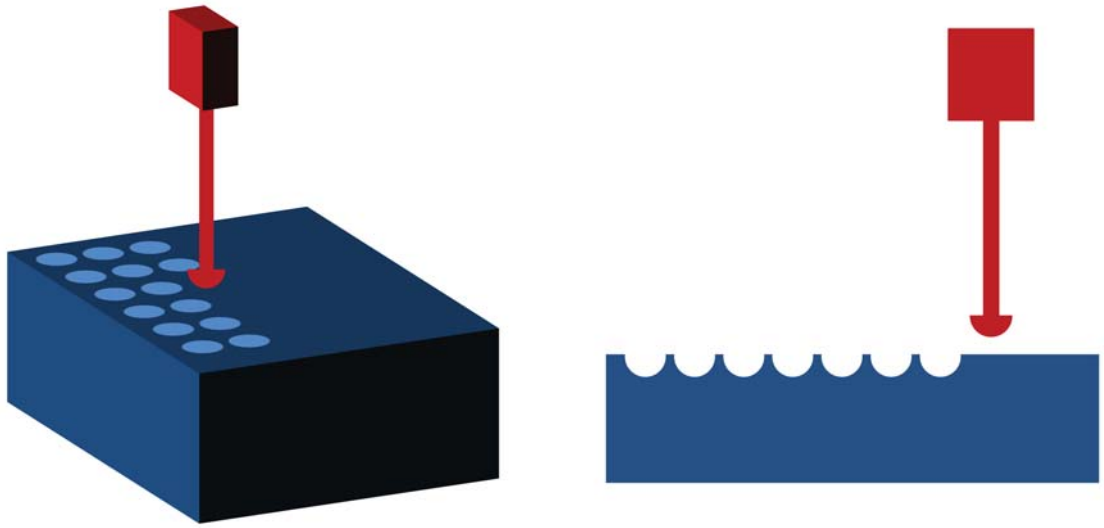
Selection of material

To create the microlenses, a suitable material had to be chosen. Polydimethylsiloxane (PDMS) was chosen as it is optically clear and generally non-toxic, non-flammable and inert [156]. It has a refractive index of ≈ 1.41, which is in the same order as glass and fused silica from which microlens arrays are commonly manufactured in industry. PDMS once set is also flexible, so will be able to be removed from moulds with relative ease. The implications of this in terms of the stability of the lenses is unknown, but it is believed initially the ease of removing the microlens array from the mould is a greater priority. It is formed by first producing silanol through the hydrolysis of dichlorodimethylsilane, where OH groups replace the chlorine atoms. The PDMS is then created through the dehydration and then distillation of silanol.

7.3.1 Hemispherical drill pressing

The first cost-effective method to produce a microlens array involved creating a mould by pressing a hemispherical drill end into a substrate. Copper and perspex were tried as the substrate materials. These were chosen because of their ability to be deformed permanently into the desired shapes, and the availability of the materials in the laboratory. The pressing is repeated sequentially with a CNC machine until the entire microlens pattern

is completed. This will produce a mould with the opposite shape to that of the desired microlens array, as shown in Figure 7.3. This method had been chosen as it is incredibly cheap and easy, and does not require the use of expensive equipment required for the photo-lithography techniques explained in Section 7.2.



(a) 3D view (b) Lateral view

Figure 7.3: Schematic diagram of the process in which the hemispherical drill end (red) is pressed into the perspex and copper (blue) to create the microlens array mould.

7.3.2 Micro-stereolithography acrylic mould

The second method to create a microlens array mould is through the use of a micro-stereolithography (MSL) machine in which photopolymerizable resins are exposed to patterned light to selectively cure (solidify) the resin [157]. The MSL machine used to create the acrylic mould was produced at the University of Warwick by Dr Simon Leigh, using a commercially available equipment produced by Envisiontec GmbH (Perfactory Mini Multilens) with a build envelope of 28 $\times$ 21 $\times$ 230 mm and a resolution of 1400 $\times$ 1050 pixels [158]. This process produces a mould by digital projections selectively curing (solidifying) a layer of a photosensitive resin. Once a layer is completely cured, the build surface is moved in the order of tens of microns, after which a new layer of resin is added.

Figure 7.4: A schematic of a single microlens mould fabricated via the MSL method

This process is then repeated until the final microlens array mould is completed. The resin used for the fabrication of this mould was the commercially available R11 manufactured by Envisiontec GmbH. A 2D schematic of a single microlens mould fabricated by this method can be seen in Figure 7.4.

The MSL was chosen to create a microlens array mould mainly due to the availability of the equipment. As the method involves curing a layer of photoresist followed by a small translation of the build surface, it is possible that the PDMS microlens arrays cast from the mould could be stepped. It is therefore important to assess the microlens arrays through a variety of methods to see if these errors occur.

7.4 Analysis of microlens arrays

The methods described previously were performed to fabricate a microlens array for each method. It is now important to find suitable methods to accurately assess the feasibility of each method and quality of the microlens arrays.

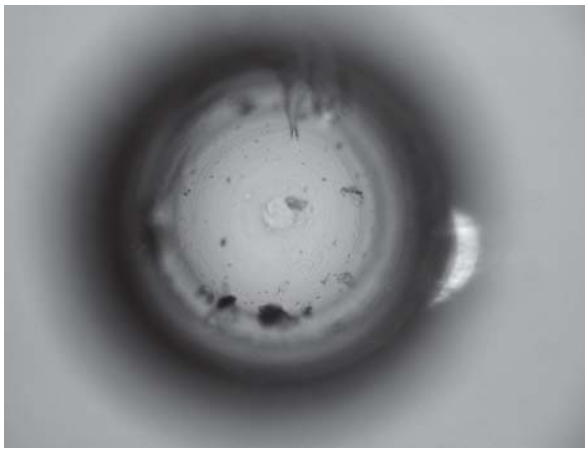
7.4.1 Bright field microscopy

The first method to assess the self fabricated microlens arrays is with a bright field microscope. The microscope used is a Zeiss Axiovert 200M, with a 10x 0.3NA Zeiss plan-neofluar objective, imaged with a Hamamatsu Orca AG CCD camera. This has been chosen due to the ease of imaging individual lenses with this microscope and the availability of the equipment at the University of Birmingham. The microscope was focussed on different heights of the lenses and each image has been assessed qualitatively to look at the quality of the lens. Images of the top, middle and bottom of a single microlens for the arrays fabricated from the perspex, copper and micro-stereolithography moulds can be found in Figures 7.5, 7.6 and 7.7 respectively. For reference, a commercial microlens array from SÜSS MicroOptics has also been imaged and can be seen in Figure 7.8.

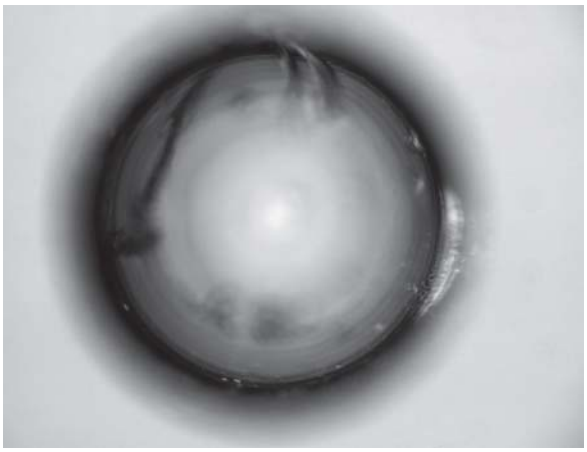
From assessing the images in Figure 7.5 and 7.6, qualitatively the lenses fabricated from the perspex mould do appear to produce fewer deformities than copper, but significantly more than the commercial microlens array in Figure 7.8. The mould that produces the fewest deformities is that from the MSL mould, however circular lines do appear throughout the lens. This is expected, as the mould itself follows a stepped pattern rather than a continuous curve. Further analysis of the lenses will help to confirm this and to gain some quantitative information about the microlenses. A grid pattern is also visible on the MSL microlens array. This is due to the mould itself containing the pattern as an inherent effect of the fabrication method. It is not know the effect this will have on the optical performance of the lens.

7.4.2 Interferometry

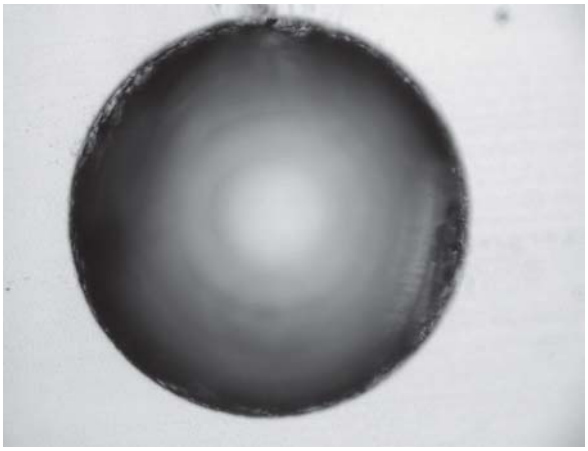
One method to gain quantitative information about the size and shape of microlenses is through interferometry [159]. Due to availability in the laboratory, images were taken using a commercial interferometer (MicroXAM2,KLA Tencor (USA)), with the results of the microlens arrays fabricated from the perspex, copper, MSL moulds and the com-



(a) Image taken whilst focusing on the top of the lens

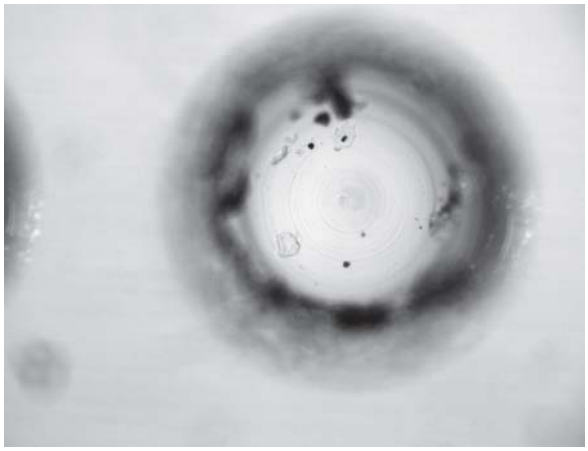


(b) Image taken whilst focusing in the middle of the lens

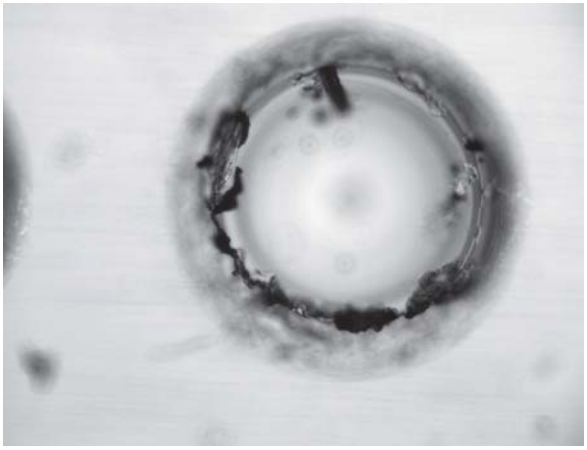


(c) Image taken whilst focusing at the base of the lens

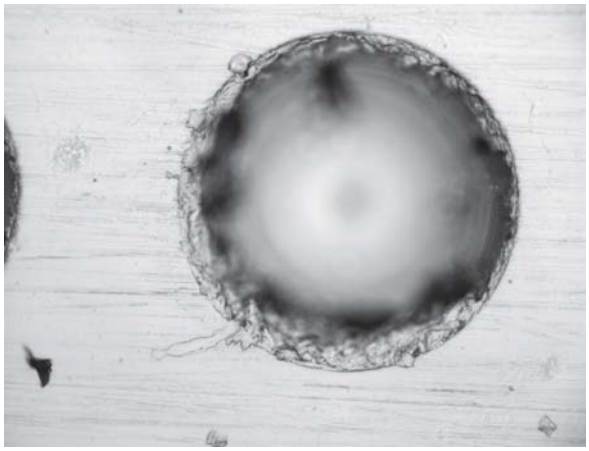
Figure 7.5: Bright field microscopy images of the fabricated microlenses created via pressing a hemispherical drill into a block of clear plastic (see Section 7.3.1).



(a) Image taken whilst focusing on the top of the lens

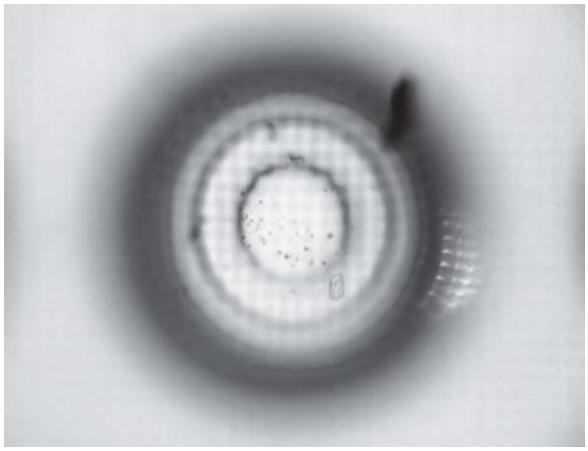


(b) Image taken whilst focusing in the middle of the lens

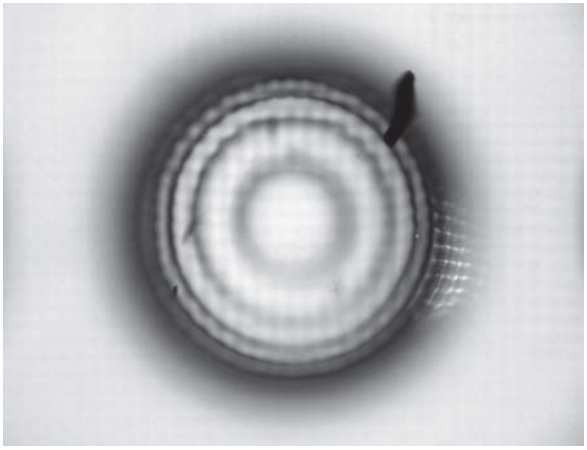


(c) Image taken whilst focusing at the base of the lens

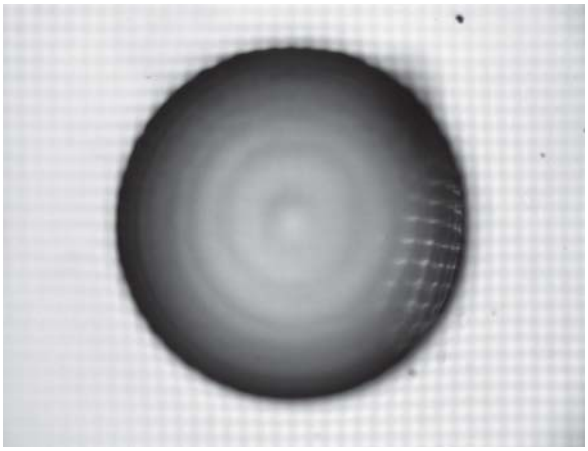
Figure 7.6: Bright field microscopy images of the fabricated microlenses created via pressing a hemispherical drill into a block of copper (see Section 7.3.1).



(a) Image taken whilst focusing on the top of the lens

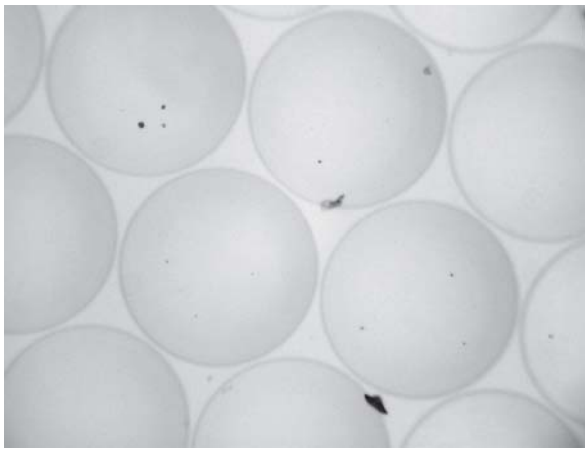


(b) Image taken whilst focusing in the middle of the lens

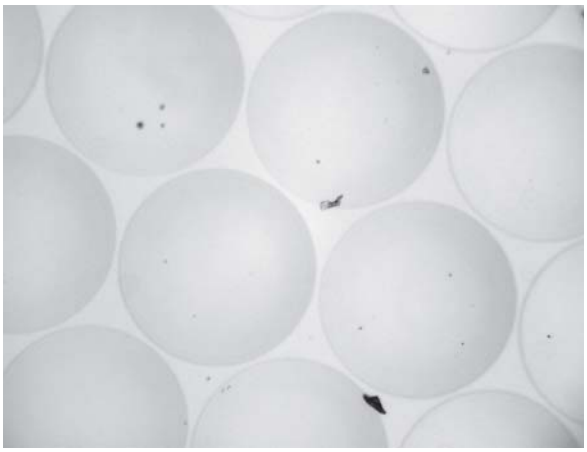


(c) Image taken whilst focusing at the base of the lens

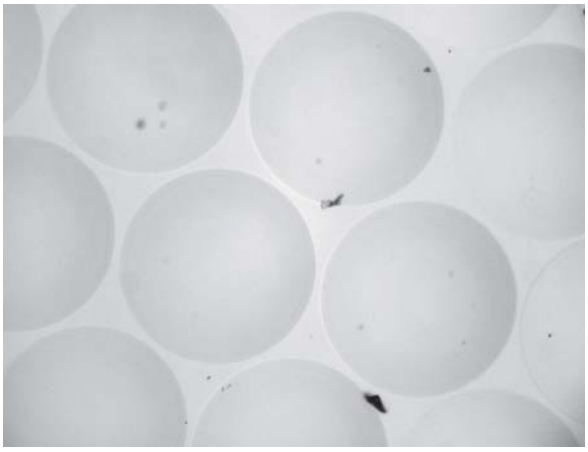
Figure 7.7: Bright field microscopy images of the fabricated microlenses created via the MLA acrylic mould (see Section 7.3.2).



(a) Image taken whilst focusing on the top of the lens



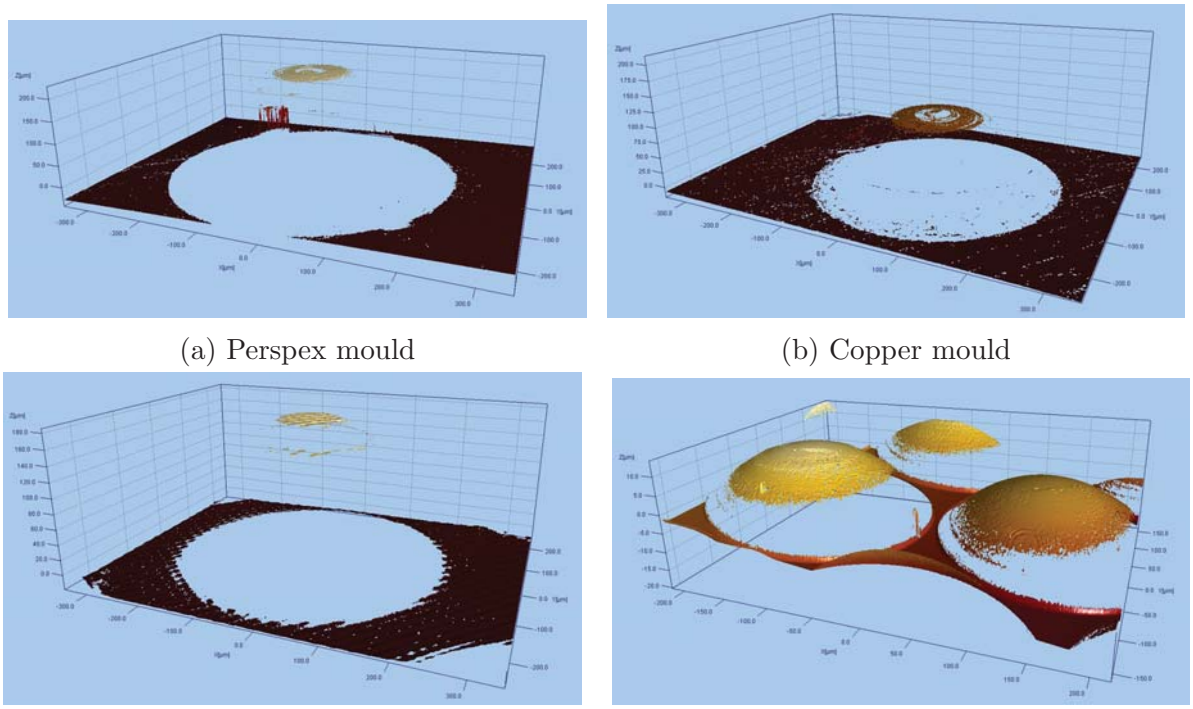
(b) Image taken whilst focusing in the middle of the lens



(c) Image taken whilst focusing at the base of the lens

Figure 7.8: Bright field microscopy images of commercial microlenses purchased from SÜSS MicroOptics (SÜSS MicroTech AG).

mercially bought microlens array in Figure 7.9. Images are produced by analysing the interference pattern of a beam of light incident on the object (in this case the microlens array), and an identical beam which has reflected off of a reference material. In the case of the MicroXAM2, the reference material is a flat mirror perpendicular to the incident beam. This is not ideal for imaging spherical lenses, as angles greater than $\approx 10°$ will not produce any results. Therefore, only areas of the lenses which are within this threshold, i.e. the top of the lens and the base surface, will produce any results.



(a) Perspex mould

(b) Copper mould

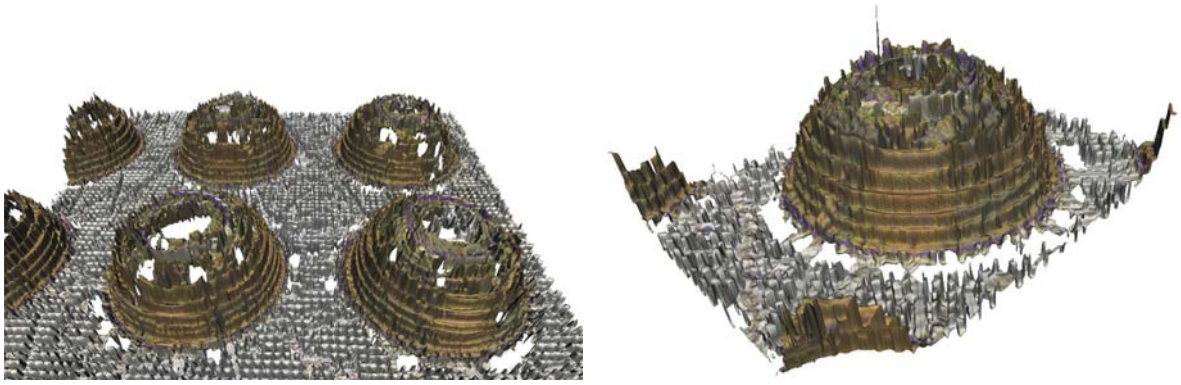
(c) Micro-stereolithography mould

(d) Commercially bought microlens array

Figure 7.9: Images taken using a MicroXAM2,KLA Tencor (USA) interferometer of three self fabricated microlens arrays and one commercially bought array.

On first inspection of the interferometry images in Figure 7.9 it is apparent that, as predicted, only the top of the lens and the base produce any reading. This is due to the reference surface in the interferometer being flat, so only surfaces within 10° to the plane which is parallel to the base of the lens array will produce results. As the reference surface cannot be changed in this particular interferometer, it will not be possible to inspect the surface of the lenses with this method. However, examination of the lens

fabricated from the MSL mould suggests that it may be stepped, in agreement with the bright field microscopy images in Figure 7.7. In an attempt to verify the if these steps are present, images of the MSL microlens array were taken using a surface profiler available in the laboratory built by Alicona. The results of this can be seen in Figure 7.10.



(a) View of multiple lenses (b) View of single lens

Figure 7.10: Images taken with the Alicona of the microlens array fabricated from the MSL mould

These images do clearly show the steps present in this particular microlens array. As these would likely cause optical aberrations in an image focussed by them, a new method to smooth the lenses was devised. Due to the time constraints on this part of the project, only the MSL microlens array was imaged using the surface profiler. This was chosen as it appeared qualitatively to produce the most consistent lenses with fewer inaccuracies. If more time was available each microlens array would have been assessed through this analysis technique.

7.5 PMMA casting to improve surface profile

In order to smooth the stepped lenses produced from the micro-stereolithography acrylic mould, an additional step was added to the process. Firstly, an opposite mould to that used in Section 7.3.2 was created, and can be seen schematically in Figure 7.11. Polymethyl methacrylate (PMMA) is then cast over the top of this mould. This resin has been chosen as when it is heated, it undergoes reflow. This would therefore mean

Figure 7.11: A schematic of a single microlens mould fabricated via the MSL to be used in the PMMA casting method to smooth the stepping created by MSL.

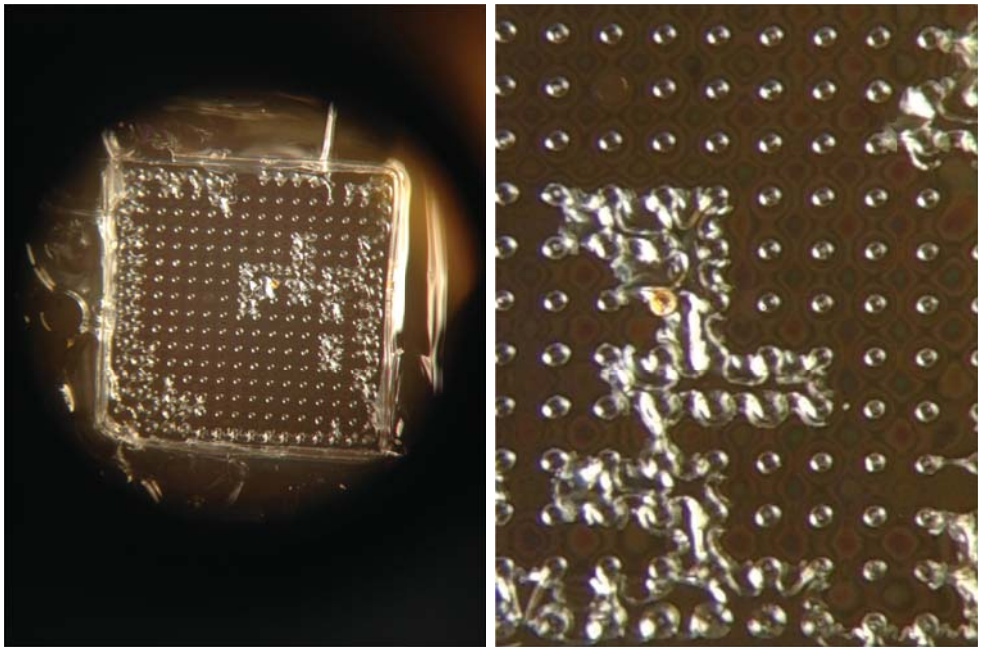
that if a stepped mould can be cast into PMMA, when heated the stepped features will create the desired smooth surface.

An initial attempt was made to spin coat PMMA onto the slide from which the mould would be pressed, however after initial experiments it was found that the maximum thickness of PMMA that could practically be coated the slide was $\approx 10\mu$m, which is too thin for the desired height of each microlens in the order of 100μm. Because of this the PMMA was poured onto the slide until above the required thickness. Images showing the results of this method can be seen in Figure 7.12.

From analysing the images in Figure 7.12 it is clear that many errors have occurred in the pressing process. This is due to the acrylic mould not releasing from the PMMA easily, causing some significant deformations into the shape of the mould. Many attempts were made to remove these errors, but none were successful. Future work to remove the stepping of the microlens arrays from the MSL acrylic mould cannot be achieved through the current proposed method using PMMA. It may be possible to insert a lubrication layer between the mould and the PMMA to stop the PMMA from sticking to the mould, however due to time restrictions this was not investigated.

7.6 Conclusion

Throughout this chapter attempts have been made to develop novel, cheap methods of fabricating custom microlens arrays to be used in plenoptic cameras. The initial methods



(a) An image showing the PMMA mould after pressing with the mould shown in Figure 7.11.

(b) A zoomed in photograph of Figure 7.12a highlighting the errors present in the mould.

Figure 7.12: Photographs showing the PMMA mould with an area enlarged to highlight the errors.

proposed used a hemispherical drill piece being pressed into blocks of copper and perspex, and the use of a micro-stereolithography printer to make an acrylic mould from which the lenses could be fabricated. The first results showed numerous errors in the lens arrays fabricated via the use of the hemispherical drill piece, and far fewer with the MSL mould, so this method was pursued further. Ideally all of the methods would have been analysed to the same extent, but as this investigation had such strong time constraints this was not possible.

The array fabricated via the MSL mould had a stepping to the lens shape, inherent from the production method. Attempts were made to remove these through the use of reflow of PMMA. Unfortunately this method did not manage to remove the errors as it added many more abnormalities to the lens array than before. These could potentially be removed by the use of a lubricated layer between the mould and the PMMA, however due to time

constraints this was not tested. It is believed this method showed the most promise, so to advance this project further this avenue is the one which should be progressed.

Although initial results did seem promising from this chapter, a much greater amount of work would need to be conducted before the lens arrays fabricated from these cheap methods could be achieved. This shows that fabricating cheap microlens arrays for specific applications such as retinal imaging is not straight forward. Future work could involve looking into fabricating lenses of different sizes which may make the fabrication process easier, such as lenses with a smaller lens height. This would enable spin coated PMMA to be used, and if the diameter of the lens did not change, more of the lens visible in the interferometer results used throughout this chapter. This would greatly aid in the assessment of the surface profile of the lenses.

Chapter 8

Discussion and future directions

The principal aim of this thesis was to investigate how the fields of plenoptic and retinal imaging can be combined, and whether it could ultimately help with the diagnosis of retinal diseases through 3D imaging of the retina without the need for an OCT imaging device. The main novel contributions presented in this thesis were:

Computational simulations to show how plenoptic imaging is not applicable for layered imaging of the retina

This study was important in determining the future direction of the thesis. This was a novel contribution as it used Monte Carlo simulations to look at light propagation in the retina to determine the functionality of imaging the retina with a plenoptic camera. It was hypothesised that if enough photons reflect off the retinal pigment epithelium (RPE) without any scattering, then neural retinal thickness could be calculated from the plenoptic image. It was found, however, that only 0.35% of light even reach the RPE with the same orientation as the initial photon, which was deemed too small for the hypothesis to be correct. Therefore, layered imaging of the retina would not be possible with a plenoptic camera [50].

Application of textured projection and depth analysis for plenoptic retinal imaging

Initial images of a sheet with an image of a retina (see Figure 4.1) showed only having recoverable depth from the small amount of detail present. Protocols were then developed, with the used of projected texture, to recover depth from an angled plane to an accuracy of 0.13° with the intended application of recovering 3D surface measurements of the retina. By applying projected texture onto a scene with limited features such as a retinal surface, more accurate depth measurements can be made.

Metric recovery based on a developed work-flow rather than relative

A protocol was developed to convert the virtual depth into metric depth values based on a novel method. This method is easy to implement, and is independent of the plenoptic camera used to capture the images.

Investigation of characteristics and concepts of plenoptic based imaging

Thorough characterisation of different projection patterns to improve depth recovery from a plenoptic image. It was found that the largest contributing factor was the density of the pattern, which if this becomes too dense the accuracy drops off significantly, with errors of over 27° on a 45° plane under these circumstances. Protocols were also developed for imaging spherical objects, and recovering metric depth from a plenoptic depth map outputting relative virtual depth values.

Novel cost effective production of microlens arrays

With the aim of producing an alternative to optical coherence tomography (OCT) for retinal imaging, it was imperative to be cheaper, as the functionality and resolution of images was unlikely to better OCT. With the cost of purchasing custom microlens arrays nearing £10,000, cheaper methods of production are necessary. Three novel methods were devised for a fraction of the current cost. Further work is necessary to make these

functional in plenoptic cameras, but a strong foundation for future work has been developed.

8.1 Discussion

The investigation took on this format to try to look at the problem of combining plenoptic and retinal imaging from many different directions, whilst also advancing current methods and proposing new ones. By first investigating computationally the potential capabilities of imaging the retina with a plenoptic camera, this helped to drive the project towards the more achievable goal of imaging the surface and not attempting to refocus through the different layers. As one of the initial ideas was to refocus through the layers, the simulations proved this would not be possible so the focus of the project shifted to producing a surface profile of the retina.

The next section of the study involved assessing the accuracy of the Raytrix R11 plenoptic camera. Initial images were taken of objects in the laboratory, and it was quickly noted that the areas in scenes which had little or no detail did not produce an accurate depth. This did cause some worry, as large portions of the retina have little detail. However, by knowing that the depth algorithms worked by using the disparity of objects in neighbouring microimages, this firstly explained why only areas with detail recorded depth, but also led to the idea of projecting a pattern to artificially create texture.

Once the accuracy of the calculated depth was shown to increase by projecting a pattern, a more systematic approach was taken to optimise the method. A test into using either a laser or LED projector was performed, with the results being in favour of the LED projector. The use of this projector was then validated against the same pattern being printed onto a sheet, with a negligible difference between the two.

This work led on to one of the most significant parts of the thesis. A novel method for gaining quantitative data from a plenoptic camera was devised. By performing a

calibration set of images at known distances, a relationship between the virtual and real depths was found. This method was then used to quantitatively calculate the angle of a sloped scene with a pattern projected onto it. Further factors had to be accounted for such as correcting the field of view and removing the high frequency noise. A sloped scene was chosen for the initial experiments due to the relative ease in calculating how accurately the depth has been reconstructed and that it has a continuously changing depth.

The investigation then led on to assessing the effect of different patterns on the accuracy of the depth reconstruction. The density and regularity of the pattern were tested and found that the regular dense pattern produced the most accurate results. An interesting point to note was that when the angles became large, the dense patterns become significantly less accurate. This is thought to be due to the density with reference to the sensor plane becoming too large, and different features falling onto a single pixel.

The final investigation using the commercial Raytrix R11 camera involved replacing the simple sloped scene with a sphere. This was performed to see the effect of a more complex shape had on the accuracy of the depth reconstruction, and to be more relevant to imaging the eye. A method was devised to assess the accuracy of the depth map. The conclusion on which pattern yielded the most accurate results was the same as with the slope, the regular dense pattern.

With the principal aim of the thesis being to develop a more cost effective method of gaining 3D images of the retina, using a Raytrix which costs similar to an OCT does not fulfil the aim. Therefore a more cost effective plenoptic camera needs to be developed. Although microlens arrays can be bought commercially, finding one with the correct parameters proved very difficult, so a custom array was needed. However, these are only available commercially in volume and at great cost and are unsuitable for prototype development. Therefore work began into devising novel, cheap methods of fabricating custom microlens arrays to build a customised plenoptic camera. The methods tested

involved creating a mould by pressing a hemispherical drill piece into a block of copper and perspex, and also through the use of a micro-stereolithography (MSL) acrylic mould. Initial results showed fewer errors with the PDMS cast microlens arrays from the MSL mould, so this method was progressed further.

The errors which did occur in the MSL microlens array were a stepping, which is caused from the process by which the mould was fabricated. Attempts were then made to remove these stepping artefacts through the use of reflow mechanisms on a PMMA mould. Initial results were unsuccessful as the mould could not be removed form the PMMA without introducing significant artefacts, so more research will be needed in order to fabricate custom microlens arrays for a limited expense.

8.2 Future directions

This thesis has presented an initial study into the plausibility of combining plenoptic imaging with the field of retinal imaging. To further the project and move closer to the overall aim of imaging the retina in 3D with a single snapshot, a few more areas will need to be advanced. One of these is the work into fabricating custom microlens arrays at a reasonable cost. Although preliminary experiments have produced encouraging results, it is believed that a much larger project will be necessary to develop the microlens arrays to a workable quality. Another avenue for the future research could be into different types of microlens arrays other than refractive optics. One example of this would be to use Fresnel zone plate arrays, which have been developed by Firester *et al* [160], but not for the application of plenoptic imaging.

Another future direction this area of research could go into would be to use a custom Raytrix camera and attach this to a fundus imaging device in order to test how well a plenoptic camera can image a retina. This would have been the desirable option for this project, however attempts to use the R11 on a fundus camera were unsuccessful due to it

not being optimised for the required field of view or depth of field of a retina, and there were insufficient funds to purchase one with the required parameters.

It is the authors belief that the major future potential for plenoptics lies outside of field of retinal imaging. A major area of research which would benefit from such a technology would be in robotics. The ability to gain a 3D representation of the world from a single image could offer huge benefits for robots, self-driving cars and even in the field of robotic surgery. As was shown in Chapter 2, plenoptic imaging does not fair very well in environments with scattering media. However, quantitative depth of objects in free space can already be achieved. Therefore, using plenoptic imaging for applications which deal completely in free space and non-scattering environments, there is huge potential and should be investigated further.

Fundamentally this project was designed to test the feasibility of plenoptic retinal imaging, and to determine if future research into the field would be worthwhile. Even though the results presented did advance the field in terms of quantitative imaging, with current available technology it can be concluded that further research into plenoptic retinal imaging is not one which should be pursued.

Bibliography

[1] Christian Perwass and Lennart Wietzke. Single lens 3d-camera with extended depth-of-field. In *Proc. SPIE*, volume 8291, page 829108, 2012.

[2] Rupert RA Bourne, Seth R Flaxman, Tasanee Braithwaite, Maria V Cicinelli, Aditi Das, Jost B Jonas, Jill Keeffe, John H Kempen, Janet Leasher, Hans Limburg, et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. *The Lancet Global Health*, 5(9):e888–e897, 2017.

[3] Ren Ng, Marc Levoy, Mathieu Brédif, Gene Duval, Mark Horowitz, and Pat Hanrahan. Light field photography with a hand-held plenoptic camera. *Computer Science Technical Report CSTR*, 2:11, 2005.

[4] Michael R Hee, Caroline R Baumal, Carmen A Puliafito, Jay S Duker, Elias Reichel, Jason R Wilkins, Jeffery G Coker, Joel S Schuman, Eric A Swanson, and James G Fujimoto. Optical coherence tomography of age-related macular degeneration and choroidal neovascularization. *Ophthalmology*, 103(8):1260–1270, 1996.

[5] A. Gershun. The light field. moscow. *Journal of Mathematics and Physics*, 18, 1936.

[6] Ian Ashdown. Near-field photometry: A new approach. *Journal of the Illuminating Engineering Society*, 22(1):163–180, 1993.

[7] M. Levoy and P. Hanrahan. Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 31–42. ACM, 1996.

[8] Steven J Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F Cohen. The lumigraph. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 43–54. ACM, 1996.

[9] E.H. Adelson and J.Y.A. Wang. Single lens stereo with a plenoptic camera. *IEEE transactions on pattern analysis and machine intelligence*, 14(2):99–106, 1992.

[10] M. Levoy, Z. Zhang, and I. McDowall. Recording and controlling the 4d light field in a microscope using microlens arrays. *Journal of Microscopy*, 235(2):144–162, 2009.

[11] Robert C Bolles, H Harlyn Baker, and David H Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *International Journal of Computer Vision*, 1(1):7–55, 1987.

[12] Michael W Halle. Holographic stereograms as discrete imaging systems. In *IS&T/SPIE 1994 International Symposium on Electronic Imaging: Science and Technology*, pages 73–84. International Society for Optics and Photonics, 1994.

[13] Bennett S Wilburn, Michal Smulski, Hsiao-Heng K Lee, and Mark A Horowitz. Light field video camera. In *Electronic Imaging 2002*, pages 29–36. International Society for Optics and Photonics, 2001.

[14] Gabriel Lippmann. Epreuves reversibles. photographies integrals. *Comptes-Rendus Academie des Sciences*, 146:446–451, 1908.

[15] Gabriel Lippmann. Epreuves reversibles donnant la sensation du relief. *J. Phys. Theor. Appl.*, 7(1):821–825, 1908.

[16] Herbert E Ives. A camera for making parallax panoramagrams. *JOSA*, 17(6):435–437, 1928.

[17] Todor G Georgiev and Andrew Lumsdaine. Resolution in plenoptic cameras. In *Computational Optical Sensing and Imaging*. Optical Society of America, 2009.

[18] Andrew Lumsdaine and Todor Georgiev. The focused plenoptic camera. In *Computational Photography (ICCP), 2009 IEEE International Conference on*, pages 1–8. IEEE, 2009.

[19] Todor Georgiev and Andrew Lumsdaine. The multifocus plenoptic camera. In *IS&T/SPIE Electronic Imaging*, pages 829908–829908. International Society for Optics and Photonics, 2012.

[20] Aaron Isaksen, Leonard McMillan, and Steven J Gortler. Dynamically reparameterized light fields. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 297–306. ACM Press/Addison-Wesley Publishing Co., 2000.

[21] Ren Ng. Fourier slice photography. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 735–744. ACM, 2005.

[22] Andrew Lumsdaine and Todor Georgiev. Full resolution lightfield rendering. *Indiana University and Adobe Systems, Tech. Rep*, 2008.

[23] Todor Georgiev and Andrew Lumsdaine. Superresolution with plenoptic camera 2.0. *Adobe Systems Incorporated, Tech. Rep*, 2009.

[24] Tom E Bishop, Sara Zanetti, and Paolo Favaro. Light field superresolution. In *Computational Photography (ICCP), 2009 IEEE International Conference on*, pages 1–9. IEEE, 2009.

[25] Todor Georgiev and Andrew Lumsdaine. Focused plenoptic camera and rendering. *Journal of Electronic Imaging*, 19(2):021106–021106, 2010.

[26] T Georgiev and A Lumsdaine. Reducing plenoptic camera artifacts. In *Computer Graphics Forum*, volume 29, pages 1955–1968. Wiley Online Library, 2010.

[27] Zhan Yu, Jingyi Yu, Andrew Lumsdaine, and Todor Georgiev. Color demosaicing in plenoptic cameras.

[28] Todor Georgiev. Plenoptic camera resolution. In *Computational Optical Sensing and Imaging*, pages JTh4A–2. Optical Society of America, 2015.

[29] Tom E Bishop and Paolo Favaro. Plenoptic depth estimation from multiple aliased views. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 1622–1629. IEEE, 2009.

[30] Tom E Bishop and Paolo Favaro. Full-resolution depth map estimation from an aliased plenoptic light field. In *Computer Vision–ACCV 2010*, pages 186–200. Springer, 2011.

[31] Tom E Bishop and Paolo Favaro. The light field camera: Extended depth of field, aliasing, and superresolution. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(5):972–986, 2012.

[32] Ren Ng. *Digital light field photography*. PhD thesis, stanford university, 2006.

[33] Todor Georgiev, Zhan Yu, Andrew Lumsdaine, and Sergio Goma. Lytro camera technology: theory, algorithms, performance analysis. In *IS&T/SPIE Electronic Imaging*, pages 86671J–86671J. International Society for Optics and Photonics, 2013.

[34] Ulrich Perwass and Christian Perwass. Digital imaging system, plenoptic optical device and image data processing method, December 31 2013. US Patent 8,619,177.

[35] Richard J Marshall, Chris J Meah, Massimo Turola, Ela Claridge, Alex Robinson, Kai Bongs, Steve Gruppetta, and Iain B Styles. Improving depth estimation from

a plenoptic camera by patterned illumination. In *SPIE Optical Metrology*, pages 952815–952815. International Society for Optics and Photonics, 2015.

[36] Chris Meah, Massimo Turola, Richard J Marshall, Ela Claridge, Alex Robinson, Kai Bongs, Steve Gruppetta, and Iain Styles. Towards plenoptic multi-view imaging. In *Propagation through and Characterization of Distributed Volume Turbulence and Atmospheric Phenomena*, pages JT5A–42. Optical Society of America, 2015.

[37] Ryan Decker, Azad Shademan, Justin Opfermann, Simon Leonard, Peter CW Kim, and Axel Krieger. Performance evaluation and clinical applications of 3d plenoptic cameras. In *SPIE Sensing Technology+ Applications*, pages 94940B–94940B. International Society for Optics and Photonics, 2015.

[38] Marc Levoy, Ren Ng, Andrew Adams, Matthew Footer, and Mark Horowitz. Light field microscopy. *ACM Transactions on Graphics (TOG)*, 25(3):924–934, 2006.

[39] Robert Prevedel, Young-Gyu Yoon, Maximilian Hoffmann, Nikita Pak, Gordon Wetzstein, Saul Kato, Tina Schrödel, Ramesh Raskar, Manuel Zimmer, Edward S Boyden, et al. Simultaneous whole-animal 3d imaging of neuronal activity using light-field microscopy. *Nature methods*, 2014.

[40] JM Rodríguez-Ramos, JG Marichal-Hernández, JP Luke, J Trujillo-Sevilla, M Puga, M López, JJ Fernández-Valdivia, C Dominguez-Conde, JC Sanluis, F Rosa, et al. New developments at cafadis plenoptic camera. In *Information Optics (WIO), 2011 10th Euro-American Workshop on*, pages 1–3. IEEE, 2011.

[41] Eduardo Magdaleno, Manuel Rodríguez, and José Manuel Rodríguez-Ramos. An efficient pipeline wavefront phase recovery for the cafadis camera for extremely large telescopes. *Sensors*, 10(1):1–15, 2009.

[42] Kyle Lynch, Tim Fahringer, and Brian Thurow. Three-dimensional particle image velocimetry using a plenoptic camera. In *AIAA Aerospace Sciences Meeting*, 2012.

[43] Brian S Thurow and Timothy Fahringer. Recent development of volumetric piv with a plenoptic camera. In *PIV13; 10th International Symposium on Particle Image Velocimetry, Delft, The Netherlands, July 1-3, 2013*. Delft University of Technology, Faculty of Mechanical, Maritime and Materials Engineering, and Faculty of Aerospace Engineering, 2013.

[44] Timothy W Fahringer and Brian S Thurow. 3d particle position reconstruction accuracy in plenoptic piv. In *52nd Aerospace Sciences Meeting: American Institute of Aeronautics and Astronautics*, 2014.

[45] Timothy W Fahringer and Brian S Thurow. Comparing volumetric reconstruction algorithms for plenoptic-piv. *AIAA SciTech*, (53rd):1–10, 2015.

[46] Massimo Turola and Steve Gruppetta. Wave optics simulations of a focused plenoptic system. In *Frontiers in Optics*, pages JTu3A–24. Optical Society of America, 2014.

[47] Massimo Turola, Chris J Meah, Richard J Marshall, Iain B Styles, and Stephen Gruppetta. Effect of wavefront aberrations on a focused plenoptic imaging system: a wave optics simulation approach. In *SPIE Optical Metrology*, pages 95260X–95260X. International Society for Optics and Photonics, 2015.

[48] Bin Liu, Yuan Yuan, Sai Li, Yong Shuai, and He-Ping Tan. Simulation of light-field camera imaging based on ray splitting monte carlo method. *Optics Communications*, 355:15–26, 2015.

[49] Massimo Turola and Steve Gruppetta. 4d light field ophthalmoscope: a study of plenoptic imaging of the human retina. *Frontiers in Optics, JW3A*, 26, 2013.

[50] Richard Marshall, Chris Meah, Iain Styles, Ela Claridge, and Kai Bongs. Investigating the potential for plenoptic imaging of the retina. *5 th Annual*, page 17.

[51] Shenchang Eric Chen and Lance Williams. View interpolation for image synthesis. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 279–288. ACM, 1993.

[52] Chester C Slama, Charles Theurer, Soren W Henriksen, et al. *Manual of photogrammetry*. Number Ed. 4. American Society of photogrammetry, 1980.

[53] Ivo Ihrke, Timo Stich, Heiko Gottschlich, Marcus Magnor, and Hans-Peter Seidel. Fast incident light field acquisition and rendering. 2008.

[54] Yuichi Taguchi, Amit Agrawal, Srikumar Ramalingam, and Ashok Veeraraghavan. Axial light field for curved mirrors: Reflect your perspective, widen your view. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 499–506. IEEE, 2010.

[55] Chia-Kai Liang, Tai-Hsu Lin, Bing-Yi Wong, Chi Liu, and Homer H Chen. Programmable aperture photography: multiplexed light field acquisition. *ACM Transactions on Graphics (TOG)*, 27(3):55, 2008.

[56] Yoav Y Schechner, Shree K Nayar, and Peter N Belhumeur. Multiplexing for optimal lighting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(8):1339–1354, 2007.

[57] Todor Georgiev, Ke Colin Zheng, Brian Curless, David Salesin, Shree Nayar, and Chintan Intwala. Spatio-angular resolution tradeoffs in integral photography. *Rendering Techniques*, 2006:263–272, 2006.

[58] Kensuke Ueda, Takafumi Koike, Keita Takahashi, and Takeshi Naemura. Adaptive integral photography imaging with variable-focus lens array. In *Electronic Imaging 2008*, pages 68031A–68031A. International Society for Optics and Photonics, 2008.

[59] Kensuke Ueda, Dongha Lee, Takafumi Koike, Keita Takahashi, and Takeshi Naemura. Multi-focal compound eye: liquid lens array for computational photography. In *ACM SIGGRAPH 2008 new tech demos*, page 28. ACM, 2008.

[60] Ashok Veeraraghavan, Ramesh Raskar, Amit Agrawal, Ankit Mohan, and Jack Tumblin. Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Transactions on Graphics*, 26(3):69, 2007.

[61] Ashok Veeraraghavan, Amit Agrawal, Ramesh Raskar, Ankit Mohan, and Jack Tumblin. Non-refractive modulators for encoding and capturing scene appearance and depth. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[62] Bennett Wilburn, Neel Joshi, Vaibhav Vaish, Eino-Ville Talvala, Emilio Antunez, Adam Barth, Andrew Adams, Mark Horowitz, and Marc Levoy. High performance imaging using large camera arrays. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 765–776. ACM, 2005.

[63] Jason C Yang, Matthew Everett, Chris Buehler, and Leonard McMillan. A real-time distributed light field camera. *Rendering Techniques*, 2002:77–86, 2002.

[64] Yoshikuni Nomura, Li Zhang, and Shree K Nayar. Scene collages and flexible camera arrays. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*, pages 127–138. Eurographics Association, 2007.

[65] Warren J Smith. *Modern optical engineering*. Tata McGraw-Hill Education, 1966.

[66] Max Born and Emil Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[67] Marc Levoy. Stanford light field microscopy. `http://graphics.stanford.edu/projects/lfmicroscope/`, 2010.

[68] Marc Levoy. Stanford light field microscopy. `http://graphics.stanford.edu/projects/lfmicroscope/2006.html`, 2009.

[69] Todor Georgiev. Qualcomm. `www.tgeorgiev.net`, June 2013.

[70] Tom E Bishop and Paolo Favaro. Full-resolution depth map estimation from an aliased plenoptic light field. In *Asian Conference on Computer Vision*, pages 186–200. Springer, 2010.

[71] Michael D Abràmoff, Mona K Garvin, and Milan Sonka. Retinal imaging and image analysis. *Biomedical Engineering, IEEE Reviews in*, 3:169–208, 2010.

[72] Helga Kolb. How the retina works. *American Scientist*, 91(1):28–35, 2003.

[73] David C Klonoff and Daniel M Schwartz. An economic analysis of interventions for diabetes. *Diabetes care*, 23(3):390–404, 2000.

[74] ETDRS Research Group et al. Early photocoagulation for diabetic retinopathy. *Ophthalmology*, 98:766–785, 1991.

[75] Diabetic Retinopathy Study Research Group et al. Photocoagulation treatment of proliferative diabetic retinopathy. clinical application of diabetic retinopathy study (drs) findings. Technical report, DRS report, 1981.

[76] James M Tielsch, Joanne Katz, Kuldev Singh, Harry A Quigley, John D Gottsch, Jonathan Javitt, and Alfred Sommer. A population-based evaluation of glaucoma screening: the baltimore eye survey. *American journal of epidemiology*, 134(10):1102–1110, 1991.

[77] David S Friedman, BJ Ocolmain, Beatriz Munoz, SC Tomany, Cathy McCarty, PT De Jong, B Nemesure, Paul Mitchell, John Kempen, et al. Prevalence of age-related macular degeneration in the united states. *Arch ophthalmol*, 122(4):564–572, 2004.

[78] Macular Photocoagulation Study Group et al. Argon laser photocoagulation for senile macular degeneration. results of a randomized clinical trial. *Arch Ophthalmol*, 100(6):912–918, 1982.

[79] Mircea Mujat, R Daniel Ferguson, Nicusor Iftimia, and Daniel X Hammer. Compact adaptive optics line scanning ophthalmoscope. *Optics express*, 17(12):10242–10258, 2009.

[80] David Huang, Eric A Swanson, Charles P Lin, Joel S Schuman, William G Stinson, Warren Chang, Michael R Hee, Thomas Flotte, Kenton Gregory, Carmen A Puliafito, et al. Optical coherence tomography. *Science*, 254(5035):1178–1181, 1991.

[81] Eric A Swanson, JA Izatt, CP Lin, JG Fujimoto, JS Schuman, MR Hee, D Huang, and CA Puliafito. In vivo retinal imaging by optical coherence tomography. *Optics letters*, 18(21):1864–1866, 1993.

[82] IB Styles, A Calcagni, Ela Claridge, Felipe Orihuela-Espina, and JM Gibson. Quantitative analysis of multi-spectral fundus images. *Medical image analysis*, 10(4):578–597, 2006.

[83] Lihong Wang, Steven L Jacques, and Liqiong Zheng. Mcmlmonte carlo modeling of light transport in multi-layered tissues. *Computer methods and programs in biomedicine*, 47(2):131–146, 1995.

[84] Tos TJM Berendschot, Peter Jaap DeLint, and Dirk van Norren. Fundus reflectancehistorical and present ideas. *Progress in retinal and eye research*, 22(2):171–200, 2003.

[85] D Van Norren and LF Tiemeijer. Spectral reflectance of the human eye. *Vision research*, 26(2):313–320, 1986.

[86] François C Delori and Kent P Pflibsen. Spectral reflectance of the human ocular fundus. *Applied optics*, 28(6):1061–1077, 1989.

[87] Von P Kubelka. Ein beintrag zur optik der farbanstriche. *Zeits. fur techn. Physik*, 12:593–601, 1931.

[88] M Hammer, A Roggan, D Schweitzer, and G Muller. Optical properties of ocular fundus tissues-an in vitro study using the double-integrating-sphere technique and inverse monte carlo simulation. *Physics in medicine and biology*, 40(6):963, 1995.

[89] Martin Hammer and Dietrich Schweitzer. Quantitative reflection spectroscopy at the human ocular fundus. *Physics in medicine and biology*, 47(2):179, 2002.

[90] Scott A Prahl, Marleen Keijzer, Steven L Jacques, and Ashley J Welch. A monte carlo model of light propagation in tissue. *Dosimetry of laser radiation in medicine and biology*, 5:102–111, 1989.

[91] Stephen J Preece and Ela Claridge. Monte carlo modelling of the spectral reflectance of the human eye. *Physics in Medicine and Biology*, 47(16):2863, 2002.

[92] Bernard L Horecker. The absorption spectra of hemoglobin and its derivatives in the visible and near infra-red regions. *Journal of biological chemistry*, 148(1):173–183, 1943.

[93] R Rox Anderson and John A Parrish. The optics of human skin. *Journal of investigative dermatology*, 77(1):13–19, 1981.

[94] Malvin H Kalos and Paula A Whitlock. *Monte carlo methods*. John Wiley & Sons, 2008.

[95] Edmond D Cashwell and Cornelius Joseph Everett. A practical manual on the monte carlo method for random walk problems. 1959.

[96] Ivan Lux and László Koblinger. *Monte Carlo particle transport methods: neutron and photon calculations*, volume 102. CRC press Boca Raton, Florida, 1991.

[97] Louis G Henyey and Jesse Leonard Greenstein. Diffuse radiation in the galaxy. *The Astrophysical Journal*, 93:70–83, 1941.

[98] M. Wolffe. How safe is the light during ophthalmic diagnosis and surgery. *Eye*, 30(2):186–188, 2015.

[99] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[100] N Zeller, F Quint, and U Stilla. Calibration and accuracy analysis of a focused plenoptic camera. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(3):205, 2014.

[101] Niclas Zeller, Franz Quint, and Uwe Stilla. Depth estimation and camera calibration of a focused plenoptic camera for visual odometry. *ISPRS Journal of Photogrammetry and Remote Sensing*, 118:83–100, 2016.

[102] Jason Geng. Structured-light 3d surface imaging: a tutorial. *Advances in Optics and Photonics*, 3(2):128–160, 2011.

[103] Robert J Valkenburg and Alan M McIvor. Accurate 3d measurement using a structured light system. *Image and Vision Computing*, 16(2):99–110, 1998.

[104] Idaku Ishii, Kenkichi Yamamoto, Kensuke Doi, and Tokuo Tsuji. High-speed 3d image acquisition using coded structured light projection. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pages 925–930. IEEE, 2007.

[105] Hans Sagan. *Space-filling curves*. Springer Science & Business Media, 2012.

[106] Peisen S Huang and Song Zhang. Fast three-step phase-shifting algorithm. *Applied optics*, 45(21):5086–5091, 2006.

[107] Sai Siva Gorthi and Pramod Rastogi. Fringe projection techniques: whither we are? *Optics and lasers in engineering*, 48(IMAC-REVIEW-2009-001):133–140, 2010.

[108] Z Jason Geng. Rainbow three-dimensional camera: new concept of high-speed three-dimensional vision systems. *Optical Engineering*, 35(2):376–383, 1996.

[109] Zheng J Geng. Method and system for three-dimensional imaging using light pattern having multiple sub-patterns, March 2 2004. US Patent 6,700,669.

[110] Kim L Boyer and Avinash C Kak. Color-encoded structured light for rapid active ranging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1):14–28, 1987.

[111] Sergio Fernandez, Joaquim Salvi, and Tomislav Pribanic. Absolute phase mapping for one-shot dense pattern projection. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 64–71. IEEE, 2010.

[112] Minoru Maruyama and Shigeru Abe. Range sensing by projecting multiple slits with random cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(6):647–651, 1993.

[113] Nelson G Durdle, Jaishankar Thayyoor, and VJ Raso. An improved structured light technique for surface reconstruction of the human trunk. In *Electrical and Computer Engineering, 1998. IEEE Canadian Conference on*, volume 2, pages 874–877. IEEE, 1998.

[114] F Jessie MacWilliams and Neil JA Sloane. Pseudo-random sequences and arrays. *Proceedings of the IEEE*, 64(12):1715–1729, 1976.

[115] Heinz Hügli and Gilbert Maitre. Generation and use of color pseudo random sequences for coding structured light in active ranging. *Proceedings of Industrial Inspection*, 1010:75–82, 1989.

[116] TP Monks and John N Carter. Improved stripe matching for colour encoded structured light. In *International Conference on Computer Analysis of Images and Patterns*, pages 476–485. Springer, 1993.

[117] Harold Fredricksen. A survey of full length nonlinear shift register cycle algorithms. *SIAM review*, 24(2):195–221, 1982.

[118] Jacqueline Le Moigne and Allen M Waxman. *Multi-resolution grid patterns for building range maps*. Society of Manufacturing Engineers, 1985.

[119] Hiroyoshi Morita, Kaanyasn Yajima, and Shojiro Sakata. Reconstruction of surfaces of 3-d objects by m-array pattern projection method. In *Computer Vision., Second International Conference on*, pages 468–473. IEEE, 1988.

[120] Pierre Payeur and Danick Desjardins. Structured light stereoscopic imaging with dynamic pseudo-random patterns. *Image Analysis and Recognition*, pages 687–696, 2009.

[121] EM Petriu, Z Sakr, HJW Spoelder, and A Moica. Object recognition using pseudo-random color encoded structured light. In *Instrumentation and Measurement Technology Conference, 2000. IMTC 2000. Proceedings of the 17th IEEE*, volume 3, pages 1237–1241. IEEE, 2000.

[122] Jordi Pages, Joaquim Salvi, and Carles Matabosch. Robust segmentation and decoding of a grid pattern for structured light. In *IbPRIA*, volume 2652, pages 689–696. Springer, 2003.

[123] Ali Osman Ulusoy, Fatih Calakli, and Gabriel Taubin. Robust one-shot 3d scanning using loopy belief propagation. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 15–22. IEEE, 2010.

[124] Danick Desjardins and Pierre Payeur. Dense stereo range sensing with marching pseudo-random patterns. In *Computer and Robot Vision, 2007. CRV'07. Fourth Canadian Conference on*, pages 216–226. IEEE, 2007.

[125] Matthew J Baker, Joe F Chicharo, Jiangtao Xi, and Enbang Li. Accuracy limitations introduced by digital projection sources in profilometric optical metrology systems. In *Optoelectronic and Microelectronic Materials and Devices, 2004 Conference on*, pages 261–264. IEEE, 2004.

[126] Zewei Cai, Xiaoli Liu, Xiang Peng, Yongkai Yin, Ameng Li, Jiachen Wu, and Bruce Z Gao. Structured light field 3d imaging. *Optics express*, 24(18):20324–20334, 2016.

[127] Bipm. `https://web.archive.org/web/20110927012931/http://www.bipm.org/en/convention/wmd/2004/`, 2004.

[128] David J Whitehouse. Surface metrology. *Measurement Science and Technology*, 8(9):955, 1997.

[129] Ken J Stout and Liam Blunt. *Three dimensional surface topography*. Elsevier, 2000.

[130] Richard Edmund Reason. *The measurement of surface texture*. Cleaver-Hume Press, 1960.

[131] Gerd Binnig and Heinrich Rohrer. Scanning tunneling microscopy. *Surface science*, 126(1-3):236–244, 1983.

[132] Gerd Binnig, Calvin F Quate, and Ch Gerber. Atomic force microscope. *Physical review letters*, 56(9):930, 1986.

[133] Robert H Webb. Confocal optical microscopy. *Reports on Progress in Physics*, 59(3):427, 1996.

[134] RJ Hocken, N Chakraborty, and C Brown. Optical metrology of surfaces. *CIRP Annals-Manufacturing Technology*, 54(2):169–183, 2005.

[135] J Simon. New noncontact devices for measuring small microdisplacements. *Applied optics*, 9(10):2337–2340, 1970.

[136] S-N Uchida, H Sato, and M O-hori. Two dimensional measurement of surface roughness by the light sectioning method. *SPIE MILESTONE SERIES MS*, 129:3–8, 1996.

[137] Morley M Blouke and James R Janesick. Charge-coupled device image sensors. *Arms Control Verification: The Technologies that Make it Possible, Washington, DC: Pergamon/Brasseys*, pages 109–10, 1986.

[138] Glenn E Healey and Raghava Kondepudy. Radiometric ccd camera calibration and noise estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(3):267–276, 1994.

[139] Leonard GC Hamey. Radiometric camera calibration. *Wiley Encyclopedia of Computer Science and Engineering*, 2008.

[140] David Forsyth and Jean Ponce. *Computer vision: a modern approach*. Upper Saddle River, NJ; London: Prentice Hall, 2011.

[141] Michael D Grossberg and Shree K Nayar. Modeling the space of camera response functions. *IEEE transactions on pattern analysis and machine intelligence*, 26(10):1272–1282, 2004.

[142] James A Guggenheim, Hamid Dehghani, Hector Basevi, Iain B Styles, and Jon Frampton. Development of a multi-view, multi-spectral bioluminescence tomography small animal imaging system. In *European Conference on Biomedical Optics*, page 80881K. Optical Society of America, 2011.

[143] Clyde W Oyster. *The human eye: structure and function*. Sinauer Associates, 1999.

[144] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.

[145] Robert M Haralick, Stanley R Sternberg, and Xinhua Zhuang. Image analysis using mathematical morphology. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (4):532–550, 1987.

[146] Alan Jennings. Sphere fit (least squared). `https://uk.mathworks.com/matlabcentral/fileexchange/34129-sphere-fit--least-squared-`, 2013.

[147] Renato Ambrósio Jr and Michael W Belin. Imaging of the cornea: topography vs tomography. *J refract Surg*, 26(11):847–9, 2010.

[148] Ph Nussbaum, Reinhard Voelkel, Hans Peter Herzig, Martin Eisner, and Stefan Haselbeck. Design, fabrication and testing of microlens arrays for sensors and microsystems. *Pure and Applied Optics: Journal of the European Optical Society Part A*, 6(6):617, 1997.

[149] Zoran D Popovic, Robert A Sprague, and GA Neville Connell. Technique for monolithic fabrication of microlens arrays. *Applied optics*, 27(7):1281–1284, 1988.

[150] Stefan Sinzinger and Jürgen Jahns. *Front Matter*. Wiley Online Library, 2003.

[151] D Daly, RF Stevens, MC Hutley, and N Davies. The manufacture of microlenses by melting photoresist. *Measurement Science and Technology*, 1(8):759, 1990.

[152] T Krupenkin, S Yang, and P Mach. Tunable liquid microlens. *Applied Physics Letters*, 82(3):316–318, 2003.

[153] Norio Sugiura and Shinzo Morita. Variable-focus liquid-filled optical lens. *Applied Optics*, 32(22):4181–4186, 1993.

[154] De-Ying Zhang, Victor Lien, Yevgeny Berdichevsky, Jaehyuck Choi, and Yu-Hwa Lo. Fluidic adaptive lens with high focal length tunability. *Applied Physics Letters*, 82(19):3171–3172, 2003.

[155] Yu Lei, Qing Tong, Xinyu Zhang, Hongshi Sang, An Ji, and Changsheng Xie. An electrically tunable plenoptic camera using a liquid crystal microlens array. *Review of Scientific Instruments*, 86(5):053101, 2015.

[156] Florian Schneider, Jan Draheim, Robert Kamberger, and Ulrike Wallrabe. Process and material properties of polydimethylsiloxane (pdms) for optical mems. *Sensors and Actuators A: Physical*, 151(2):95–99, 2009.

[157] Ferry PW Melchels, Jan Feijen, and Dirk W Grijpma. A review on stereolithography and its applications in biomedical engineering. *Biomaterials*, 31(24):6121–6130, 2010.

[158] Simon J Leigh, J Bowen, CP Purssell, James A Covington, DR Billson, and David A Hutchins. Rapid manufacture of monolithic micro-actuated forceps inspired by echinoderm pedicellariae. *Bioinspiration & biomimetics*, 7(4):044001, 2012.

[159] Dan Daly. *Microlens arrays*. CRC Press, 2000.

[160] Arthur H Firester. Properties and fabrication of micro fresnel zone plates. *Applied optics*, 12(7):1698–1702, 1973.