Plenoptic Camera

Summary Report

Korey Nicholas Rouse[1]

University of New South Wales at the Australian Defence Force Academy

The addition of a microlens array in front of an image sensor transforms a normal camera into a plenoptic camera capable of capturing the 4D light field information contained within a scene. To make effective use of the image sensor the working focal numbers of the main lens is designed to match the microlens. Reconstructing an image from a plenoptic camera can be achieved by sampling a number of pixels from the centre of each microimage and tiling them together. The sample size can then equate to a distance relative to the main lens through the use of geometric optics and the thin lens equation. The relationship between the microlens focal length and the distance from the image sensor determines the application of the plenoptic camera. Matching the microlens focal length to the microlens distance from the image sensor results in a standard plenoptic camera that favours image resolution and digital refocussing. A focussed plenoptic camera has a slightly larger distance to image sensor than focal length and has the advantage of better depth estimation abilities. Combining experimental results with research demonstrates the concepts behind the plenoptic camera.

Contents

I. Introduction 2
II. Background 2
III. Design Considerations 2
A. Light Fields 2
B. Plenoptic Camera Overview 3
C. Focal Number Matching 3
D. Depth of Field 4
IV. Operational Abilities 6
A. Intrinsic Characteristics 6
B. Image Synthesis 7
C. Depth Estimation 8
V. Future Applications 10
VI. Conclusion 10
VII. Reference 10

[1] SBLT, RAN. School of Engineering & Information Technology.

Nomenclature

a	=	Image formation in front of microlens	a_L	=	Image formation in front of main lens
b	=	Image formation behind microlens	b_L	=	Image formation behind main lens
f	=	Microlens focal length	f_L	=	Main lens focal length
B	=	Distance of microlens from image plane	B_L	=	Distance of main lens from image plane
D	=	Diameter/Aperture of microlens	D_L	=	Aperture of main lens
F_w	=	Working focal number of microlens	m	=	Magnification
$F_\#$	=	Focal number of main lens	ε	=	Effective resolution ratio
S	=	Square sample size	v	=	Virtual depth
u	=	Pitch size	λ	=	Wavelength of light
ERR	=	Effective resolution ratio	s_0	=	Minimum projection of pixel sensor
L	=	Radiance of a ray			

I. Introduction

Photography first arose with the works of Joseph Nicéphore Niépce and his creation of the first permanent photograph in 1826. Since then modern cameras have evolved to use digital sensors and computational processing to create photographs in ever higher demanding quality and quantity. The one thing that has remained the same is the information cameras produce: a two-dimensional representation [1]. With the capabilities of modern digital cameras and the computational resources at our disposal, new ways of capturing photographic information are being explored. Lippmann [2] worked with pinholes to describe the theory behind integral photography. This theory has been expanded with work in light fields. Light fields, as described by Levoy [6], represent the flow of light through an empty 3D space. Attempts to capture and understand the information contained within light fields has led to the creation of a new type of camera known as the 'plenoptic camera'. The plenoptic camera, derived from the words "complete" and "view" as described by Adelson [3], is created by inserting an array of microlens in front of the image sensor. The image from a single shot of a plenoptic camera contains a continuum of viewpoints and once recorded, one can select an image from a particular sub-region. Taking this region and comparing it with its neighboring regions allows for virtual displacements to be measured [3]. This ability allows the plenoptic camera to capture light fields which will have a multitude of practical uses. In 2012, two plenoptic cameras became commercially available. Both camera's involve a complex means of computational imaging and post processing algorithms to achieve practical applications. The Illum by Lytro is the product of the work by CEO Ren Ng [7] who has been significantly impacting the work on light fields since 2005. The other is the R11 from Raytrix [8]. Christian Perwas is credited to the research behind the success of the Raytrix company. Similarly to Ng, Perwas's research has forwarded the advancement of plenoptic cameras and computational imaging. Their research will be heavily cited in this report and these cameras can be used as comparable industry standards.

II. Background

This project involved researching the concepts at work in the plenoptic camera. To achieve this, the project initially investigated the design concepts and physical properties of a plenoptic camera. This allowed insight into how the plenoptic camera operates and has the capability to conduct image reconstruction and depth estimation. Areas such as calibration and light distortion through the camera are not covered in this paper. Two commercial cameras that have been modified to become plenoptic cameras were available to aid in research. The first camera is an inexpensive Z14 Traveler, modified to contain a Tessera microlens array which produced an approximate 12 MP image containing 5x4 microimages. Initially information regarding the dimensions such as distances and focal lengths of lenses in this camera were unknown and were required to be confirmed. The other plenoptic camera available was the digital back of a 39MP Phase One P45+ with an APO-Q-P300-F2(633) fused silica microlens array of diameter 0.3µm and focal length of 2mm.

III. Design Considerations

A. Light Fields

The theory of light fields has its basis in geometric optics. Light travels through space via a fundamental carrier known as a ray. The amount of light travelling in one direction along all possible lines through a cylinder with a given cross-sectional area and angle is known as the radiance *L* of a ray, as depicted in figure 1 left. Radiance is measured in watts (W) per steridan (sr) per metre squared (m^2). The plenoptic function is the name given to the radiance travelling along all rays in a region of three dimensional space under unchanging lighting conditions [3]. This gives rise to a five dimensional representation as space contains three dimensions and there are two angles required to describe the radiance. The angles are shown in figure 1 middle. By making the assumption that the radiance from one point to another remains constant, which occurs in occlude free space, then the plenoptic

function contains redundant information. This redundancy was found to be equivalent to one dimension and once removed can parameterize the four dimensional light field [13]. The parameterization that is most commonly used is by Levoy where the ray's intersection of two planes defines the light field, given in figure 1 right. The light field can then be described as a collection of perspective views on the *st* plane from observation positions on the *uv* plane [14].

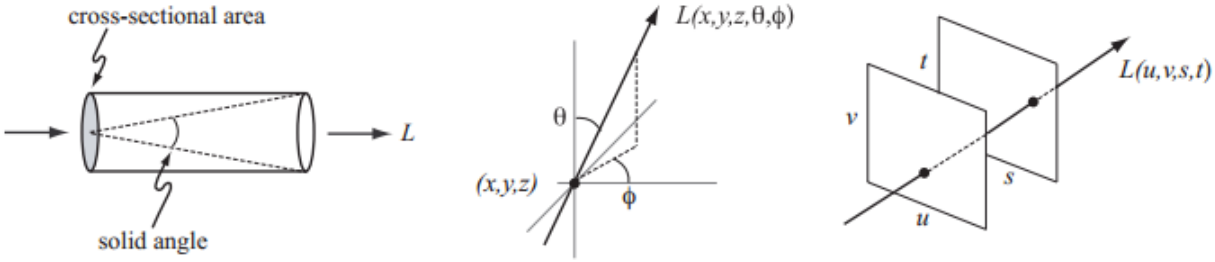

Figure 1: Radiance (left), 5D Plenoptic Function (middle) and 4D Light Field (right) [14]

B. Plenoptic Camera Overview

To exploit the concept of integral photography [2], a camera can add a microlens array between the main lens and the image sensor to create a plenoptic camera [1,3,4,9,10,11,12]. Figure 2 shows a ray tracing example of the plenoptic imaging process, where light from an object creates a virtual image that each microlens array views from a slightly different perspective. This perspective captures information about rays entering at different angles, which contains information regarding the four dimensional light field [14]. The unique light gathering properties of the plenoptic camera opens possibilities for post-processing capabilities such as digital refocussing and depth estimation. Finding information about the distances of lenses and focal lengths of each camera was essential in order to apply this imaging concept further.

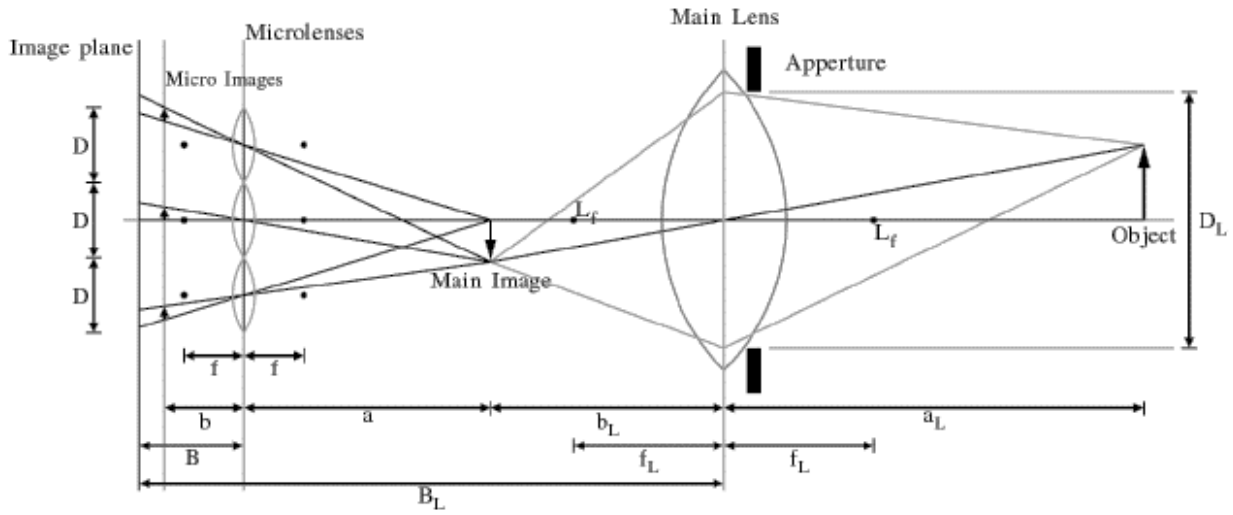

Figure 2: Plenoptic Camera Image Process[4]

C. Focal Number Matching

A consistent design aspect among almost all plenoptic camera designs [4,9,12,15,16] is the process of matching the working focal number of the main lens to the microlens as described in equation 1.

$$F_w = (1 + |m|)F_{\#} = \frac{B}{D} \approx \frac{B_L - B}{D_L} \tag{1}$$

As the microlens position is fixed this means that the main lens no longer has the freedom to change its focal length. The purpose of matching working focal numbers is to make effective use of the image sensor. This is seen in the images where no overlap exists between microimages as shown in figure 3 right. The working focal numbers

of the Z14 main lens and microlens were found to be 8.8 and 10.83 respectively. Whilst the Z14 can be matched so the microimages do not overlap, the working focal numbers are not equal which has consequences in other areas to be discussed.

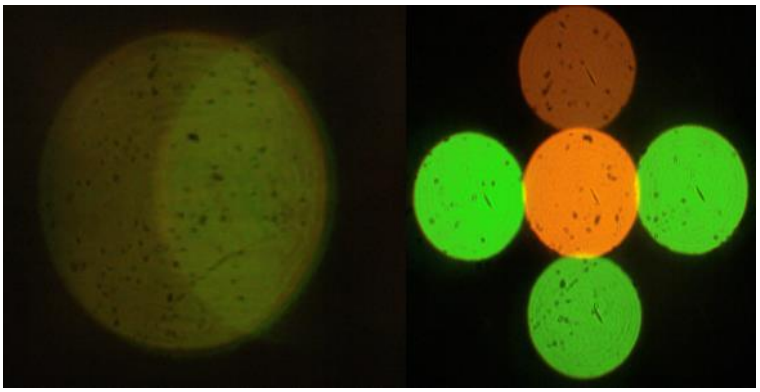

Figure 3: F-number: mismatch (left) and matched (right)

D. Depth of Field

Depth of field (DOF) can be thought of as the range in which an object can remain in focus in an imaging system. A way of quantising the depth of field is in the form of the Effective Resolution Ratio (ERR) [4] where the DOF is the range between a specified ERR value. The ERR is a ratio between the pixel size of the sensor and the maximum size that light can be projected through the imaging system, where s_0 represents the minimum projection point that can be resolved by a pixel sensor. s_0 can be quantified as the maximum between the pixel size and the Point Spread Function (PSF) of the system (approximated as $1.22\lambda N$). This is where the working focal number mismatch will cause a mismatch in distortion between the main lens and microlens. Seeing the microlens as an array of pinholes with infinite depth of field yields an ideal plenoptic camera to have an ERR of the inverse of the virtual depth (v=a/B). Multiplication of the microlens ERR with an ideal pinhole plenoptic camera ERR provides an approximation for the ERR for the plenoptic camera. Note, that this value then needs to be projected through a main lens for an object side DOF. An image side ERR for the plenoptic camera[4] is given in equation 2:

$$ERR = \frac{1}{|v|}\frac{p}{max\left[\left|D\left(\frac{\mathrm{B}}{\mathrm{f}}-\frac{1}{v}-1\right)\right|, s_0\right]} \tag{2}$$

Using plenoptic dimensions from Perwas [4] as a comparison tool, the ERR and DOF can be seen in figure 4, where the Z14 is seen in blue and the reference in red. This figure shows where the virtual projection of an image through the microlens without main lens effects is in focus when projected behind the image sensor. The trend between the two cameras is quite similar, however, the clear difference lies in the position of the virtual image. This difference in virtual image distance is caused by the difference in microlens focal length and positioning. The reference camera uses larger pixel sensor sizes which increases the ERR in comparison to the Z14. This highlights the importance of microlens position with respect to the microlens focal length. Better resolution will be given if a virtual image is being produced closer to the image sensor. The Z14 could be improved by decreasing the distance of the microlens to the image plane (B).

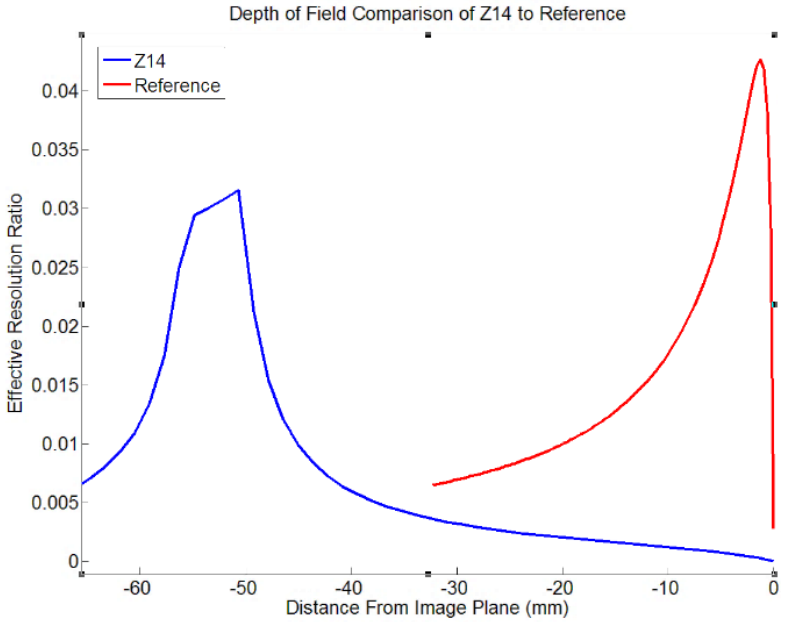

Figure 4: Image Side Depth of Field Comparison between Z14 and Reference[4]

Thin Lens Equation:

$$\frac{1}{f} = \frac{1}{a} + \frac{1}{b} \tag{3}$$

The thin lens equation can be implemented to find the object side DOF of a camera, which includes the substitution of equation 3 into equation 2. A DOF analysis of the Phase One camera was under taken to further investigate the effects of the microlens position from the image plane. This results in the position of the microlens significantly effecting the DOF of the plenoptic camera as shown in figure 5. This shows a variation of the distance of the microlens with values less than, equal to and greater than the focal length of the microlens. When B=f, the microlens is focussed to infinity and yields a constant ERR. Under these conditions, a plenoptic camera is referred to as a standard plenoptic camera [11,15]. From figure 5, a standard plenoptic camera will be better suited to synthetic aperture photography applications due to the constant ERR and a large DOF. Focusing the microlens at optical infinity maximizes resolution as the microlens will focus to the sharpest point at the image plane. The industry standard of standard plenoptic cameras is the Lytro Illium [17]. The larger variation in ERR occurs when B>f, which is known as a focused plenoptic or plenoptic 2.0 [4,10]. This design of plenoptic camera allows for greater depth estimation due to the variation in ERR at the cost of reduced DOF, shown in figure 5. The Raytrix R11 is the industry standard for plenoptic 2.0 with depth estimation [18]. The Raytrix R11 also implements a microlens array which includes multiple focal lengths, which results in multiple ERR peaks. Designing the focal lengths correctly can then effective increase the DOF as more virtual depth planes are created [4]. The Z14 falls into the focused plenoptic design and the Phase One is a standard plenoptic design.

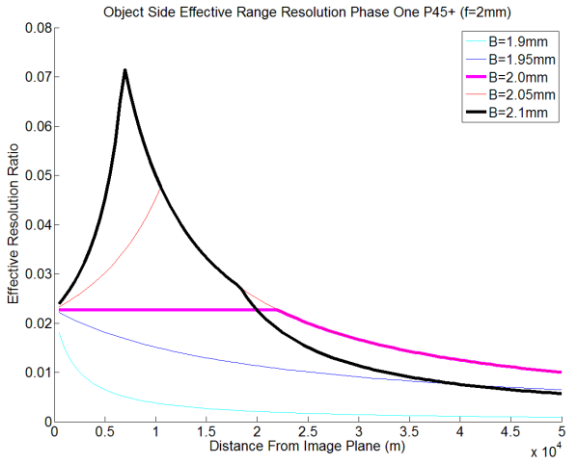

Figure 5: Phase One Depth of Field

IV. Operational Abilities

A. Intrinsic Characteristics

A plenoptic camera captures a light field representation of a scene/object. This manifests in a collection of rays entering the camera at certain angles and positions. Following the convolution theorem for image formation, an image is the convolution of light from the object and the PSF of an optical system. If we require information about the object based on information gained from the image, the inverse PSF of the camera is needed. This can be thought of as the intrinsic characteristics of the Z14 under certain conditions and can be achieved by the following process:

1. Focus an LED so that light is contained within a single microlens and take an image.
    a. Repeat for each microlens.
2. Segment a portion of each microimage.
    a. Repeat for all images.
3. Sum intensities of each segment to create a vector.
    a. Repeat for each image, adding rows to create the intrinsic characteristic matrix.

A test image, shown in figure 6, contains two LEDs where their respective positions are to be found. The overlay seen in the top left corner represents two different segmentation schemes utilised. The centre method shown in red samples a portion from the centre of each microimage where the quadrant method uses all pixels and breaks up the microimage into four areas. To find the LED positions, the image is segmented and multiplied by the pseudoinverse of the intrinsic matrix. This results in a relative microlens match to the information contained in the intrinsic matrix. Testing an image that was used to create the intrinsic matrix results in a 100 percent match at the correct location with zeros at all locations. Any image of an LED that was not used to create the matrix will contain some difference in angle which will displace/bias the intensity in a direction. This will result in a number less than 100 in the LED position with non-zero numbers given for neighbouring microlenses. The neighbouring values gives insight into the angle at which the rays entered the camera in comparison to the reference LED. A limitation of this method is highlighted when using this test image as two LED's have been used. Figure 7 shows the result of the pseudoinverse multiplication using the two segmentation methods. The two LED positions were identified correctly at the peaks (2,1) & (4,1), however, the relative match is approximately 50% for each LED. As a result extracting information regarding the angle at which rays entered the system is harder to determine due to the multiple LED sources. The quadrant method provides a better contrast between microlenses at the cost of a larger matrix size and processing time. The quadrant method can also then map to quadrant positions within the microimage producing enhanced microlens position information. Ideally to extract precise angle information, comparison between intensities would be required on a pixel level between corresponding microlens positions. Using the Z14 as an example, pixel comparison would result in a characteristic matrix containing 280 million intensities. The computational expense will quickly get out of hand with higher quality cameras and over multiple depth planes. This process proved to be an important step in understanding the concepts of how the plenoptic camera views light whilst not being a practical method of extracting detailed information.

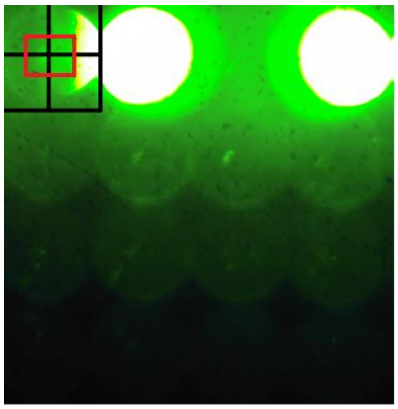

Figure 6: Microlens Position Test Image

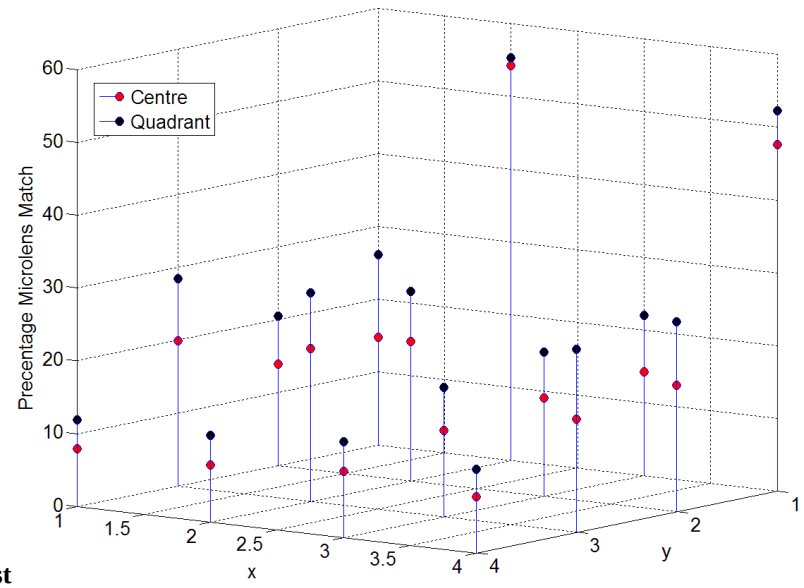

Figure 7: Microlens Position Interpretation of Test Image

B. Image Synthesis

The raw image produced from a plenoptic camera does not display useful information without post-processing. However, with knowledge of how the plenoptic camera views a scene, post-processing methods can take a plenoptic image and conduct digital refocussing to reconstruct an image that is in focus. Methods of image reconstruction vary depending on the design of the plenoptic camera and is best explained in figure 10. The concept behind image reconstruction lies in the concept where a microlens projecting a portion of the scene will overlap with its neighbouring microlenses. Therefore one way an image could be reconstructed is by taking a square sample (S) from the centre of each microlens and stitching it together. Reconstructions with different sample sizes is essentially digital refocussing. Figure 8, shows the raw image from the Z14 of the synthesised images shown in figure 9. Loss of information occurs when sampling is too low and artefacts occur at higher sampling sizes. These results agree with the results of Georgiev and Lumsdaine [16]. Each image shown in figure 9 is constructed of 16 blocks of their respective sampling pixel sizes. The resulting image has reduced in resolution in comparison to the quality that the camera could produce without the microlens array. This finding confirms the reduction in resolution limitation of the plenoptic camera [4,10,11,16].

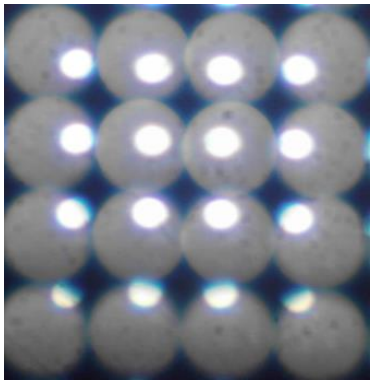

Figure 8: Raw Plenoptic Image from Z14

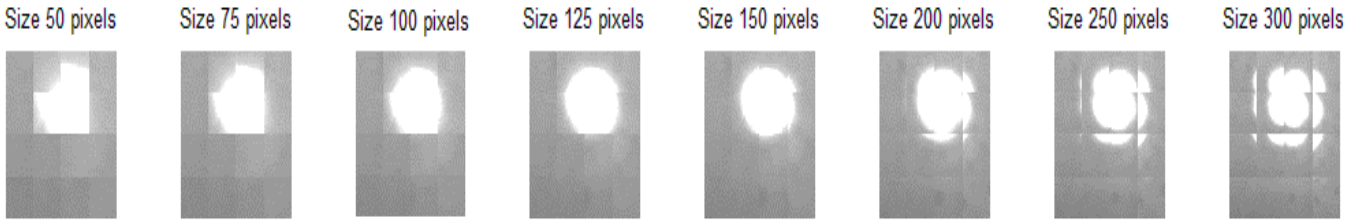

Figure 9: Synthesised Image with Different Sample Sizes

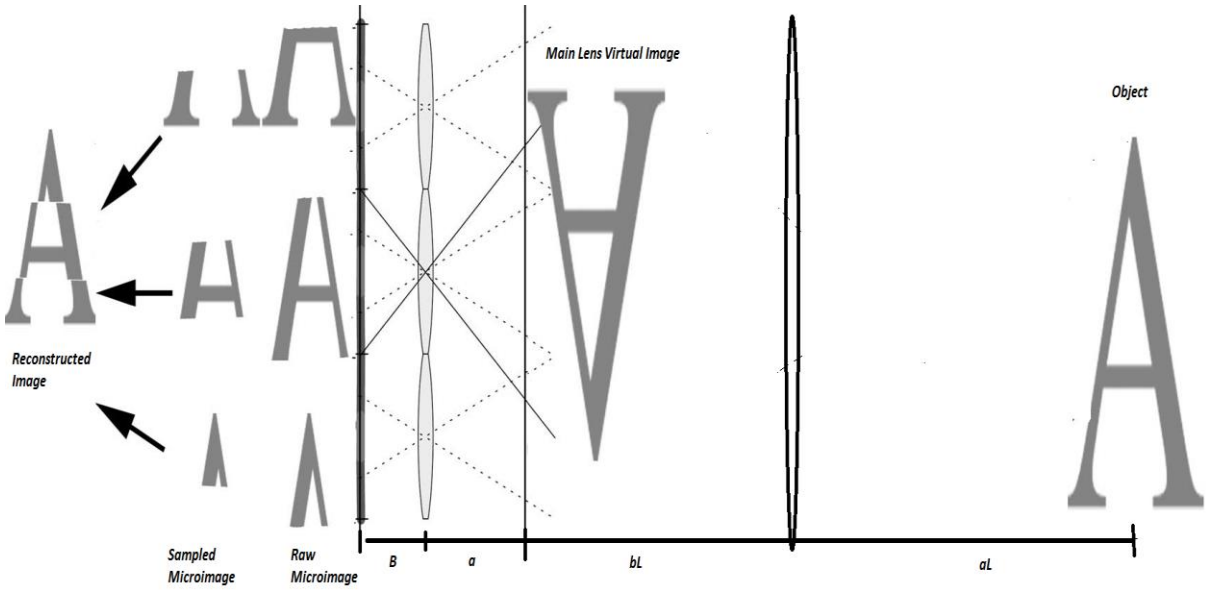

Figure 10: Image Synthesis and Depth Estimation Diagram (Expanded from [16])

C. Depth Estimation

One of the practical applications of the plenoptic camera is the ability to estimate the distance to the object from the camera. One method of depth estimation extends from the image synthesis process previously mention. An object will have a sample size (S) that will best represent an object in focus. In figure 9, this sample size is approximately 125 pixels as the circle is best recreated. Figure 10 shows the imaging process and how it can be used in depth estimation. Taking this sample size and using a magnification equation with the known size of the microlens diameter, the position of the virtual image (b) of the microlens can be estimated [16] using equation 4. This position can be determined through the use of similar triangles and geometric optics as shown in figure 11. A similar equation is derived from the use of geometric optics from these references [3,16].

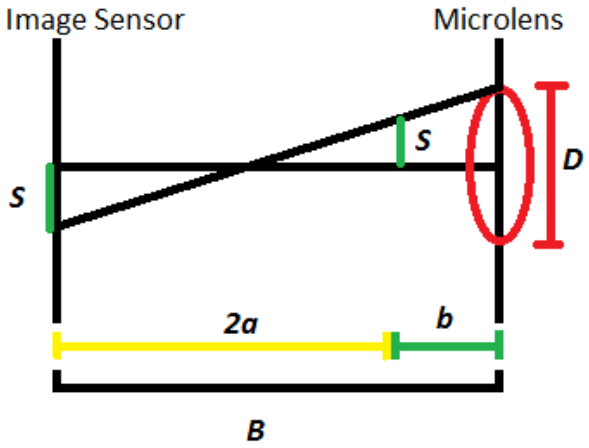

Figure 11: Method of Similar Triangles and Geometric Optics to find Virtual Position of Microlens Image

$$b = \frac{D \times (B - 1)}{S} \tag{4}$$

Through the application of the thin lens equation (Eq 3), the virtual image of the main lens can be found in front of the main lens (a). Another iteration of the thin lens equation (Eq 3) finds the object distance from the main lens in real space as shown in equation 5.

$$a = \frac{fb}{b - f}$$
$$a_L = \frac{f_L(B_L - B - a)}{B_L - B - a - f_L} \tag{5}$$

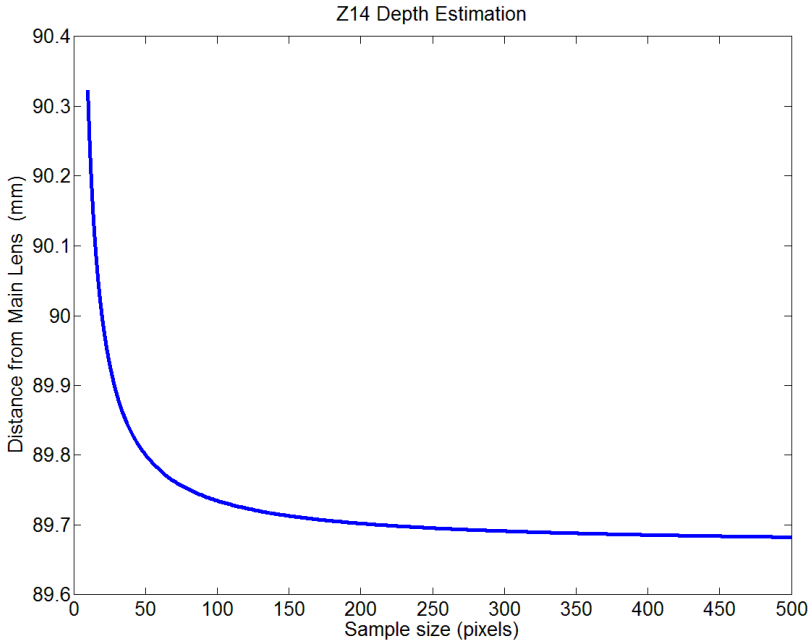

Figure 12: Z14 Depth Estimation

Figure 12 shows depth estimation for the Z14 plenoptic camera. This plot clearly shows that the Z14 is not a practical tool for depth estimation for most real world applications using this depth estimation method as the depth range is less than 1mm. This is caused by the large distance that the microlens is away from the image sensor and similarly their large focal length. However, it does show a curve in the resolution of the estimated depth. Larger sample sizes increases the resolution of the depth estimate to the micrometre scale. One could then suggest that plenoptic cameras with microlens arrays of similar size to the Z14 could be better suited for applications in microscopic levels. Adelson and Wang [3] provide a similar conclusion with more detail into the use of microlens arrays in microscopes. When imaging a scene of varying depths, different microlenses will view different areas and therefore depths. As a result a single sample size will only capture one area/object in focus. The depth estimation and imaging reconstruction process can be applied on individual microlens, where correlations in the x and y directions produce a specific sample size in which that microlens should use in the reconstruction algorithm. Finding specific microlens sample sizes and reconstructing an image with those values can produce an image where all objects are in focus. Applying the depth estimation algorithm to the 'all in focus' image can render a full depth map [16]. The Z14 was unable to produce all in focus images and full depth maps following the referenced algorithms due to camera mismatches and a lack of understanding of the correlation process. Despite this the concept behind the all in focus image and depth mapping was found to be sound and could be implemented with the Phase One with further research into correlation algorithms and how this applies to the microlenses.

Other methods of depth estimation involved the use of Fourier transforms [11], virtual depth [4,5] and energy minimization [9,19]. All of these methods are far more technical in nature requiring in depth knowledge of light fields. The significant step that was not achieved to follow these methods was the ability to convert the raw plenoptic image into a 4D light field for processing. The advantages of methods utilizing the light field is the possibility to achieve a depth estimation for each pixel rather than for each microlens as the algorithm in this paper is able to achieve. However, from the pseudoinverse experiment it was found that calculating accurate angular information is computationally expensive. In comparison to the depth estimation algorithm mentioned in this paper the computational expense would come from the correlation algorithm, which involves pixel level comparison. One example of a limitation that the depth estimation algorithm which does not find the angle of rays entering the optical system is to determine where a distortion/ray direction change occurs. A possible test could involve the projection of a light source through a clear Perspex sheet and calculating the depth that the distortion occurred. The accuracy in the depth estimation process ultimately determines its suitability for real world applications.

V. Future Applications

Many of the applications of the plenoptic camera can be achieved using multiple cameras in a stereo-vision configuration. The advantage to using a plenoptic camera avoids the difficulty in calibrating multiple cameras and capturing an image at a single instance of time. To photographers, synthetic aperture photography will allow post-processing to resolve an object that is not in focus through an image reconstruction algorithm. The main hurdle that plenoptic cameras need to overcome in this area is the inability to meet the large image resolutions that conventional cameras are capable of [11,15]. Depth estimation is motivated by needs of 3D image reconstruction which could be used in facial reconstruction for security purposes [5]. Areas such as quality control in production lines could make use of 3D reconstructions for visual processing of products. Plenoptic cameras could also be used for fluid flow analysis in the 3D space [18]. Plenoptic technology has been extended into other optical applications including microscopes, where images can be rotated over a range of angles to alleviate the superposition of organisms [3]. The astronomical community could also make use of plenoptic concepts by incorporating microlens arrays into telescopes with aims to find areas of light distortion. Whilst it was not covered in this paper, calibration and distortion effects on the plenoptic camera are areas of future research.

VI. Conclusion

This paper has covered the proof of concept behind the plenoptic camera in areas of microlens design and capabilities. The standard and focussed plenoptic cameras are the result of a design decision based on the position of microlens in relation to its focal length. The standard plenoptic camera was found to have a large DOF with a constant ERR whist the focussed had a variable ERR. Both forms of plenoptic cameras have a main lens designed so that the working focal number matches the microlens. This allows effective use of the imaging sensor and avoids overlap of the microimages. A basic method of image reconstruction was explored allowing a raw image from a plenoptic camera to be rendered into an image that could be digitally refocussed. Refocussing an image to a correct focus produces a samples size that can be equated into a distance in which the microlens produces a virtual image. Use of the thin lens equation can then produce a depth estimation of the object in real space. The trade-off using these methods is not utilising the true potential of the plenoptic camera which is to capture the angle in which rays from and object enter. A pseudoinverse multiplication of an image with the intrinsic characteristics was found to be a possible method of determining such angles. The method presented is not practical for plenoptic applications due to the large computational expense for accurate information. However, experiment provided sound insight into the concepts mentioned by reference material. Whilst the detailed algorithms into their depth estimation methods were beyond the scope of this project.

VII. Acknowledgements

I would like to thank Dr. Andrew Lambert UNSW@ADFA for his support and aid in my research.

VIII. Reference

1. Wetzstein,G, et al., “On Plenoptic Multiplexing and Reconstruction” MIT Media Lab.
2. Lippmann, M. G., “Reversible Prints. Integral Photographs”, Academy of the Sciences. (1908)
3. Adelson, E. H and Wang, J., “Single Lens Stereo with a Plenoptic Camera”, IEEE transactions on pattern analysis and machine intelligence, vol. 14, no. 2, (1992)
4. Perwas, C and Wietzke, L., “Single Lens 3D-Camera with Extended Depth-of-Field”, Raytrix GmbH, Schauenburgerstr. Kiel, Germany. (2010)
5. Perwas, C and Wietzke,L., “The next generation of photography”, Raytrix GmbH, Schauenburgerstr. Kiel, Germany. (2010)
6. Levoy, M., “Light Fields and Computational Imaging” IEEE Computer Society. (2006)
7. Lytro, Inc., “Lytro Illum”. www.lytro.com (2014)
8. Raytrix.,(2013) “3D Light Field Camera Technology” http://www.raytrix.de/index.php/R11_en.html US-Pat.-No.: 2012/0050562 A1

9. Bishop, T and Favaro, P., “Plenoptic Depth Estimation from Multiple Aliased Views” Heriot-Watt University, United Kingdom. (2012)
10. Lumsdaine, A. and Georgiev, T., “The Focused Plenoptic Camera”. International Conference on Computational Photography. (2009)
11. Ng, R., “Fourier Slice Photography”. Stanford University. (2005)
12. Johannsen, O. Heinze, O. Goldluecke, B. Perwas, C., “On the Calibration of Focused Plenoptic Cameras”. (2012)
13. Gershun, A., “The Light Field”. Translated by P. Moon and G. Timoshenko in J. Mathematics and Physics XVIII(1939), 51-151. (1936)
14. Levoy, M. Zhang, Z. and McDowall, I., “Recording and Controlling the 4D Light Firled in a Microscope using Microlens Arrays”. Computer Science Department, Stanford University, Stanford, CA 94305, USA (2009)
15. Ng, R, et al., “Light Field Photography with a Hand-held Plenoptic Camera”, Stanford Tech Report. (2005)
16. Georgiev T, Lumsdaine, A., “Focused Plenoptic Camera and Rendering”, Journal of Electronic Imaging 19(2), (2010)
17. Lytro Inc., “Lytro Illium” https://www.lytro.com/ (2014)
18. Raytrix GmbH,. “Raytrix R11”. http://raytrix.de/index.php/R11_en.html (2014)
19. Bishop, T and Favaro, P., “Full-Resolution Depth Map Estimation from an Aliased Plenoptic Light Field”, Heriot-Watt University, United Kingdom.(2010)