Title: The Standard Plenoptic Camera: Applications of a Geometrical Light Field Model

Name: Christopher Hahne

The Standard Plenoptic Camera

— ≺ ∘ ≻ —

Applications of a Geometrical Light Field Model

by

Christopher Hahne

Department of Computer Science and Technology
University of Bedfordshire

A thesis submitted in accordance with the requirements
for the degree of

Doctor of Philosophy

in the Institute for Research in Applicable Computing.

January, 2016

Abstract

The plenoptic camera is an emerging technology in computer vision able to capture a light field image from a single exposure which allows a computational change of the perspective view just as the optical focus, known as refocusing. Until now there was no general method to pinpoint object planes that have been brought to focus or stereo baselines of perspective views posed by a plenoptic camera.

Previous research has presented simplified ray models to prove the concept of refocusing and to enhance image and depth map qualities, but lacked promising distance estimates and an efficient refocusing hardware implementation. In this thesis, a pair of light rays is treated as a system of linear functions whose solution yields ray intersections indicating distances to refocused object planes or positions of virtual cameras that project perspective views. A refocusing image synthesis is derived from the proposed ray model and further developed to an array of switch-controlled *semi-systolic* FIR convolution filters. Their real-time performance is verified through simulation and implementation by means of an FPGA using VHDL programming.

A series of experiments is carried out with different lenses and focus settings, where prediction results are compared with those of a real ray simulation tool and processed light field photographs for which a blur metric has been considered. Predictions accurately match measurements in light field photographs and signify deviations of less than 0.35 % in real ray simulation. A benchmark assessment of the proposed refocusing hardware implementation suggests a computation time speed-up of 99.91 % in comparison with a state-of-the-art technique.

It is expected that this research supports in the prototyping stage of plenoptic cameras and microscopes as it helps specifying depth sampling planes, thus localising objects and provides a power-efficient refocusing hardware design for full-video applications as in broadcasting or motion picture arts.

To my parents

Herzlichen Dank für Euren steten Zuspruch.

Acknowledgements

First and foremost, it has to be mentioned that I am deeply indebted to my doctoral advisor Prof. Amar Aggoun. It was a privilege to be supervised by Amar as he has been available for consultations at any time. His belief in my ideas and his exceptional support laid the foundations for this work, which would not exist without him. It has been a great delight to have had Dr. Vladan Velisavljevic as a mentor since his suggestions regarding technical aspects were of primary importance, when my research reached the experimental stage.

Special thanks are due to Dr. Susanne Fiebig and Dr. Matthias Pesch from *ARRI Cine Technik GmbH & Co. KG* who kindly took time off from their own work to help measure objective lens parameters. At this point, I also would like to express my sincere gratitude towards Dr. Johannes Steurer for introducing me to the research field of plenoptic cameras.

I am thankful to Dr. Shyqyri Haxha, Dr. Nikolaos Boulgouris, Dr. Emmanuel Tsekleves and Dr. Renxi Qiu who were partly guiding me through academic and organisational matters during my time as a research student. Many thanks are due to my research fellows Dr. Obaid Abdul Fatah and Juan Carlos Jácome Fernández for providing assistance with the laboratory equipment, sharing thoughts on optics and the disputes we had, which kept me walking that extra mile. I further acknowledge the university's *Department of Media Arts & Production* and particularly Michael Brennan and his team for providing professional light equipment.

I consider myself to be lucky that I got the opportunity to have met Dr. Andrew Lumsdaine as well as Dr. Héctor Navarro Fructuoso who inspired and encouraged me to pursue my research ideas. Besides, I feel grateful for all the volunteers developing and contributing to LaTeX, who made writing this thesis a pure joy. Furthermore, I thank my close friends who gave advice and support to me throughout the past four years and were enriching my life in the United Kingdom.

Lastly, I give thanks to Dr. Martin Charlton and Prof. Eran Edirisinghe for their suggestions as they were examining this thesis.

Declaration

Hereby, it is declared that the work in this thesis was carried out in line with the requirements of the University's Research Degree Regulations and that it has not been submitted for any other academic award. Unless indicated by specific reference in the text, the presented study is the candidate's own work. Research done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in this thesis are those of the author.

Christopher Hahne

13^{th} January, 2016

Table of Contents

List of Abbreviations

AND	Logical AND gate
ASIC	Application-Specific Integrated Circuit
BRAM	Block Random-Access Memory
CAC	Coded-Aperture Camera
CPU	Central Processing Unit
DoF	Depth of Field
FIR	Finite Impulse Response
FP	Focal Plane
FPGA	Field Programmable Gate Array
GPU	Graphics Processing Unit
HDMI	High-Definition Multimedia Interface
IIR	Infinite Impluse Response
LUT	Look-Up Table
MIC	Micro Image Centre
MLA	Micro Lens Array
NN	Nearest-Neighbour
PCLK	Pixel clock frequency
RoI	Region of Interest
SAD	Sum of Absolute Differences
SPC	Standard Plenoptic Camera
TFC	Time of Flight Camera
TMDS	Transition-Minimised Differential Signalling
VHDL	VHSIC Hardware Description Language

List of Figures

Figure **Page**

List of Tables

Table **Page**

Chapter 1

Introduction

'The art of perspective is of such a nature as to make what is flat appear in relief and what is in relief flat.'

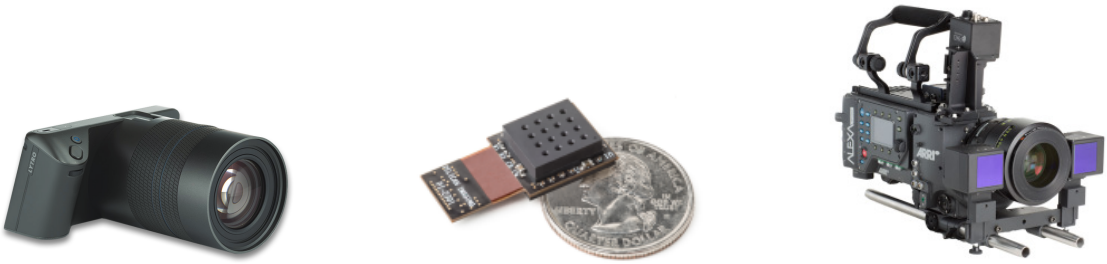
Leonardo da Vinci

The above quotation from possibly the most influential 'image technician' of his time gives an idea of how people in earlier centuries were intrigued by the phenomenon of three dimensions in space. About 500 years after da Vinci, imaging science has taken on such a significance in the field of computer science that a sub-field called 'computer vision' has emerged, which particularly addresses the use of machines to solve problems like the statement above. Though the field of computer vision was first defined only a few decades ago, it appears that its fundamental observations and related questions can be traced back to at least the Renaissance era.

1.1 Motivation

While the first single lens imaging system, the *Camera Obscura*, was used by artists simply as an aid for drawing images, the development of photo-sensitive recording media enabled the capture and reproduction of realistic pictures of actual objects. With such a conventional camera, the depth and thus the angular information of light rays is lost at

the moment of image acquisition, since the irradiance of all rays striking a sensor element is averaged regardless of the rays' incident angle. Light rays originating from an object point that is out of focus will be scattered across many sensor elements. This becomes visible as a blurred region and cannot be satisfyingly resolved afterwards. To overcome this problem, an optical imaging system is required to enable detection of the light rays' direction. Figure 1.1 shows a selection of cutting-edge technologies which aim to solve this task in different ways.



(a) Plenoptic camera [55] (b) Multi-camera array [78] (c) Time of flight camera [37]

Figure 1.1: **State-of-the-art camera technologies.**

Probably, the oldest and best-studied attempt to infer depth from real-world object scenes is to imitate our human natural vision and utilise two cameras to form a stereoscopic setup. This allows a manageable complexity and can be easily adapted to certain demands such as in the automotive industry or cinematography. Over the time, this setup has advanced to a multi-camera system which consists of more than two cameras and therefore enhances the quality of depth information. Examples of such camera arrays are nowadays found in mobile phones [78]. Besides their capability of capturing depth, mobile camera arrays allow a phone to be manufactured thinner than single lens cameras; the camera's width constrains the mobile phone's thickness and this is important for marketing the phone.

Currently, several other technologies are the subject of research, with some already on the market. Among them is the *Time-of-Flight Camera* (TFC) which exploits the infrared light spectrum and measures the time between light emittance and the arrival of reflected light waves. Due to the fact that infrared light is invisible to the human eye, the image output simply provides a depth map without any texture or pattern information. Consequently, the TFC is often used in conjunction with conventional cameras, having the visible colour information mapped onto the reconstructed depth relief [37]. An alternative to this is the hand-held plenoptic camera, also called a light

field camera, which employs a non-sequential array of tiny lenses placed in front of the image sensor to acquire the light rays' angular direction. Since it also makes use of the visible light spectrum, lenses and image sensors, it can be seen as a derivative of the multi-camera setup. Thus, the majority of knowledge gained about camera arrays may be similarly applied to the plenoptic camera. One of those commonalities is that the plenoptic camera enables the computational reconstruction of photos with varying focus, as well as a change of the spectator's viewing angle after image capture [45, 61]. Because the plenoptic camera unifies these capabilities in a compact size, it has been claimed that its release will cause a revolution in photography, with an impact similar to that of digitisation. Different types of plenoptic camera have already been created, but its commercial breakthrough is still to come. In the meantime, plenoptic camera research has been conducted in various fields such as microscopy [15, 51], endoscopy [80], motion picture arts [5, 72] and particle image velocimetry [32, 63]. While researchers are devising setup modifications, there still exists a little knowledge about the geometrical–optics behaviour of the plenoptic camera. Therefore, this study aims to complement to the research done so far by investigating the impact of the lens design on the image quality.

1.2 Contributions

When scrutinising publications in the field of plenoptic cameras [35, 61, 66], it became apparent that an uncertainty exists about their light field geometry. Unlike in a camera array, the distance between observers' viewpoints, so-called baseline, is not obvious in a plenoptic camera. Can the well known concept of stereo triangulation be applied here? If so, it raises the next question about which parameters affect the plenoptic camera's baseline. For time being, a lack of promising research studies leaves this open. The baseline is essential in cases where plenoptic camera content has to be displayed on autostereoscopic monitors.

Apart from this, the physical depth range to which a plenoptic camera is able to refocus is unknown at this stage. A professional photographer might wonder whether a light field camera's settings enable to equally refocus on an object at 30 cm distance and another one at about 30 m distance. What kind of impact does the main lens focus and zoom settings have on the refocusing distance? Is there a non-refocusable area and where is it located? What is the depth of field of a single refocused photograph and how is it controlled?

This doctoral dissertation introduces a new light field model describing rays travelling through the elements of a plenoptic camera to help answer these questions. A summary of the major contributions is provided in the following:

1. **Light field ray model** – partially published in [39–41]

   With the help of the proposed ray model, it is now possible to pinpoint the object distance at which refocused photographs exhibit least blur. In addition to that, the position of virtual cameras providing perspective views can be determined enabling the stereo triangulation scheme to be applied to a plenoptic camera. As an alternative to existing methods, the model also facilitates to derive efficient equations in spatial domain to synthesise refocused or perspective images in the light field. A side benefit is the model's illustrative presentation which only requires basic understanding of geometrical optics and thus may be intuitively accessible.

2. **Plenoptic camera calibration** – publication pending

   A part of this thesis deals with the assembly of a plenoptic camera and thereby presents novel optical calibration methods to ensure micro and main lenses are positioned appropriately.

3. **Refocusing hardware implementation** – partially published in [38]

   It has been the goal to develop the first hardware architecture enabling a *Field Programmable Gate Array* (FPGA) to accomplish a computational focus variation in real-time. The design consists of switch-controlled *Finite Impulse Response* (FIR) filter modules organised such that incoming data streams can be pipelined and parallelised. It is shown that a computation time speed-up is achieved by using underexposed light field photographs.

The proposed ray model assists in the object distance estimation through refocusing if combined with a blur metric or elaborated triangulation using the well studied disparity analysis. This will be a key feature for depth maps acquired from a plenoptic camera.

Provided novelties will also support in the prototyping stage of a plenoptic camera. For instance, optical engineers can benefit from the model's capability to estimate the synthetic focusing distance of refocused photographs and virtual baseline of perspective views in order to optimise the optical design for certain demands. Furthermore, the presented real-time refocusing hardware implementation is needed in light field cinematography to allow a plenoptic camera to preview refocused images on a monitor.

It would be the writer's pleasure to see that the presented work enlightens interested peers, inspires to dig deeper and carry it further to any direction.

1.3 Outline

This section aims to provide a general map to the content structure of this thesis. To serve as a groundwork, Chapter 2 covers relevant theoretical principles of three-dimensional (3-D) image capture such as stereo and multi-view vision as well as related ray models. In order to narrow down the subject, this is followed by an overview of the plenoptic camera development which is made up of a brief historic analysis examining early concepts right up to all types of state-of-the-art setups.

The elementary work in this thesis is presented in Chapter 3 which is split into two major parts, namely theoretical models and practical applications. Thereby, the first section introduces a novel method to trace rays through a plenoptic camera by means of paraxial optics enabling to deduce image synthesis algorithms. A second section contrives attempts to predict the refocusing distance of depth filtered photographs as well as positions of virtual cameras that project viewpoint images onto the plenoptic sensor.

Since plenoptic camera based image processing requires accurate light field photographs and reference points as an orientation in the captured image data, Chapter 4 concerns with the question of how to properly calibrate the optical components of a plenoptic camera. The chapter is subdivided into two parts explaining the calibration on an optical bench just as the detection of reference image points.

All experimental work validating the suggested ray model is provided in Chapter 5. This includes an evaluation by actual photographs taken from a customised plenoptic camera just as by real ray simulation using *Zemax* [89]. Throughout this chapter, *Zemax* is either referred to as 'optical design software' or 'real ray simulation' for reasons of neutrality and comprehensibility. The validation of actual light field images is carried out by employing a blur metric and disparity analysis.

On the basis of the contrived ray model, Chapter 6 demonstrates the development of a power-efficient hardware architecture in the form of *Finite Impulse Response* (FIR) filters to perform refocusing in real-time. The filter behaviour is validated in an experimental section by VHDL simulation and implementation using a *Field Programmable Gate Array* (FPGA). Lastly, Chapter 7 summarises results by glancing at what has been achieved and what can be solved in future. Further research ideas that possibly lead to encouraging improvements are presented at the very end.

CHAPTER 2

Background

The core tasks in the field of computer vision have been to digitally capture, recover and interpret a real object scene with the aid of computer systems. By attempting to reproduce human visual perception, computer vision advanced to a level at which it can acquire relative depth information.

This chapter reviews relevant concepts of 3-D image retrieval. At the start, the stereoscopy theory is discussed as it strongly relates to human vision due to the fixed number of only two positions from which a scene is observed. This is followed by an introduction of multi-view imaging models that aim to obtain depth information from more than two standpoints. Bearing these explanations in mind, the second half of this chapter turns the focus on the image acquisition side and more specifically on the history and recent development of the plenoptic camera. The fact that cameras consist of optical elements signifies that the research subject is interdisciplinary as it spans from physical optics, such as geometrical and *Fourier* optics, to multi-dimensional signal-processing.

2.1 Three-Dimensional Imaging

2.1.1 Stereoscopy

According to Wheatstone [81], a set of solely two adjacent cameras facilitates imitating our human's binocular vision. Using these two images in conjunction with a matching stereoscopic display technique, e.g. 3-D glasses [43], allows for the reproduction of depth as perceived by the human eyes. With regards to the location in object space, however,

this stereo camera setup concedes much more freedom than the human's perception as the separation between cameras may vary. Hence, the flexibility in camera stereoscopy makes it possible to adapt to particular depth scenarios which will be subject to scrutiny in what follows.

2.1.1.1 Coplanar Stereo Cameras

Figure 2.1 visualises the geometry of a stereoscopic system. The setup may be parameterised by the spacing of the cameras' axes, denoted as B for baseline, the cameras' image distance b and the optical centres O_L, O_R for each camera, respectively. As seen in the diagram, an object point is projected onto both camera sensors as indicated by orange dots. With regards to corresponding image centres, the position of the image point in the left camera clearly differs from that in the right. This phenomenon is known as *parallax* and results in a relative displacement of image points at different viewpoints. To measure this displacement, the horizontal disparity Δx is introduced which can be obtained by $\Delta x = x_R - x_L$ where x_R and x_L denote horizontal distances from each projected image point with respect to the optical image centre. Nowadays, image detectors are composed of discrete photosensitive cells making it easy to locate and measure Δx. The disparity computation is a well studied task [10, 56, 86]. Often, the problem has been extended to not only a single but rather many image points which yields a map of Δx values indicating the depth of a captured scene.

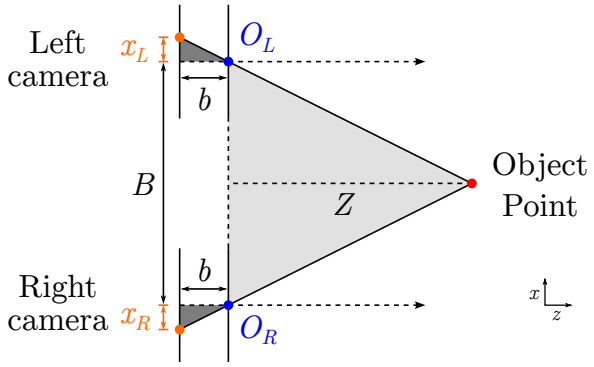


Figure 2.1: **Stereoscopic triangulation with parallel cameras.** A point is projected through the optical centres O_L, O_R yielding two image points (orange) in each camera. The relative displacement of these points returns the horizontal disparity $\Delta x = x_R - x_L$. The baseline B and image distance b affect the measured disparity.

The point distance Z can be directly fetched from parameters in Fig. 2.1. As highlighted with a dark tone of grey, Δx may represent the base of any acute scalene triangle with b as its height. Another triangle spanned by the base B and height Z is a scaled version of

it and shown in light grey. This relationship relies on the method of similar triangles and can be written as an equality of ratios

$$\frac{Z}{B} = \frac{b}{\Delta x}\,. \tag{2.1}$$

To infer the depth distance Z, Eq. (2.1) may be rearranged to

$$Z = \frac{b \times B}{\Delta x}\,. \tag{2.2}$$

As seen by these equations, it is feasible to retrieve information about the depth location Z. Likewise, if Δx is fixed, it may be obvious that by decreasing the baseline B the object distance Z shrinks. Given a case where the depth range is located at a far distance, it is thus recommended to aim for a large baseline. Note that this relationship and corresponding mathematical statements only hold for cases where optical axes of O_L and O_R are aligned in parallel.

2.1.1.2 Tilted Stereo Cameras

Reasonable scenarios exist in which a cameras' optical axis is rotated with respect to the other. In such case, the principle of similar triangles does not apply in the same manner as in Eq. (2.1).

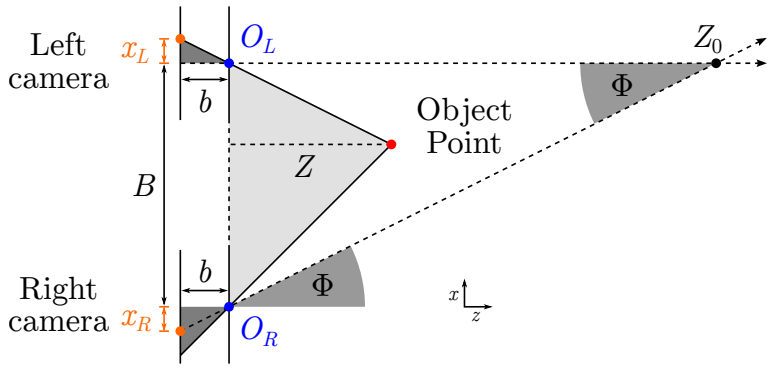


Figure 2.2: **Stereoscopic triangulation with non-parallel cameras.** Whilst sensors are still seen to be coplanar, Φ denotes the tilt angle of the right camera's main lens O_R as related to that of the left camera O_L.

Taking the left lens O_L as the orientation reference, the right lens O_R is seen to be tilted as shown in Fig. 2.2. If the rotation occurs around the y-axis, Iocchi [44] concludes that the optical axes intersect in a point Z_0 as both axes lie on the x, z plane. Under the assumption of a small angle rotation Φ, the image planes of both cameras are still seen

to be parallel. Let Φ be the rotation angle, then laws of trigonometry allow to put

$$Z_0 = \frac{B}{\tan(\Phi)} \tag{2.3}$$

and

$$Z = \frac{b \times B}{\Delta x + \dfrac{b \times B}{Z_0}} \tag{2.4}$$

which may be shortened to

$$Z = \frac{b \times B}{\Delta x + b \times \tan(\Phi)} \tag{2.5}$$

after substituting for Z_0. This approximation satisfies to fairly estimate the depth Z for small angles Φ in non-coplanar stereoscopic systems without the need of an image rectification.

2.1.2 Multi-View Imaging

As the term *multi-view* suggests, an object scene may also be observed from more than two viewpoints and thus it can be seen as concatenating multiple stereoscopic setups. Therefore, the geometrical laws of stereoscopic systems are maintained in multi-view imaging whereas the latter attempts to contrive parameterisations to satisfy 3-D image capture in a much broader sense by analysing the quality of depth information from a set of observers.

A popular study on human and machine vision addressing this was undertaken by Adelson and Bergen [3] in 1991. With that, the authors coined the term *plenoptic* composed of the Latin roots 'plen' and 'optic' which can be translated as 'full view' to propose the plenoptic function P which is given by

$$P = P(x, y, \lambda, T, V_x, V_y, V_z) \tag{2.6}$$

which consists of seven parameters making it often referred to as the seven-dimensional (7-D) plenoptic function. In that, parameters (x, y) denote the Cartesian coordinates of the plane from which light has been emanated, λ represents the light's wavelength, T the time at which images are taken and the three space coordinates V_x, V_y, V_z indicate the observer's position. For conciseness, the 7-D plenoptic function can be shortened to a

five-dimensional (5-D) function when capturing still monochromatic photographs leaving out the time T and wavelength λ terms.

With this in mind, Levoy and Hanrahan [50] devised a four-dimensional (4-D) notation in 1996 to achieve 3-D image acquisition by adequately describing a real object scene as a set of directional light rays. Thereby, the direction of each light ray is uniquely determined by its intersection with two two-dimensional (2-D) planes located behind one another as depicted in Fig. 2.3.

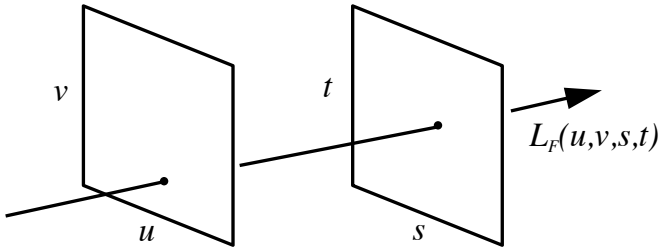


Figure 2.3: **4-D light field notation** (Ref. [50], Fig. 1 with minor modifications).

During the capturing process, a single ray can be seen as piercing through one plane (u, v) giving two coordinates at the respective intersection. Provided that the same ray passes through the second plane (s, t) placed at a known distance, its angular direction in space may be recovered from both intersections. This yields four coordinates in total to fulfil the condition of 3-D image retrieval from a 4-D light field L_F where

$$L_F = L_F(u, v, s, t)\,. \tag{2.7}$$

The experimental work in [50] has been accomplished with a single digital camera at a variable position along plane (u, v) and whose image sensor represents plane (s, t). The following years, the 4-D light field established popularity among scientists in the field of multi-view imaging as it outperforms the 7-D and 5-D plenoptic function due to the reduced number of dimensions. It is worth mentioning that the terms *light field camera* and *plenoptic camera* may be used interchangeably.

2.1.3 Refocusing

One of the many research outcomes that arose from the 4-D light field notation is the discovery of how to refocus photographs after image capture by using an array of cameras [46]. Therein, Isaksen *et al.* exposed the captured light field's capability of computing 2-D images focused at different depths for the first time. The synthetic focus variation has been accomplished by recording a scene with cameras at different viewpoint

positions (s, t) seen as the *entrance plane* and a rendering process called *aperture filtering* able to reconstruct photographs focused on selected *exit planes* (u, v) in object space. The suggested *aperture filtering* process finds spatial samples (u, v) in each view (s, t) that all correspond to the same object point at the given *exit plane*. Then, it merges the intensities of found samples (u, v) of each view (s, t) to create a spatial point of the synthetically refocused image. Further details on the image synthesis can be found in [45].

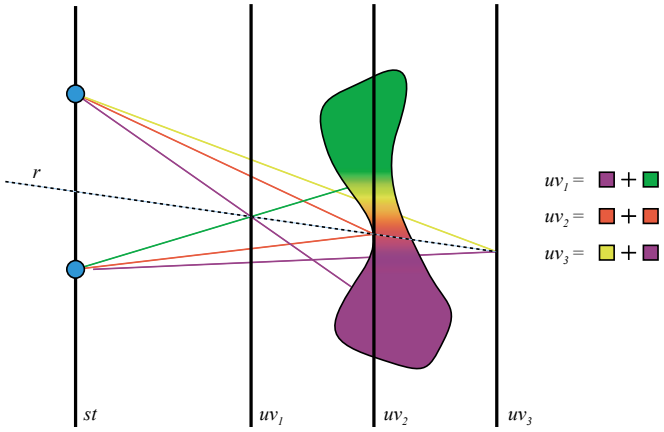


Figure 2.4: **Refocusing based on a camera array** (Ref. [46], Fig. 1).

With respect to Fig. 2.4, potential image points along the depth of a synthetic ray r are recovered by integrating image samples whose rays coincide at the intersection of r with the uv planes. Thus, parameterising the *aperture filter* to refocus on plane uv_1 and uv_3 results in a blurred image region whereas uv_2 exhibits best focus for the respective object. As pointed out in [46], one advantage of the synthetic aperture system is that it can 'look through' an object, even though this particular object occludes the background scene in each taken viewpoint image. Besides, the proposed reparameterisation allows lenslet displays to render autostereoscopic content. More detailed information on Isaksen's work can be found in his master's dissertation [45]. Interestingly, a later study performed by Vaish [76] revealed that the (u, v) plane does not necessarily have to be coplanar to the (s, t) camera plane. This may be easily comprehensible when letting the *aperture filtering* procedure to alter the combination of samples involved to form a desired (u, v) parameter whilst generating a single synthetically focused image.

2.2 Plenoptic Cameras

2.2.1 History

Commonly, it often remains disputable which person may be referred to as the originator of a certain research field. In case of the over 100 years old history of integral imaging, light field and plenoptic cameras, the majority of literature acknowledges French Nobel laureate in physics Gabriel Lippmann to have been the scientist constituting the research field of plenoptic cameras with his publication *Épreuves Réversibles Donnant La Sensation du Relief* in 1908 [52]. Indeed, Lippmann was the first person to use an array of micro lenses to gather depth information without the presence of any additional optical element. However, it is worth mentioning that his imaging system does not fully comply with the later definition of a plenoptic camera. Another noticeable fact is that the American physicist Frederic E. Ives already filed a patent of a camera setup [47] in 1903 consisting of a pinhole array whilst having a main lens attached in front of it.

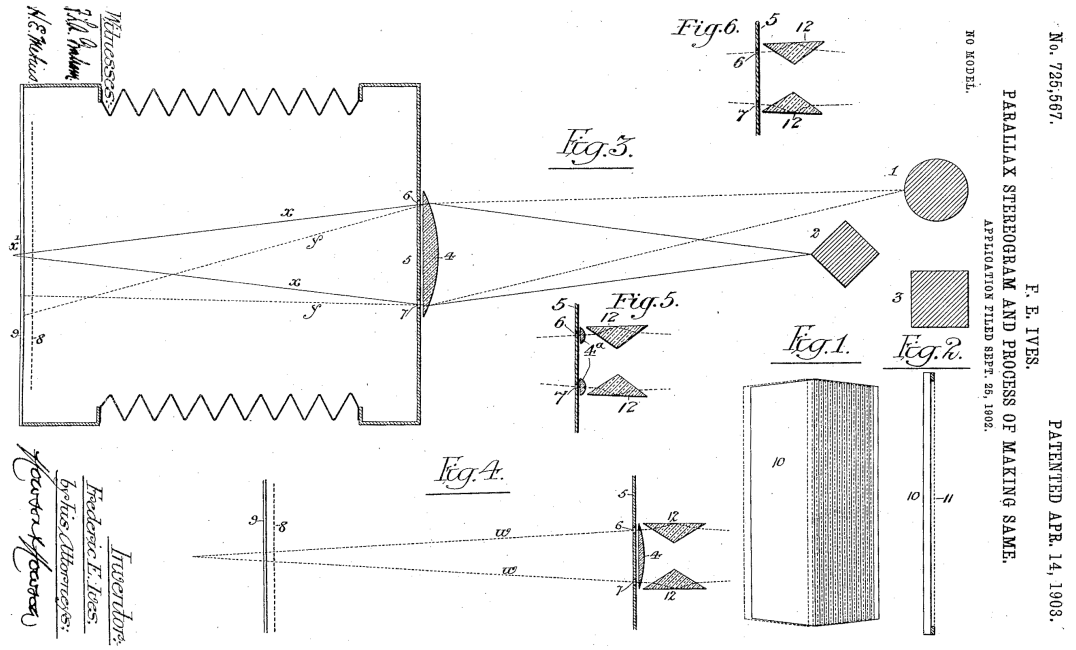


Figure 2.5: **Patent showing first known draft of a plenoptic camera** (Ref. [47]).

An excerpt of F. E. Ives' submitted patent is shown in Fig. 2.5. As seen in the patent drawings, F. E. Ives recognised the plenoptic camera's capability of distinguishing depth from different objects. In 1911, Sokolov [71] discussed geometrical ray tracing methods to achieve stereoscopic image capture and display based on the propositions Lippmann

made. Many similar camera designs have been proposed over the subsequent decades of the previous century including H. E. Ives [48], Coffey [20] and Dudnikov who published several articles [23–28] in partial collaboration with Rozhkov and Antipova. From the historical perspective, the digitisation of imaging capture devices revived the research on plenoptic cameras as it allowed for storing a large amount of data and cost-effective, quick methods in post-processing such as reordering or blending sampled illuminance values.

The term *Plenoptic Camera* is credited to Adelson and Wang [4] named after the publication *Single Lens Stereo with a Plenoptic Camera* presented in 1992. This study was groundbreaking in several ways. It provided the first plenoptic camera setup consisting of a digital image sensor, a lenticular array and a main lens which made it possible to generate sub-aperture images for the first time. Moreover, Adelson and Wang were the first to perform a computer-aided analysis on image data obtained by their plenoptic camera to find disparities Δx.

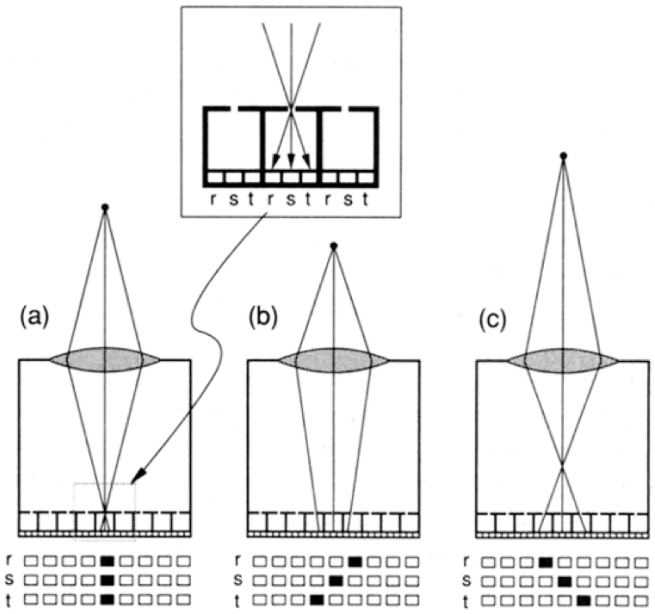


Figure 2.6: **Digital plenoptic camera** (Ref. [4], Fig. 5) illustrating rays for different object point distances in (a), (b) and (c) where r, s and t refer to pixel positions in each micro image.

The authors came to the conclusion that the plenoptic camera's major restriction is that its baseline is confined to the main lens' aperture size which would cause depth inaccuracies [4]. However, quantitative baseline results and methods to support this assumption were not provided. It is believed that a small baseline limits the capability

of resolving depth at far distances, but improves depth resolution for nearby objects. Figure 2.6 depicts a basic ray model of a digital plenoptic camera indicating the orientation of directional information within so-called micro images. A micro image can be either projected by a single micro lens or pinhole that is part of a larger array. It is shown in (a) that a single object point focused on the array is scattered along a single micro image. In turn, image points converging behind (b) or in front (c) of the array are distributed over different micro images and flipped over with respect to each other. Several optical setups have been proposed in this work including pinhole just as lenslet arrays. Besides that, a relay lens was utilised between the lenticular lens sheet and the imaging device to overcome the problem of mounting the array directly on the sensor.

2.2.2 Standard Plenoptic Camera

A major study that combines the idea of refocusing with the physical concept of a plenoptic camera was firstly undertaken by Ng *et al.* [61] in 2005. In his doctoral dissertation [59], Ng examines several image distances b_s of a *Micro Lens Array* (MLA) with respect to the sensor in a range smaller than and equal to the focal length f_s of the micro lenses ($b_s \leq f_s$). Furthermore, the author suggests to separate the sensor from the micro lenses by one focal length ($b_s = f_s$) stating that this MLA position offers maximal u directional resolution and minimal s spatial resolution [59]. Later, this setup was named the *Standard Plenoptic Camera* (SPC) [66].

The SPC system is an optimisation with priority on directional resolution meaning that the number of samples in spatial domain have been traded off against a larger number of samples in directional domain. In fact, the amount of micro lenses per dimension represent the effective spatial resolution of an SPC. Hence, a *full frame* sensor format of 36 mm width and 24 mm height having an array of micro lenses specified by 0.125 mm pitch size in front of it merely amounts to a 288 by 192 resolution viewpoint image. Nonetheless, much higher sample rates can be achieved with larger MLA's containing more micro lenses with a smaller diameter. To investigate the SPC's resolution limitations, it is worth referring to Abbe's diffraction limit at this point in which the resolvable feature size D is given by

$$D = \frac{\lambda}{2n \sin \theta_{\max}} \tag{2.8}$$

where n is the medium's refraction index and $\theta_{\max}$ denotes the half-angle of the maximum cone of a light beam that enters the lens [53]. Term $n \sin \theta_{\max}$ is known as the *Numerical Aperture* NA. Considering red light having a wavelength $\lambda = 700$ nm and

$NA = n \sin \theta_{\max} = 1.4$, the smallest resolvable image point would be $D = 250$ nm. Suppose a monochromatic instrument-limited system where the pixel width $p_p = 1\ \mu\text{m}\ > D$, then a fair number of directional samples u may be 10 which yields a micro lens pitch of about 10 μm. Placing an array of this micro lens type in front a *full frame* sensor provides 3600 by 2400 pixel viewpoint images without any kind of superresolution techniques. This would match the 4 k resolution of present-day television devices. The company founded by Ng which made SPCs commercially available has practically proven that manufacturing an MLA of approximately 13.9 μm micro lens pitch is feasible [62].

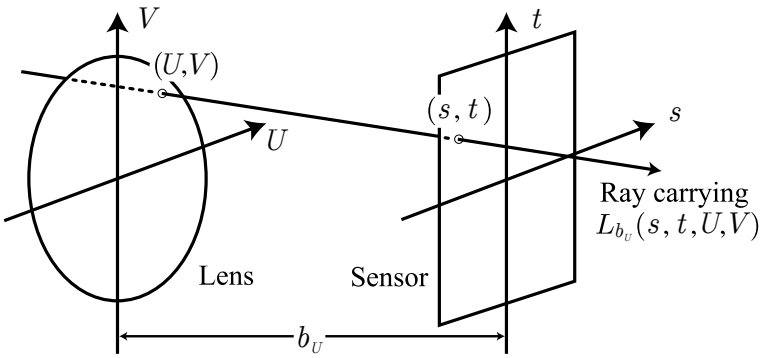


Figure 2.7: **Light field camera parameterisation** (Ref. [58], Fig. 1 with minor modifications). In accordance with the 4-D light field definition, a single ray intersects two 2-D planes, namely main lens (U, V) and sensor plane (s, t).

To mathematically demonstrate image rendering based on the SPC, Ng deploys the 4-D light field parameterisation illustrated in Fig. 2.7. Thereby, the main lens is treated as a thin lens whose 2-D principal plane is given by variables (U, V) and the 2-D sensor plane by coordinates (s, t) such that each light ray may be parameterised by $L_{b_U}(s, t, U, V)$. Note that notations have been slightly modified for the sake of consistency with declarations made in following chapters.

2.2.2.1 Sub-Aperture Image Extraction

Initially, the extraction of digital multi-view images from an SPC was achieved by Adelson and Wang [4]. In the literature of SPC's [59, 61], these synthetically generated multi-view images are commonly referred to as *sub-aperture images* and correspond to perspective views in the light field. Figure 2.8 depicts the relationship between micro images and sub-aperture images. To compose a sub-aperture image, the extraction process collects one sample at a fixed location (U, V) from every micro image (s, t) in the captured light field L_{b_U} and strings them together.

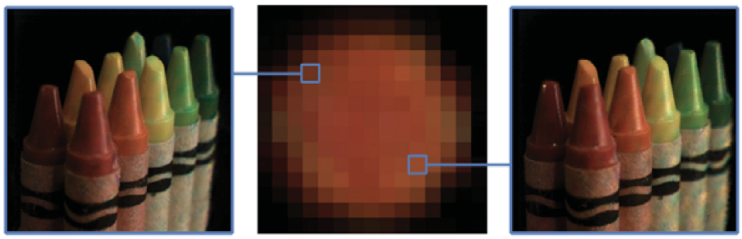


Figure 2.8: **Sub-aperture image extraction** (Ref. [59], Fig. 4.2) generates sub-aperture images from the raw SPC image capture. Each spatial image point (s, t) of a particular sub-aperture image (e.g. left or right) is taken from the same respective sample (U, V) in each micro image. Here, a single micro image (middle) highlights the (U, V) position of the two given sub-aperture images.

Consequently, the spatial resolution of a sub-aperture image corresponds to the number of micro lenses. The resolution loss is due to the fact that portions of the sensor space are sacrificed to capture depth information. Bishop *et al.* [9] exposed superresolved sub-aperture images taken from an SPC by deploying lens deconvolution algorithms.

2.2.2.2 Shift and Integration Synthesis

Probably, the easiest way to grasp digital refocusing by means of an SPC is by starting from previously extracted sub-aperture or conventional viewpoint images. Figure 2.9 illustrates overlaid sub-aperture images from Fig. 2.8 that are blended whilst varyingly shifted with respect to each other. On the left, pixels of both images were averaged without shifting. Closer inspection reveals that ghosting artefacts occur for the red and green crayons except for the yellow ones. These artefacts become visible due to the difference in directional information. Since this simple example solely utilises two sub-apertures, the directional domain can be seen as undersampled. Thus, using more adjacent sub-apertures (higher directional sample rate) leads to a distribution of these artefacts in nearby areas resulting in a blur as if captured by a conventional misfocusing camera. As seen in the middle column of Fig. 2.9, focus of a desired object (e.g. red crayon) is retained by shifting its images to identical positions such that it remains sharp when overlaid whereas blurry artefacts emerge for all surrounding objects (e.g. green crayon).

Figure 2.9: **Shift and integration process** (Ref. [59], Fig. 4.2) implies that a sub-aperture image (left) is overlaid with another (middle) and shifted whilst blending pixel values. The shift variation moves the synthetic focus from red crayons in the foreground (middle) to green crayons in the background (right).

The geometrical derivation of refocusing according to Ng [59] is more laborious and explained hereafter. In a conventional camera, the irradiance I_{b_U} of a light field L_{b_U} measured on the sensor at (s,t) is a weighted integral of directional samples (U,V) given by

$$I_{b_U}(s,t) = \frac{1}{{b_U}^2} \int\int L_{b_U}(s,t,U,V)\, A(U,V)\, \cos^4\theta \,\mathrm{d}U\,\mathrm{d}V \tag{2.9}$$

where $A(\cdot)$ is the aperture and b_U the spacing between main lens and image sensor. The combination of ray angle θ and the roll-off factor $\cos^4\theta$ operates as optical vignetting [42].

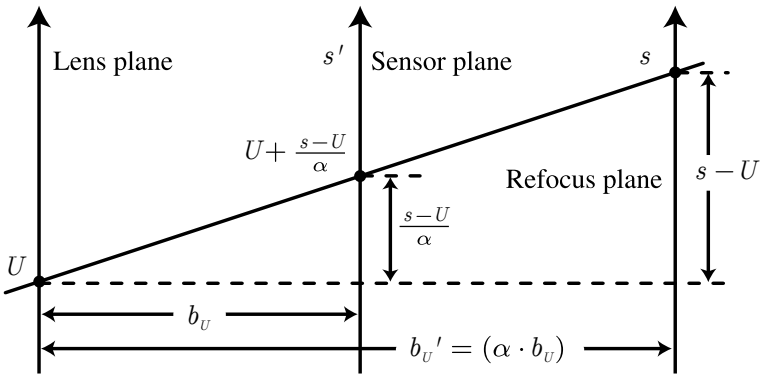


Figure 2.10: **Cross-sectional view of light field reparameterisation** (Ref. [58], Fig. 2 with minor modifications). By the method of similar triangles and a reparameterisation in which plane s now moves behind the sensor, the sensor position of the depicted ray travelling through U and s may be retrieved via $U + \frac{s-U}{\alpha}$.

Similar to Isaksen's approach, Ng demonstrates refocusing by a reparameterisation shown as a cross-section in Fig. 2.10. When reparameterising the 4-D light field such that the (s,t) plane is seen to be shifted by factor $\alpha \cdot b_U$ behind the sensor at ${b_U}'$, the height of a ray is given by $s - U$. The overall height of the corresponding ray on the actual sensor

plane s' is now given by $U + \frac{s-U}{\alpha}$. This can be easily proven by the method of similar triangles. Following to the reparameterisation where L_{b_U} produces $L_{b_U\prime}$, variable s in L_{b_U} may be substituted by $U + \frac{s-U}{\alpha}$ which, if also applied to the vertical dimension, gives

$$\begin{aligned} L_{b_U\prime}\left(s', t', U, V\right) &= L_{b_U}\left(U + \frac{s-U}{\alpha}, V + \frac{t-V}{\alpha}, U, V\right) \\ &= L_{b_U}\left(U\left(1 - \frac{1}{\alpha}\right) + \frac{s}{\alpha}, V\left(1 - \frac{1}{\alpha}\right) + \frac{t}{\alpha}, U, V\right). \end{aligned} \tag{2.10}$$

We now substitute the reparameterised light field term $L_{b_U}(\cdot)$ in Eq. (2.9) and obtain

$$I_{(\alpha \cdot b_U)}(s,t) = \frac{1}{\alpha^2 {b_U}^2} \int\int L_{b_U}\left(U\left(1 - \frac{1}{\alpha}\right) + \frac{s}{\alpha}, V\left(1 - \frac{1}{\alpha}\right) + \frac{t}{\alpha}, U, V\right) \mathrm{d}U\,\mathrm{d}V \tag{2.11}$$

to accomplish refocusing with parameter α in spatial domain. Note that $A(U,V)\cos^4\theta$ have been neglected for simplicity.

2.2.2.3 Fourier Slice Photography Theorem

The *Fourier Slice Photography Theorem* has been contrived by Ng [58] as an alternative signal processing approach to achieve refocusing in frequency domain. As a starting point, Ng employs the *Fourier Projection-Slice Theorem* [13] which originated from the *Radon Transformation* [68]. The classical *Fourier Projection-Slice Theorem* states that the *Fourier*-transformed signal $\mathfrak{G}_X$ of a from Y to X dimensions integral-projected signal G_X is equivalent to a slice $\mathfrak{G}_X$ parallel to the projection direction and going through the origin of the same Y-dimensional signal $\mathfrak{G}_Y$ in frequency domain. This may be put as

$$\mathcal{F}^X \circ \mathcal{P}_X^Y = \mathcal{S}_X^Y \circ \mathcal{F}^Y \tag{2.12}$$

where $\mathcal{P}$ denotes an integral projection operator of a Y- to an X-dimensional signal. $\mathcal{F}$ represents the *Fourier* transform and $\mathcal{S}$ the slicing operator reducing the dimensions of the signal from Y to X in frequency domain by zeroing-out. In accordance with Ng's notation, calligraphic letters describe operators which are successively performed from the right to the left.

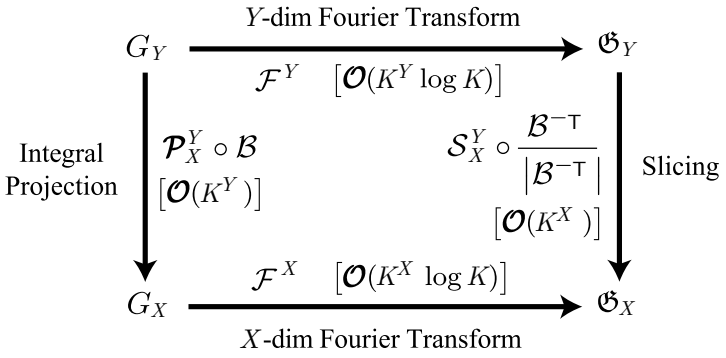


Figure 2.11: **Generalised Fourier Slice Theorem** (Ref. [58], Fig. 4 with minor modifications) as a visualisation of Eq. (2.13) where the signal G_Y can be integral-projected down to X dimensions yielding G_X or *Fourier*-transformed to $\mathfrak{G}_Y$ whose spectrum may be sliced to X dimensions giving $\mathfrak{G}_X$. Both procedures require a change of basis $\mathcal{B}$ to be applied before dimension reduction. $\mathcal{O}(\cdot)$ indicates the computation time for each operation where K is the number of samples per dimension.

An extended version of Eq. (2.12) is the *Generalised Fourier Slice Theorem* depicted in Fig. 2.11. It states that when changing the basis of an Y-dimensional signal G_Y which is integral-projected down to X dimensions and *Fourier*-transformed afterwards, the resulting signal G_X corresponds to *Fourier*-transforming the original signal G_Y, changing its basis with the normalised inverse transpose of the projection basis and slicing it down to X dimensions. Using the previously introduced operators, this process is algebraically expressed by

$$\mathcal{F}^X \circ \mathcal{P}_X^Y \circ \mathcal{B} = \mathcal{S}_X^Y \circ \frac{\mathcal{B}^{-\mathsf{T}}}{|\mathcal{B}^{-\mathsf{T}}|} \circ \mathcal{F}^Y \tag{2.13}$$

with $\mathcal{B}$ as a change of basis indicating a passive transformation that can be seen as a shear transformation while $\mathcal{B}^{-\mathsf{T}}$ is its inverse transpose and $|\mathcal{B}^{-\mathsf{T}}|$ the respective determinant. Bearing in mind that refocusing is based on a 4-D light field, Ng describes the change of basis $\mathcal{B}_\alpha$ as

$$\mathcal{B}_\alpha = \begin{pmatrix} \alpha & 0 & 1-\alpha & 0 \\ 0 & \alpha & 0 & 1-\alpha \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{B}_\alpha^{-1} = \begin{pmatrix} \frac{1}{\alpha} & 0 & 1-\frac{1}{\alpha} & 0 \\ 0 & \frac{1}{\alpha} & 0 & 1-\frac{1}{\alpha} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.14}$$

As a consequence of the 4-D notation, the input signal G_Y is of four dimensions and the refocused output image G_X spans two dimensions such that $Y = 4$ and $X = 2$. When

considering this, a refocusing operator $\mathfrak{R}_\alpha$ is defined as

$$\mathfrak{R}_\alpha [L_F] \equiv \frac{1}{\alpha^2 {b_U}^2} \mathcal{P}_2^4 \circ \mathcal{B}_\alpha [L_F] \tag{2.15}$$

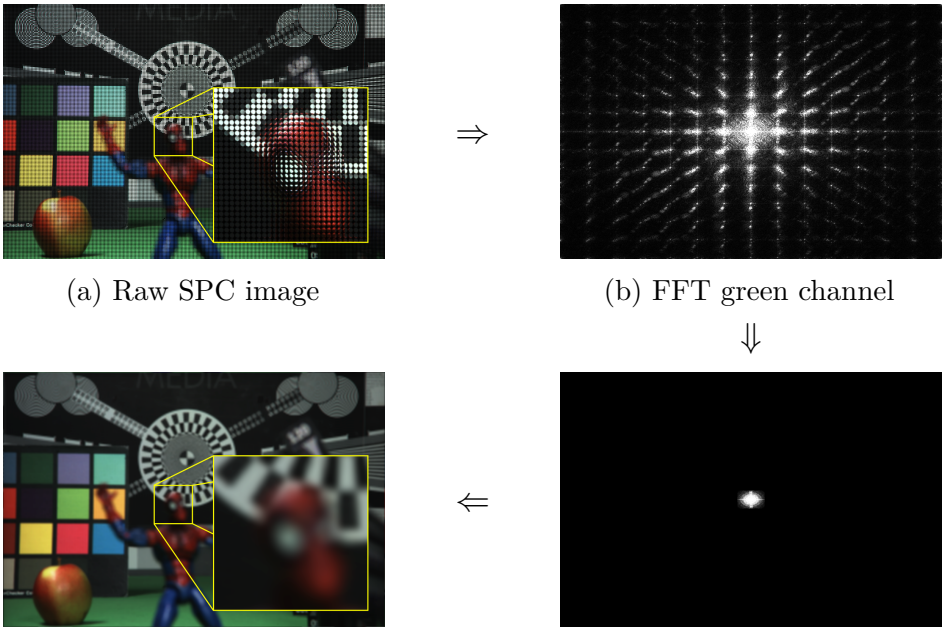
when using the spatial integration. According to Eq. (2.13), the refocusing operator can be rewritten as follows

$$\mathfrak{R}_\alpha \equiv \frac{1}{\alpha^2 {b_U}^2} \mathcal{F}^{-2} \circ \mathcal{S}_2^4 \circ \frac{\mathcal{B}_\alpha^{-\mathsf{T}}}{|\mathcal{B}_\alpha^{-\mathsf{T}}|} \circ \mathcal{F}^4 \tag{2.16}$$

where $\mathcal{F}^{-2}$ is the 2-D inverse *Fourier* transform. Since $|\mathcal{B}_\alpha^{-\mathsf{T}}| = |\mathcal{B}_\alpha^{-1}| = 1/\alpha^2$, Eq. (2.16) may be reduced to

$$\mathfrak{R}_\alpha \equiv \frac{1}{{b_U}^2} \mathcal{F}^{-2} \circ \mathcal{S}_2^4 \circ \mathcal{B}_\alpha^{-\mathsf{T}} \circ \mathcal{F}^4 \tag{2.17}$$

which gives the final procedure of how refocusing is achieved in frequency domain when starting from a raw light field camera image.



(a) Raw SPC image (b) FFT green channel

(c) IFFT where $\alpha = 1$ (d) FFT Kaiser window green channel

Figure 2.12: **Fourier slice photography**. The rectangular arrangement of depicted images is compliant with Fig. 2.11. A raw SPC image is illustrated in (a) whose shifted magnitude per frequency component is seen in (b). The central tile carries the DC component and all surrounding tiles contain the angular information. (d) shows the magnitude after using the slicing operator $\mathcal{S}$. The remaining central tile is weighted with a Kaiser-Bessel filter having a width of 3.5 to prevent aliasing in (c).

Figure 2.12 illustrates results of the employed *Fourier Slice Photography Theorem* using Eq. (2.17). Note that in this example $\alpha = 1$ and hence $\mathcal{B}_{\alpha}^{-\mathsf{T}}$ becomes an identity matrix making the change of basis operation redundant.

With reference to the *Generalised Fourier Slice Theorem* in Fig. 2.11, the shearing and slicing operation $\left(\mathcal{S}_2^4 \circ \mathcal{B}_{\alpha}^{-\mathsf{T}}\right)$ in frequency domain results in a quicker computation time $\mathcal{O}$ than shearing and integrating via $\left(\mathcal{P}_2^4 \circ \mathcal{B}_{\alpha}\right)$ in spatial domain since $\mathcal{O}\left(K^2\right) < \mathcal{O}\left(K^4\right)$. However, it may be obvious that a single refocused photograph based on the *Fourier* approach includes the $\mathcal{F}^4$ and $\mathcal{F}^{-2}$ transforms, which in total takes more time than the spatial integration. Nonetheless, Ng argues that $\mathcal{F}^4$, taking $\mathcal{O}\left(K^4 \log K\right)$ time, merely needs to be performed once since all remaining refocused photographs can be extracted from its resulting frequency signal $\mathfrak{G}_Y$ yielding a computation time of only $\mathcal{O}\left(K^2\right) + \mathcal{O}\left(K^2 \log K\right)$ for each of them. For comparison, the integral projection takes $\mathcal{O}\left(K^4\right)$ time per photograph and therefore longer when a stack of computationally focused images is required.

2.2.2.4 Calibration

Even though an SPC facilitates a variety of 3-D image reconstruction techniques, calibrating a plenoptic camera is the premise of acquiring and processing accurate output images. Ng extensively studied lens design in 2006 to suppress lens aberrations in extracted sub-apertures [60]. However, procedures to find reference points in the 4-D raw image remained unpublished until Cho *et al.* [19] and Dansereau *et al.* [22] addressed this. From then, micro image centres serve as reference points and their detection became a crucial part in the SPC calibration. The work carried out by Dansereau *et al.* exposes quantitative results on the basis of a thin lens model tracing rays from micro image centres as seen in Fig. 2.13.

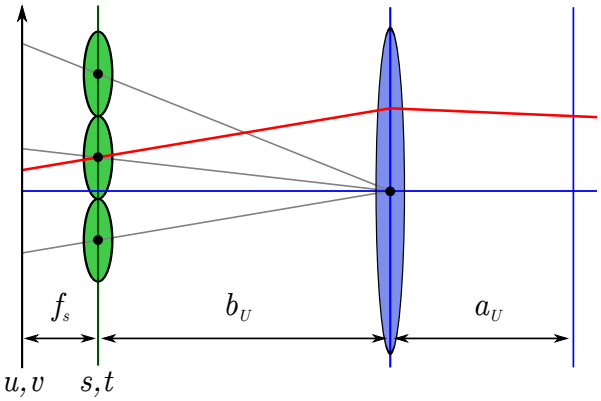


Figure 2.13: **Pinhole and thin lens model** (Ref. [22], Fig. 3 with minor modifications). Gray lines originating from the main lens (blue-coloured) represent micro lens chief rays indicating their corresponding image centres on the sensor plane (u, v).

Despite the fact that Dansereau proposed a geometrical approach to estimate micro image centre positions, he deployed a method to calculate the illuminance peak of each micro image starting from a raw SPC image as an input signal. To eliminate lens distortions, Dansereau applies the Brown-Conrady model [14]. More details on the transformations and applications of his developed intrinsic and extrinsic matrices as well as results can be found in [22]. Recent research outcomes dealing with the calibration of a plenoptic camera have also been presented by Johannsen *et al.* [49] and Zeller *et al.* [88]. The authors employ a disparity analysis in combination with the intercept theorem yielding an intersection of rays coming from the same point in object space. Given the distance of this intersection with respect to the sensor, the corresponding object distance may be calculated by means of the thin lens equation. The obtained object distances have been utilised as ground truth references in experimental work to verify correct calibration methods. It is noteworthy that both studies were based on a slightly different optical setup that will be discussed in the following section.

2.2.3 Focused Plenoptic Camera

In 2009, Lumsdaine and Georgiev [54] substantially improved the effective spatial resolution of refocused photographs by proposing a new rendering technique for a slightly different optical setup. Their so-called *Focused Plenoptic Camera* (FPC), also referred to as the *Plenoptic Camera 2.0*, allows the spacing between sensor and MLA to be larger than the micro lens focal length $(b_s > f_s)$. As can be seen in Fig. 2.14, considering the main lens to focus on an object at the Focal Plane (FP), a partial number of its object points (along the green bar) are projected through a single micro lens and covered by the

respective micro image. In contrast to the SPC where a single micro image corresponds to one sample in spatial domain, the FPC contains more than one spatial sample per micro image. Hence, the FPC trumps the SPC in terms of the spatial resolution. In return, the FPC inherently causes a loss of angular information resulting in a trade-off between angular and spatial resolution [36]. This may be easily understood by the fact that all samples within an SPC micro image represent the angular information whereas the micro image space in an FPC is partially occupied to sample more spatial points which inevitably diminishes the amount of angular samples and therefore the depth information.

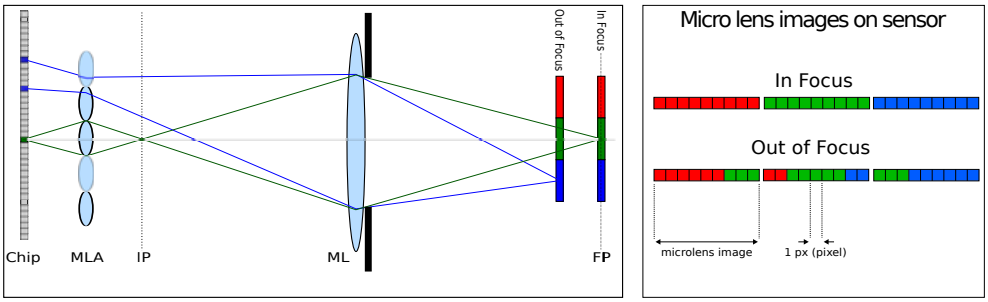


Figure 2.14: **Focused plenoptic camera ray model** (Ref. [79], Fig. 2). On the left, light rays emanating from object points at the Focal Plane (FP) are projected onto the Main Lens' (ML) Image Plane (IP) in front of the MLA and travel through a micro lens. Thereby, each micro image covers a portion of contiguous object points in FP as seen on the right. By contrast, the same portion of contiguous object points at a plane closer than FP is distributed in an interlaced manner among different micro images. Figure 2.15 illustrates the reconstruction process which brings an out-of-focus area into focus.

To render an image focused at a certain plane, the basic image process requires to crop micro images around their centre and place these patches aside to each other. This rendering procedure is well illustrated in Fig. 2.15. There, the upper left part depicts refocused images exhibiting block artefacts instead of a blur for out-of-focus objects. Similar to SPC refocusing, these artefacts can be turned into a natural blur by blending pixels which is basically an integration in the angular dimension. Thereby, all raw micro images are shifted with respect to each other creating a certain overlap. The intensity of overlapped areas is then added up and divided by the number of overlapping pixels to prevent clipping. This results in low image frequencies that are visually perceived as smooth transitions known as a blur. The literature rarely stresses that the effective resolution massively drops (by factor 2 per descending patch size per number of micro lenses per dimension) when computationally varying the depth focus in images taken by an FPC. A method for improving the effective resolution of refocused FPC captures

was initially conceived by Georgiev *et al.* [33] and advanced by Yu *et al.* [87] who deploy colour demosaicing after plenoptic rendering.

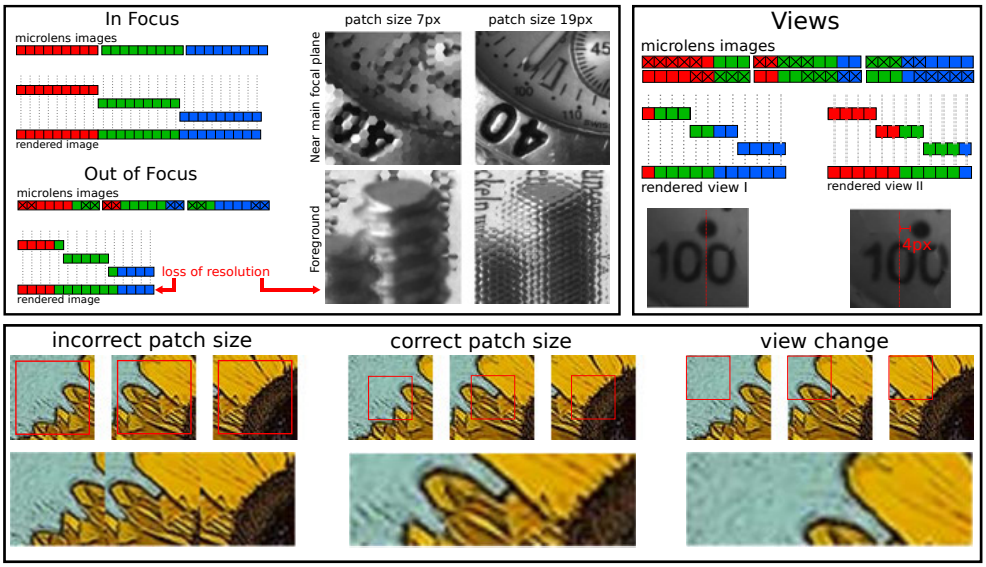


Figure 2.15: **Focused plenoptic camera syntheses** (Ref. [79], Fig. 4). As indicated in the upper left scheme, to reconstruct an image focusing on FP, largest possible patches of the micro images are simply strung together. Refocusing is realised by varying the patch size. This implies leaving out pixels in each micro image and hence results in a resolution loss. Whilst a smaller patch size brings close objects (e.g. the screw) into focus, objects placed at FP (e.g. the watch) exhibit severe artefacts and vice versa. Overlaying and integrating overlapped micro image areas resolves artefact blocks yielding a blur. A change of the perspective view is achieved by moving patches away from the centre as seen on the right.

In [35], Georgiev *et al.* identified the principal planes of plenoptic cameras in which the optical system is modeled as a composition of two thin lenses representing main lens and a single micro lens. For time being, however, this approach fails to provide the baseline which is a necessary parameter when displaying 3-D content on autostereoscopic monitors.

The currently most advanced type of an FPC is the *multi-focus camera* which was simultaneously published by Georgiev and Lumsdaine [34] and by Perwass and Wietzke [66] whereas the latter brought this camera type on the market and hold a corresponding patent [67]. By composing three varyingly specified micro lens types, the authors in [66] claim that the resolution of rendered photographs from their camera amounts to a quarter of the overall resolution. However, this statement merely refers to images computationally focused at the furthest distance. Since the image resolution shrinks when refocusing

closer, it is expected that the effective resolution of photographs focused at close objects is much less than that.

2.2.4 Coded-Aperture Camera

For the sake of completeness, this section concerns with *Dappled Photography* which deploys an optically deviant setup to capture 3-D information with a hand-held camera making it part of the plenoptic camera family.

Unlike MLA-based camera types, a *Coded-Aperture Camera* (CAC) gathers directional light ray information with the aid of a film mask. Thereby, high-frequency sinusoidal patterns are shaped into aperture discs that are placed between main lens and sensor of the camera. As opposed to MLAs, an advantage of the coded aperture is its cost-effective fabrication. Due to the non-refractive nature of the masks, 4-D images taken by the CAC do not feature the presence of any optical abberations. Moreover, similar to lens array based cameras, the design of the mask facilitates to select a desired trade-off between spatial and angular resolution in the CAC [77]. However, since the mask attenuates light, the downside is that about 50% of the entered light is blocked. Besides this, a coded mask lowers the aperture and thus causes severe diffraction making it impractical to implement this technology in microscopes.

In a recent study conducted by Xu *et al.* [84], a dual-mask CAC design composed of two consecutively placed masks has been examined. Therein, one mask replaces the aperture stop of the main lens whereas the other remains in the optical path of the camera. The authors conclude that 'fewer pixels are needed to achieve the same resolution as what one can achieve with a conventional lightfield camera' [84]. Although the authors claim to achieve a higher spatial resolution than previous attempts [61, 77], supportive data from a benchmark comparison with single-mask CACs is, however, not provided in a quantitative manner.

As indicated by the thesis title, the two latter plenoptic camera models fall outside the scope of this study.

CHAPTER 3

Standard Plenoptic Ray Model

While Chapter 2 gives insights into established methods describing the conceptual design of a plenoptic camera and existing image processing syntheses, the geometrical optics behaviour of the captured light field remained unclear. Knowledge about the light field geometry is needed in computational photography when there is a demand for the physical distance to which an image is refocused and capabilities to constrain the depth of field. Moreover, locations of virtual cameras providing perspective views have not been investigated in previous research. This chapter proposes a platform-independent model to initially solve these problems.

The first section introduces an innovative approach to geometrically specify a light field taken from an SPC by tracing rays through the optical system. It will be demonstrated how the devised model assists to derive refocusing and sub-aperture rendering equations. After this, mathematical solutions are elaborated to find the refocusing distance and depth as well as the position of virtual cameras. An evaluation of the hypotheses made in this chapter is presented separately in Chapter 5.

3.1 Derivation

A classical attempt in modelling plenoptic cameras is based on the method of similar triangles [35, 61, 66]. Contrary to the previous work, this section aims to model the SPC light field by tracing a sufficient number of distinct light rays. As a starting point, the

proposed approach deploys the well known thin lens equation which can be written as

$$\frac{1}{f_s} = \frac{1}{a_s} + \frac{1}{b_s}\,, \tag{3.1}$$

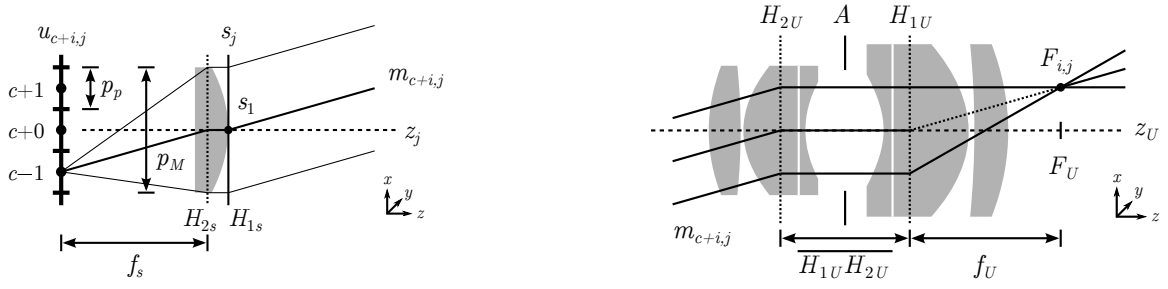
where f_s denotes the focal length, b_s the image distance and a_s the object distance in respect of a micro lens s. Since micro lenses are placed at a stationary distance f_s in front of the image sensor of an SPC, f_s always equals the micro lens image distance ($f_s = b_s$). Therefore, f_s may be substituted for b_s in Eq. (3.1) giving

$$\frac{1}{f_s} = \frac{1}{a_s} + \frac{1}{f_s}\,, \tag{3.2}$$

which yields

$$0 = \lim_{a_s \to \infty} \left(\frac{1}{a_s} \right) \tag{3.3}$$

after subtracting the term $1/f_s$. The interesting observation in Eq. (3.3) is that $a_s \to \infty$ when $b_s = f_s$ which demonstrates that rays converging on a distance f_s behind the lens have emanated from a point at an infinitely far distance a_s. As a consequence, rays coming from infinity travel parallel to each other in front of the lens which is known as the effect of collimation. To support this, it is assumed that image spots focusing at a distance f_s are infinitesimally small. In addition, we regard micro image sampling positions u to be discrete from which light rays are traced back through lens components. Figure 3.1 shows collimated light rays entering a micro lens and leaving main lens elements.



(a) Micro lens (b) Main lens

Figure 3.1: **Lens components of a plenoptic camera.** Cross-sectional views of (a) micro lens s_j and chief ray $m_{c+i,j}$; (b) collimated rays travelling through an objective lens. Note that H_{1U} and H_{2U} may be interchanged and thus $\overline{H_{1U}H_{2U}}$ can be negative.

At the micro image plane, a *Micro Image Centre* (MIC) operates as a reference point $c = {}^{(M\,-\,1)}/_{2}$ where M denotes the one-dimensional (1-D) micro image resolution, which

is seen to be consistent. Horizontal micro image samples are then indexed by $c + i$, where $i \in [-c, c]$. Horizontal micro image positions are given as $u_{c+i,j}$, where j denotes the 1-D index of the respective micro lens s_j. In the case of a digital image sensor, the space that separates two adjacent positions $u_{c+i,j}$ represents the pixel pitch p_p. Physical pixels, however, are not placed directly next to each other, but instead some space is left between them to accommodate the electrical circuitry. Each pixel of a traditional sensor, therefore, has a tiny lens attached with a diameter of p_p to focus incident light on the actual light-sensitive pixel area [31]. This keeps the proximity of p_p whilst increasing the measured light intensity and thus enhancing noise suppression. These lenses should not be confused with those of the MLA.

In contrast, a plenoptic micro lens illuminates several pixels $u_{c+i,j}$ and requires its lens pitch, denoted as p_M, to be greater than p_p. Each chief ray arriving at any $u_{c+i,j}$ exhibits a specific slope $m_{c+i,j}$. For example, object side rays of a micro lens which focus at $u_{c-1,j}$ have a slope $m_{-1,j}$ in common. Hence, all rays $m_{-1,j}$ form a collimated light beam in front of the micro lenses. As seen in Fig. 3.1(b), due to the collimating effect of light, parallel rays of a light beam $m_{-1,j}$ are refracted at the system's principal planes H_{1U} and H_{2U} in a manner that they originated from a point along the main lens focal plane F_U. The position of this respective point, denoted as $F_{i,j}$, varies depending on the chief ray slope $m_{c+i,j}$. The depth is generally given by z and more specifically by z_U with respect to the main lens optical axis and z_j for each micro lens s_j. Figure 3.2 illustrates a model which combines micro and main lenses. Therein, chief rays have been highlighted in colours to emphasise the angular information.

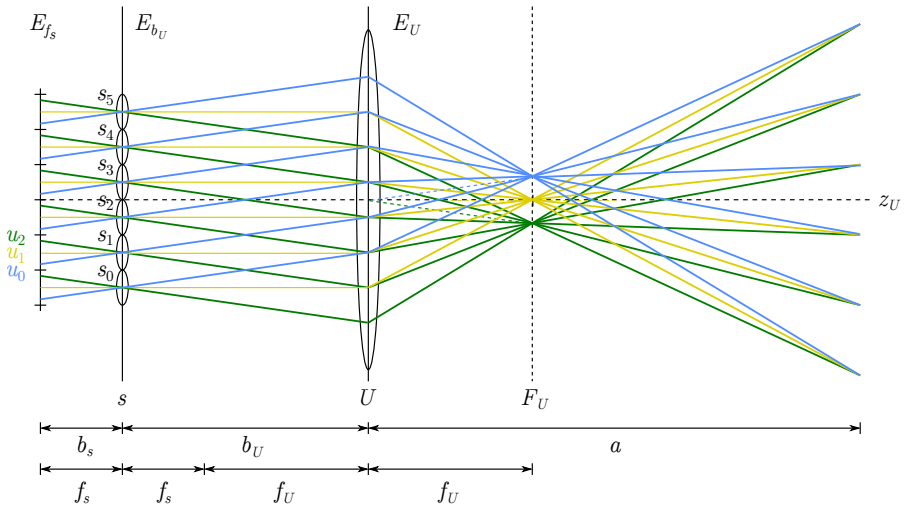


Figure 3.2: **Ideal SPC ray model.** For simplicity, the main lens U is depicted as a thin lens with a fully open aperture. The micro image size amounts to $M = 3$.

In the ideal model in Fig. 3.2, it is assumed that each MIC lies on the optical axis of its corresponding micro lens. In Dansereau's doctoral dissertation [21] it is mentioned that this hypothesis would only be true where the main lens is at an infinite distance from the MLA. Because of the finite separation distance between the main lens and the MLA, the centres of micro images deviate from their micro lens optical axes.

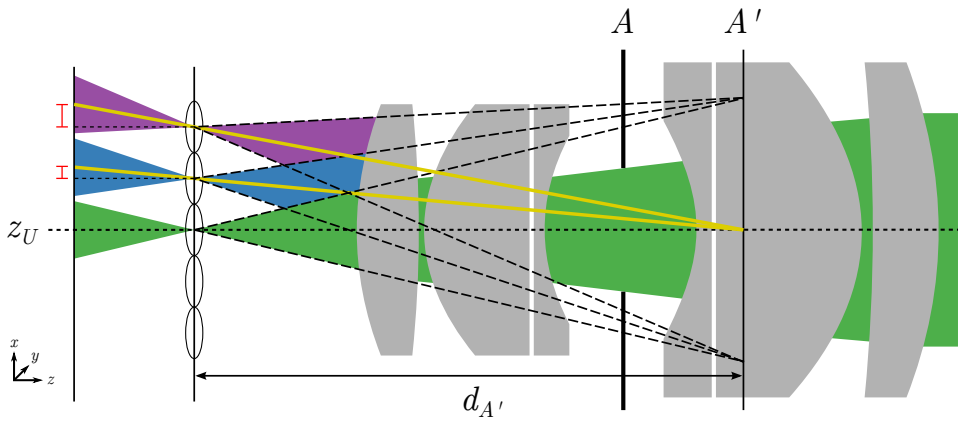


Figure 3.3: **MIC chief ray trajectories** where red coloured crossbars signify gaps between MICs and respective micro lens optical axes. Rays arriving at MICs arise from the exit pupil's centre A'. It can be seen that crossbars grow towards image edges.

According to Dansereau's suggestions presented in Chapter 2, more accurate MICs would be obtained by tracing chief rays through the optical centres of main lens and micro lens. The intersection of these rays with the sensor plane yield new MICs that deviate from those in the ideal model. First experimental examination justifying this theory has been published by Dansereau *et al.* [22]. As an optical centre is only present when simplifying the main lens to be a thin lens, Dansereau's proposition lacks of indicating the exact origin of MIC chief rays in real objective lenses, which consist of more lens elements as well as an aperture. An extension of Dansereau's assumption is proposed in Fig. 3.3, where the centre of the aperture's exit pupil A' is seen to be the MIC chief rays' origin. The exit pupil is a geometrical representation of an objective's aperture with respect to the image side [42]. Experimental work attempting to verify whether chief rays that lead to MICs originate from the centre of the exit pupil A', aperture A or principal plane H_{2U} is conducted in Chapter 5.

It is of particular importance to detect MICs correctly since they are taken as reference origins in the image synthesis process, which is discussed in Chapter 4 in more detail. Hence, the model posed in Fig. 3.2 does not take a displacement in MIC positions into account. Therefore, it is the aim to refine the previous model considering actual MICs.

Contrary to the previous approach, all chief rays impinging on the MIC positions originate from the exit pupil centre which, for simplicity, coincides with the main lens optical centre in Fig. 3.4. All chief ray positions that are adjacent to MICs can be ascertained by a corresponding multiple of the pixel pitch p_p.

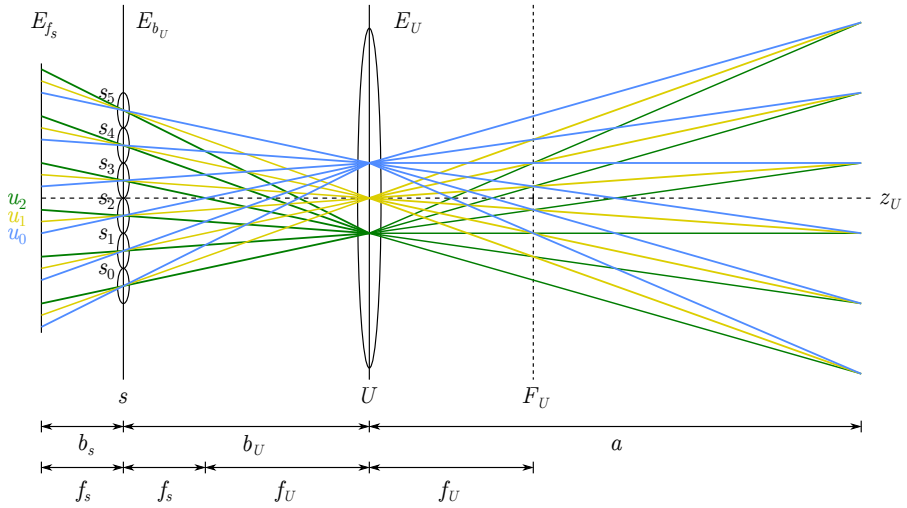


Figure 3.4: **Realistic SPC ray model.** The refined model considers more accurate MICs obtained from chief rays crossing the micro lens optical centres and the exit pupil centre of the main lens (yellow coloured rays). For convenience, the main lens is depicted as a thin lens where aperture pupils and principal planes coincide.

A more in-depth analysis showing the influence of principal planes as well as aperture pupils on the chief rays' path is presented later in this chapter. Subsequent sections discuss practical applications and potential benefits using the contrived optical model.

3.2 Refocusing

The first systematic study revealing a refocused image synthesis on the basis of a plenoptic camera was reported by Ng *et al.* [61] in 2005. Therein, it has been pointed out that the irradiance I_{b_U} at a film plane (s,t) of a conventional camera is obtained by

$$I_{b_U}(s,t) = \frac{1}{{b_U}^2} \int\int L_{b_U}(s,t,U,V)\, A(U,V)\, \cos^4\theta \,\mathrm{d}U\,\mathrm{d}V \tag{3.4}$$

where $A(\cdot)$ denotes the aperture, (U,V) the main lens plane coordinate space and b_U the separation between the main lens and the film plane (s,t). The factor $1/{b_U}^2$ is often referred to as the inverse-square law [42]. If θ is the incident ray angle, the roll-off factor $\cos^4\theta$ describes the gradual decline in irradiance from object points at an

oblique angle impinging on the film plane, also known as *natural vignetting* [42]. It is implied that coordinates (s, t) represent the spatial domain in horizontal and vertical dimensions whilst (U, V) denote the angular light field domain. To simplify Eq. (3.4), a horizontal cross-section of the light field is regarded hereafter so that $L_{b_U}(s, t, U, V)$ becomes $L_{b_U}(s, U)$. Thereby, subsequent declarations build on the assumption that camera parameters are equally specified in horizontal and vertical dimensions allowing propositions to be applied to both dimensions in the same manner. Since the overall measured irradiance I_{b_U} is scalable (e.g. on electronic devices) without affecting the light field information, the inverse-square factor $1/{b_U}^2$ may be omitted at this stage. On the supposition that the main lens aperture is seen to be completely open, the aperture term becomes $A(\cdot) = 1$. To further simplify, $\cos^4\theta$ will be neglected given that pictures do not expose *natural vignetting*. Provided these assumptions, Eq. (3.4) can be shortened yielding

$$I_{b_U}(s) = \int L_{b_U}(s, U)\,\mathrm{d}U\,. \tag{3.5}$$

Let us suppose that the irradiance $I_{b_U}(s)$, and therefore the information of the entire light field L_{b_U}, is located at plane U in the form of $I_U(s, U)$ since all rays of a potentially captured light field travel through U. From this it follows that

$$L_{b_U}(s, U) = I_U(s, U)\,. \tag{3.6}$$

as it preserves a distinction between spatial and angular information. Figure 3.5 visualises the irradiance planes, whilst introducing a new physical sensor plane $I_{f_s}(s, u)$ located one focal length f_s behind I_{b_U}, with u as a horizontal and v as a vertical angular sampling domain in the 2-D case. The former spatial image plane $I_{b_U}(s, t)$ is now replaced by an MLA, enabling light to pass through and strike the new sensor plane $I_{f_s}(s, u)$. When applying the method of similar triangles to Fig. 3.5, it becomes apparent that $I_U(s, U)$ is directly proportional to $I_{f_s}(s, u)$, which gives

$$I_U(s, U) \propto I_{f_s}(s, u) \tag{3.7}$$

where $\propto$ designates the equality up to scale. When ignoring the scale factor in Eq. (3.7), which simply lowers the overall irradiance, $I_{f_s}(s, u)$ and $I_U(s, U)$ become equal. From this it follows that Eq. (3.5) can be written as

$$I_{b_U}(s) = \int I_{f_s}(s, u)\,\mathrm{d}u\,. \tag{3.8}$$

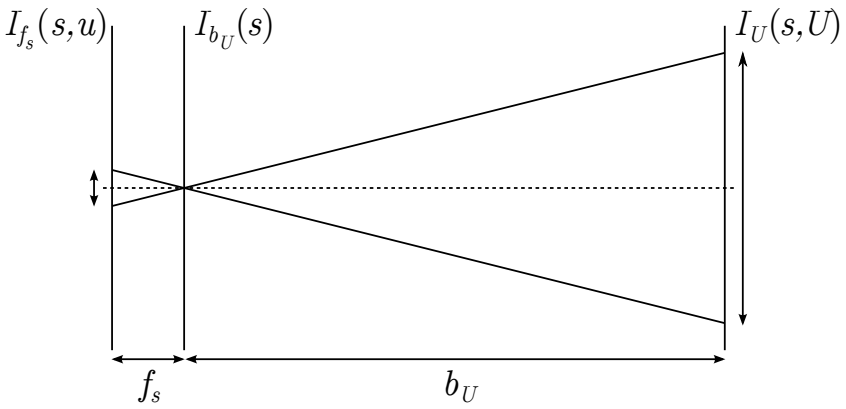


Figure 3.5: **Irradiance planes**. If light rays emanate from an arbitrary point in object space, the measured energy $I_U(s,U)$ at the main lens' aperture is seen to be concentrated on a focused point $I_{b_U}(s)$ at the MLA and distributed over the sensor area $I_{f_s}(s,u)$. Neglecting the presence of light absorptions and reflections, $I_{f_s}(s,u)$ is proportional to $I_U(s,U)$, which may be proven by comparing similar triangles.

Due to the human visual perception, photosensitive sensors limit the irradiance signal spectrum to the visible wavelength range. For this purpose, bandpass filtersare placed in the optical path of present-day cameras which prevents infrared and ultraviolet radiation from being captured. The resulting spectral sensitivity suffices to represent the luminosity function (also called luminous efficacy) [11, 70]. Therefore, Eq. (3.8) will be rewritten as

$$E_{b_U}(s) = \int E_{f_s}(s,u)\,\mathrm{d}u \tag{3.9}$$

in order that photometric illuminances E_{b_U} and E_{f_s} substitute irradiances I_{b_U} and I_{f_s} in accordance with the luminosity function. Besides, it is assumed that $E_{f_s}(s,u)$ is a monochromatic signal being represented as a grey scale image. Recalling index notations of the derived model, a discrete equivalent of Eq. (3.9) may be given by

$$E_{b_U}\left[s_j\right] = \sum_{i=-c}^{c} E_{f_s}\left[s_j\,,\,u_{c+i}\right] \tag{3.10}$$

provided that the sample width p_p is neglected here as it simply scales the overall illuminance $E_{b_U}\left[s_j\right]$ whilst preserving relative brightness levels. It is further implied that indices in the vertical angular domain v_{c+g} and vertical spatial domain t_h are constant meaning that only a single horizontal row of sampled s_j and u_{c+i} is regarded in the following. Nonetheless, subsequent formulas can be identically applied in the vertical direction under the assumption that indices are interchangeable and thus of the same size. Note that suggested derivations ignore several phenomena including light absorption

and reflection just as all kinds of optical aberrations. As these do not contribute to the light field, Eq. (3.10) serves as a basis for refocusing syntheses in spatial domain.

Invoking the *Lambertian* reflectance, an object point scatters light in all directions uniformly, meaning that each ray coming from that point carries the same energy. With this, an object placed at plane $a = 0$ reflects light with a luminous emittance M_a. An example which highlights the rays' path starting from a spatial point s' at object plane M_0 is shown in Fig. 3.6.

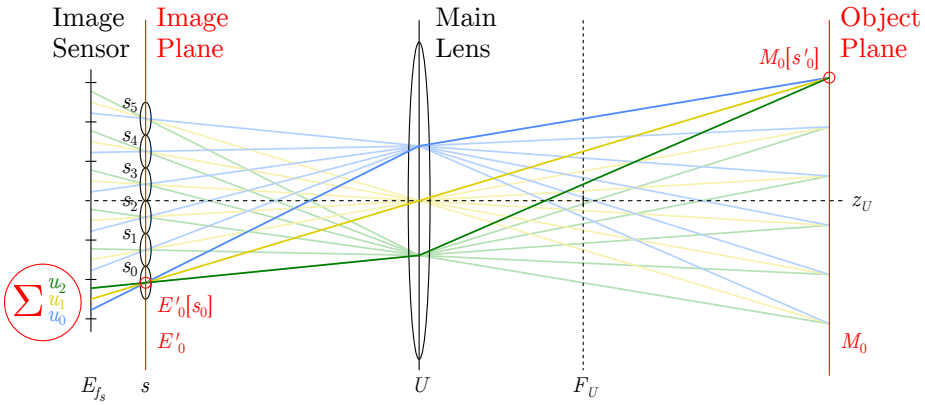


Figure 3.6: **Refocusing from raw capture where $a = 0$.**

Closer inspection of Fig. 3.6 reveals that the luminous emittance M_0 at a discrete point s'_0 may be seen as projected onto a micro lens s_0 and scattered across micro image pixels u. In the absence of reflection and absorption at the lens material, a synthesised image $E'_a\left[s_j\right]$ at the MLA plane ($a = 0$) is recovered by integrating all illuminance values u_{c+i} for each s_j. Taking $E'_0\left[s_0\right]$ as an example, this is mathematically given by

$$E'_0\left[s_0\right] = E_{f_s}\left[s_0\,,\,u_0\right] + E_{f_s}\left[s_0\,,\,u_1\right] + E_{f_s}\left[s_0\,,\,u_2\right]\,. \tag{3.11}$$

Similarly, an adjacent spatial point s_1 in E'_0 can be retrieved by

$$E'_0\left[s_1\right] = E_{f_s}\left[s_1\,,\,u_0\right] + E_{f_s}\left[s_1\,,\,u_1\right] + E_{f_s}\left[s_1\,,\,u_2\right]\,. \tag{3.12}$$

Developing this concept further makes it obvious that

$$E'_0\left[s_j\right] = \sum_{i=-c}^{c} E_{f_s}\left[s_j\,,\,u_{c+i}\right] \tag{3.13}$$

reconstructs an image $E'_0\left[s_j\right]$ as it appeared on the MLA by summing up all pixels within each micro image to form a respective spatial point of that particular plane. As

claimed, refocusing allows more than only one focused image plane to be recovered. Figure 3.7 depicts rays emitted from an object point located closer to the camera device ($a = 1$).

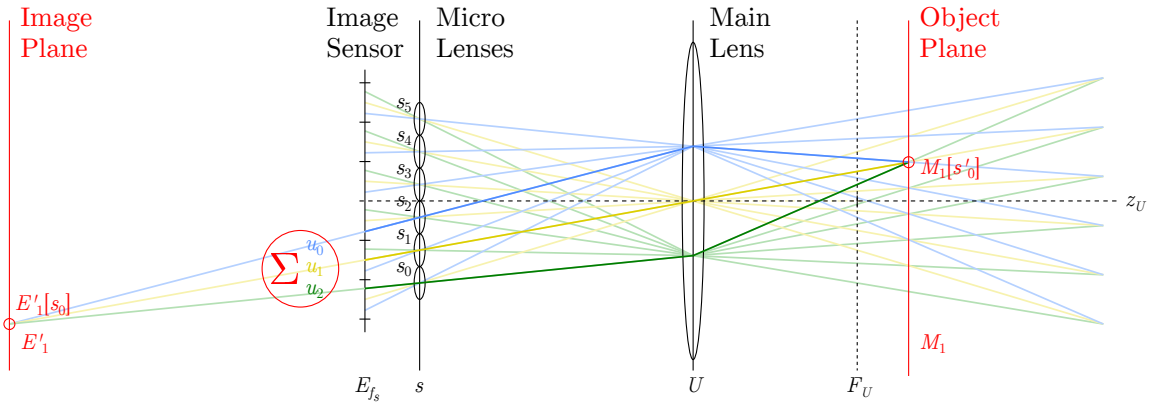


Figure 3.7: **Refocusing from raw capture where $a = 1$.**

For comprehensibility, light rays have been extended on the image side in Fig. 3.7 yielding an intersection at a distance where the corresponding image point would have focused without the MLA and image sensor. The presence of both, however, enables the illuminance of an image point to be retrieved as it would have appeared with a conventional sensor at E'_1. Further analysis of light rays in Fig. 3.7 unveils coordinate pairs $\left[s_j\,,\,u_{c+i}\right]$ that have to be considered in an integration process synthesising E'_1. Accordingly, the illuminance E'_1 at point s_0 can be obtained as follows

$$E'_1\left[s_0\right] = E_{f_s}\left[s_2\,,\,u_0\right] + E_{f_s}\left[s_1\,,\,u_1\right] + E_{f_s}\left[s_0\,,\,u_2\right]\,. \tag{3.14}$$

The adjacent image point s_1 is formed by calculating

$$E'_1\left[s_1\right] = E_{f_s}\left[s_3\,,\,u_0\right] + E_{f_s}\left[s_2\,,\,u_1\right] + E_{f_s}\left[s_1\,,\,u_2\right]\,. \tag{3.15}$$

The observation that the index j has simply been incremented by 1 from Eq. (3.14) to Eq. (3.15) allows conclusions to be drawn about the final refocusing synthesis equation which reads

$$E'_a\left[s_j\right] = \sum_{i=-c}^{c} E_{f_s}\left[s_{j+a(c-i)}\,,\,u_{c+i}\right] \tag{3.16}$$

and satisfies any plane a to be recovered. In Eq. (3.16) it is supposed that synthesised intensities $E'_a\left[s_j\right]$ ignore *clipping* which occurs when quantised values exceed the maximum amplitude of the given bit depth range. Thus, Eq. (3.16) only applies to

underexposed plenoptic camera images on condition that peaks in $E'_a\left[s_j\right]$ do not surpass the quantisation limit. To prevent *clipping* during the refocusing process, one can simply average intensities E_{f_s} prior to summing them up as provided by

$$E'_a\left[s_j\right] = \sum_{i=-c}^{c} \frac{1}{M} E_{f_s}\left[s_{j+a(c-i)}\,,\, u_{c+i}\right] \quad , \quad a \in \mathbb{Q} \tag{3.17}$$

which, on the downside, requires an additional computation step to perform the division. Letting $a \in \mathbb{Q}$ involves an interpolation of micro images which increases the spatial and angular resolution at the same time. In such a scenario, a denominator in a fraction number a represents the upsampling factor for the number of micro images.

3.3 Focus Range Estimation

In geometrical optics, light rays are viewed as straight lines with a certain angle in a given interval. These lines can be represented by linear functions of z possessing a slope m. By regarding the rays' emission as an intersection of ray functions, it may be viable to pinpoint their local origin. This position is seen to indicate the focusing distance of a synthetically focused photograph. In order for it to function, the proposed concept requires the geometry and thus the parameters of the camera system to be known. This section develops a theoretical approach based on the realistic SPC model to estimate the distance and *Depth of Field* (DoF) of objects that have been computationally brought into focus.

3.3.1 Refocusing Distance

First of all, it is necessary to define the optical centre of an SPC image by letting the micro lens index be $j = o$ where

$$o = \frac{J-1}{2}\,. \tag{3.18}$$

Here, J is the total number of micro lenses in the horizontal direction. Given the micro lens diameter p_M, the horizontal position of a micro lens' optical centre is given by

$$s_j = (j-o) \times p_M \tag{3.19}$$

where j is seen to start counting from 0 at the leftmost micro lens with respect to the main lens optical axis. An accurate method to retrieve a chief ray slope $m_{c,j}$ that impinges on an MIC would be to measure the centroid $\bar{c}$ of a respective 1-D micro image.

In Chapter 4 the centroid calculation is explained in more detail. Given a measured MIC position $u_{\bar{c}}$, a MIC chief ray slope $m_{\bar{c},j}$ of the j-th micro lens is obtained by

$$m_{\bar{c},j} = \frac{s_j - u_{\bar{c}}}{f_s} \,. \tag{3.20}$$

In case, however, a calibration capture is not present, e.g. in a plenoptic camera design process, the slope may be obtained by means of the proposed ray model. As rays impinging on MICs are seen to connect an optical centre of a micro lens s_j and the exit pupil A', their respective slope $m_{c,j}$ may be given by

$$m_{c,j} = -\frac{s_j}{d_{A'}} \tag{3.21}$$

where $d_{A'}$ denotes the separation between exit pupil plane and the MLA's front vertex. An investigation whether $d_{A'}$ or b_U, the separation between s and the principal plane H_{2U}, acts as the origin distance is exposed in Chapter 5. Provided the MIC chief ray slope $m_{c,j}$, an MIC position $u_{c,j}$ is estimated by extending $m_{c,j}$ until it intersects the sensor plane which is calculated by

$$u_{c,j} = -m_{c,j} \times f_s + s_j \,. \tag{3.22}$$

Central positions of adjacent pixels $u_{c+i,j}$ are given by the number of pixels i separating $u_{c+i,j}$ from the centre $u_{c,j}$. To calculate $u_{c+i,j}$, we simply compute

$$u_{c+i,j} = u_{c,j} + i \times p_p \tag{3.23}$$

which requires the pixel width p_p. The slope $m_{c+i,j}$ of a ray that crosses a micro image at $u_{c+i,j}$ is obtained by

$$m_{c+i,j} = \frac{s_j - u_{c+i,j}}{f_s} \,. \tag{3.24}$$

The main lens deflects light rays before they enter the MLA and thus changes their slope in object and image space, respectively. If modelled as a thin lens, the intersection of rays with U can be received by

$$U_{i,j} = m_{c+i,j} \times b_U + s_j \,. \tag{3.25}$$

where c has been left out in the subscript of $U_{i,j}$ as it is a constant and will be omitted in following ray functions for simplicity. Due to the effect of collimation, each ray landing on an infinitesimally small spot beneath a micro lens is seen to yield a beam of parallel

light rays in the range between s and U. Such a collimated light beam passed through a point $F_{i,j}$ along the main lens focal plane F which can be found by use of the beam's slope $m_{c+i,j}$. With the main lens focal length f_U, the calculation of $F_{i,j}$ at which a beam converges writes as follows

$$F_{i,j} = m_{c+i,j} \times f_U \, . \tag{3.26}$$

Consequentially, a chief ray slope $q_{i,j}$ of that beam in object space is given by

$$q_{i,j} = \frac{F_{i,j} - U_{i,j}}{f_U} \tag{3.27}$$

as it depends on the intersections at U, F_U and the ray's travelling distance which is f_U in this particular case. With reference to preliminary remarks, a ray's path may be provided as a linear function $\widehat{f}_{i,j}$ of the depth z which is written as

$$\widehat{f}_{i,j}(z) = q_{i,j} \times z + U_{i,j} \ , \quad z \in [U, \infty) \ . \tag{3.28}$$

At this stage, it may be worth discussing the selection of appropriate rays for the intersection. A set of two chief ray functions meets the requirements to locate an object plane a because all adjacent ray intersections lie on the same planar surface parallel to the sensor. It is of key importance, however, to select a chief ray pair that intersects at a desired plane a. In respect of the refocusing synthesis in Eq. (3.17), a system of linear ray functions is found by letting the index subscript in $\widehat{f}_{i,j}(z)$ be $\mathbf{A} = \{i, j\} = \{-c, e\}$ for the first chief ray where e is an arbitrary, but valid micro lens s_e and $\mathbf{B} = \{i, j\} = \{c, e - a(M-1)\}$ for the second ray. Given the synthesis example depicted in Fig. 3.7, parameters would be $e = 2$, $a = 1$, $M = 3$, $c = 1$ such that corresponding ray functions are $\widehat{f}_{-1,2}(z)$ for $E_{f_s}[u_0, s_2]$ and $\widehat{f}_{1,0}(z)$ for $E_{f_s}[u_2, s_0]$. After choosing appropriate chief ray functions, a distance to their intersection is obtained by setting them equal to each other, so that

$$\widehat{f}_{\mathbf{A}}(z) = \widehat{f}_{\mathbf{B}}(z) \ , \quad z \in [U, \infty) \tag{3.29}$$

and solving the equation system to get a result which represents a new refocused object distance ${a_U}'$. It is important to note that ${a_U}'$ only indicates the distance from U to the respective intersection at plane a. The overall distance d_a from the ray intersection in object space to the MLA front vertex plane requires to add all optical distances that separate them from each other. This is accomplished by

$$d_a = b_U + \overline{H_{1U} H_{2U}} + {a_U}' \, . \tag{3.30}$$

where $\overline{H_{1U}H_{2U}}$ is the distance between principal planes (see Fig. 3.1(b)).

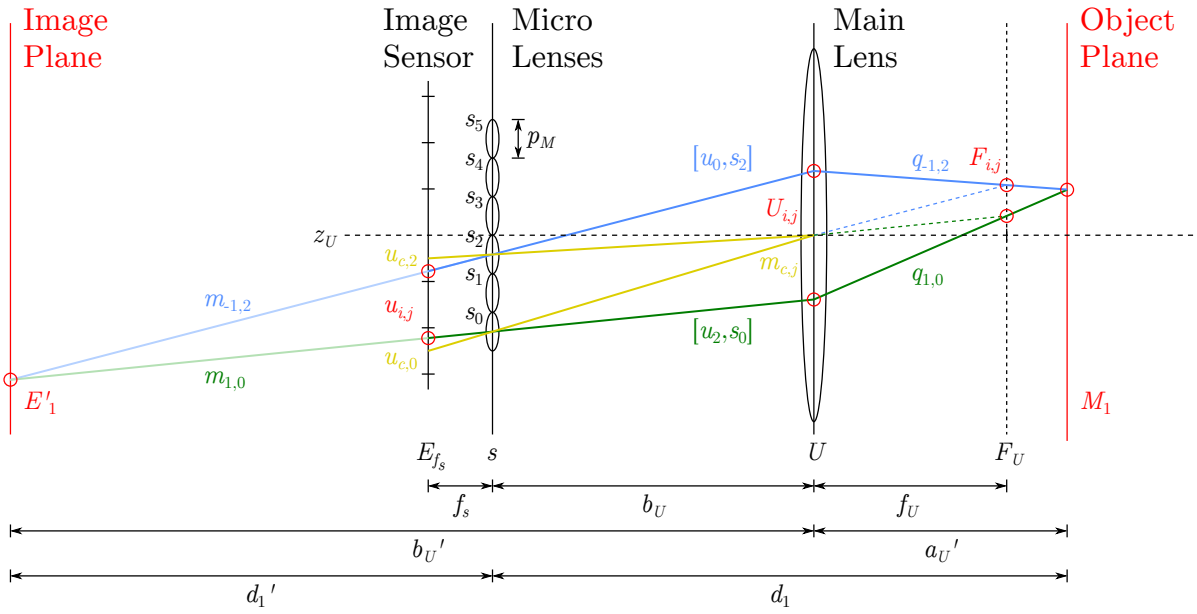


Figure 3.8: **Refocusing distance estimation.** Taking the example from Fig. 3.7, the diagram illustrates parameters that help to find the distance at which refocused photographs exhibit best focus. The proposed model offers two ways to accomplish this by regarding rays as intersecting linear functions in object and image space. Note that $d_a{}'$ is negative if the intersection occurs behind the MLA.

The suggested ray model implementation merely builds on basic mathematics making the computation time negligibly quick as even hand-held calculators could provide such results instantly. The number of provided equations, however, may be shortened to simplify the implementation process. A shortcut may be achieved by locating the intersection of rays involved in the integration process behind or in front of the MLA and employing the thin lens equation for the main lens to project a refocused image point back onto the corresponding object distance. To start off, rays intersecting at the potential image point are again given as a linear function $\mathring{f}_{i,j}(z)$ such that

$$\mathring{f}_{i,j}(z) = m_{c+i,j} \times z + s_j \ , \quad z \in (-\infty \, , \, U \,] \ . \tag{3.31}$$

Similar to an intersection in object world, the distance $d_a{}'$ from MLA front vertex to the thought image plane is recovered by solving a system of two linear functions

$$\mathring{f}_{\mathbf{A}}(z) = \mathring{f}_{\mathbf{B}}(z) \, , \quad z \in (-\infty \, , \, U \,] \ , \tag{3.32}$$

where $\mathbf{A}$ and $\mathbf{B}$ denote the same intersecting chief ray pair as previously used for the object-space-based approach and the zero point in z refers to the MLA position. In order

to project the image point to the real world, the new overall image distance ${b_U}'$ needs to be found by

$$ {b_U}' = b_U - {d_a}' \,. \tag{3.33} $$

With the new image distance ${b_U}'$ it is possible to calculate the refocused object distance of that point by means of the thin lens equation which then reads

$$ {a_U}' = \left(\frac{1}{f_U} - \frac{1}{{b_U}'} \right)^{-1} \,. \tag{3.34} $$

In the end, the alternatively obtained ${a_U}'$ also requires to be added up to all distances of the optical path as shown with Eq. (3.30) to retrieve the refocusing distance d_a. It is expected that the results from object and image side approach are identical. However, the advantage of the latter is that it obviously reduces the number of steps whilst yielding the same result as Eq. (3.29). This makes up a neat solution which may be preferred over the more laborious attempt to trace rays through all elements.

The preceding developments of a refocusing distance estimator neglect the fact that projection lenses produce image spots of a finite size. Although points $u_{c+i,j}$ are seen to be infinitesimally small, the proposed system succeeds to calculate the distance of best focus as $u_{c+i,j}$ per definition denote centres of image points exhibiting some spatial width. The subsequent section discusses the effect of finite sampling on the refocusing performance.

3.3.2 Depth of Field

A focused image spot of a finite size, by implication, causes the focused depth range in object space to be finite as well. In conventional photography, this range is abbreviated to DoF (*Depth of Field*). Optical phenomena such as aberrations or diffraction are known to limit the spatial extent of projected image points. However, most kinds of lens aberrations can be nominally eliminated through optical lens design (e.g. aspherical lenses, glasses of different dispersion). In that case, the circle of least confusion solely depends on diffraction making an imaging system called *diffraction-limited*. Thereby, light waves that encounter a pinhole, aperture or slit of a size comparable to the wavelength λ propagate in all directions and interfere at an image plane inducing wave superposition due to the ray's varying path length and corresponding difference in phase. A diffracted image point is made up of a central disc possessing the major energy surrounded by rings with alternating intensity. This is often referred to as *Airy* pattern named after George Biddell Airy. According to Hecht [42], the radius r_A of an *Airy* pattern's central peak disc is approximately given by

$$r_A \approx 1.22 \frac{f\lambda}{A} . \tag{3.35}$$

To assess the optical resolution limit of a lens, it is straightforward and sufficient to refer to the *Rayleigh* criterion. The *Rayleigh* criterion states that two image points of equal irradiance in the form of an *Airy* pattern need to be separated by a minimum distance $(\Delta\,\ell)_{\min} = r_A$ to be visually distinguishable. Let us suppose a *non-diffraction-limited* camera system in which the pixel pitch p_p is larger than or equal to $(\Delta\,\ell)_{\min}$ at the smallest aperture diameter A. In this case, the DoF merely depends on the pixel pitch p_p. To distinguish between different pixel positions, we define three types of rays that are class-divided into:

- *central rays* at pixel centres $u_{c+i,j}$
- *inner rays* at pixel borders $u_{\{c+i,j\}-}$ towards the MIC
- *outer rays* at pixel borders $u_{\{c+i,j\}+}$ closer to the micro image edge

For conciseness, the image-side based intersection method nearby the MLA is applied hereafter. Nonetheless, it is feasible to derive DoF distances from an intersection in object space on the basis of Eqs. (3.19) to (3.30).

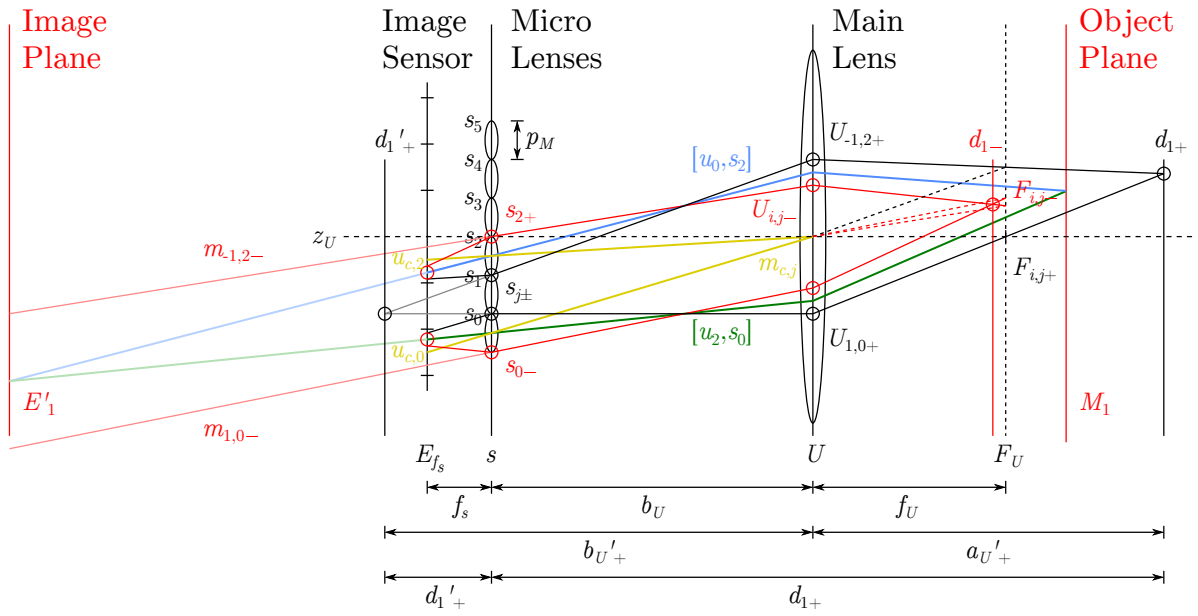


Figure 3.9: **Refocusing DoF estimation** where $a = 1$. DoF border d_{1-} cannot be attained via image-based intersection as *inner rays* do not converge on the image side which is a consequence of $a_U'_- < f_U$. Distances $d_{a'_\pm}$ are negative in case they are located behind the MLA and positive otherwise.

Similar to the acquisition of *central ray* positions $u_{c+i,j}$ in Section 3.3.1, pixel border positions $u_{\{c+i,j\}\pm}$ may be obtained as follows

$$u_{\{c+i,j\}\pm} = u_{c,j} + i \times p_p \pm \frac{p_p}{2} \,. \tag{3.36}$$

where $u_{c,j}$ is taken from Eq. (3.22). Given $u_{\{c+i,j\}\pm}$ as spatial points at pixel borders, chief ray slopes $m_{\{c+i,j\}\pm}$ starting from these respective locations are given by

$$m_{\{c+i,j\}\pm} = \frac{s_j - u_{\{c+i,j\}\pm}}{f_s} \,. \tag{3.37}$$

Since border points are assumed to be infinitely small and positioned at the distance of one micro lens focal length, light rays ending up at $u_{\{c+i,j\}\pm}$ form collimated beams between s and U propagating with respective slopes $m_{\{c+i,j\}\pm}$ in that particular interval. The range that spans from the furthest to closest intersection of these beams defines the DoF. Closer inspection of Fig. 3.9 reveals that the nearer intersection in object space is made by *inner rays* which, however, pass through the external micro lens edges. This similarly applies to *outer rays*, whose intersection provides the furthest DoF boundary whereas *outer rays* cross internal micro lens edges. Therefore, it is of importance to

determine micro lens edges $s_{j\pm}$ which is accomplished by

$$s_{j\pm} = s_j \pm \frac{p_M}{2} \,. \tag{3.38}$$

Outer and *inner rays* converging on the image side are seen to disregard the refraction at micro lenses and continue their path with $m_{\{c+i,j\}\pm}$ from the micro lens edge as depicted in Fig. 3.9. Hence, a linear function representing a light ray at a pixel border is given by

$$\mathring{f}_{i,j\pm}(z) = m_{\{c+i,j\}\pm} \times z + s_{j\pm} \;, \quad z \in (-\infty \,,\, U\,] \;. \tag{3.39}$$

Image side intersections at d_{a-}' for nearby and d_{a+}' for far-away DoF borders are found where

$$\mathring{f}_{\mathbf{A}\pm}(z) = \mathring{f}_{\mathbf{B}\pm}(z) \;, \quad z \in [U, \infty) \;, \tag{3.40}$$

by recalling that $\mathbf{A} = \{i, j\}$, $\mathbf{B} = \{i, j\}$ and $\mathbf{A}\pm$, $\mathbf{B}\pm$ select a desired DoF ray pair $\mathbf{A}\pm = \{-c, e\}$, $\mathbf{B}\pm = \{c, e - a(M-1)\}$ as discussed in Section 3.3.1. We get new image distances ${b_U}_{\pm}'$ of the particular refocused DoF boundaries when calculating

$${b_U}_{\pm}' = b_U - d_{a\pm}' \,. \tag{3.41}$$

Related DoF object distances ${a_U}_{\pm}'$ are retrieved by deploying the thin lens equation such that

$${a_U}_{\pm}' = \left(\frac{1}{f_U} - \frac{1}{{b_U}_{\pm}'} \right)^{-1} \;. \tag{3.42}$$

With respect to the MLA location, the DoF boundary distances $d_{a\pm}$ can be acquired by summing up all parameters separating the MLA from the principal plane H_{1U} as demonstrated in

$$d_{a\pm} = b_U + \overline{H_{1U}H_{2U}} + {a_U}_{\pm}' \,. \tag{3.43}$$

where $\overline{H_{1U}H_{2U}}$ is the distance that separates principal planes from each other (see Fig. 3.1(b)). Finally, the difference of the near limit d_{a-} and far limit d_{a+} yield the DoF_a that reads

$$DoF_a = d_{a+} - d_{a-} \,. \tag{3.44}$$

The contrived model implies that the micro image size directly affects the refocusing and DoF performance. A reduction of M, for example via cropping each micro image,

causes depth aliasing due to downsampling in the angular domain (see Figure 6.12(d)). This consequently lowers the number of refocused image slices and increases their DoF. Upsampling M, in turn, raises the number of refocused photographs and shrinks the DoF per slice. An evaluation of these statements is carried out in Chapter 5 where results are presented.

3.4 Sub-Aperture Images

It has been shown in [4, 21, 59] that extracting viewpoints from an SPC can be attained by collecting all pixels sharing the same respective micro image position. In compliance with provided notations, a 1-D sub-aperture image $E_i\left[s_j\right]$ with index i is computed with

$$E_i\left[s_j\right] = E_{f_s}\left[s_j\,,\, u_{c+i}\right] \tag{3.45}$$

where u and c have been omitted in the subscript of E_i since i is a sufficient index for sub-aperture images in the 1-D row. Equation (3.45) implies that the effective resolution equals the number of micro lenses. Figure 3.10 depicts the reordering process producing 2-D sub-aperture images $E_{(i,g)}$ by means of index variables $\left[s_j\,,\, t_h\right]$ and $\left[u_{c+i}\,,\, v_{c+g}\right]$ for spatial and directional domain, respectively. As can be seen by colour-highlighted pixels, samples at a specific micro image position correspond to the respective viewpoint location in the camera array.

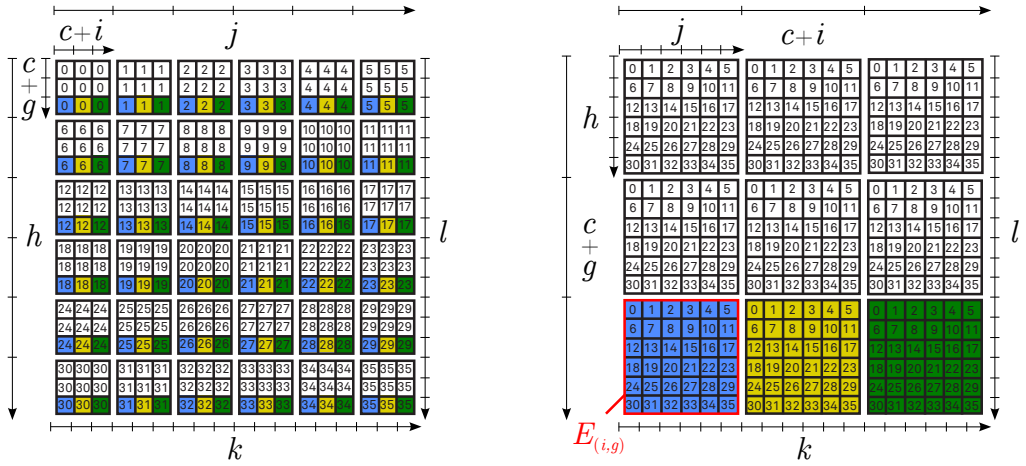


Figure 3.10: **Multiple sub-aperture image extraction.** (a) Calibrated raw image obtained by an SPC. Micro image samples are indexed by $\left[s_j\,,\, t_h\right]$ and pixels within micro images by $\left[u_{c+i}\,,\, v_{c+g}\right]$ where $M = 3$. (b) 2-D sub-aperture images $E_{(i,g)}$ after extraction where each colour represents a different perspective view. Coordinates $\left[u_{c+i}\,,\, v_{c+g}\right]$ index viewpoint images and $\left[s_j\,,\, t_h\right]$ their related spatial pixels.

Since raw SPC captures do not naturally feature the $E_{f_s}\left[s_j\,,\,u_{c+i}\right]$ index notation, it is convenient to define an index translation formula considering the light field photograph to be of two regular sensor dimensions $[x_k\,,\,y_l]$ as taken by a conventional sensor. Indices are converted by

$$k = j \times M + c + i \tag{3.46}$$

in the horizontal dimension meaning that $[x_k]$ is formed by

$$[x_k] = \left[x_{j\times M+c+i}\right] = \left[s_j\,,\,u_{c+i}\right]\,. \tag{3.47}$$

Similarly, the vertical index translation may be

$$l = h \times M + c + g \tag{3.48}$$

and therefore

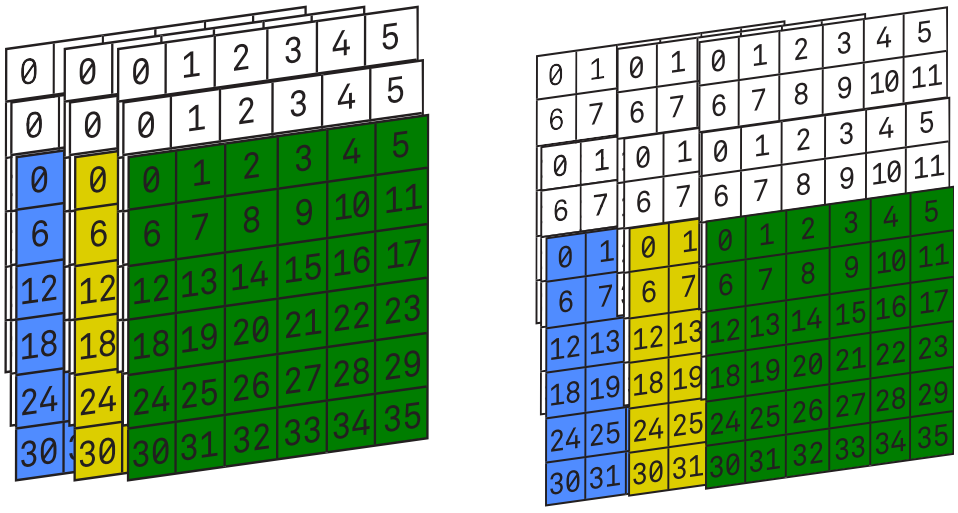
$$[y_l] = \left[y_{h\times M+c+g}\right] = [t_h\,,\,v_{c+g}]\,. \tag{3.49}$$

These definitions comply with Fig. 3.10 and enable our 4-D light field notation $\left[s_j\,,\;u_{c+i},\right.$ $\left.t_h\,,\,v_{c+g}\,\right]$ to be applied to conventionally 2-D sampled representations $[x_k\,,\,y_l]$ where k and l start to count from index 0.

According to Isaksen's discovery [45], the acquisition of a light field facilitates to digitally change the focus of a photograph by overlaying, shifting and blending pixels of *multi-view* images. Thus, previously extracted sub-aperture images may be utilised to refocus since they represent *multi-views* in the light field. Bearing the realistic SPC ray model in mind, Fig. 3.11 depicts the refocusing process starting from sub-aperture images. Mathematically, the 1-D refocusing process from sub-aperture images is given by

$$E'_a\left[s_j\right] = \sum_{i=-c}^{c} E_i\left[s_{j-a(c+i)}\right], \quad a \in \mathbb{Q} \tag{3.50}$$

which demonstrates that refocusing from sub-apertures also includes a summation of directional samples. In case refocusing is accomplished starting from a raw image taken by an SPC, the drawback of Eq. (3.50) is that it requires the sub-aperture extraction to be performed in advance. This results in two separate image rendering processes to digitally change the focus, namely Eqs. (3.45) and (3.50). As shown, the sub-aperture extraction is simply a rearrangement of pixel values, so that it may be implemented in the integration process to accommodate this in a single process. Equation (3.16) satisfies this as it includes Eq. (3.45) and Eq. (3.50) saving buffer memory and time.



$a = 1$ $a = 2$

Figure 3.11: **Refocusing scheme based on sub-aperture images.** To achieve refocusing from *multi-view* images, rendered sub-aperture images can be seen as layered and shifted slightly with respect to each other whilst overlaid pixels are integrated. The shift parameter a affects the refocusing distance. Sub-aperture images are highlighted with colours consistent with those of the ray model in Fig. 3.4.

3.5 Virtual Camera Array

In the previous section, it was shown how to render *multi-views* from SPC photographs by means of the proposed ray model. Because a 4-D plenoptic camera image can be reorganised to a set of *multi-view* images as if taken by an array of cameras, it is supposed that each of these images possesses a virtual camera with a distinct location. The localisation of a virtual camera's optical centre is, however, not obvious. This problem was first recognised and addressed in a publication by the author of this thesis [39]. Once theoretical positions of virtual cameras are derived, the goal of this study is to examine in which way the well established concept of stereo triangulation (see Section 2.1.1) applies to the proposed SPC ray model. As the name suggests, sub-aperture images are created at the main lens' aperture. Thus, an obvious attempt would be to locate a baseline $B_{A'}$ at the exit pupil. A baseline spanning from a virtual camera at the exit pupil centre to any other virtual camera A'_i is found by

$$B_{A'} = m_{c+i,j} \times d_{A'} \,, \tag{3.51}$$

where $m_{c+i,j}$ is obtained from Eq. (3.24). Practical applications of an image-side baseline $B_{A'}$ are unclear at this stage. As it is the aim to deploy the triangulation in object space, a much more valuable plane may be the entrance pupil denoted as A''. The entrance pupil is a representative of an objective's aperture with regards to the object side [42]. To investigate rays' positions at the entrance pupil, it is worth introducing the aperture's geometrical equivalents to the proposed model, which have not been considered in [39]. Figure 3.12 offers a closer look at the rays behaviour by also showing principal planes H_{1U} and H_{2U}. It can be seen in Fig. 3.12 that all rays having i in common (e.g. blue rays) geometrically converge to the entrance pupil and diverge from the exit pupil. Intersecting chief rays at the entrance pupil can be seen as indicating object-side-related positions of virtual cameras A''_i.

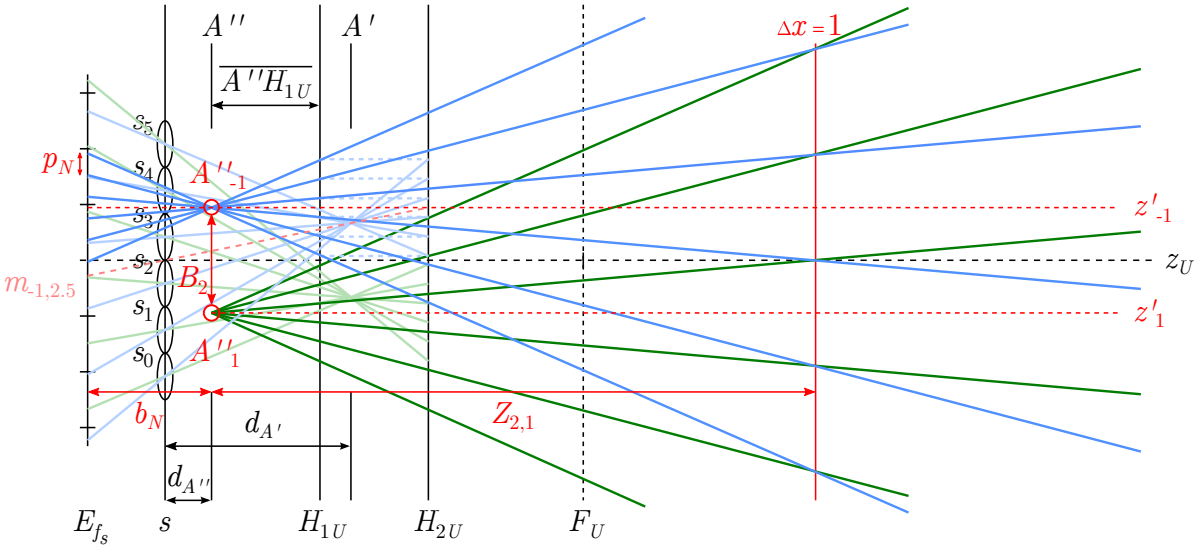


Figure 3.12: **SPC model triangulation with** $b_U = f_U$ and principal planes H_{1U}, H_{2U} just as the exit A' and entrance pupil plane A''. Red circles next to A''_i indicate virtual camera positions. Note that virtual cameras A''_{-1} and A''_1 are separated by gap $G = 2$ yielding baseline B_2.

Calculating virtual camera positions A''_i requires parameters $U_{i,j}$ from Eq. (3.25) and $q_{i,j}$ from Eq. (3.27) that have been elaborated in Section 3.3.1. Recall that the index for the central micro lens s_j is found by $j = o = {}^{(J\,-\,1)}/_2$ where o defines the image centre offset. By taking object space ray functions $\widehat{f}_{i,j}(z)$ as seen in Eq. (3.28) for two rays with different j but same i and setting them equal as given by

$$q_{i,o} \times z + U_{i,o} = q_{i,o+1} \times z + U_{i,o+1}\,, \quad z \in (-\infty\,,\, H_{1U}\,]\,, \tag{3.52}$$

we can solve for the equation system which yields a distance $\overline{A''H_{1U}}$ from virtual camera to object-side principal plane H_{1U}. To obtain an entrance pupil distance $d_{A''}$ with regards to the MLA, we calculate

$$d_{A''} = \overline{A''H_{1U}} + b_U + \overline{H_{1U}H_{2U}}\,. \tag{3.53}$$

The object-side-related position of A_i'' can be acquired by

$$A_i'' = q_{i,o} \times \overline{A''H_{1U}} + U_{i,o}\,. \tag{3.54}$$

With this, a baseline that spans from one A_i'' to another is denoted by B_G which can be obtained as follows

$$B_G = A_i'' + A_{i+G}''\,. \tag{3.55}$$

For example, a baseline B_1 ranging from A_0'' to A_1'' is identical to that from A_{-1}'' to A_0''. This relies on the principle that virtual cameras are separated by a continuous width. To apply the stereoscopic concept, rays are virtually extended towards the image space by

$$N_{i,j} = -q_{i,j} \times b_N + A_i''\,, \tag{3.56}$$

where b_N is an arbitrary scalar, which can be thought of as virtual image distance and $N_{i,j}$ as a spatial position at the virtual sub-aperture image plane. This scalable variable linearly affects a virtual pixel pitch p_N which is found by

$$p_N = \left|N_{i,o} - N_{i,o+1}\right|. \tag{3.57}$$

Setting $b_U = f_U$ aligns optical axes z_i' of virtual cameras to be parallel to the main optical axis z_U (see Fig. 3.12). For all other cases where $b_U \neq f_U$ (e.g. Fig. 3.13), the rotation angle Φ_i of a virtual optical axis z_i' is obtained by

$$\Phi_i = \arctan\left(q_{i,o}\right). \tag{3.58}$$

The relative tilt angle Φ_G from one camera to another can be calculated with

$$\Phi_G = \Phi_i + \Phi_{i+G}\,, \tag{3.59}$$

which completes the characterisation of virtual cameras.

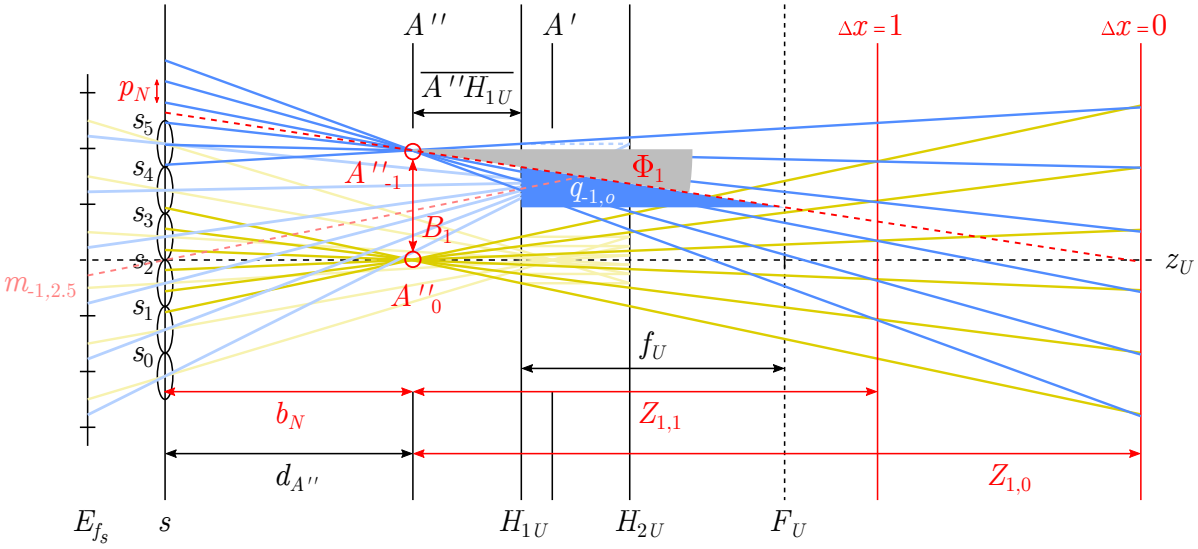


Figure 3.13: **SPC model triangulation with** $b_U > f_U$ and principal planes H_1, H_2 just as the exit A' and entrance pupil plane A''. Red circles next to A''_i indicate virtual camera positions. Note that the gap $G = 1$ and therefore B_1 and Φ_1.

Figure 3.13 visualises rays paths' in the light field when focusing the objective lens such that $b_U > f_U$. In this case, z'_i intersects with z_U at the plane the objective lens is focusing at. Objects placed at this plane expose a disparity $\Delta x = 0$ and thus are expected to be located at the same relative 2-D position in each sub-aperture image. As a consequence, objects reasonably located behind the $\Delta x = 0$ plane expose negative disparity.

Establishing the triangulation in an SPC would allow object distances to be retrieved just as in a stereoscopic camera system. On the basis of Eq. (2.5), a depth distance $Z_{G,\Delta x}$ of an object with certain disparity Δx is obtained by

$$Z_{G,\Delta x} = \frac{b_N \times B_G}{\Delta x \times p_N + b_N \times \tan\left(\Phi_G\right)} \tag{3.60}$$

and can be shortened to

$$Z_{G,\Delta x} = \frac{b_N \times B_G}{\Delta x \times p_N}, \quad \text{if } \Phi_G = 0 \tag{3.61}$$

which is only the case where $b_U = f_U$.

3.6 Summary

This chapter has introduced a light field ray model for an SPC that is exclusively based on geometrical optics. The proposed model implies two novelties, namely, a more accurate MIC approximation considering the exit pupil and a method to pinpoint object planes that have been computationally brought to focus. The latter approach simply relies on the idea of tracing rays that impinge on two corresponding pixels and finding their intersection in space. This not only helps localising distances to refocused object planes, but also detects positions of virtual cameras in the light field. Thereby, it was uncovered that virtual cameras can be regarded as lined up on the entrance pupil with baselines smaller than the pupil's diameter.

There is a wide range of use cases for the presented model. Regardless of whether a depth map is obtained from a set of photographs with varying focus [6, 18], stereo correspondence using disparity analysis [1, 2] or a combination of both [73, 74], the distance predication can contribute to the depth information by providing physical object distances. Industry branches such as the automotive market, robotics or machine vision can clearly benefit from measuring such distances. Besides this, the model's capability to precisely estimate an SPC's baselines and distances to which images can be refocused may prove to be useful in professional photography or motion picture arts.

The fact that the proposed model does not take aberrations into account may cause deviations when using simple lens systems such as a spheric lens. However, the research development carried out by lens makers during the last century enabled objective lenses to suppress aberrations to a negligible degree. This in turn validates the assumptions of geometrical optics and thus should make the proposed model applicable to any sort of SPC.

An extensive model verification is performed in Chapter 5. Prior to this, the calibration of an SPC is discussed in Chapter 4 constituting an essential requirement for the subsequent experimental work.

CHAPTER 4

Camera Configuration

According to Chapter 3, an SPC carries an array of micro lenses separated by one focal length in front of the sensor plane. To ensure the location of the MLA is ideally positioned in a customised camera, the optical elements of the SPC have to be aligned in an appropriate manner. Hence, an optical calibration is of tremendous necessity to comply with the declarations made in Chapter 3. If done incorrectly, the miscalibrated SPC deteriorates all subsequent processes and their results. A novel optical calibration procedure is discussed in the first section of this chapter.

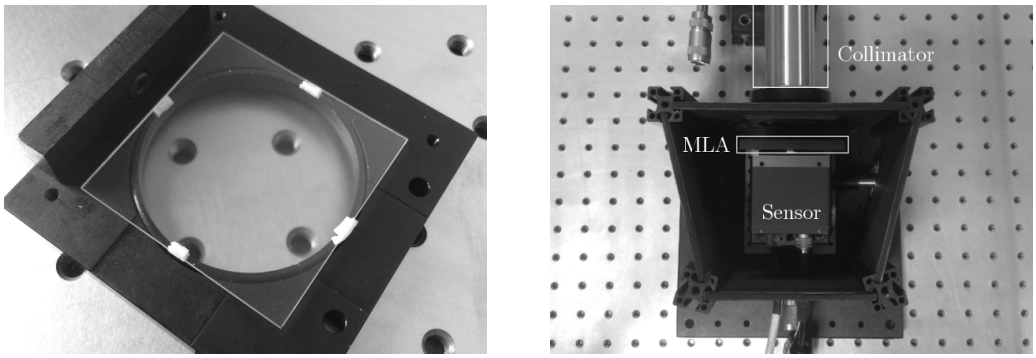
This is followed by a computational image rectification covering the identification of MIC reference points just as a compensation for physical misalignments in the optical setup. First, an algorithmic procedure is presented to detect MICs. Based on the obtained data, an image processing approach is proposed thereafter to counteract MLA rotation alignment errors. A subsequent section presents an image rendering process, which includes an interpolation to make use of the fractional digits in detected MIC coordinates. These computational calibration problems have been addressed by other researchers in the field [19, 22], but are documented here in more detail since ignoring them may bias the experimental work in Chapter 5. Results exposing image quality enhancements are provided at the very end of this chapter.

4.1 Optical Calibration

The calibration of a plenoptic camera can be considered as part of the assembly or follows just right after that. For instance, care needs to be taken when choosing optical components as it might be that an image sensor's cover glass blocks a desired MLA position. Due to the refracting behaviour of lens elements, parallel light may be useful for the lens alignment as it visually indicates the focal point. A collimated light emitting source, so-called collimator, is therefore employed.

4.1.1 Micro Lens Array Calibration

With respect to the sensor, the physical MLA positioning includes a rotation around z axis and a tilt around x and y axes to assure MLA and image sensor plane are parallel to each other. Besides, the MLA has to be placed at an appropriate distance from the sensor in z direction. Whilst signal processing can correct slight deviations in the rotation angle around z, it is impossible to compensate for a tilted or misfocused MLA due to inconsistent image distances between MLA and sensor ($b_s \neq fs$). Section 4.2 presents an attempt to computationally counter-rotate a light field photograph around z-axis to rectify a misaligned MLA. Figure 4.1 depicts an MLA and the collimator placed in front of it without the presence of an objective lens. This way, we ensure that only parallel light rays enter the MLA.

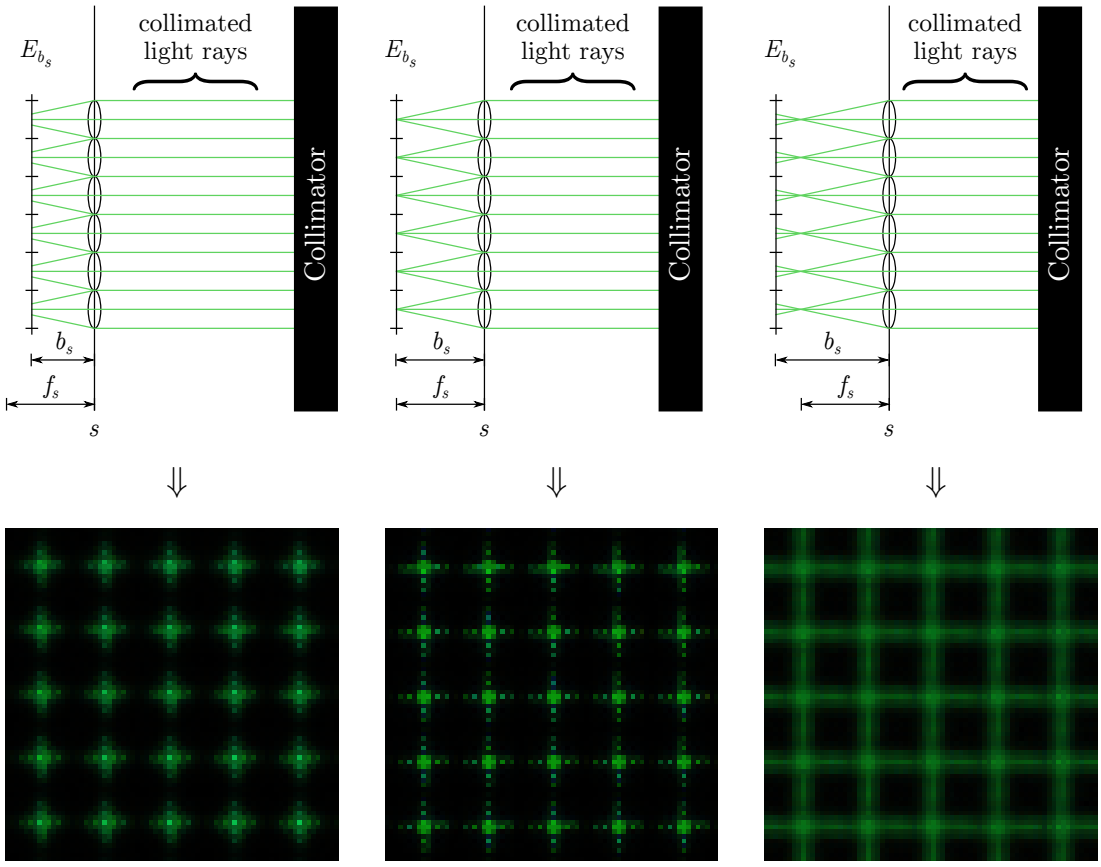


(a) MLA fixation (b) Collimator in front of camera body

Figure 4.1: **Photographs from MLA setup.**

Using collimated light, the focal distance of a lens is found where image points reach smallest spot sizes. This is anticipated to occur when the MLA image distance equals its focal length ($b_s = f_s$). The upper row of diagrams in Fig. 4.2 illustrates collimated ray paths at different MLA positions. Below, a plot of respective output images is shown.

Therein, it is seen that the central capture Fig. 4.2(b) exhibits best focus of the cross-shaped object in the micro images whereas the settings in Figs. 4.2(a) and 4.2(c) result in blurred cross shapes. The projected cross shape in Figs. 4.2(a) to 4.2(c) is caused by the LED arrangement inside the collimator and varies depending on the device type.



(a) $b_s < f_s$ (b) $b_s = f_s$ (c) $b_s > f_s$

Figure 4.2: **MLA positioning.** Captured images E_{b_s} exhibit out-of-focus crosses in (a) and (c). Best focus of an orthographic projection of the green cross is achieved in (b).

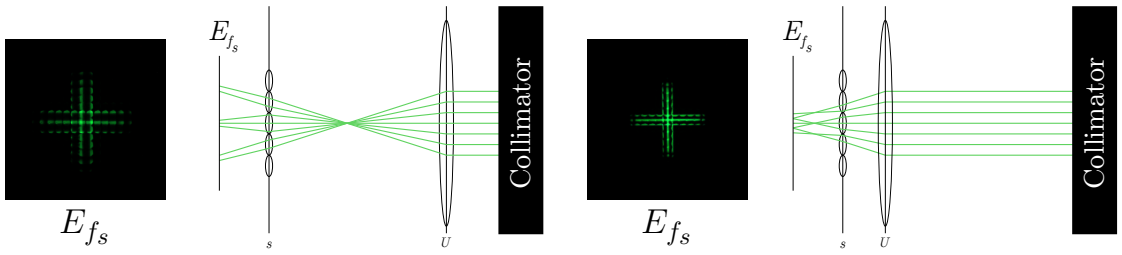
Despite convincing results, it needs to be mentioned that the MLA shift calibration potentially entails a minor misfocus of the MLA. This is due to the fact that the judgement relies on the visual perception of the person calibrating, which may cause a measurement inaccuracy. In cases where high-precision setups are needed, it thus may be suitable to deploy computer-aided systems to measure the spot size and send signals to the user when best focus is achieved.

A rather neat solution to this problem is to design micro lenses such that the back focal plane of a plano-convex MLA coincides with the sensor's image plane just when its

cover glass touches the planar back vertex of the MLA. This facilitates to stick the MLA directly on the sensor without the need for a physical calibration. A side benefit of this ideal scenario is that it also ensures that sensor and MLA plane are parallel. However, all other cases require a z-distance calibration as demonstrated above.

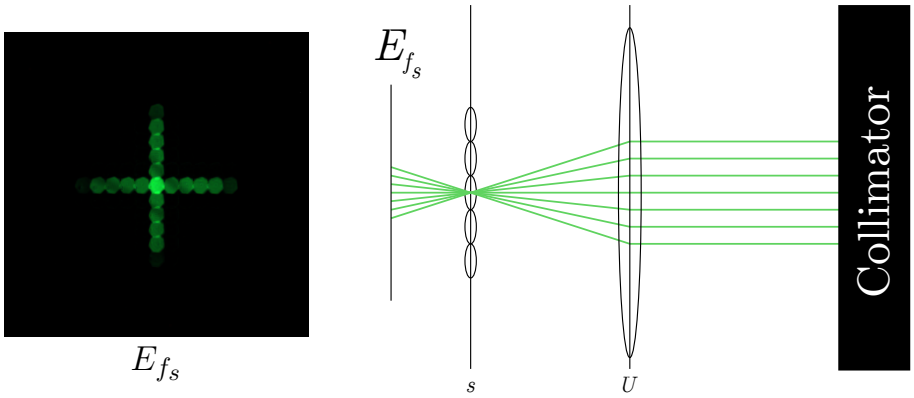
4.1.2 Main Lens Calibration

Given a calibrated MLA compound sensor, an objective lens can be mounted on the plenoptic camera body. Similar to the MLA calibration, the main lens has to be placed at a known distance to comply with statements made in Chapter 3. That is, if the focus ring is set to infinity, one focal length away from the MLA. This condition suits the employment of the collimator again as it helps to optically indicate the focal plane. Figure 4.3 depicts parallel rays travelling through optical components with different separations between MLA and main lens. Respective output images E_{f_s} are provided next to the ray models. Hereby, a huge green coloured cross is projected by the main lens onto the MLA and indicates the current focusing position.



(a) $b_U > f_U$

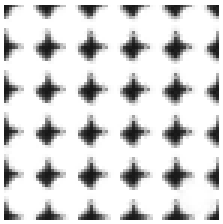
(b) $b_U < f_U$



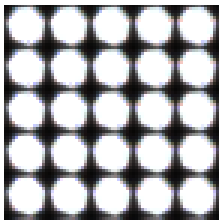
(c) $b_U = f_U$

Figure 4.3: **Main lens positioning.** Subfigures show ray models and corresponding image captures at different objective lens focus settings b_U. Optimised main lens position ($b_U = f_U$) is achieved in (c) whereas (a) and (b) indicate a misfocus ($b_U \neq f_U$).

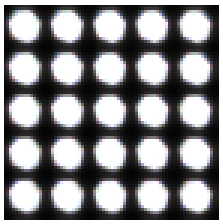
When shifting the objective lens such that $b_U \approx f_U$, an uncertainty exists for the exact image distance b_U. This is because positions $b_U \approx f_U$ illuminate the same micro image. On condition of a wide-opened aperture, the uncertainty range can be minimised.



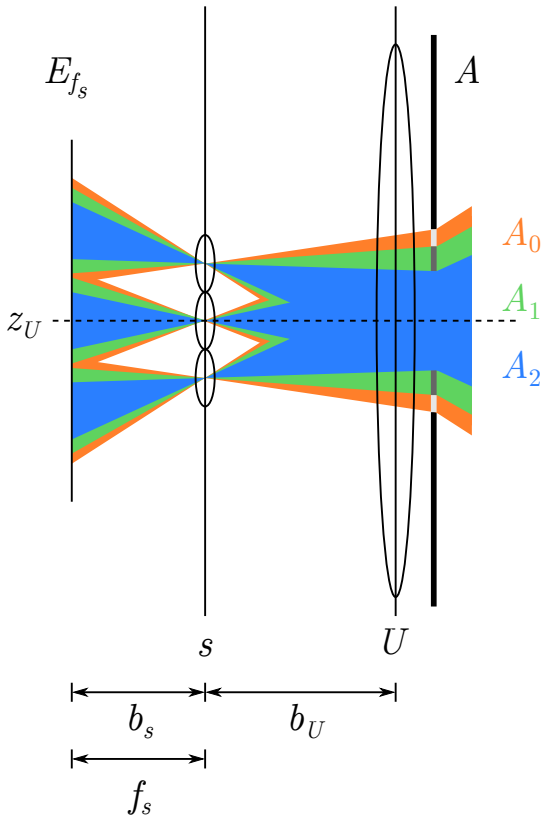
(a) A_0



(b) A_1



(c) A_2

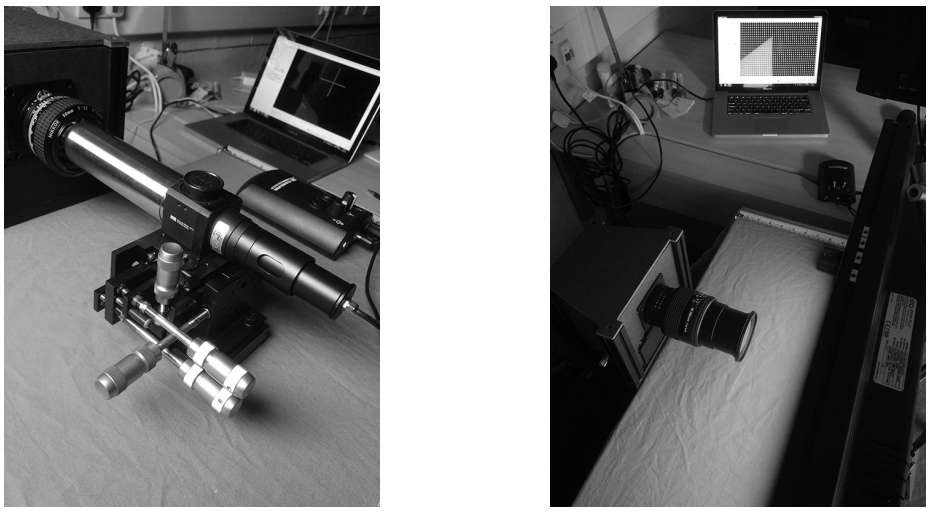


(d)

Figure 4.4: **Main lens aperture calibration.** (a) Overlapping micro images making it impossible to discern between objects in an image segmentation process; (b) properly sized micro images allowing for an MIC detection; (c) undersized micro images leaving sensor space unused; (d) light beams under varying aperture stops with alternating micro image size.

The main lens' aperture is the final optical parameter that has to be adjusted to get the customised SPC calibrated. To do so, it has been proven to utilise a soft white light source in front of the camera in order to illuminate each micro lens. As shown in Fig. 4.4, a variation in the aperture diameter linearly controls the size of micro images. It is important to ensure their boundaries do not touch each other as seen in Fig. 4.4(a) since cross-talk occurs among samples u. Besides, overlapping micro images impede to detect MICs at a later stage. Nevertheless, an optimal aperture diameter is found by letting micro images be as large as possible because it increases the number of samples

u that would be lost otherwise. Thus, an ideal compromise is to find an aperture size that yields a high ratio of occupied sensor space which is often referred to as *fill factor*. It may be obvious that utilising square-shaped apertures enhances the *fill factor* as micro images also appear as squares on the sensor. Alternatively, a hexagonal micro lens arrangement may be employed to enhance the *fill factor* [66]. On the contrary, hexagonal sampling requires an additional image interpolation step to convert barycentric to Cartesian coordinates if the content is displayed on a monitor with non-hexagonal pixel packing [19]. A solution for this conversion is presented in Section 4.2.2. Pictures in Fig. 4.5 show practical examples of the objective lens calibration.



(a) Main lens calibration (b) White image capture

Figure 4.5: **Photographs from main lens setup.** (a) Calibration of main lens in z-direction using a collimator; (b) plenoptic camera in front of a diffused light source.

4.2 Image Calibration

It has been shown in previous chapters that lens-array-based light field photographs require image processing treatments to be perceived as images taken by traditional cameras. Here, MICs play the key role to serve as reference points for an orientation in the 4-D image. Hence, it is the aim of this section to first develop an MIC localisation procedure and then present an image rendering technique based on detected MICs. For the purpose of detecting MICs, calibration images have been captured when the customised SPC was exposed to diffused white light. Respective micro images therefore contain no information other than smoothly distributed white light. Following explanations build on the 4-D to 2-D image translation as given by Eq. (3.46) in Section 3.4.

4.2.1 Micro Image Array Alignment

First of all, colour channels of a white calibration capture are converted to a single greyscale channel image $E_{f_s}[x_k, y_l]$ having $\mathcal{L}$ grey levels $[1, 2, ..., \mathcal{L}]$ which is then treated with a linear normalisation to obtain $E_{\text{norm}}[x_k, y_l]$ as given by

$$E_{\text{norm}}[x_k, y_l] = \left(E_{f_s}[x_k, y_l] - \mathcal{L}_{\text{min}}\right) \frac{\mathcal{L} - 1}{\mathcal{L}_{\text{max}} - \mathcal{L}_{\text{min}}} + 1 \tag{4.1}$$

with $\mathcal{L}_{\text{min}} := \min\left\{E_{f_s}[x_k, y_l]\right\}$ and $\mathcal{L}_{\text{max}} := \max\left\{E_{f_s}[x_k, y_l]\right\}$. Subsequently, a binary image $E_{\text{bin}}[x_k, y_l]$ is generated from the normalised greyscale image $E_{\text{norm}}[x_k, y_l]$ using a threshold value τ obtained by Otsu's method [64]. Hence, the binary conversion is performed as follows

$$E_{\text{bin}}[x_k, y_l] = \begin{cases} 1, & \text{if } E_{\text{norm}}[x_k, y_l] \geq \tau\,; \\ 0, & \text{if } E_{\text{norm}}[x_k, y_l] < \tau\,. \end{cases} \tag{4.2}$$

Based on this, $E_{\text{bin}}[x_k, y_l]$ contains a number of binary regions equal to the number of micro images in the raw picture. To distinguish between regions, each region $\mathcal{R}$ is labelled in compliance with its parent micro lens index $[j, h]$. From this, MICs can be detected by finding the centroid of each labelled region $\mathcal{R}_{j,h}$. According to Burger and Burge [16], a centroid $\bar{\mathbf{c}}_{j,h} = \left\{\bar{k}_{j,h}, \bar{l}_{j,h}\right\} \in \mathbb{R}^2$ of a region $\mathcal{R}_{j,h}$ is computed via

$$\bar{k}_{j,h} = \frac{1}{|\mathcal{R}_{j,h}|} \sum_{(k,l)\,\in\,\mathcal{R}_{j,h}} k \qquad \text{and} \qquad \bar{l}_{j,h} = \frac{1}{|\mathcal{R}_{j,h}|} \sum_{(k,l)\,\in\,\mathcal{R}_{j,h}} l \tag{4.3}$$

where $|\cdot|$ denotes the cardinality which provides the total number of elements within a particular region $\mathcal{R}_{j,h}$. Figure 4.6 shows an example of detected MICs using Eq. (4.3).

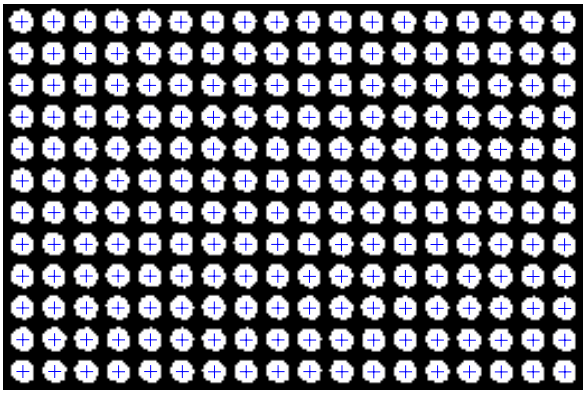


Figure 4.6: **Algorithmic MIC detection.** Detailed view of a binary image E_{bin} overlaid with detected MICs based on centroids $\bar{\mathbf{c}}_{j,h}$ (blue markers).

As mentioned earlier in the optical calibration section, an MLA can be possibly rotated with respect to the image sensor grid. In the post-processing stage, it is feasible to compensate for that rotation. It is a well known fact that images tend to be distortion-free around their central axes. This suggests that a central row of MICs may be a sufficient indicator for the MLA rotation angle. Such a central row of MICs is obtained by $\bar{\mathbf{c}}_{j,\delta} = \{\bar{k}_{j,\delta},\, \bar{l}_{j,\delta}\}$ coordinates when $\delta := \lfloor {}^{(H\,-\,1)}\!/_{2} \rceil$ where $\lfloor \cdot \rceil$ is the nearest integer operation and H the total number of micro lenses in vertical direction. Applying the linear regression model may be suitable to acquire the slope of MIC data set $\bar{\mathbf{c}}_{j,\delta}$. Using the ordinary least squares estimation, the regression coefficients $\boldsymbol{\beta}$ are obtained by

$$\boldsymbol{\beta} = (\mathbf{x}^{\mathsf{T}}\,\mathbf{x})^{-1}\;\mathbf{x}^{\mathsf{T}}\,\mathbf{y} \tag{4.4}$$

where $^{\mathsf{T}}$ denotes the matrix transpose and

$$\mathbf{x} = \begin{pmatrix} 1 & \bar{k}_{0,\delta} \\ 1 & \bar{k}_{1,\delta} \\ \vdots & \vdots \\ 1 & \bar{k}_{J-1,\delta} \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} \bar{l}_{0,\delta} \\ \bar{l}_{1,\delta} \\ \vdots \\ \bar{l}_{J-1,\delta} \end{pmatrix} \tag{4.5}$$

contain respective MIC coordinates $\bar{\mathbf{c}}_{j,\delta} = \{\bar{k}_j,\, \bar{l}_j\}$. We put $\boldsymbol{\beta} = \{\beta_n\,, \beta_m\}$ as $\boldsymbol{\beta}$ consists of two coefficients reflecting the offset and slope, respectively. Converting β_m to radians β is achieved by

$$\beta = \arctan\left(\beta_m\right). \tag{4.6}$$

Rotating a normalised image $E_{\text{norm}}\left[x_k, y_l\right]$ as well as MICs $\bar{\mathbf{c}}_{j,h}$ is accomplished via multiplying coordinates with a rotation matrix and subsequent pixel interpolation for the image process. The counter-clockwise rotation matrix $\mathbf{R}_z$ around z-axis can be written as

$$\mathbf{R}_z = \begin{pmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.7}$$

and the translation matrix $\mathbf{T}$ is given by

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & -{}^{K}\!/_{2} \\ 0 & 1 & -{}^{L}\!/_{2} \\ 0 & 0 & 1 \end{pmatrix} \tag{4.8}$$

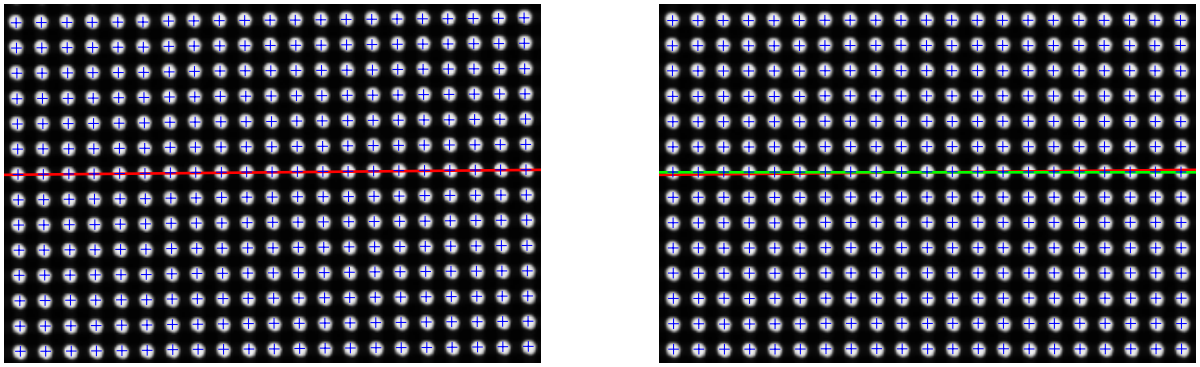
shifting all $\bar{\mathbf{c}}_{j,h}$ by half the image resolution to rotate about the image centre with the aid of total pixel numbers K and L for horizontal and vertical direction, respectively. To perform the rotation, transformation matrices are successively multiplied with one another. In order to obtain the rotated centroid data set $\bar{\mathbf{c}}'_{j,h}$, this reads

$$\bar{\mathbf{c}}'_{j,h} = \mathbf{T}^{-1} \times \mathbf{R}_z \times \mathbf{T} \times \bar{\mathbf{c}}_{j,h} \,, \tag{4.9}$$

with

$$\bar{\mathbf{c}}_{j,h} = \begin{pmatrix} \bar{k}_{0,0} & \bar{k}_{1,0} & \dots & \bar{k}_{J-1,0} & \bar{k}_{0,1} & \bar{k}_{1,1} & \dots & \bar{k}_{J-1,H-1} \\ \bar{l}_{0,0} & \bar{l}_{1,0} & \dots & \bar{l}_{J-1,0} & \bar{l}_{0,1} & \bar{l}_{1,1} & \dots & \bar{l}_{J-1,H-1} \\ 1 & 1 & \dots & 1 & 1 & 1 & \dots & 1 \end{pmatrix} \tag{4.10}$$

An image rotation requires to interpolate pixels values E_{norm} and map them to a new image given as $E_{\text{rot}}\left[x_k, y_l\right]$. An example showing results of the proposed rotation procedure is given in Fig. 4.7.



(a) (b)

Figure 4.7: **Micro image array rotation.** (a) Magnified view of a micro image array $E_{\text{norm}}\left[x_k, y_l\right]$ taken from a rotated MLA; (b) computationally rectified image $E_{\text{rot}}\left[x_k, y_l\right]$ after applying transformation matrices. Straight lines in colours represent $\boldsymbol{\beta}$ whereas red indicates the untreated and green the aligned result. Comparison of image edges or straight lines indicate the performance of the computational rotation alignment.

In a subsequent stage, it would be worth to countervail lens distortions. Usually, Browne's model [14] satisfies to rectify optically distorted images. However, in this work it is supposed that lenses do not exhibit severe lens distortions which is why they are neglected. Further details on plenoptic lens distortion removal can be found in publications from Dansereau *et al.* [22] and Bok *et al.* [12]. Apart from this, vignetting may be treated by brightening pixels values gradually towards micro image borders as well as among micro

images closer to the image edges. This, however, is left out due to scarce vignetting in our plenoptic system. A related article that successfully attempts to combat general lens aberrations has been proposed by Ng and Hanrahan [60].

4.2.2 Image Extraction

This section concerns with extracting a sub-aperture image to demonstrate the use of detected centroids in a plenoptic image rendering process. According to Eq. (3.45) and Fig 3.10, a central sub-aperture image is obtained by collecting all pixels at MICs and relocating them consecutively in a new image array.

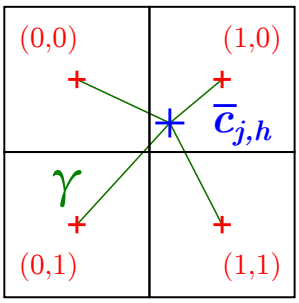


Figure 4.8: **MIC interpolation weighting scheme.** Green lines show distances γ from MIC to adjacent pixel centres. The length of each γ represents the weighting factor to form the intensity of pixels at the real MIC position.

Since $\bar{\mathbf{c}}_{j,h} \in \mathbb{R}^2$, MICs are likely to be found at sub-pixel positions. Instead of simply taking pixels from nearest integer coordinates that lead to errors and thus to image artefacts, it is the idea to preserve fractional digits of MIC coordinates and resample $E_{\text{rot}}\left[x_k\,,\,y_l\right]$ at $\bar{\mathbf{c}}_{j,h}$. A similar approach has been proposed by Cho *et al.* [19], however, not provided in much detail. Therefore, it may be worth introducing an interpolation scheme which retains the accuracy of detected MIC coordinates. Figure 4.8 depicts an MIC $\bar{\mathbf{c}}_{j,h}$ and the corresponding distance lengths γ to surrounding pixel centres indexed by (i, g).

Taking $\gamma_{(i,g)}$ as a weighting coefficient, the sub-pixel precise illuminance E_{int} at each centroid is computed via

$$E_{\text{int}}\left[x_{\lfloor\bar{k}\rfloor}, y_{\lfloor\bar{l}\rfloor}\right] = \sum_{g=0}^{1}\sum_{i=0}^{1} E_{\text{rot}}\left[x_{\lfloor\bar{k}\rfloor+i}\,,\, y_{\lfloor\bar{l}\rfloor+g}\right] \times \gamma_{(i,g)}\,,$$

$$\gamma_{(i,g)} \begin{cases} (\bar{k} - \lfloor\bar{k}\rfloor) \times (\bar{l} - \lfloor\bar{l}\rfloor), \text{ if } (\bar{k} - \lfloor\bar{k}\rfloor + i) \geq 0.5 \wedge (\bar{l} - \lfloor\bar{l}\rfloor + g) \geq 0.5 \\ (\lceil\bar{k}\rceil - \bar{k}) \times (\bar{l} - \lfloor\bar{l}\rfloor), \text{ if } (\bar{k} - \lfloor\bar{k}\rfloor + i) \geq 0.5 \wedge (\bar{l} - \lfloor\bar{l}\rfloor + g) < 0.5 \\ (\bar{k} - \lfloor\bar{k}\rfloor) \times (\lceil\bar{l}\rceil - \bar{l}), \text{ if } (\bar{k} - \lfloor\bar{k}\rfloor + i) < 0.5 \wedge (\bar{l} - \lfloor\bar{l}\rfloor + g) \geq 0.5 \\ (\lceil\bar{k}\rceil - \bar{k}) \times (\lceil\bar{l}\rceil - \bar{l}), \text{ if } (\bar{k} - \lfloor\bar{k}\rfloor + i) < 0.5 \wedge (\bar{l} - \lfloor\bar{l}\rfloor + g) < 0.5 \end{cases} \tag{4.11}$$

with $\lceil\cdot\rceil$ as ceiling and $\lfloor\cdot\rfloor$ as floor operator. For implementation purposes, a pseudo code notation can be found in Appendix B. Since the proposed interpolation only applies for pixels at MICs, Eq. (4.11) is extended in the algorithm to include adjacent pixels u_{c+i} in the interpolation process.

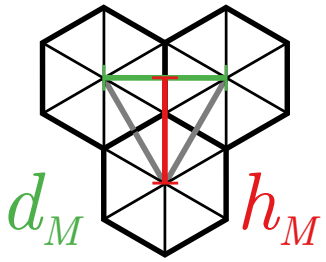


Figure 4.9: **MIC spacing in hexagonal grid.**

In case micro lenses are arranged in a hexagonal grid, image rendering procedures require an additional process to translate the barycentric to Cartesian coordinates which is addressed in the following. It can be seen in Fig. 4.9 that 3 straight lines connecting adjacent micro image centres in an hexagonally structured MLA form an equilateral triangle. If d_M represents the side length of the equilateral triangle, the height h_M of it amounts to

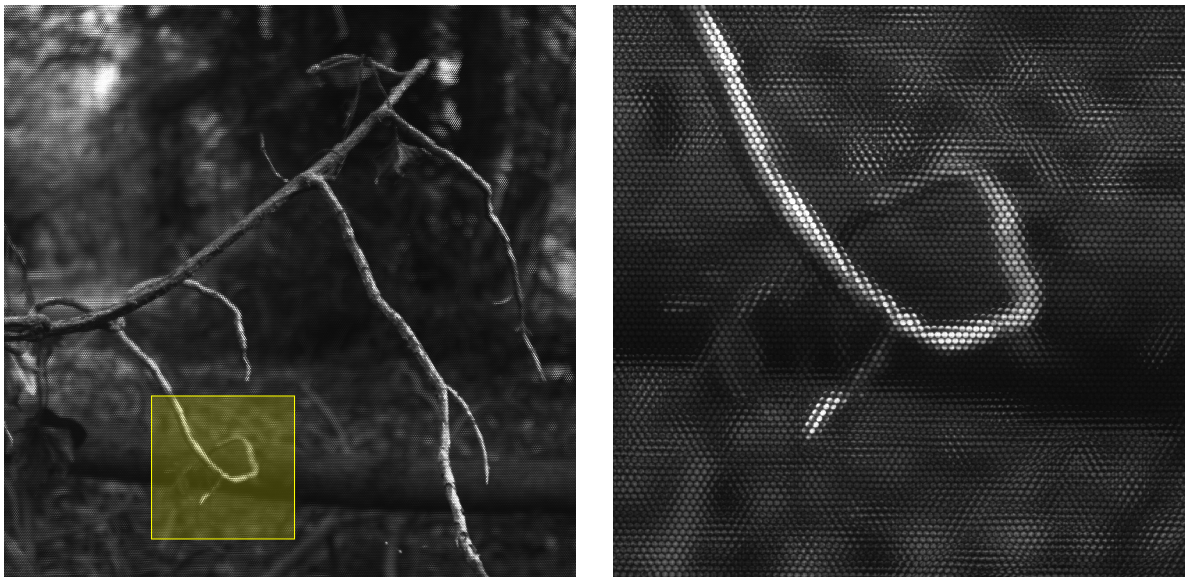
$$h_M = \frac{\sqrt{3}}{2} \times d_M\,. \tag{4.12}$$

Equation (4.12) suggests that a compensation for the hexagonal micro lens structure could be achieved if an extracted sub-aperture $E_{(i,g)}\left[s_j\,,\,t_h\right]$ (obtained via 2-D application of Eq. (3.45)) is first upsampled by factor 2 in one dimension and then downsampled in the other by $\sqrt{3}$ afterwards. The upsampling part is accomplished via

$$E_{\text{hex}}\left[s_{2j+\varepsilon}\,,\,t_h\right] = E_{(i,g)}\left[s_j\,,\,t_h\right]\,, \quad \varepsilon = h \bmod 2 \tag{4.13}$$

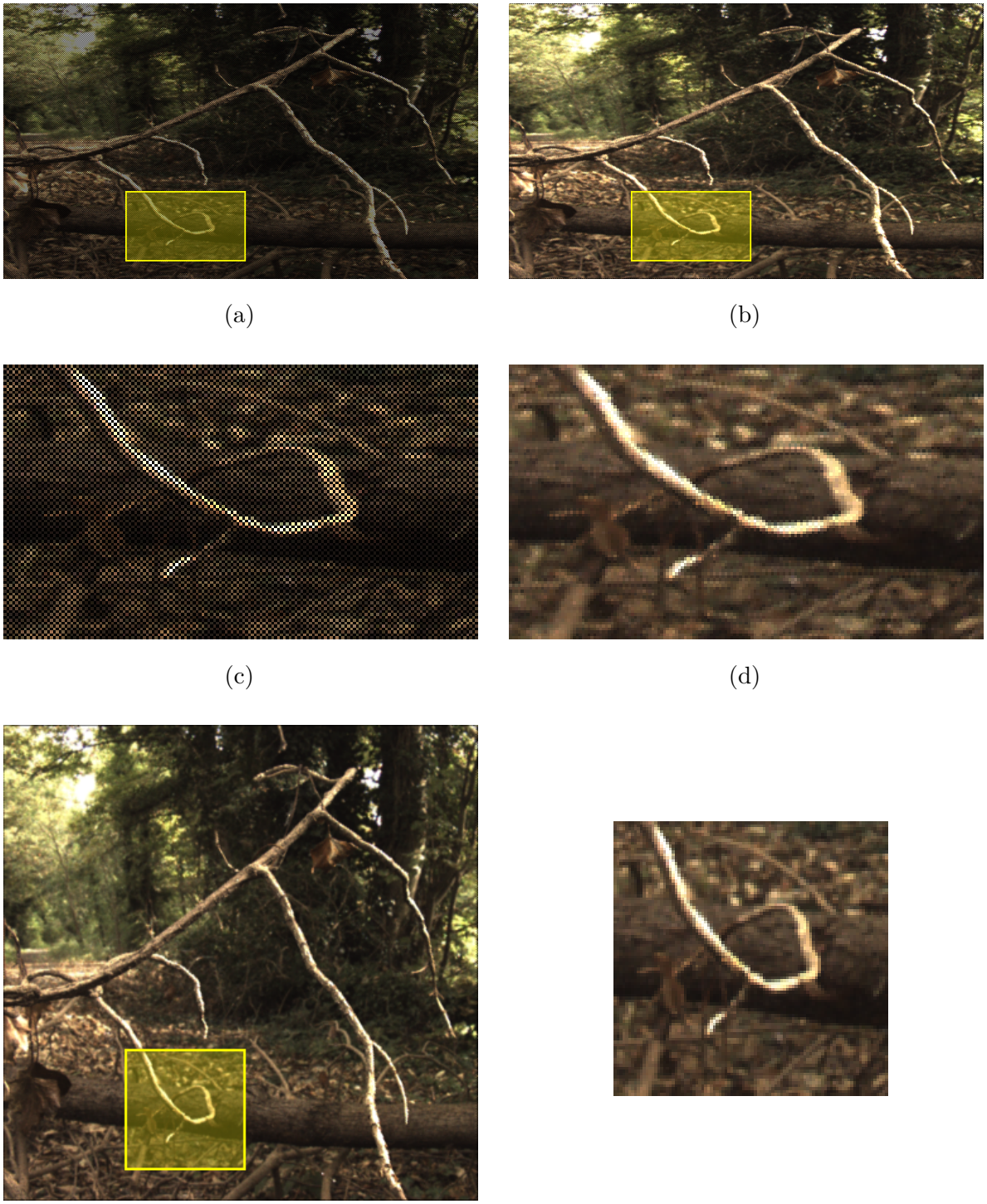
where mod is the *modulo* operation, which helps rearrange pixels by inserting empty (black) pixels in a chequerboard fashion as an intermediate step.

Based on a raw light field image with hexagonal MLA arrangement (see Fig. 4.10), a typical chequerboard image is provided in Fig. 4.11(a). The image stretch by factor 2 is then completed by interpolating missing (black) pixels which can be seen in Fig. 4.11(b). Due to the upsampling by factor 2 in horizontal direction, Fig. 4.11(b) still exhibits a wrong aspect ratio. According to Eq. (4.12), the resolution of Fig. 4.11(b) needs to be downsampled by factor $\sqrt{3}$ in the horizontal dimension to compensate for the previous horizontal stretching by factor 2. Alternatively, the stretched image may be vertically upsampled by $\sqrt{3}$. Figure 4.11(e) shows the central sub-aperture image when horizontally downscaling by $\sqrt{3}$ and hence exposes the effective resolution which amounts to 376 by 376 pixels.



(a) (b)

Figure 4.10: **Raw SPC capture.** (a) Entire SPC capture and (b) magnified region showing hexagonal micro image arrangement before demosaicing.



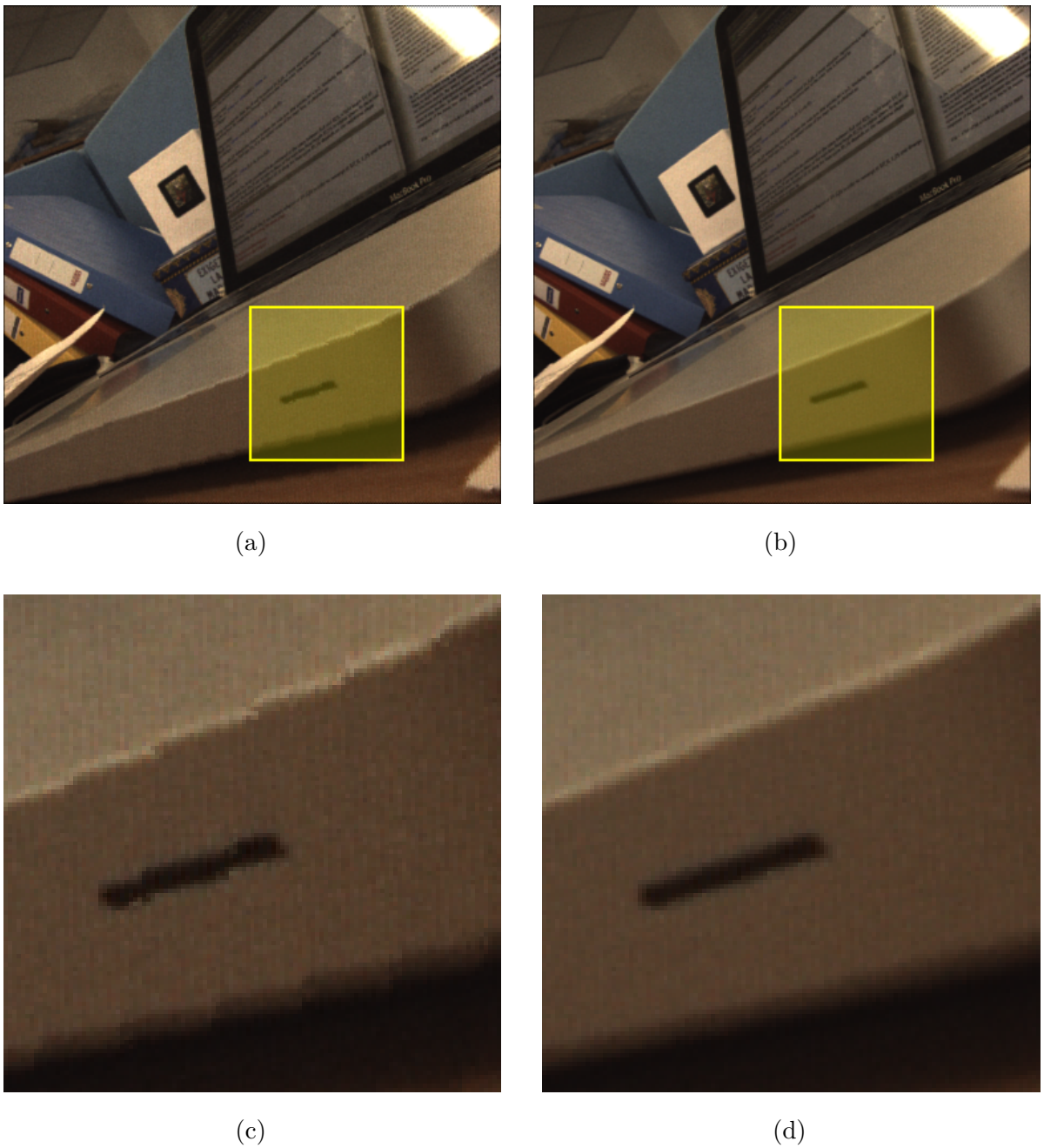
(a) (b)

(c) (d)

(e) (f)

Figure 4.11: **Sub-aperture extraction from hexagonal MLA structure.** (a) extracted pixel values at MICs in chequerboard pattern with magnified view in (c); (b) picture and (d) its magnification of same sub-aperture image after interpolating missing pixels in (a); (e) final view with correct aspect ratio and (f) its magnified region.

The presence and absence of image artefacts can be seen in Fig. 4.12. As intended, artefacts are prevented by taking sub-pixel MIC coordinates into account. A positive side effect of the MIC sub-pixel interpolation is that it reduces the image noise level.



(a) (b)

(c) (d)

Figure 4.12: **Sub-aperture image** (a) **without and** (b) **with MIC interpolation.** (a) exposes visible artefacts, e.g. stairway transition at straight diagonal lines which are due to rounding errors of the MIC; (b) suppresses artefacts by preserving the fractional digits in MIC coordinates; (c) and (d) are magnified views of (a) and (b) respectively.

4.3 Summary

This chapter demonstrated approaches to properly calibrate an SPC in accordance with definitions made in Chapter 3. It has presented novel ways to ensure MLA, main lens and aperture are set-up as intended by using a collimator. However, it has been pointed out that the MLA positioning can be simplified in case MLA focal plane and SPC image sensor coincide once the MLA back vertex touches the sensor's surface. This incidentally assures parallelism between sensor and MLA plane.

Moreover, alternative signal processing attempts have been devised to detect MICs, counter-rotate a light field photograph from a plenoptic camera and interpolate MICs at sub-pixel positions. Thereby, this study has confirmed the findings of Cho *et al.* [19] who discovered that ignoring sub-pixel precision in MICs yields image artefacts. In other words, it has been proven that preserving decimal places of MIC coordinates in plenoptic image rendering is essential. The sub-pixel precision equally applies to the refocusing synthesis and is deployed in the experimentation in Chapter 5.

CHAPTER 5

Model Validation

This chapter supplies an extensive examination assessing the light field geometry of an SPC as modelled in Chapter 3. The verification is split into two parts: First, the SPC model is evaluated by comparing the prediction system to a real ray tracing simulator [89] which, in contrast, takes optical aberrations into account. A second study is carried out using raw light field images captured with a calibrated SPC to further test the reliability of the model. Both validations are expected to clarify which optical parameters affect the light field geometry. For the evaluation, measured errors ERR are given as a percent deviation calculated as follows

$$ERR = \frac{\text{prediction} - \text{measurement}}{\text{prediction}} \times 100\,. \tag{5.1}$$

To allow the reader for a reproduction of experimental work, the sensor and optical design of our SPC is discussed in a preceding section.

5.1 Camera Design

Our customised camera contains a *full frame* sensor with $K = 4008$ by $L = 2672$ active pixels and $p_p = 9$ μm pixel pitch meaning that its size amounts to 36.072 mm by 24.048 mm. Lens specifications are presented hereafter.

5.1.1 Micro Lens Specification

Information about the substrate of the MLA integrated to our camera remains undisclosed which makes it difficult to model the micro lenses in an optical design software [89].

However, the material's refractive index $n(\lambda)$ with wavelength λ was provided by the manufacturer (see Appendix C.1) and is of the form

$$n(\lambda) = 1.5375 + \frac{8290.45}{\lambda^2} - \frac{2.11046 \times 10^8}{\lambda^4} \tag{5.2}$$

being compliant with Cauchy's equation [8] which is

$$n(\lambda) = \Gamma_0 + \sum_{\epsilon=1}^{2} \frac{\Gamma_\epsilon}{\lambda^{2\epsilon}} \tag{5.3}$$

with coefficients Γ_ϵ. Cauchy's formula reveals a limitation, since it tends to have an insufficient accuracy for non-visible wavelength spectra (infra-red and ultraviolet). As the camera sensor used in the experiments solely considers visible light ($\lambda \,|\, 390 \text{ nm} \leq \lambda \leq 700 \text{ nm}$), the shortcoming in Cauchy's equation could be ignored in the following. Nevertheless, its limitation might be a reason for developers of an optical design software [89] to omit support for Cauchy functions which still makes it impossible to enter our MLA's substrate data. Alternatively, the simulation tool provides a wide range of more advanced dispersion functions such as the Schott equation [8] which is given by

$$n^2(\lambda) = \Gamma_0 + \Gamma_1\lambda^2 + \Gamma_2\lambda^{-2} + \Gamma_3\lambda^{-4} + \Gamma_4\lambda^{-6} + \Gamma_5\lambda^{-8}\,. \tag{5.4}$$

Although it appears that the literature did not address this yet, it may be an obvious and neat attempt to perform the square on the manufacturer's provided Cauchy Eq. (5.2) to ascertain the equivalent Schott representation of the lens material's dispersion. In doing so, we obtain

$$\begin{aligned} n^2(\lambda) = \;& 2.36390625 + 0\lambda^2 + 2.549313375 \times 10^{-2}\lambda^{-2} \\ & - 5.802348887975 \times 10^{-4}\lambda^{-4} - 3.4993326214 \times 10^{-6}\lambda^{-6} \\ & + 4.454041411599998 \times 10^{-8}\lambda^{-8}\,. \end{aligned} \tag{5.5}$$

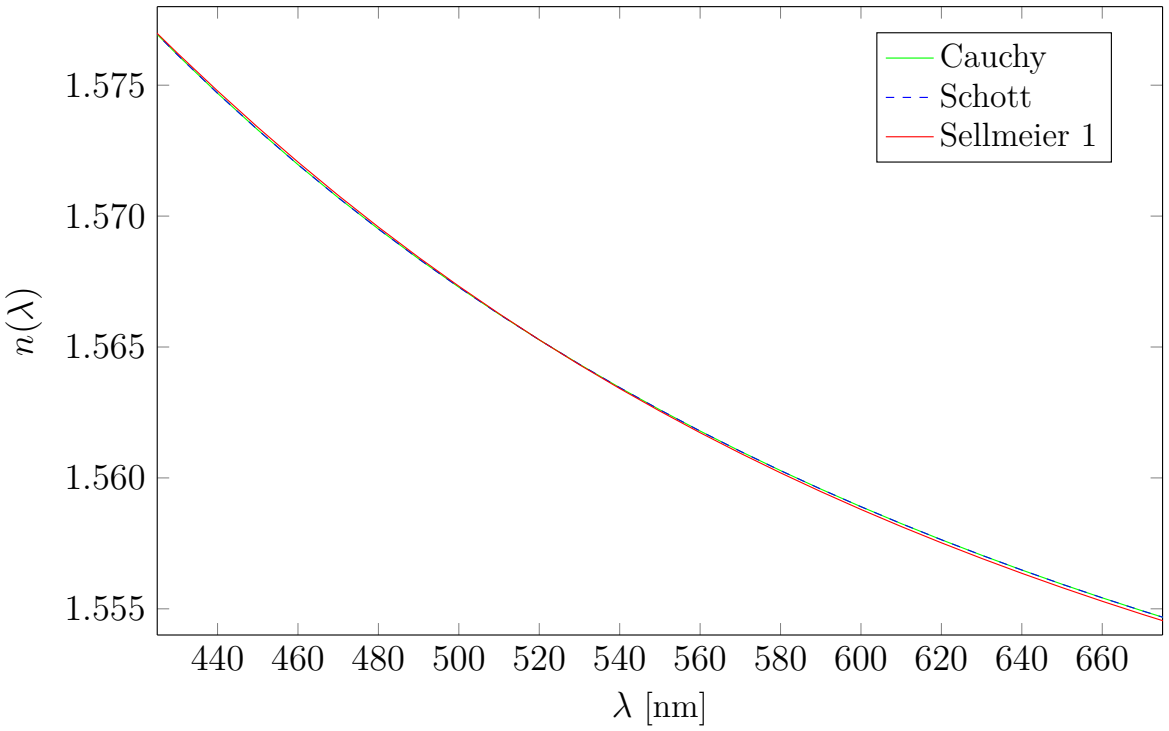
The optical design software [89] also offers a tool to insert a data list of given refractive index values at several wavelengths and compute the Sellmaier coefficients from that which are denoted by Υ and Λ. The Sellmeier 1 formula [8] reads

$$n^2(\lambda) = 1 + \frac{\Upsilon_1\lambda^2}{\lambda^2 - \Lambda_1} + \frac{\Upsilon_2\lambda^2}{\lambda^2 - \Lambda_2} + \frac{\Upsilon_3\lambda^2}{\lambda^2 - \Lambda_3}\,. \tag{5.6}$$

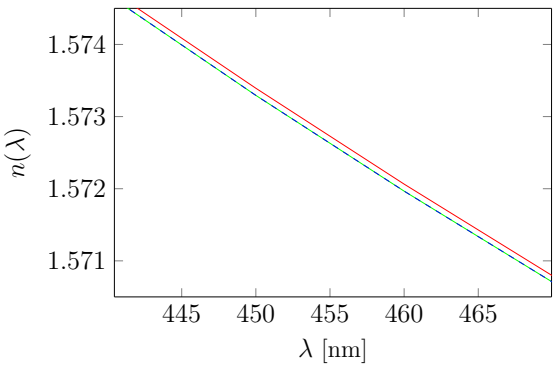
Accordingly, the Sellmaier 1 coefficients obtained by the simulation tool are as follows

$$\begin{aligned} n^2(\lambda) = 1 &+ \frac{-8.86496761 \times 10^5\lambda^2}{\lambda^2 - 3.49477799 \times 10^{-2}} + \frac{8.86498127 \times 10^5\lambda^2}{\lambda^2 - 3.49477525 \times 10^{-2}} \\ &+ \frac{4.54358371 \times 10^3\lambda^2}{\lambda^2 - 6.30419436 \times 10^6}\,. \end{aligned} \tag{5.7}$$

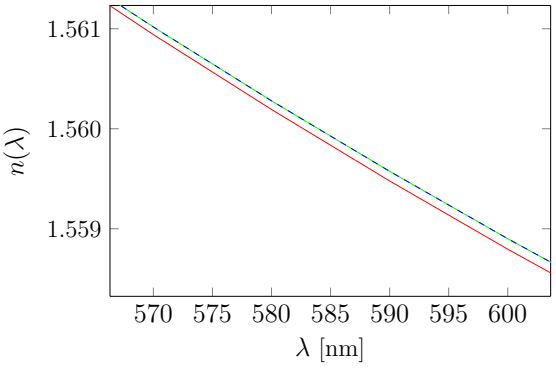
Since the original Cauchy formula has been manually converted to the Schott representation, it is expected that both functions yield an identical refractive index behaviour whereas the approximation of Sellmaier 1 coefficients may provide deviating results.



(a)



(b)



(c)

Figure 5.1: **Dispersion functions.** (a) Eqs. (5.2), (5.5) and (5.7); (b) and (c) depict magnified sections from (a) revealing the deviation of Eq. (5.7).

Figure 5.1 consists of plots comparing the results of Eqs. (5.2) and (5.5) depicted in green and dashed blue, respectively. As anticipated, these two equations are identical whereas the red coloured graph of Eq. (5.7), which represents the Sellmeier function computed

by the simulation tool, slightly differs. Hence, statements hereafter are based on the derived Schott coefficients as the manual conversion from the Cauchy Eq. (5.2) retains the original dispersion data. To find a micro lens' principal plane, its radii of curvature need to be examined. With the aid of the lens maker's formula [42] which reads

$$\frac{1}{f_s} = (n(\lambda) - 1) \times \left(\frac{1}{R_{1s}} - \frac{1}{R_{2s}} \right) , \tag{5.8}$$

the radius of curvature R_{1s} of a plano-convex micro lens ($R_{2s} \to \infty$) may be obtained by simplifying and rearranging the equation to

$$R_{1s} = f_s \times (n(\lambda) - 1) \ . \tag{5.9}$$

Similarly, Eq. (5.8) can be solved for R_{s2} instead, if $R_{1s} \to \infty$, so that

$$R_{2s} = -\left(f_s \times (n(\lambda) - 1) \right) \ . \tag{5.10}$$

Let t_s denote the MLA thickness, its principal plane positions H_{1s} and H_{2s} are obtained via

$$H_{1s} = -\frac{f_s \times t_s \times (n(\lambda) - 1)}{n(\lambda) \times R_{2s}} , \tag{5.11}$$

and

$$H_{2s} = -\frac{f_s \times t_s \times (n(\lambda) - 1)}{n(\lambda) \times R_{1s}} . \tag{5.12}$$

With reference to the plano-convex lens in Fig. 3.1(a), $\overline{H_{1s}H_{2s}}$ denotes the separation between principal plane positions and is given by

$$\overline{H_{1s}H_{2s}} = t_s + H_{2s} . \tag{5.13}$$

The distance d_s from the micro lens back vertex V_{2s} to the sensor image plane equals the focal length f_s in case the plano side of the lens is placed towards object space. Otherwise, d_s is obtained by

$$d_s = f_s + \overline{H_{1s}H_{2s}} - t_s . \tag{5.14}$$

Table 5.1 lists parameters of two micro lens specifications, denoted MLA (I.) and (II.), used in subsequent experimentations.

Table 5.1: **Micro lens specifications for $\lambda = 550$ nm.**

MLA	f_s	p_M	N_s	t_s	$n(\lambda)$	R_{s1}	R_{s2}	$\overline{H_{1s}H_{2s}}$	d_s
(I.)	1.25 mm	125 μm	$f/10$	1.1 mm	1.5626	0.70325	-∞	0.396045 mm	0.546045 mm
(II.)	2.75 mm	125 μm	$f/22$	1.1 mm	1.5626	1.54715	-∞	0.396045 mm	2.046045 mm

A given micro lens pitch p_M and active sensor resolution K by L lead to $J = p_p \times K/p_M = 288.576$ and $H = p_p \times L/p_M = 192.384$ samples for the spatial image resolution whereas the theoretical directional resolution M amounts to $M = p_M/p_p = 13.889$ samples for both dimensions equally. To avoid vignetting, the micro image resolution is reduced to $M = 11$ in real light field photographs which involves cropping each micro image around its centre.

5.1.2 Main Lens Specification

Modelling the objective lens of an SPC is as important as the MLA specification. To cover different zoom and focus settings in the experiment, a variety of main lens designs found in [17, 85, 89] are provided in the following. The simulation tool [89] offers principal plane positions as distances with respect to front and back vertexes as well as image and object plane, respectively. The principal plane separation $\overline{H_{1U}H_{2U}}$ may be retrieved by applying

$$\overline{H_{1U}H_{2U}} = \overline{V_{1U}V_{2U}} - \overline{V_{1U}H_{1U}} - \overline{H_{2U}V_{2U}} \tag{5.15}$$

where V_{1U} denotes the front vertex and V_{2U} the back vertex of the lens system. Detailed measurement data for main lenses, given as f_{100}, f_{193} and f_{90}, are presented in Figs. 5.2 to 5.4. Parameter b_U remains variable, which in practice means to adjust the focus by shifting the entire objective lens.

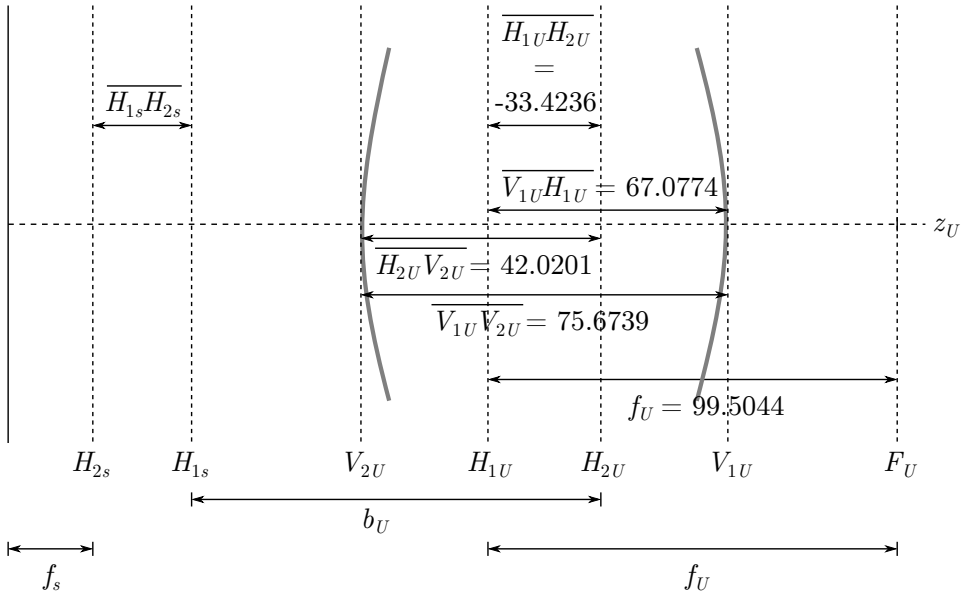


Figure 5.2: **Main lens specification for Double Gauss objective f_{100} from [89].** All measures are given in mm.

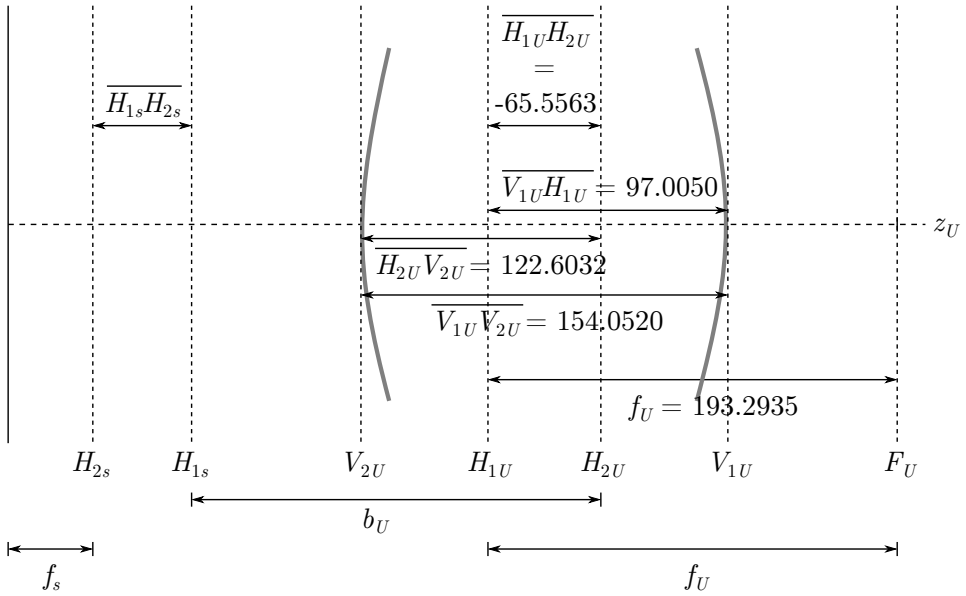


Figure 5.3: **Main lens specification for f_{193} objective lens based on [17].** All measures are given in mm.

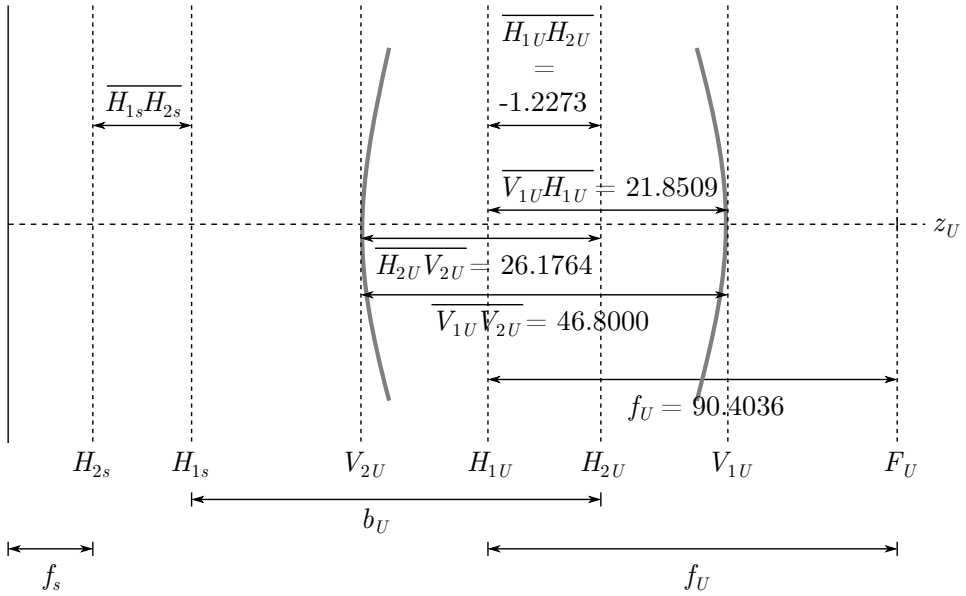


Figure 5.4: **Main lens specification for f_{90} objective lens based on [85].** All measures are given in mm.

Because Figs. 5.2 to 5.4 are conceptual simulation models which do not physically exist, another objective lens, denoted f_{197}, has been mounted to our customised SPC. A major problem with specifications of available objective lenses is that manufacturers have no interest in disclosing technical details about their products. Susanne Fiebig and Matthias Pesch from *ARRI Cine Technik GmbH & Co. KG* aided in the investigation of the f_{197} lens design by measuring the focal length and exit pupil with a collimator system that gradually changes the angle of incidence [75]. Principal planes were identified by Abbe's method [30]. Figure 5.5 depicts $\overline{H_{1U}H_{2U}}$ and f_U accordingly. An alternative method to measure real lens pupil positions is presented in Section 5.3.1.

Modern objective lenses are known to change the optical focus by shifting particular lens groups while other elements are static, which in turn alters cardinal point positions of that lens system. To preserve previously elaborated principal plane locations, a variation of the image distance b_U is achieved by shifting the MLA compound sensor away from the objective lens while its focus ring remains at infinity. The only limitation is, however, that the space inside our customised camera confines the shift range of the sensor system to an overall focus distance of $d_f \approx 4$ m. Due to this, solely two focus settings ($d_f \to \infty$ and $d_f \approx$ 4 m) are subject to examination in the following experiment.

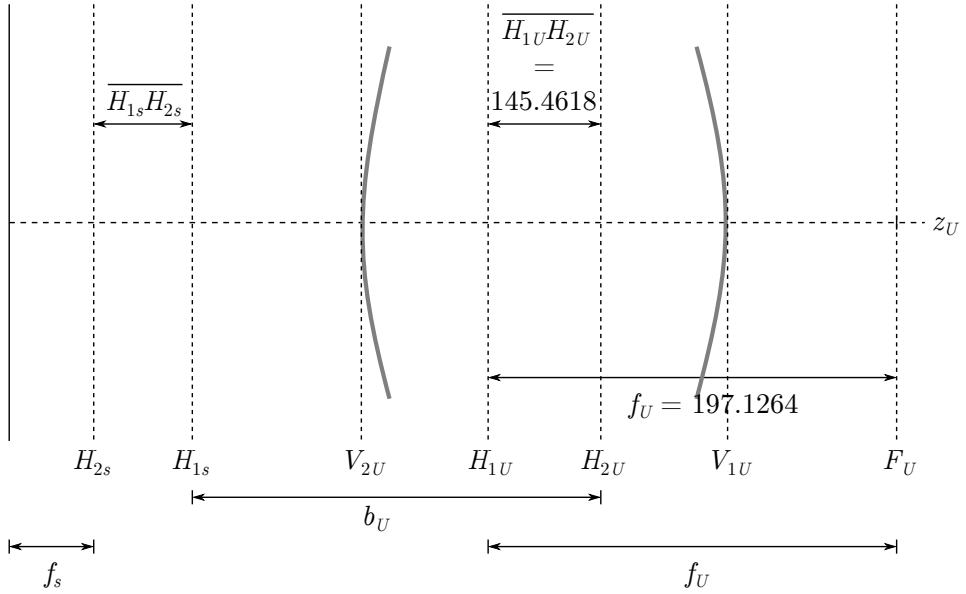


Figure 5.5: **Main lens specification for f_{197} objective lens measured with [75].** All measures are given in mm.

For validation purposes, the SPC ray model is evaluated under different main lens positions b_U. When shifting the main lens, b_U varies such that $b_U > f_U$. With respect to the thin lens equation, b_U is obtained via

$$b_U = \left(\frac{1}{f_U} - \frac{1}{a_U} \right)^{-1} , \tag{5.16}$$

where $a_U = d_f - b_U - \overline{H_{1U}H_{2U}}$ and d_f is the distance from MLA to the plane the main lens is focused on. By substituting for a_U, however, it becomes obvious that b_U is an input and output variable at the same time which gives a classical *chicken-and-egg* problem. To solve this, we initially set the input $b_U := f_U$, substitute the output b_U for the input variable and iterate this procedure until both b_U are identical. Calculated image and exit pupil distances for the presented objective lens specifications are found in Table 5.2.

Table 5.2: **Main lens parameters**.

Focus	**Image distance**				**Exit pupil position**			
d_f	b_U [mm]				$d_{A'}$ [mm]			
	f_{100}	f_{193}	f_{90}	f_{197}	f_{100}	f_{193}	f_{90}	f_{197}
∞	99.5044	193.2935	90.4036	197.1264	108.1293	111.0324	85.1198	100.5000
4 m	–	–	–	208.3930	–	–	–	111.7666
3 m	103.0019	207.3134	93.3043	–	111.6268	125.0523	88.0205	–
1.5 m	106.9659	225.8852	96.6224	–	115.5908	143.6241	91.3386	–

5.2 Ray Tracing Simulation

In order to prove claims made about the SPC ray model and its ability to predict the light ray geometry in 3-D space, the proposed principles may be analysed by modelling an SPC in a real ray simulation tool [89]. Due to the geometrical nature of the suggested aberration-free model, it is expected that results slightly deviate from the real ray tracing simulation which, by contrast, takes aberrations into account.

5.2.1 MIC Chief Ray Origin

The simplified model depicted in Fig. 3.4 implies that the main lens' principal plane and its aperture coincide. In real objective lenses, however, the position of their corresponding planes differ from each other. Thus, it raises the question whether chief rays impinging on a micro image centre rather originate from the principal plane or from an aperture plane. To discern between these possibilities, an experiment is conducted by calculating potential MICs and comparing the results with those obtained from the real ray simulation as reference. According to Chapter 3, MIC positions are denoted by $u_{c,j}$. To get an MIC u_{b_U} based on the principal plane position, we compute

$$u_{b_U} = \frac{s_j \times (b_U + f_s)}{b_U} \tag{5.17}$$

and do the same for the exit pupil dependent MICs $u_{d_{A'}}$ by calculating

$$u_{d_{A'}} = \frac{s_j \times (d_{A'} + f_s)}{d_{A'}}\,, \tag{5.18}$$

where $d_{A'}$ may be substituted by d_A to get MICs originating from the actual aperture A. In the experimental results shown in Table 5.3, the simulated light ray's wavelength amounts to $\lambda = 550$ nm and main objective lenses are focused to infinity ($b_U = f_U$) as rays emanate from there.

Table 5.3: **Evaluation of MIC chief ray origin where $b_U = f_U$.**

(a) f_{100} objective lens and MLA (II.) with $f_s = 2.75$ mm

$\frac{s_j}{p_{\hat{m}}}$	Simulation	$b_U = 99.5044$ mm		$d_A = 91.6030$ mm		$d_{A'} = 108.1293$ mm	
		u_{b_U} [mm]	ERR [%]	u_{d_A} [mm]	ERR [%]	$u_{d_{A'}}$ [mm]	ERR [%]
14	1.7945	1.7984	0.2169	1.8025	0.4438	1.7945	0.0000
28	3.5890	3.5967	0.2141	3.6051	0.4466	3.5890	0.0000
56	7.1773	7.1935	0.2252	7.2101	0.4549	7.1780	0.0098

(b) f_{193} objective lens and MLA (II.) $f_s = 2.75$ mm

$\frac{s_j}{p_{\hat{m}}}$	Simulation	$b_U = 193.2935$ mm		$d_A = 123.8701$ mm		$d_{A'} = 111.0324$ mm	
		u_{b_U} [mm]	ERR [%]	u_{d_A} [mm]	ERR [%]	$u_{d_{A'}}$ [mm]	ERR [%]
14	1.7933	1.7749	-1.0367	1.7889	-0.2460	1.7933	0.0000
28	3.5866	3.5498	-1.0367	3.5777	-0.2488	3.5867	0.0028
56	7.1730	7.0996	-1.0339	7.1554	-0.2460	7.1734	0.0056

(c) f_{193} objective lens and MLA (I.) with $f_s = 1.25$ mm

$\frac{s_j}{p_{\hat{m}}}$	Simulation	$b_U = 193.2935$ mm		$d_A = 123.8701$ mm		$d_{A'} = 111.0324$ mm	
		u_{b_U} [mm]	ERR [%]	u_{d_A} [mm]	ERR [%]	$u_{d_{A'}}$ [mm]	ERR [%]
14	1.7697	1.7613	-0.4769	1.7677	-0.1131	1.7697	0.0000
28	3.5394	3.5226	-0.4769	3.5353	-0.1160	3.5394	0.0000
56	7.0786	7.0453	-0.4727	7.0706	-0.1131	7.0788	0.0028

To ensure that experimental results do not rely on the focus setting of the objective lens, another testing has been carried out in which the objective lenses were shifted towards the object space focusing at 3 m. Results are provided in Table 5.4.

Table 5.4: **Evaluation of MIC chief ray origin where $b_U > f_U$.**

(a) f_{100} objective lens and MLA (II.) with $f_s = 2.75$ mm

$\frac{s_j}{p_{\hat{m}}}$	Simulation	$b_U = 103.0019$ mm		$d_A = 95.1005$ mm		$d_{A'} = 111.6268$ mm	
		u_{b_U} [mm]	ERR [%]	u_{d_A} [mm]	ERR [%]	$u_{d_{A'}}$ [mm]	ERR [%]
14	1.7931	1.7967	0.2004	1.8006	0.4165	1.7931	0.0000
28	3.5862	3.5934	0.2004	3.6012	0.4165	3.5862	0.0000
56	7.1718	7.1869	0.2101	7.2024	0.4249	7.1725	0.0084

(b) f_{193} objective lens and MLA (II.) with $f_s = 2.75$ mm

$\frac{s_j}{p_{\hat{m}}}$	Simulation	$b_U = 207.3134$ mm		$d_A = 137.8900$ mm		$d_{A'} = 125.0523$ mm	
		u_{b_U} [mm]	ERR [%]	u_{d_A} [mm]	ERR [%]	$u_{d_{A'}}$ [mm]	ERR [%]
14	1.7885	1.7732	-0.8629	1.7849	-0.2017	1.7885	0.0000
28	3.5769	3.5464	-0.8600	3.5698	-0.1989	3.5770	0.0028
56	7.1536	7.0929	-0.8558	7.1396	-0.1961	7.1539	0.0042

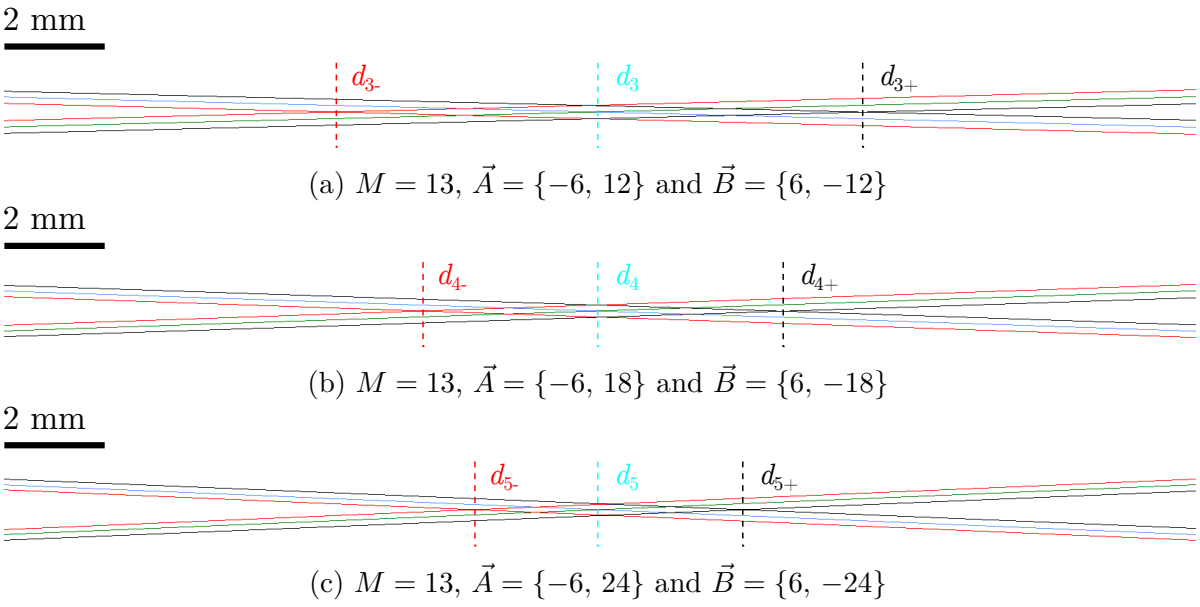
(c) f_{193} objective lens and MLA (I.) with $f_s = 1.25$ mm

$\frac{s_j}{p_{\hat{m}}}$	Simulation	$b_U = 207.3134$ mm		$d_A = 137.8900$ mm		$d_{A'} = 125.0523$ mm	
		u_{b_U} [mm]	ERR [%]	u_{d_A} [mm]	ERR [%]	$u_{d_{A'}}$ [mm]	ERR [%]
14	1.7675	1.7606	-0.3919	1.7659	-0.0906	1.7675	0.0000
28	3.5350	3.5211	-0.3948	3.5317	-0.0934	3.5350	0.0000
56	7.0698	7.0422	-0.3919	7.0635	-0.0892	7.0700	0.0028

Tables 5.3 and 5.4 signify a clear trend towards a lower error in predicting MICs at $u_{c,j}$ when micro lens chief rays originate from the exit pupil at A'. This matches the observation made in Fig. 4.4 where a varying aperture size A directly affects the micro image size by retaining the position of its central point. Remaining deviations are believed to be due to lens aberrations. Generally, the outcome of the experiments reinforces the hypothesis that chief rays coming from any position in object space cross the optical axis z_U at the exit pupil A'.

5.2.2 Distance Estimation

The validation of distance predictions using the real ray tracing simulation is achieved by firing off *central rays* from the sensor side into object space. However, *inner* and *outer rays* start from micro lens edges with a slope calculated from the respective pixel borders. The given pixel and micro lens pitch entail a micro image resolution of $M = 13$. Due to the paraxial approximation, rays starting from samples at the micro image border cause largest possible errors. To testify predictions limits, simulation results base on $\mathbf{A} = \{-6,\, e\}$ and $\mathbf{B} = \{6,\, e - a(M-1)\}$ unless specified otherwise. To align rays, e is dimensioned such that $\mathbf{A}$ and $\mathbf{B}$ are symmetric with an intersection close to the optical axis z_U (e.g. $e = 0, 6, 12, ...$). DoF rays $\mathbf{A}\pm$ and $\mathbf{B}\pm$ are fired from micro lens edges. Ray slopes build on MIC predictions obtained by Eq. (5.18). Simulated distances are measured by intersections of corresponding rays in object space.



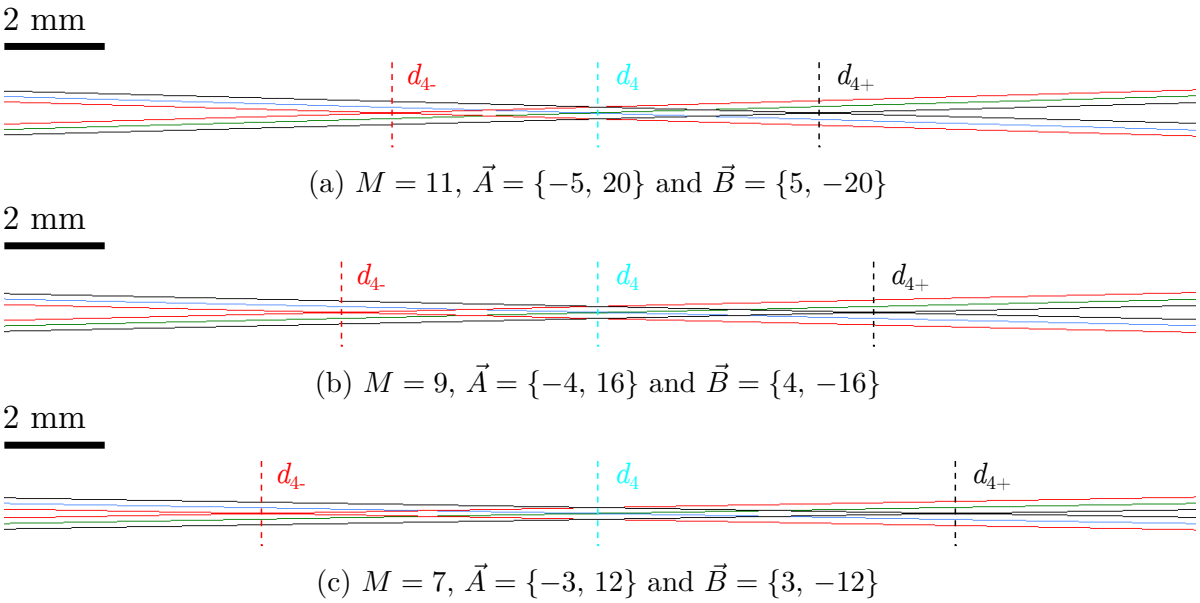
(a) $M = 13$, $\vec{A} = \{-6,\, 12\}$ and $\vec{B} = \{6,\, -12\}$

(b) $M = 13$, $\vec{A} = \{-6,\, 18\}$ and $\vec{B} = \{6,\, -18\}$

(c) $M = 13$, $\vec{A} = \{-6,\, 24\}$ and $\vec{B} = \{6,\, -24\}$

Figure 5.6: **Refocusing in real ray simulation.** (a) to (c) Intersecting *inner*, *outer* and *central rays* at different refocusing distances $d_{a\pm}$. Results are taken from the f_{90} lens with $d_f \to \infty$ focus, $M = 13$ and MLA (II.) with varying a. The consistent scale of the 3 examples suggests that the DoF shrinks with ascending a.

Exemplary screenshots are seen in Figs. 5.6 and 5.7. It is the observation in Fig. 5.6 that the DoF shrinks with increasing parameter a which reminds of the focus behaviour in traditional cameras. Figure 5.7 contains simulation results with a fixed a, but varying M. As anticipated in Chapter 3, a DoF becomes larger with less directional samples u and vice versa.



(a) $M = 11$, $\vec{A} = \{-5, 20\}$ and $\vec{B} = \{5, -20\}$

(b) $M = 9$, $\vec{A} = \{-4, 16\}$ and $\vec{B} = \{4, -16\}$

(c) $M = 7$, $\vec{A} = \{-3, 12\}$ and $\vec{B} = \{3, -12\}$

Figure 5.7: **Micro image size variation in real ray simulation** with f_{90} lens, $d_f \rightarrow \infty$ and MLA (II.). A DoF grows for a fix refocusing plane $(a = 4)$ by reducing samples in M as seen in (a) to (c).

A plot showing the geometrical ray model of the integrated distance prediction system is seen in Fig. 5.8. Proposed predictions are verified hereafter. To examine predicted distances, the validation is carried out for three main lens specifications at three different positions each and two MLAs.

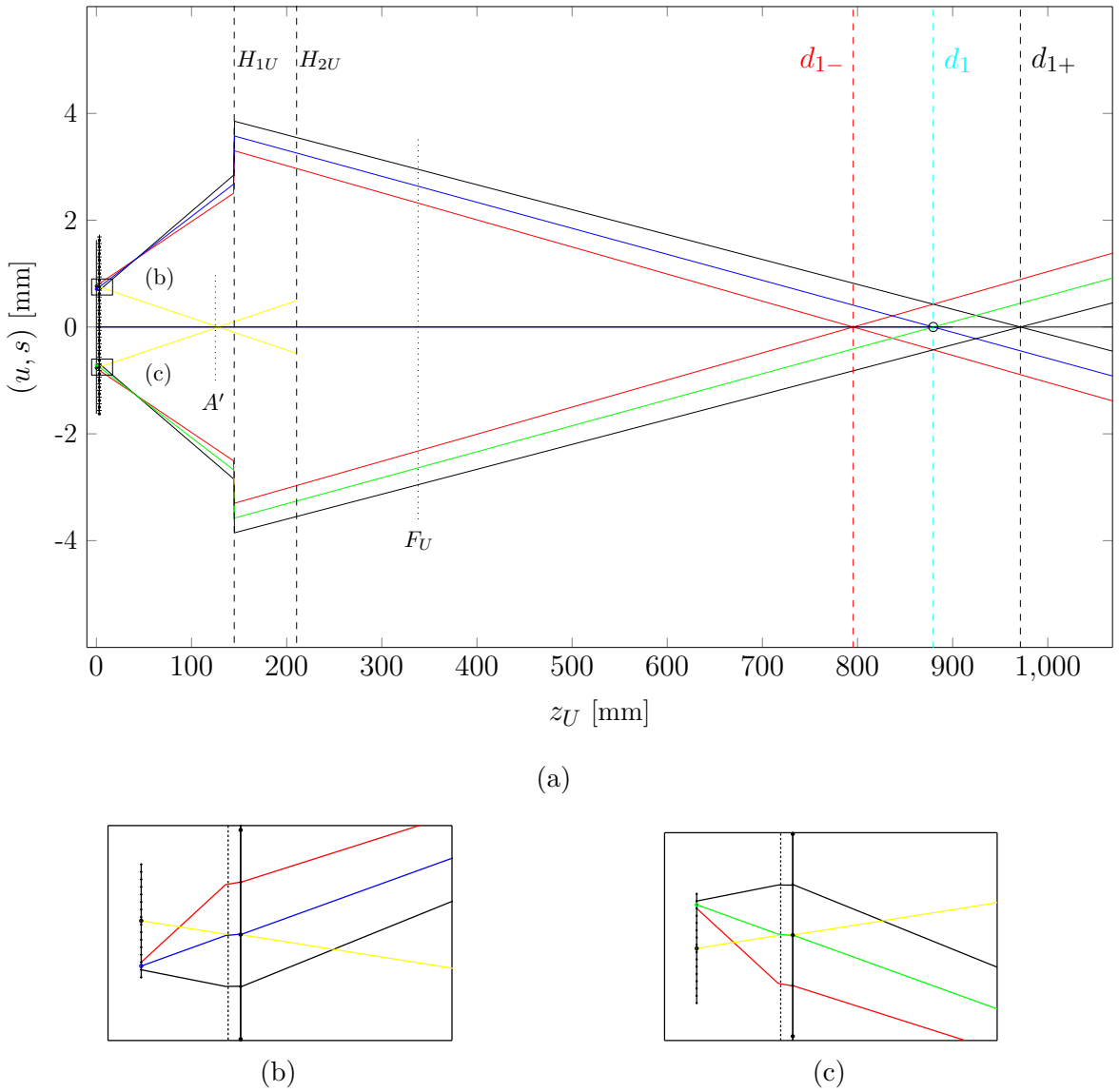


Figure 5.8: **Implemented refocusing distance prediction system.** (a) Paraxial ray tracing based on refraction at principal planes using f_{193} whilst $d_f = 3$ m and $a = 1$; (b) Magnified view of *central*, *inner* and *outer rays* at a micro lens; (c) Magnified view of ray intersection in object space.

Tables 5.5 to 5.7 show that each relative error of the refocusing distance prediction remains below 0.35 % ($|ERR_{a\pm}| < 0.35$ %). This is a significant improvement compared to previous results [40] which were up to 10 %. The main reason for the enhancement relies on the more accurate MIC prediction. Whilst [40] was based on the ideal SPC ray model where MICs are seen to be at the height of s_j, the refined model takes actual MICs into consideration by connecting chief rays from the exit pupil's centre to micro lens centres.

Refocusing on narrow planes is achieved with a successive increase in a. Thereby, prediction results move further away from the simulation which is reflected in slightly increasing errors. This may be explained by the fact that short distances d_a and $d_{a\pm}$ force light ray slopes to become steeper which counteracts the paraxial approximation in geometrical optics. As a consequence, aberrations come into play that are not taken into account, which in turn deteriorates the prediction accuracy.

Table 5.5: **Refocusing distance comparison for f_{193} with MLA (II.) and $M = 13$.**

Focus		**Prediction** [mm]			**Simulation** [mm]			**Deviation** [%]		
d_f	a	d_{a-}	d_a	d_{a+}	d_{a-}	d_a	d_{a+}	ERR_{a-}	ERR_a	ERR_{a+}
Inf	0	11 081.3954	∞	∞	11 080.4243	∞	∞	0.0088	–	–
	1	838.1359	962.7459	1110.0123	838.2854	962.8738	1110.1197	−0.0178	−0.0133	−0.0097
	2	428.4055	473.6385	522.8047	428.7008	473.9417	523.1184	−0.0689	−0.0640	−0.0600
	3	284.4462	310.6026	338.2537	284.8310	311.0122	338.6859	−0.1353	−0.1319	−0.1278
	4	210.9976	229.0847	247.9415	–	229.5672	248.4568	–	−0.2106	−0.2078
3 m	0	2471.0051	3000.0038	3706.5386	2470.8317	2999.7143	3706.0756	0.0070	0.0096	0.0125
	1	792.3090	876.4830	968.0428	792.4930	876.6591	968.2108	−0.0232	−0.0201	−0.0174
	2	471.3353	512.7312	556.2164	471.6550	513.0663	556.5697	−0.0678	−0.0654	−0.0635
	3	335.2111	362.1106	389.9630	335.6241	362.5552	390.4386	−0.1232	−0.1228	−0.1220
	4	259.9543	279.7459	300.0776	260.4367	280.2727	300.6458	−0.1856	−0.1883	−0.1894
1.5 m	-1	12 865.6844	15 787.9956	19 760.8327	12 864.2020	15 781.7890	19 759.8150	0.0115	0.0393	0.0052
	0	1389.6267	1499.9995	1613.5806	1389.6910	1500.0380	1613.5844	−0.0046	−0.0026	−0.0002
	1	742.0961	795.1380	848.9834	742.4041	795.4576	849.3141	−0.0415	−0.0402	−0.0390
	2	509.7028	544.4714	579.5960	510.1519	544.9504	580.1081	−0.0881	−0.0880	−0.0884
	3	390.1328	415.9684	442.0038	390.6861	416.5664	442.6483	−0.1418	−0.1438	−0.1458

When the objective lens is set to $d_f \to \infty$ ($a_U \to \infty$) and the refocusing value amounts to $a = 0$, *central rays* travel in a parallel manner whereas *outer rays* even diverge and therefore never intersect each other. In this case, only the distance to the nearby DoF border, also known as hyperfocal distance, can be obtained from the *inner rays*. This is given by d_{a-} in the first row of Table 5.5. The 4-th row of the measurement data where $a = 4$ and $d_f \to \infty$ for f_{193} contains an empty field in the d_{a-} simulation column. This is due to the fact that corresponding *inner rays* lead to an intersection inside the objective lens which turns out to be an invalid refocusing result. Since the new image distance is smaller than the focal length (${b_U}' < f_U$), results of this particular setting prove to be impractical as they exceed the natural focusing limit.

Despite promising results, the first set of analyses merely examined the impact of the focus distance d_f (a_U). In order to assess the effect of the MLA focal length parameter f_s, the simulation process has been repeated using MLA (I.) with results provided in Table 5.6. Comparing the outcomes with Table 5.5, distances $d_{a\pm}$ suggest that a reduction in f_s moves refocused object planes further away from the camera when $d_f \to \infty$. Interestingly, the opposite occurs when focusing with $d_f = 1.5$ m.

Table 5.6: **Refocusing distance comparison for f_{193} with MLA (I.) and $M = 13$.**

Focus		**Prediction** [mm]			**Simulation** [mm]			**Deviation** [%]		
d_f	a	d_{a-}	d_a	d_{a+}	d_{a-}	d_a	d_{a+}	ERR_{a-}	ERR_a	ERR_{a+}
Inf	0	23 993.8329	∞	∞	24 005.3662	∞	∞	−0.0481	–	–
	1	1831.4004	2136.6039	2497.2990	1831.2929	2136.7159	2496.7267	0.0059	−0.0052	0.0229
	2	944.9031	1060.5674	1186.2896	945.5058	1061.2475	1186.7985	−0.0638	−0.0641	−0.0429
	3	633.4310	701.8886	774.2581	634.4282	702.9615	775.3241	−0.1574	−0.1529	−0.1377
	4	474.5167	522.5492	572.6256	475.8015	523.9335	574.0820	−0.2708	−0.2649	−0.2543
3 m	0	2730.3538	3000.0038	3280.5813	2729.6673	2999.1804	3279.4685	0.0251	0.0274	0.0339
	1	1313.1473	1428.1764	1545.4024	1313.4978	1428.6282	1545.6197	−0.0267	−0.0316	−0.0141
	2	864.1231	936.8289	1010.4430	864.9995	937.8155	1011.3707	−0.1014	−0.1053	−0.0918
	3	643.7518	696.8422	750.4255	644.9925	698.2172	751.8347	−0.1927	−0.1973	−0.1878
	4	512.8249	554.6198	596.7234	514.3536	556.2956	598.5153	−0.2981	−0.3022	−0.3003
1.5 m	-1	2519.1508	2538.1912	2554.8669	2518.5193	2536.7834	2554.0951	0.0251	0.0555	0.0302
	0	1447.0493	1499.9995	1548.9802	1447.4099	1500.3766	1544.3287	−0.0249	−0.0251	0.3003
	1	1018.1263	1067.4629	1114.1507	1019.0813	1068.5581	1115.2412	−0.0938	−0.1026	−0.0979
	2	787.1482	830.2762	871.6006	788.5313	831.8363	873.2308	−0.1757	−0.1879	−0.1870
	3	642.7800	680.4780	716.8850	644.5049	682.4038	718.9492	−0.2683	−0.2830	−0.2879

According to the data in Tables 5.5 and 5.6, we can infer that $d_a \approx d_f$ if $a = 0$ which means that synthetically focusing with $a = 0$ results in a focusing distance as with a conventional camera having a traditional sensor at the position of the MLA.

A third experimental validation was undertaken to investigate the impact of the main lens focal length parameter f_U. As Table 5.7 shows, using a shorter f_U implies a rapid decline in $d_{a\pm}$ with ascending a. From this observation it follows that the depth sampling rate of refocusing slices is much denser for large focal lengths f_U.

Table 5.7: **Refocusing distance comparison for f_{90} with MLA (II.) and $M = 13$.**

Focus		**Prediction [mm]**			**Simulation [mm]**			**Deviation [%]**		
d_f	a	d_{a-}	d_a	d_{a+}	d_{a-}	d_a	d_{a+}	ERR_{a-}	ERR_a	ERR_{a+}
Inf	0	2533.3495	∞	∞	2539.1267	∞	∞	−0.2280	–	–
	1	272.0095	297.5436	327.7202	272.0365	297.5765	327.7656	−0.0099	−0.0111	−0.0139
	2	181.5560	190.5541	200.3347	181.5743	190.5751	200.3589	−0.0101	−0.0110	−0.0121
	3	149.7750	154.8909	160.2992	149.7854	154.9052	160.3165	−0.0069	−0.0092	−0.0108
	4	133.5602	137.0593	140.7074	133.5574	137.0607	140.7129	0.0021	−0.0010	−0.0039
3 m	0	1464.9082	3000.0112	∞	1466.7026	3009.8165	∞	−0.1225	−0.3268	–
	1	274.7098	298.6157	326.3330	274.7386	298.6508	326.3803	−0.0105	−0.0118	−0.0145
	2	187.4853	196.5484	206.3334	187.5053	196.5711	206.3597	−0.0107	−0.0115	−0.0127
	3	155.4984	160.7901	166.3628	155.5109	160.8065	166.3825	−0.0080	−0.0102	−0.0118
	4	138.8965	142.5667	146.3829	138.8966	142.5712	146.3916	−0.0001	−0.0032	−0.0059
1.5 m	0	1029.1371	1500.0090	2676.9278	1029.9600	1502.2530	2685.9575	−0.0800	−0.1496	−0.3373
	1	277.3997	299.5258	324.6638	277.4306	299.5632	324.7134	−0.0111	−0.0125	−0.0153
	2	193.9990	203.0798	212.8082	194.0208	203.1047	212.8369	−0.0112	−0.0123	−0.0135
	3	161.9285	167.3927	173.1209	161.9433	167.4116	173.1432	−0.0091	−0.0113	−0.0129
	4	144.9515	148.8024	152.7939	144.9548	148.8103	152.8060	−0.0023	−0.0053	−0.0079

Considering simulated testings, it can be stated that the refocusing distance $d_{a\pm}$ drops with

- decreasing main lens focusing distance d_f
- ascending refocusing parameter a
- enlarging MLA focal length f_s
- reducing objective lens focal length f_U

and vice versa. A visualisation of the measurement data from Tables 5.5 to 5.7 is shown in Fig. 5.9.

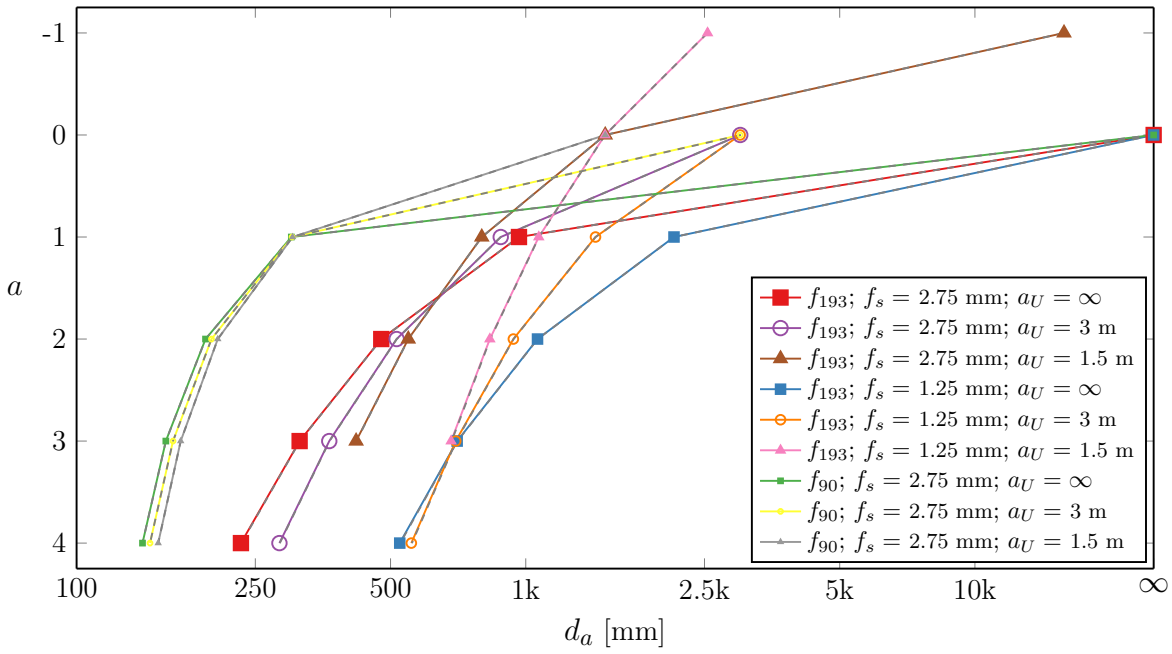


Figure 5.9: **Plot of predicted and simulated distances d_a.** Semi-log scale diagram based on data taken from Tables 5.5 to 5.7. Coloured graphs visualise the impact of varying lens parameters (f_U, f_s, d_f) on the refocusing performance. Grey dashed lines (- - -) represent simulation measures.

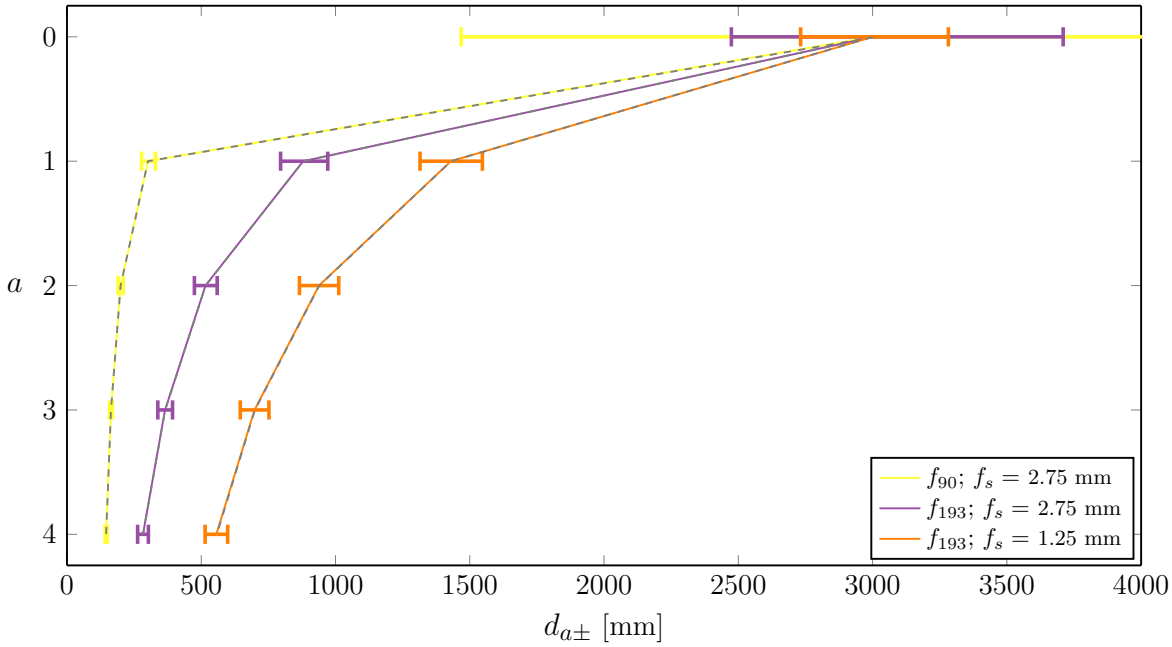


Figure 5.10: **DoF comparison** with different f_U and f_s when $a_U = 3$ m. For main lens f_{90}, DoF border $d_{0+} \to \infty$.

A linear plot in Fig. 5.10 contains confidence intervals illustrating the DoF. It becomes apparent that there are gaps between refocused DoF slices a suggesting that objects placed at these distances could not be refocused. However, these depth ranges can be retrieved by letting $a \in \mathbb{Q}$ as proposed in Eq. (3.16) which takes sub-pixels into account. The simulation experiment has been confined to integer values ($a \in \mathbb{Z}$) as $a \in \mathbb{Q}$ requires an interpolation of micro lenses which proved to be impractical when simulating light ray trajectories.

It can be concluded that tracing rays according to our model yields more accurate results in the optical design software [89] than by solving Eq. (3.40). However, deviations of less than 0.35 % appear to be insignificant. Implementing the model with a high-level programming language outperforms the real ray simulation in terms of computation time. Using a timer, the image-side based method presented in Section 3.3 takes about 0.002 to 0.005 seconds to compute d_a and $d_{a\pm}$ for each a on an Intel Core i7-3770 CPU @ 3.40 GHz system whereas modelling a plenoptic lens design and measuring distances by hand can take up to a business day. The time difference in object- and image-side based approach is negligible, however, in favour of the latter.

5.2.3 Virtual Camera Array

The positioning of the virtual camera array under different focal length and focus settings is examined in this section. Table 5.8 reveals a comparison of predicted and simulated virtual camera positions just as their baseline B_G and relative tilt angle Φ_G. Thereby, the distance from front vertex V_{1U} to entrance pupil A'' is given by

$$\overline{V_{1U}A''} = b_U + \overline{H_{1U}H_{2U}} + \overline{V_{1U}H_{1U}} - d_{A''} \tag{5.19}$$

bearing in mind that $d_{A''}$ is the distance from MLA to entrance pupil and $\overline{V_{1U}H_{1U}}$ separates the objective's front vertex V_{1U} from its object side principal plane H_{1U}. Simulated $\overline{V_{1U}A''}$ are obtained by extending ray slopes $q_{i,j}$ towards the sensor whilst these virtually elongated rays are seen to ignore lenses and finding the intersection of $q_{i,j}$ and $q_{i,j+1}$.

Table 5.8: **Baseline and tilt angle simulation with $G = 6$ and $i = 0$.**

Setup			**Prediction**			**Simulation**			**Deviation** [%]		
d_f	f_U	f_s	$\overline{V_{1U}A''}$ [mm]	B_G [mm]	Φ_i [°]	$\overline{V_{1U}A''}$ [mm]	B_G [mm]	Φ_i [°]	$ERR_{\overline{V_{1U}A''}}$	ERR_{B_G}	ERR_{Φ_i}
Inf	f_{193}	(II.)	240.2113	3.7956	0.0000	240.1483	3.7949	0.0000	0.0262	0.0184	–
	f_{90}	(II.)	27.4627	1.7752	0.0000	27.4081	1.7748	0.0001	0.1988	0.0225	–
	f_{193}	(I.)	240.2113	8.3503	0.0000	239.3988	8.3450	0.0000	0.3382	0.0635	–
3 m	f_{193}	(II.)	240.2113	4.2748	−0.0816	239.8612	4.2738	−0.0816	0.1457	0.0234	0.0000
	f_{90}	(II.)	27.4627	1.8357	−0.0361	27.3309	1.8352	−0.0360	0.4799	0.0272	0.2770
	f_{193}	(I.)	240.2113	9.4047	−0.1795	238.9043	9.3964	−0.1795	0.5441	0.0883	0.0000
1.5 m	f_{193}	(II.)	240.2113	4.9097	−0.1897	239.6932	4.9078	−0.1897	0.2157	0.0387	0.0000
	f_{90}	(II.)	27.4627	1.9049	−0.0774	27.2150	1.9042	−0.0773	0.9020	0.0367	0.1292
	f_{193}	(I.)	240.2113	10.8014	−0.4173	238.1212	10.7866	−0.4173	0.8701	0.1370	0.0000

Observations in Table 5.8 give indication that the baseline grows with

- larger main lens focal length f_U
- shorter micro lens focal length f_s
- decreasing focusing distance d_f (a_U)

given that the entrance pupil diameter is large enough to accommodate the baseline. Besides, it has been proven that tilt angle rotations become larger with decreasing d_f. Baselines have been estimated appropriately with errors below 0.1 % on average except for one example. The key problem causing the largest error is that MLA (I.) features a shorter focal length f_s than MLA (II.) which produces steeper light ray slopes $m_{c+i,j}$ and hence severe aberration effects. Tilt angle errors remain below 0.3 % although results deviate by only 0.001° for f_{90} and are even non-existent for f_{193}. However, entrance pupil location errors of about ≤ 1 % are larger than in any other simulated validation. One reason for these inaccuracies is that the entrance pupil A'' is an imaginary vertical plane which in reality may exhibit a slight spherical shape around the optical axis.

Table 5.9: **Disparity simulation and distance with $G = 6$ and $i = 0$.**

Setup		**Prediction**			**Simulation**			**Deviation**		
d_f	Δx	$Z_{6,\Delta x}$ [mm]			$Z_{6,\Delta x}$ [mm]			$ERR_{Z_{6,\Delta x}}$ [%]		
		f_{193} & (II.)	f_{90} & (II.)	f_{193} & (I.)	f_{193} & (II.)	f_{90} & (II.)	f_{193} & (I.)	f_{193} & (II.)	f_{90} & (II.)	f_{193} & (I.)
Inf	0	∞	∞	∞	∞	∞	∞	–	–	–
	1	978.2150	213.9790	2152.0729	978.2797	213.9573	2151.2840	−0.0066	0.0101	0.0367
	2	489.1075	106.9895	1076.0365	489.1026	106.9431	1075.1177	0.0010	0.0434	0.0854
3 m	0	3001.4530	2913.5460	3001.4530	3000.8133	2923.2193	2999.3120	0.0213	−0.3320	0.0713
	1	877.9068	212.1505	1429.6116	877.4653	212.0285	1427.8084	0.0503	0.0575	0.1261
	2	514.1456	110.0831	938.2541	513.8952	109.9610	937.1572	0.0487	0.1109	0.1169
1.5 m	-1	15 770.8729	–	2521.0686	15 764.1482	–	2517.6509	0.0426	–	0.1356
	0	1482.8768	1410.2257	1482.8768	1482.3969	1412.2221	1481.1620	0.0324	−0.1416	0.1156
	1	778.0154	209.7424	1050.3402	777.8168	209.5320	1049.3327	0.0255	0.1003	0.0959
	2	527.3487	113.2965	813.1535	527.0279	113.0602	811.8298	0.0608	0.2086	0.1628

An experiment assessing the relationship between disparity Δx and distance $Z_{G,\Delta x}$ using different objective lenses is presented in Table 5.9. From this, it can be concluded that denser depth sampling is achieved with larger main lens focal length f_U. Moreover, it is seen that a tilt in virtual cameras yields a negative disparity Δx for objects further away than d_f which is a phenomenon that also applies to tilted cameras in stereoscopy. The reason why $d_f \approx Z_{G,\Delta x}$ when $\Delta x = 0$ is that $Z_{G,\Delta x}$ reflects the separation between ray intersection and entrance pupil A'' which lies nearby the sensor and d_f is the spacing between intersection and sensor plane. There are only rare cases in which A'' directly coincides with the sensor's plane. Overall, it can be stated that distance estimates based on the stereo triangulation behave similar to those in the refocusing distances with errors of up to ± 0.33 %.

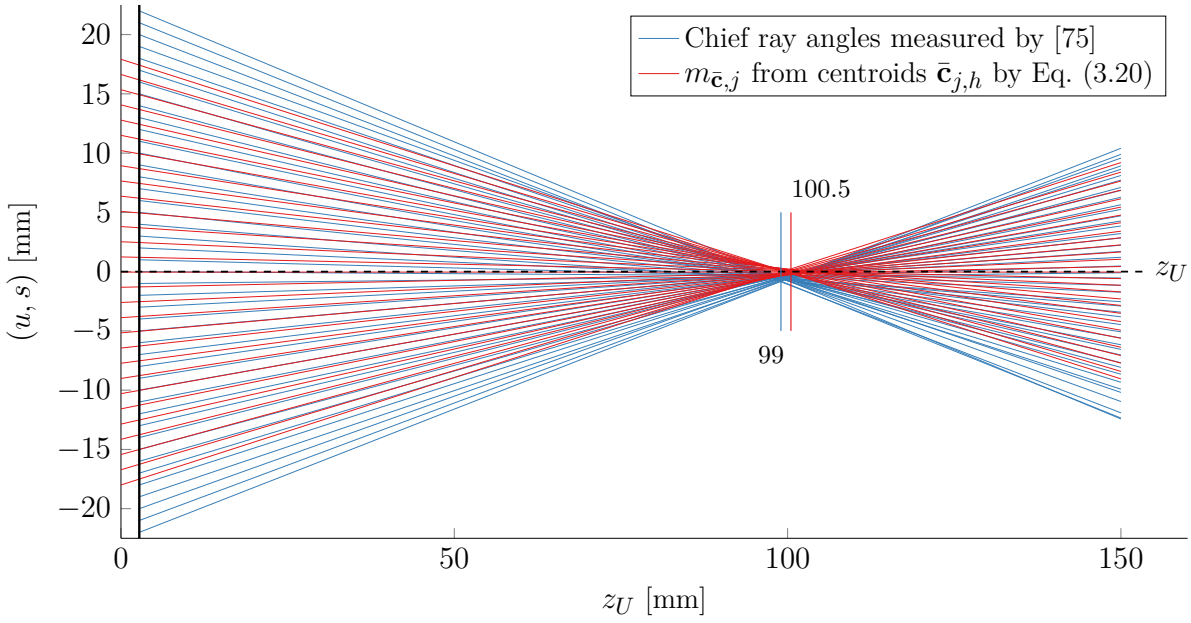
5.3 Experimental Results

This part aims to experimentally evaluate estimates of the refocusing distance as well as those of the sub-apertures' baseline and tilt angle by comparing proposed predictions and real image captures taken by a customised SPC. All experiments are carried out with MLA (II.) as specified in Table 5.1 and objective lens design f_{197} as provided in Section 5.1. MLA (I.) cannot be utilised for the real image validation since the thickness of the deployed sensor's cover glass (≈ 1 mm) blocks the proper back vertex position of $d_s = 0.5460$ mm < 1 mm. Supplementary data such as chief ray plots are provided at the beginning to verify their origin and give a better insight of the plenoptic lens system.

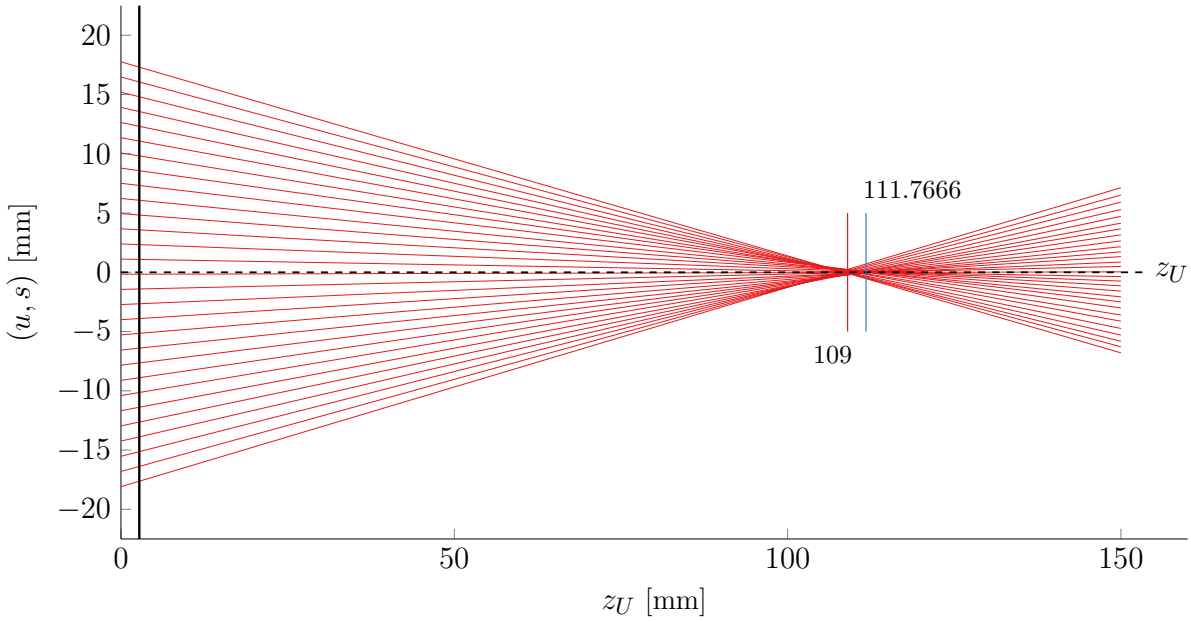
5.3.1 MIC Chief Ray Origin

In Chapter 3 it has been stated that refocusing distance estimates require knowledge about the exit pupil position $d_{A'}$. In case of an unknown exit pupil location, an ascertainment of such has to be conducted prior to predictions. As seen in Fig. 3.3, $d_{A'}$ can be obtained by chief rays of MICs which intersect the objective lens axis z_U. In the customised SPC, chief ray slopes $m_{\bar{\mathbf{c}},j}$ may be retrieved according to Eq. (3.20) using centroids $\bar{\mathbf{c}}$ from the algorithmic MIC computation in Eq. (4.3) and respective micro lens centre positions s_j. Thereby, it is supposed that the optical centre of the central micro lens and its micro image centroid $\bar{\mathbf{c}}$ share the same respective position. To the best of the writer's knowledge, plotting $m_{\bar{\mathbf{c}},j}$ for each lens j to identify the exit pupil position is the first method to do so by employing an MLA. The same applies to an entrance pupil detection when flipping the main lens before capturing the white image. The white image acquisition is discussed in Chapter 4.

Nevertheless, other attempts exist to investigate pupil locations. One of them [75] is to illuminate the objective lens with collimated light from varying angle positions. Thereby, the sensor is focused on the lens focal plane and shifted to the respective image height. Subsequently, the sensor moves towards the lens stepwise while an algorithm captures the lateral movement of the image. Based on this lateral movement, chief ray angles are retrieved. Figure 5.11 shows chief ray angles of an objective lens featuring the proposed method (red lines) and the collimation-based exit pupil measurement (blue lines) for verification purposes. Results of the latter were elaborated by Susanne Fiebig and Matthias Pesch at *ARRI Cine Technik GmbH & Co. KG* using the collimator technique [75].



(a)



(b)

Figure 5.11: **MIC chief ray origin plot for f_{197} objective lens** with (a) infinity focus ($d_f \to \infty$) and (b) close focus ($d_f \approx 4$ m). The zero point on the abscissa z_U represents the sensor position and the vertical line at 2.75 mm is the MLA. Blue lines indicate chief ray angle results from the exit pupil measurement [75]. Red lines are chief ray slopes calculated by Eq. (3.20) from centroids $\bar{\mathbf{c}}_{j,h}$ obtained by our customised SPC with MLA (II.). Slight deviations in exit pupil localisation may be due to calibration misalignments.

As seen in Fig. 5.11, results of the two different exit pupil measurements nearly coincide. Minor deviations of approximately 2.5 % are seen to be due to a slightly misaligned MLA position and hence could be improved. Whilst results of both attempts are similar, the collimation-based approach [75] appears to be more sophisticated and expensive in terms of the mechanical components being involved. Thus, the proposed method may be a cost-effective alternative to existing procedures.

5.3.2 Distance Estimation

On the basis of raw light field photographs, this section aims to evaluate the accuracy of predicted refocusing distances as proposed in Chapter 3. The challenge here is to verify whether objects placed at a known distance exhibit best focus at the predicted refocusing distance. Hence, the evaluation requires an algorithm to sweep for blurred regions in a stack of photographs with varying focus. One obvious attempt to measure the blur of an image is to analyse them in frequency domain. Mavridaki and Mezaris [57] follow this principle in a recent study to assess the blur in a single image. To employ their proposition, modifications are necessary as the distance validation requires the degree of blur to be detected for particular image portions in a refocused photograph stack. Here, the conceptual idea is to select a *Region of Interest* (RoI) surrounding the same object in each refocused image. Unlike in Chapter 3 where the vertical index h in t_h is constant for conciseness, refocused images may be regarded in vertical and horizontal direction in this section such that a refocused photograph in 2-D is given as $E''_a\left[s_j\ ,\ t_h\right]$. A RoI is a cropped version of a refocused photograph that can be selected as desired with image borders spanning from the ξ-th to Ξ-th pixel in horizontal and the ϖ-th to Π-th pixel in vertical direction. Care has been taken to ensure that a RoI's bounding box precisely crops the object at the same relative position in each image of the focal stack. When *Fourier*-transforming all RoIs of the focal stack, the key indicator for a blurred RoI is a decrease in its high frequency power. To implement the proposed concept, we first perform the 2-D *Discrete Fourier Transformation* and extract the magnitude $\mathcal{X}\left[\sigma_\omega\ ,\ \rho_\psi\right]$ as given by

(5.20)
$$\mathcal{X}\left[\sigma_\omega\ ,\ \rho_\psi\right] = \left|\sum_{j=\xi}^{\Xi-1}\sum_{h=\varpi}^{\Pi-1} E''_a\left[s_j\ ,\ t_h\right]\exp\left(-2\pi\kappa(j\omega/(\Xi-\xi)+h\psi/(\Pi-\varpi))\right)\right|$$

whilst $\kappa=\sqrt{-1}$ is the complex number and $|\cdot|$ computes the absolute value, leaving out the phase spectrum.

Provided the 2-D magnitude signal $\mathcal{X}\left[\sigma_\omega\,,\,\rho_\psi\right]$, its total energy TE is computed via

$$TE = \sum_{\omega=0}^{\Omega-1}\sum_{\psi=0}^{\Psi-1}\mathcal{X}\left[\sigma_\omega\,,\,\rho_\psi\right]^2 \tag{5.21}$$

with $\Omega = \lceil{}^{(\Xi\,-\,\xi)}/_{2}\rceil$ and $\Psi = \lceil{}^{(\Pi\,-\,\varpi)}/_{2}\rceil$ as borders cropping the first quarter of the unshifted magnitude signal. In order to identify the energy of high frequency elements HE, we calculate the power of low frequencies and subtract them from TE as seen in

$$HE = TE - \sum_{\omega=0}^{Q_H}\sum_{\psi=0}^{Q_V}\mathcal{X}\left[\sigma_\omega\,,\,\rho_\psi\right]^2 \tag{5.22}$$

where Q_H and Q_V are limits in the range of $\{1,\,..\,,\,\Omega-1\}$ and $\{1,\,..\,,\,\Psi-1\}$ separating low from high frequencies. Finally, the sharpness S of a refocused RoI is obtained by

$$S = \frac{HE}{TE} \tag{5.23}$$

serving as the blur metric. Thus, each RoI focal stack produces a set of S values which is normalised and given as a function of the refocusing variable a. The maximum in S thereby indicates best focus for a selected RoI object at the respective a.

To benchmark proposed refocusing distance estimates, an experiment is conducted similar to that from a previous publication [41]. As opposed to [41] where b_U was taken as the MIC chief ray origin, here $d_{A'}$ is given as the origin for rays that lead to MIC positions. Besides this, frequency borders $Q_H = \Omega/100$ and $Q_V = \Psi/100$ are relative to the cropped image resolution. It is also expected that the more precisely measured objective lens parameters b_U and $\overline{H_{1U}H_{2U}}$ have a positive effect on the result.

To make statements about the model accuracy, real objects are placed at predicted distances d_a. Recall that d_a is the distance from MLA to a respective refocusing plane a. As the MLA is embedded in the camera body and hence inaccessible, the objective lens' front panel was chosen to be the distance measurement origin for d_a. This causes a displacement of 12.7 cm towards object space ($d_a - 12.7$ cm), which has been accounted for in the predictions of d_a presented in Table 5.10. Moreover, Table 5.10 lists predicted DoF borders $d_{a\pm}$ at different settings M and b_U while highlighting planes a where objects were placed.

Table 5.10: **Predicted refocusing distances d_a and $d_{a\pm}$**

(a) $d_f \to \infty$, $M = 9$

a	d_{a-} [cm]	d_a [cm]	d_{a+} [cm]	Object
$^0/_9$	753	∞	∞	
$^1/_9$	400	918	∞	
$^2/_9$	273	460	1130	
$^3/_9$	207	308	529	Test chart
$^4/_9$	167	232	346	
$^5/_9$	141	186	257	Colour chart
$^6/_9$	121	155	205	
$^7/_9$	107	133	171	Striped figure
$^8/_9$	95	117	146	
$^9/_9$	86	104	128	Spiderman
$^{10}/_9$	79	94	114	
$^{11}/_9$	72	86	103	
$^{12}/_9$	67	79	93	80 cm Marker
$^{13}/_9$	62	73	86	
$^{14}/_9$	58	68	79	

(b) $d_f \to \infty$, $M = 11$

a	d_{a-} [cm]	d_a [cm]	d_{a+} [cm]	Object
$^0/_{11}$	957	∞	∞	
$^1/_{11}$	502	1122	∞	
$^2/_{11}$	341	562	1323	
$^3/_{11}$	259	376	628	
$^4/_{11}$	208	282	412	Test chart
$^5/_{11}$	175	226	307	
$^6/_{11}$	150	189	245	Colour chart
$^7/_{11}$	132	162	204	
$^8/_{11}$	118	142	175	Striped figure
$^9/_{11}$	107	127	153	
$^{10}/_{11}$	97	115	136	Spiderman
$^{11}/_{11}$	89	104	123	
$^{12}/_{11}$	83	96	112	
$^{13}/_{11}$	77	89	102	90 cm marker
$^{14}/_{11}$	72	83	95	

(c) $d_f \approx 4$ m, $M = 11$

a	d_{a-} [cm]	d_a [cm]	d_{a+} [cm]	Object
$^0/_{11}$	293	387	541	Test chart
$^1/_{11}$	240	304	398	
$^2/_{11}$	204	251	315	Colour chart
$^3/_{11}$	177	213	260	
$^4/_{11}$	156	186	222	Striped figure
$^5/_{11}$	140	165	194	
$^6/_{11}$	127	148	172	Spiderman
$^7/_{11}$	116	134	155	
$^8/_{11}$	107	123	141	120 cm marker
$^9/_{11}$	100	114	129	
$^{10}/_{11}$	93	105	119	
$^{11}/_{11}$	87	98	111	100 cm marker
$^{12}/_{11}$	82	92	104	
$^{13}/_{11}$	78	87	97	
$^{14}/_{11}$	74	82	92	

To validate the model's distance estimation performance, the main lens focus d_f and micro image resolution M vary in the experimental setup. An unprocessed light field photograph taken by our customised SPC according to the scene composition given in Table 5.10(a) is depicted in Fig. 5.12.

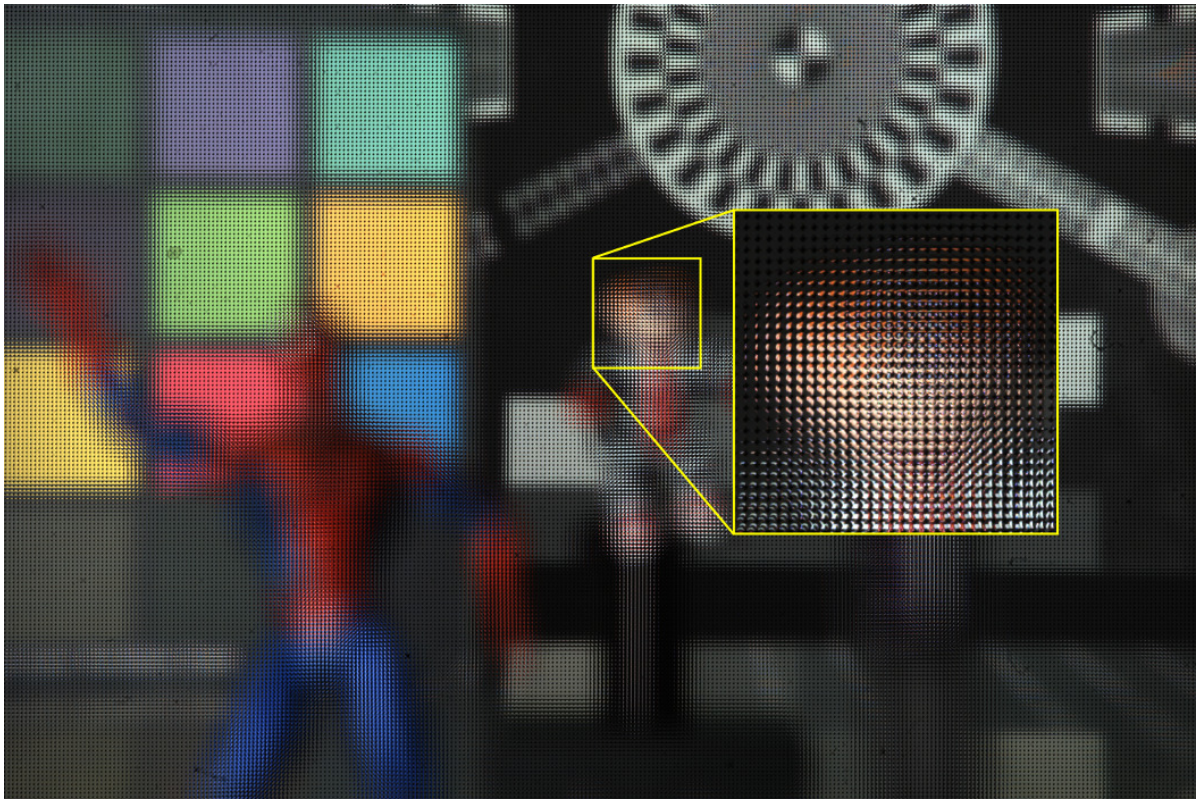


Figure 5.12: **Raw light field photograph with $b_U = f_U$.** The depicted image has been taken by our customised SPC with $d_f \to \infty$ and was used to compute refocused photographs shown in Fig. 5.13. The magnified view exhibits detailed micro images.

Figures 5.13 to 5.15 reveal outcomes of the refocusing distance validation by showing refocused images $E''_a\left[s_j\,,\,t_h\right]$ and RoIs at different slices a as well as related blur metric results S. The reason why S produces relatively large values around predicted blur metric peaks is that some objects may lie within the DoF of adjacent slices a and thus can be in focus among several refocusing slices. Taking slice $a = {}^{4}\!/_{11}$ from Table 5.10(c) as an example, it becomes obvious that its object distance $d_{4/11} = 186$ cm falls inside the DoF range of slice $a = {}^{5}\!/_{11}$ with $d_{5/11+} = 194$ cm and $d_{5/11-} = 140$ cm because $d_{5/11+} > d_{4/11} > d_{5/11-}$. Section 5.2.2 shows that reducing the micro image resolution M yields a narrower DoF which suggests to use largest possible M as this minimises the effect of wide DoFs. Experimentations given in Figs. 5.14 and 5.15 were carried out with maximum directional resolution $M = 11$ since $M = 13$ would involve pixels that start to suffer from vignetting and micro image crosstalk. Although objects are covered by DoFs of surrounding slices, the presented blur metric still detects predicted sharpness peaks as seen in Figs. 5.13 to 5.15.

(a) Refocused photograph E''_a with $a = 3/9$

(b) Refocused photograph E''_a with $a = 12/9$

(c) $a = 1/9$, $S = 0.78$

(d) $a = 2/9$, $S = 0.94$

(e) $a = 3/9$, $S = 1.00$

(f) $a = 4/9$, $S = 0.95$

(g) $a = 5/9$, $S = 0.80$

(h) $a = 3/9$, $S = 0.88$

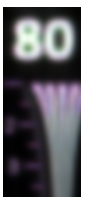
(i) $a = 4/9$, $S = 0.95$

(j) $a = 5/9$, $S = 1.00$

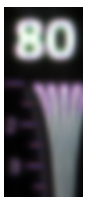
(k) $a = 6/9$, $S = 0.98$

(l) $a = 7/9$, $S = 0.92$

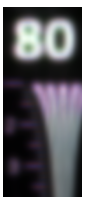
(m) $a = 10/9$, $S = 0.86$



(n) $a = 11/9$, $S = 0.96$



(o) $a = 12/9$, $S = 1.00$



(p) $a = 13/9$, $S = 0.95$

(q) $a = 14/9$, $S = 0.83$

Figure 5.13: **Regions of refocused photographs with $b_U = f_U$ and $M = 9$.** The main lens was focused at infinity ($d_f \to \infty$). S denotes the measured sharpness in (c) to (q). The denominator in a indicates the upsampling factor for the linear interpolation of whole micro images. The RoIs' illuminance values E''_a have been normalised for better blur detection.

(a) Refocused photograph E''_a with $a = {}^{8}/_{11}$

(b) Refocused photograph E''_a with $a = {}^{13}/_{11}$

(c) $a = {}^{2}/_{11}$, $S = 0.77$

(d) $a = {}^{3}/_{11}$, $S = 0.91$

(e) $a = {}^{4}/_{11}$, $S = 1.00$

(f) $a = {}^{5}/_{11}$, $S = 0.97$

(g) $a = {}^{6}/_{11}$, $S = 0.82$

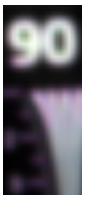
(h) $a = {}^{4}/_{11}$, $S = 0.83$

(i) $a = {}^{5}/_{11}$, $S = 0.93$

(j) $a = {}^{6}/_{11}$, $S = 1.00$

(k) $a = {}^{7}/_{11}$, $S = 0.98$

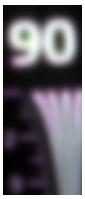
(l) $a = {}^{8}/_{11}$, $S = 0.90$



(m) $a = {}^{11}/_{11}$, $S = 0.82$

(n) $a = {}^{12}/_{11}$, $S = 0.95$

(o) $a = {}^{13}/_{11}$, $S = 1.00$



(p) $a = {}^{14}/_{11}$, $S = 0.95$

(q) $a = {}^{15}/_{11}$, $S = 0.82$

Figure 5.14: **Regions of refocused photographs with $b_U = f_U$ and $M = 11$.** The main lens was focused at infinity ($d_f \to \infty$). S denotes the measured sharpness in (c) to (q). The denominator in a indicates the upsampling factor for the linear interpolation of whole micro images. The RoIs' illuminance values E''_a have been normalised for better blur detection.

(a) Refocused photograph E''_a with $a = {}^{6}/_{11}$

(b) Refocused photograph E''_a with $a = {}^{11}/_{11}$

(c) $a = {}^{-2}/_{11}$, $S = 0.73$

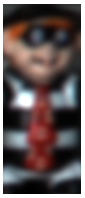
(d) $a = {}^{-1}/_{11}$, $S = 0.95$

(e) $a = {}^{0}/_{11}$, $S = 1.00$

(f) $a = {}^{1}/_{11}$, $S = 0.96$

(g) $a = {}^{2}/_{11}$, $S = 0.73$

(h) $a = {}^{2}/_{11}$, $S = 0.73$



(i) $a = {}^{3}/_{11}$, $S = 0.90$

(j) $a = {}^{4}/_{11}$, $S = 1.00$

(k) $a = {}^{5}/_{11}$, $S = 0.99$

(l) $a = {}^{6}/_{11}$, $S = 0.89$

(m) $a = {}^{9}/_{11}$, $S = 0.77$

(n) $a = {}^{10}/_{11}$, $S = 0.93$

(o) $a = {}^{11}/_{11}$, $S = 1.00$

(p) $a = {}^{12}/_{11}$, $S = 0.93$

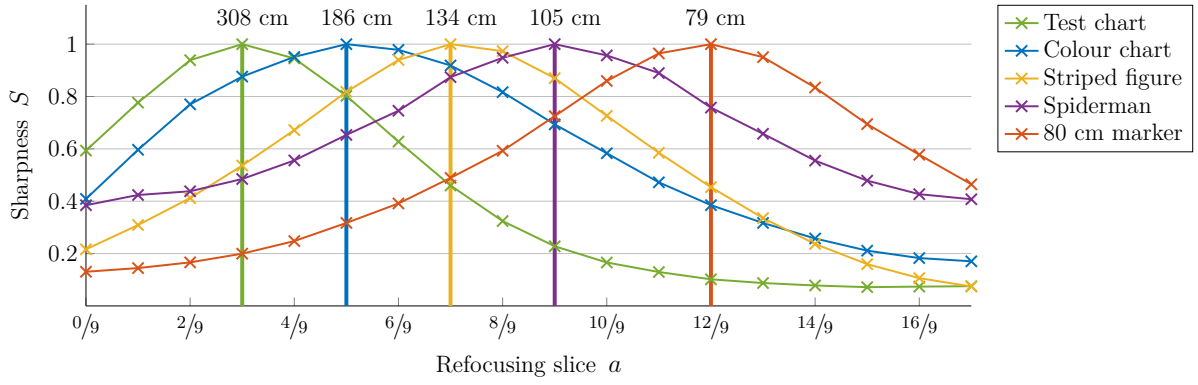
(q) $a = {}^{13}/_{11}$, $S = 0.76$

Figure 5.15: **Regions of refocused photographs with $b_U > f_U$ and $M = 11$.** The main lens was focused at 4 m distance ($d_f \approx 4$ m). S denotes the measured sharpness in (c) to (q). The denominator in a indicates the upsampling factor for the linear interpolation of whole micro images. The RoIs' illuminance values E''_a have been normalised for better blur detection.

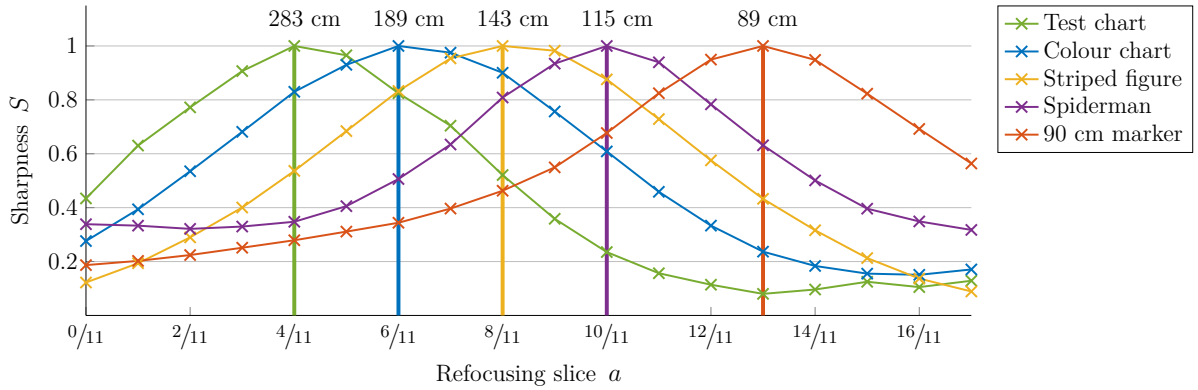
A more insightful overview illustrating the distance estimation performance of the proposed method is given in Fig. 5.16. Therein, each curve peak indicates the least blur for respective RoI of a certain object. Vertical lines represent the predicted distance d_a where objects were situated. Hence, the best case scenario is attained when a curve peak and its corresponding vertical line coincide. This would signify that predicted and measured best focus for a particular distance are in line with each other which corresponds to a 0 % error. A rather meaningful relative error should be obtained by increasing the depth resolution of an SPC. This inevitably requires more samples in the angular domain and thus more pixels per micro image. Since our camera already features an optimised micro image resolution of $M = 11$ which is further upsampled by factor 11, provided measurements are considered to be the accuracy limit at this stage. Nonetheless, a better depth resolution is expected to yield relative errors similar to that of the simulation with 0.35 %.

Results in [41] exhibit an error in predicting the distance of nearby objects whereas refocusing distance estimates in Fig. 5.16 match least blur peaks S for each object. The prediction in Fig. 5.16(c) proves that the proposed refocusing distance estimator takes alternative lens focus settings ($b_U > f_U$) into account without causing a deviation which was not investigated in [41]. The improvement in estimating the refocusing plane distance is mainly due to the correct approximation of MICs positions.

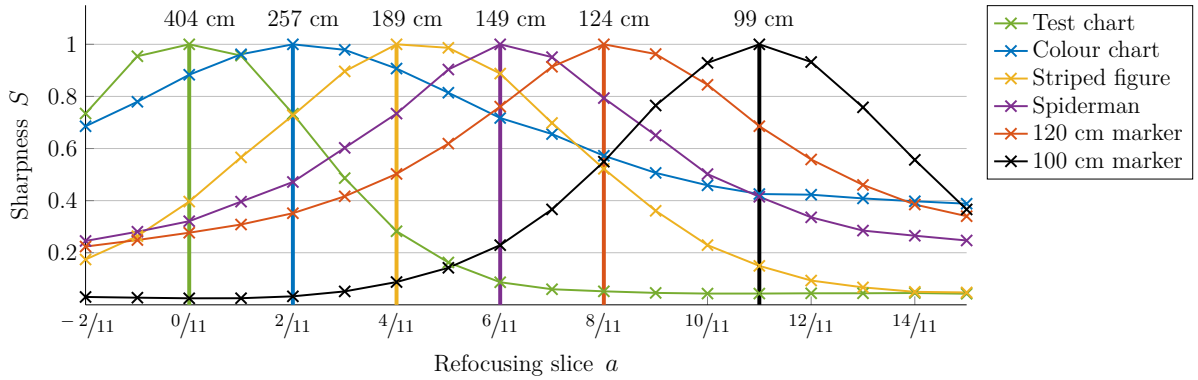
When reproducing such an experiment with a customised camera, it is important to bear in mind the possibility of biased results. Potential deviations include lens aberrations, measurement errors and calibration misalignments of MLA as well as objective lens.



(a)



(b)



(c)

Figure 5.16: **Blur metric plots.** (a) $M = 9$ and $d_f \to \infty$; (b) $M = 11$ and $d_f \to \infty$; (c) $M = 11$ and $d_f = 404$ cm. Vertical bars in colours indicate estimated position of best focus for respective objects.

5.3.3 Virtual Camera Array

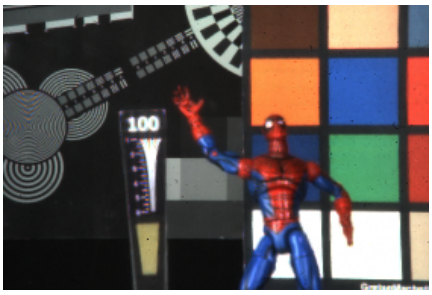
Based on the remarks in Chapter 3, an SPC carries virtual cameras lined up on the entrance pupil plane. Further, it was shown that the relationship of measured disparities Δx and baseline B_G allows conclusions to be drawn about the depth distance $Z_{G,\Delta x}$. To verify claims made about the SPC stereo triangulation, objects have been placed in the captured scene indicating predicted entrance pupil to depth plane distances $Z_{G,\Delta x}$. The extraction of a disparity map from an SPC requires at least two sub-aperture images that are obtained using Eq. (3.45). Disparity maps have been calculated by the computationally expensive *Sum of Absolute Differences* (SAD) method using Abbeloos' implementation [1, 2]. Since distances are known and disparities can be computed, the only unknown parameters that need validation are the baseline B_G and relative tilt angle Φ_G. To obtain measured baselines B_G, Eq. (3.60) has to be rearranged such that

$$B_G = \frac{Z_{G,\Delta x} \times (\Delta x \times p_N + b_N \times \tan(\Phi_G))}{b_N}\,. \tag{5.24}$$

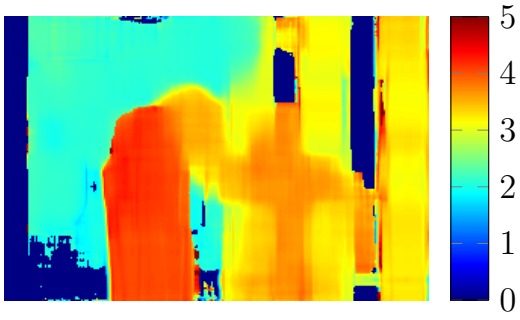
This formula can be also changed to

$$\Phi_G = \arctan\left(\frac{\frac{B_G \times b_N}{Z_{G,\Delta x}} - \Delta x \times p_N}{b_N}\right), \tag{5.25}$$

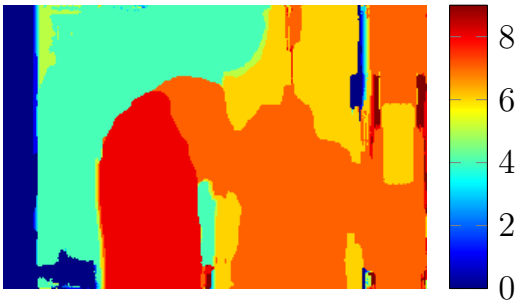
which yields a relative tilt angle Φ_G in radians that can be converted to degrees by multiplication with term $180/\pi$. Stereo triangulation experiments are conducted such that B_4 and B_8 just as Φ_4 and Φ_8 are pre-calculated based on main lens f_{197} and MLA (II.) with $d_f \to \infty$ and $d_f \approx 4$ m focus setting. From this, depth distances $Z_{G,\Delta x}$ can be ascertained which represent object distances. An exemplary sub-aperture image $E_{(i,g)}$ with infinity focus setting and related disparity maps is shown in Fig. 5.17. A sub-pixel precise disparity measurement has been applied to Figs. 5.17(b) and 5.17(d) as *Spiderman* lies between integer disparities.



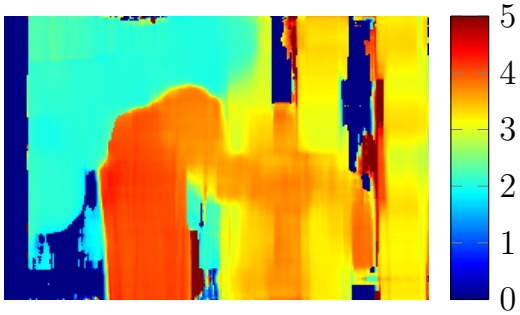
(a) Ref. image $E_{(0,0)}$ where $d_f \to \infty$



(b) Δx values from $E_{(-2,0)}$ and $E_{(2,0)}$



(c) Δx values from $E_{(-4,0)}$ and $E_{(4,0)}$



(d) Δx values from $E_{(0,0)}$ and $E_{(4,0)}$

Figure 5.17: **Disparity maps from sub-aperture images $E_{(i,g)}$ with $b_U = f_U$.** (a) Central sub-aperture image $E_{(0,0)}$ containing $K = 281$ by $L = 188$ pixels; (b) Disparity map with $G = 4$, $\max\{\Delta x\} = 5$ and window size = 29; (c) Disparity map with $G = 8$, $\max\{\Delta x\} = 9$ and window size = 39; (d) Disparity map with $G = 4$, $\max\{\Delta x\} = 5$ and window size = 29.

It may be obvious that disparities in Figs. 5.17(b) and 5.17(d) are nearly identical since both view pairs are separated by $G = 4$, however, placed at different horizontal positions. This justifies the claim that the spacing between adjacent virtual cameras is consistent. Besides, it is also apparent that objects at far distances expose lower disparity values and vice versa. Comparing Figs. 5.17(b) and 5.17(c) shows that a successive increase in the baseline B_G implies a growth in the object's disparity values, an observation also found in traditional computer stereo vision.

Table 5.11: **Baseline results B_G with infinity focus ($b_U = f_U$).**

(a) B_4 from Figs. 5.17(b) and 5.17(d)

Δx	$Z_{G,\Delta x}$ **[cm]**	**Measured** B_4 **[mm]**
2	203	2.5806
3	136	2.5806
3.5	116	2.5806
4	102	2.5806

(b) B_8 from Fig. 5.17(c)

Δx	$Z_{G,\Delta x}$ **[cm]**	**Measured** B_8 **[mm]**
4	203	5.1611
6	136	5.1611
7	116	5.1611
8	102	5.1611

(c) Comparison of predicted and measured B_G

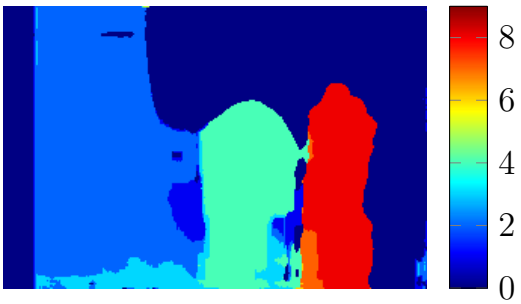
d_f	B_G	**Predicted** B_G **[mm]**	**Avg. measured** B_G **[mm]**	**Deviation** ERR_{B_G} **[%]**
∞	B_4	2.5806	2.5806	0.0000
∞	B_8	5.1611	5.1611	0.0000

Table 5.11 lists baseline measurements and corresponding deviations with respect to the predicted baseline. This table is quite revealing in several ways. First, the most striking result is that there is no significant difference between baseline predictions and measurements. The reason for this is that objects are placed at the centre of predicted depth planes $Z_{G,\Delta x}$. An experiment conducted with random object positions would yield errors that do not reflect the model's accuracy, but rather our SPC's capability to resolve depth which depends on the MLA and sensor specification. Hence, such an experiment is only meaningful when evaluating the camera's depth resolution. A more revealing percentage error is obtained by a larger number of disparities, which in turn requires the baseline to be extended. These have been maximised in our experimental setup making it difficult to further refine depth. Scaling up the size of the SPC's optical components whilst retaining the sensor's resolution can solve for the problem and may provide deviations similar to that of the real ray simulation in Section 5.2.2.

Whenever $d_f \to \infty$, tilt angles are assumed to be $\Phi_G = 0°$. Accurate baseline measurements inevitably confirm predicted tilt angles as measured baselines would deviate otherwise. To ensure this is the case, a second SPC triangulation experiment is carried out with $d_f \approx 4$ m yielding images shown in Fig. 5.18.

(a) Ref. image $E_{(0,0)}$ where $d_f \approx 4$ m

(b) Δx values from $E_{(-2,0)}$ and $E_{(2,0)}$



(c) Δx values from $E_{(-4,0)}$ and $E_{(4,0)}$

(d) Δx values from $E_{(0,0)}$ and $E_{(4,0)}$

Figure 5.18: **Disparity maps from sub-aperture images $E_{(i,g)}$ with $b_U > f_U$.** (a) Central sub-aperture image $E_{(0,0)}$ containing $K = 281$ by $L = 187$ pixels; (b) Disparity map with $G = 4$, $\max\{\Delta x\} = 5$ and window size = 33; (c) Disparity map with $G = 8$, $\max\{\Delta x\} = 9$ and window size = 39; (d) Disparity map with $G = 4$, $\max\{\Delta x\} = 5$ and window size = 33.

Disparity maps in Figs. 5.18(b) and 5.18(d) give further indication that the spacing between adjacent virtual cameras is consistent. Results in Table 5.12 demonstrate that tilt angle predictions match measurements. They further show that virtual cameras are rotated by small angles of less than a degree. Nevertheless, these tilt angles are non-negligible as they are large enough to shift the $\Delta x = 0$ disparity plane from infinity to $d_f \approx 4$ m which can be seen in Fig. 5.18.

Table 5.12: **Tilt angle results Φ_G with 4 m focus ($b_U > f_U$).**

(a) Φ_4 from Figs. 5.18(b) and 5.18(d)

Δx	$Z_{G,\Delta x}$ **[cm]**	**Measured** Φ_4 [°]
0	384	0.0429
1	218	0.0429
2	152	0.0429
4	95	0.0429

(b) Φ_8 from Fig. 5.18(c)

Δx	$Z_{G,\Delta x}$ **[cm]**	**Measured** Φ_8 [°]
0	384	0.0857
2	218	0.0857
4	152	0.0857
8	95	0.0857

(c) Comparison of predicted and measured Φ_G

d_f	Φ_G	**Predicted** Φ_G [°]	**Avg. measured** Φ_G [°]	**Deviation** ERR_{Φ_G} [%]
≈ 4 m	Φ_4	0.0429	0.0429	0.0000
≈ 4 m	Φ_8	0.0857	0.0857	0.0000

Generally, Tables 5.11 and 5.12 suggest that the adapted stereo triangulation concept proves to be viable in an SPC without measurable deviations if objects are placed at predicted distances. A maximum baseline is achieved with a short MLA focal length f_s, large micro lens pitch p_M, long main lens focal length f_U and a sufficiently large entrance pupil diameter.

A baseline approximation of the first-generation Lytro camera may be achieved with the aid of the metadata (*.json file) attached to each light field photograph as it contains information about the micro lens focal length $f_s = 0.025$ mm, pixel pitch $p_p \approx 0.0014$ mm and micro lens pitch $p_M \approx 0.0139$ mm yielding $M = 9.9286$ samples per micro image. The accommodated zoom lens provides a variable focal length in the range of $f_U = 6.45$ mm – 51.4 mm (43 mm – 341 mm as 35 mm-equivalent) [29]. Although, it is unclear whether the source refers to the main lens only or to the entire optical system including the MLA. From this, baseline estimates for the first-generation Lytro camera are calculated via Eqs. (3.52) to (3.55) and given in Table 5.13.

Table 5.13: **Baseline estimates of Lytro's 1$^{\mathbf{st}}$ generation camera.**

f_s **[mm]**	f_U **[mm]**	B_1 **[mm]**	B_8 **[mm]**
0.025	6.45	0.3612	2.8896
0.025	51.4	2.8784	23.0272

Disparity analysis of sub-aperture images taken from Lytro's first-generation camera should lead to baseline measures B_G similar to those of the prediction. However, verification is impossible as the camera's automatic zoom lens settings (current focal length, principal plane and exit pupil location) are undisclosed.

5.4 Summary

This chapter provided all experimental work verifying the SPC ray model introduced in Chapter 3 and claims made about the estimation of refocusing distances and stereo baselines. Raw light field photographs from our customised SPC have been computationally brought to focus and analysed using a blur metric which was devised for distance validation purposes. Disparity maps were computed from sub-aperture images and given in this chapter to verify the stereo triangulation concept using an SPC. In addition, an optical design software was deployed to obtain quantitative results and further ensure the model's accuracy.

The first evaluation carried out in this chapter has examined whether MIC chief rays can be seen as originating from the exit pupil or principal plane of the main lens. Results obtained from an optical design software [89] suggest that chief rays which impinge on MICs cross the optical axis at the exit pupil plane. Experiments based on real light field photographs taken by an SPC have confirmed this observation. As a consequence, an MLA imaging device combined with calibration algorithms presented in Chapter 4 is considered to be a valuable method to measure exit or entrance pupil locations of an objective lens with unknown specification.

The most important finding is that experiments based on real light field photographs neither entailed errors for refocusing distance estimates nor for the position of a virtual camera array. These findings may be somewhat limited. It could be argued that experiments with our customised camera should yield errors other than 0 %. However, there are two evident reasons for absent deviations. One is that target objects were placed at predicted distances instead of positioning them randomly in the scenery. The latter attempt would rather examine the SPC's capability to resolve depth than assessing the proposed model. Based on this, the second reason is that our customised SPC simply cannot resolve depth to the level of deviations caused by optical aberrations. Apart from this, another point could be made stating that the presented experimental work merely featured one MLA, one objective lens and two main lens focus settings. To overcome these concerns, a validation through simulation was conducted for different optical setups including three main lens and two MLA designs which backs the positive experimental

outcomes. Thereby, predictions of refocusing distances and virtual camera positions indicate deviations of up to 0.35 % compared to the simulation. Small tolerances in simulation are due to lens aberrations that cause a non-geometrical lens behaviour. Hence, the model's accuracy primarily depends on the lens' ability to suppress aberrations. Simulation results further support the assumption that DoF ranges shrink when refocusing closer, a focus behaviour similar to that of conventional cameras. Another interesting finding is the discovery of a low-cost method to measure exit pupil locations which can be also used to localise the entrance pupil when flipping the lens.

After all, the hypothetical SPC model can be deemed to accurately predict the distance range which has synthetically brought to focus. It is also possible to conclude that the geometrical laws of stereo vision equally apply to an SPC.

CHAPTER 6

Refocusing Hardware Architecture

It is the goal of this chapter to devise and demonstrate a hardware architecture that computationally changes the optical focus of serially incoming video frames in real-time. The proposed design is implemented on an FPGA as this facilitates power-efficiency and makes it easy to reconfigure and update.

A scenario demanding such a real-time refocusing computation system is a mobile video camera, e.g. at a movie set, as the *Director of Photography* potentially wishes to skip through previews to see what refocused pictures will look like. In general, the field of motion picture arts can benefit from the refocusing capability since the focus and DoF range of the footage may be adjusted in the post-production stage. This allows for more flexibility and creativity in the final production steps.

On the basis of Ng's statements [58], it has been elaborated in Chapter 2 that, in the sense of computation time, the integral projection outperforms the refocusing operation in frequency domain when it comes to synthesising a single refocused photograph. Because a video output monitor only needs to provide one refocused live view at a time, it may be a reasonable decision to achieve refocusing in spatial domain without the need for the computationally expensive *Fourier* and inverse *Fourier* transform. Therefore, the upcoming explanations concern with the development of a hardware design to accomplish refocusing based on an integral projection in spatial domain.

The main concept of computation time improvements using FPGAs builds on the principle of parallelism and pipelining [7]. A pipeline comprises chained up processor blocks that are fed with serialised data which is processed sequentially. Computation speed-up in pipelining relies on the idea of processing data chunks in one processor unit while subsequent data chunks are handled in a preceding unit. Hence, the benefit of pipelining is that data chunks are processed at the same time whereas processor units perform different tasks. While data serialisation limits a specific task to be computed with one single operation at a time, e.g. one pixel after another, parallelised data streams allow a computing system to perform at least two operations of the same type simultaneously. Parallelism can be thought of as duplicating processor pipelines which requires synchronised parallel data streams as input signals. Letting the degree of parallelism be ι, the computation time in image processing may be minimised to $\mathcal{O}\left({}^{K^2}/_{\iota}\right)$ if 2-D image dimensions consist of K samples each and provided that both computation systems run at the same clock frequency. Consequently, 1-D parallelism limit is reached where $\iota = L$ for image rows and $\iota = K$ for image columns which is the ideal scenario in terms of parallelising data processes.

Early studies in the field include a publication by Rodríguez-Ramos *et al.* [69] who employed an FPGA to process plenoptic image data with the aim to accomplish wavefront measurements. Another interesting work conceived by Wimalagunarathne *et al.* [82] proposed a design to render computationally focused photographs from a set of multi-view images using *Infinite Impluse Response* (IIR) filters. The first hardware design aiming to perform real-time refocusing from SPC captures was presented by our research group [38]. Shortly after that, Pérez *et al.* [65] published an article addressing the same problem. The authors demonstrated significant computation time improvements compared to run times based on a *Central Processing Unit* (CPU) system which was programmed using an object-oriented language. A comparison of the proposed method with that of Pérez *et al.* [65] is carried out at the end of Section 6.1.

6.1 Filter Design

An efficient hardware design which enables an FPGA to refocus in real-time may be conceptualised on the basis of the light field ray model presented in Chapter 3. The upper data line of Fig. 6.1 depicts discrete and quantised illuminance values $E_{f_s}[x_k]$ of a single horizontal row that is part of a calibrated light field image. Calibration implies a consistent 1-D micro image resolution M. The computational refocusing synthesis derived in Chapter 3 reveals that pixels involved in the integration process expose varying neighbourhood distance relations, which exclusively depend on a. This phenomenon is illustrated in the data flow diagram in Fig. 6.1 where respective pixels are highlighted for two exemplary refocusing settings $a = 0/3$ and $a = 2/3$. Here, each colour corresponds to a chief ray in the model proposed in Fig. 3.4 with $M = 3$ where yellow represents the micro image centre. It is the aim of this chapter to devise a hardware architecture that satisfies to accomplish signal processing according to Fig. 6.1.

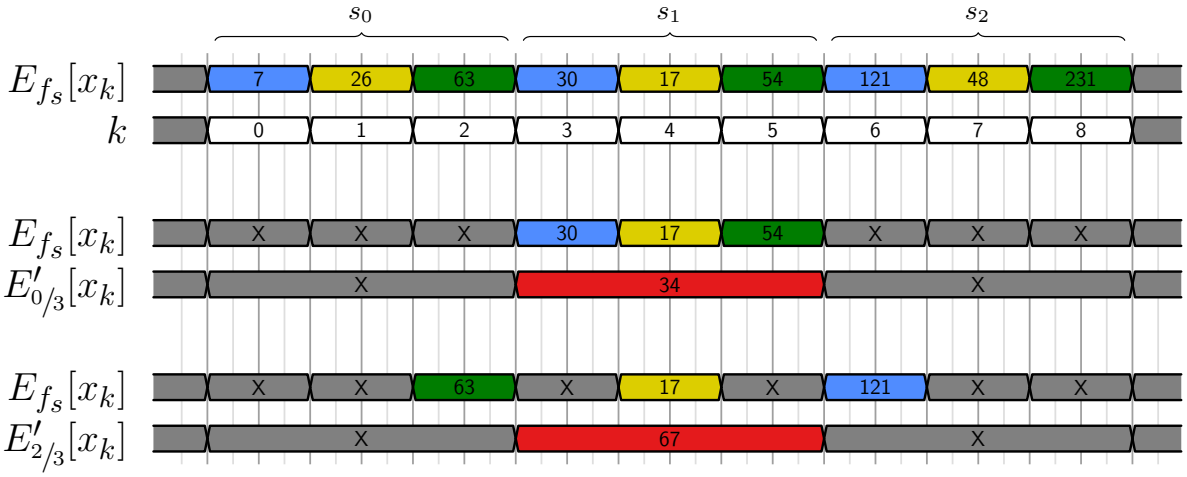


Figure 6.1: **Processing requirements for the hardware architecture.** The diagram shows exemplary input illuminance values E_{f_s} which can be subdivided into micro images s_j. and synthesised output values E'_a at a desired refocused image plane a.

As opposed to Eq. (3.17), it is the aim to produce an output image with a resolution equal to that of the incoming image. Since refocusing is based on an integration and thus a reduction of angular samples, the overall resolution can only be retained by interpolating in the spatial domain.

Following the $\left[s_j, u_{c+i}\right]$ to $[x_k]$ translation in Section 3.4, 1-D refocusing can be given by

$$E'_a[x_{k'}] = E'_a[x_{k'}] + \frac{1}{M} E_{f_s}[x_k] \quad , \quad k' = k + c + i + (k \bmod M) \times (a' - 1) \tag{6.1}$$

where $i \in [-c, c]$ and $a = {}^{a'}/_{M}$ with $a' \in \mathbb{N}$ representing sub-pixel-wise shift and integration, which implies an interpolation of refocused image slices a. On the supposition that a horizontal cross-section of a captured light field $E_{f_s}[x_k]$ is a linear, time-invariant system, the integral projection in Eq. (3.17) may be represented as a discrete convolution formula. To break down the complexity, we devise one filtering function per refocusing slice a that yields the same output $E'_a[x_k]$ as with Eq. (6.1). It is the aim to find equations that qualify for FIR filter implementation. Regardless of the micro image resolution M, a filter that computes refocusing slice $E'_0[x_k]$ in horizontal direction reads

$$E'_{0/M}[x_k] = \sum_{i=0}^{M-1} \frac{1}{M} E_{f_s}\left[x_{k-i-(k+1 \bmod M)}\right] , \tag{6.2}$$

when index k is seen to start counting from 0. Term $(k+1 \bmod M)$ comprises an NN interpolation ensuring that the output image resolution matches that of the input. A synthetically focused image with $a = 1$ is formed by

$$E'_{M/M}[x_k] = \sum_{i=0}^{M-1} \frac{1}{M} E_{f_s}\left[x_{k-i\times(a+1)}\right] . \tag{6.3}$$

Synthesis equations for different ${}^{a'}/_{M}$ are retrieved by reverse-engineering Eq. (6.1). Probably, the most straightforward refocusing filter kernel function is given by

$$E'_{1/M}[x_k] = \sum_{i=0}^{M-1} \frac{1}{M} E_{f_s}[x_{k-i}] , \tag{6.4}$$

which computes refocusing slice $a = {}^{1}/_{M}$. When implementing Eq. (6.4) as an FIR filter, it becomes obvious that the number of filter taps amounts to M. In the following, we demonstrate a refocusing hardware architecture that is adapted to an SPC with $M = 3$. Then, a photograph refocused with $a = {}^{2}/_{3}$ is computed by

$$E'_{2/3}[x_k] = \sum_{i=0}^{3-1} \frac{1}{3} E_{f_s}\left[x_{k-i+|\lceil(k+1 \bmod 3)/3\rceil-1|\times(i-1)}\right] . \tag{6.5}$$

where $\lceil\cdot\rceil$ is the ceiling operator. An exemplary step in the computation of $E'_{2/3}[x_k]$ would be

$$E'_{2/3}[x_3] = \frac{1}{3}E_{f_s}[x_3] + \frac{1}{3}E_{f_s}[x_2] + \frac{1}{3}E_{f_s}[x_1] . \tag{6.6}$$

Here, fractions ${}^{1}/_{3}$ can be regarded as multipliers, denoted as h_0, which are the same for each pixel value such that $h_0 = {}^{1}/_{M}$. It is noteworthy that multipliers are redundant and thus left out on condition that incoming images are underexposed and clipping is prevented (see Section 3.2 for more details).

6.1.1 Semi-Systolic Modules

To implement Eqs. (6.2) to (6.5), a *systolic* filter design is employed. *Systolic* arrays broadcast input data to many *Processing Elements* (PE)s. Per definition, all wired connections in a *systolic* filter contain at least one latch driven by the same clock signal. As opposed to the *systolic* fashion, filter designs presented hereafter omit these latches for simplicity what makes them so-called *semi-systolic*. Nevertheless, latches can be added to proposed designs for *systolic* FPGA implementation purposes. Descriptive information about *systolic* arrangements can be found in [83].

A positive side effect of the *systolic* filter is that it can be exploited for an NN interpolation in micro images. By letting the upsampling factor be the number of micro image samples M, the resolution loss in integral projection is compensated, since incoming and outgoing resolution are the same. Naturally, the interpolation method can be more sophisticated which in turn requires intermediate calculations causing delays and an increasing number of occupied logic gates. Closer inspection of Eq. (6.3) reveals that pixels which need to be integrated are interlaced. Thereby, gaps between merged pixels grow with ascending a and extend the filter length. The omission of pixels within gaps is realised with switches. A switch-controlled *semi-systolic* FIR filter design of Eq. (6.2) with multiplier h_0 is depicted in Fig. 6.2.

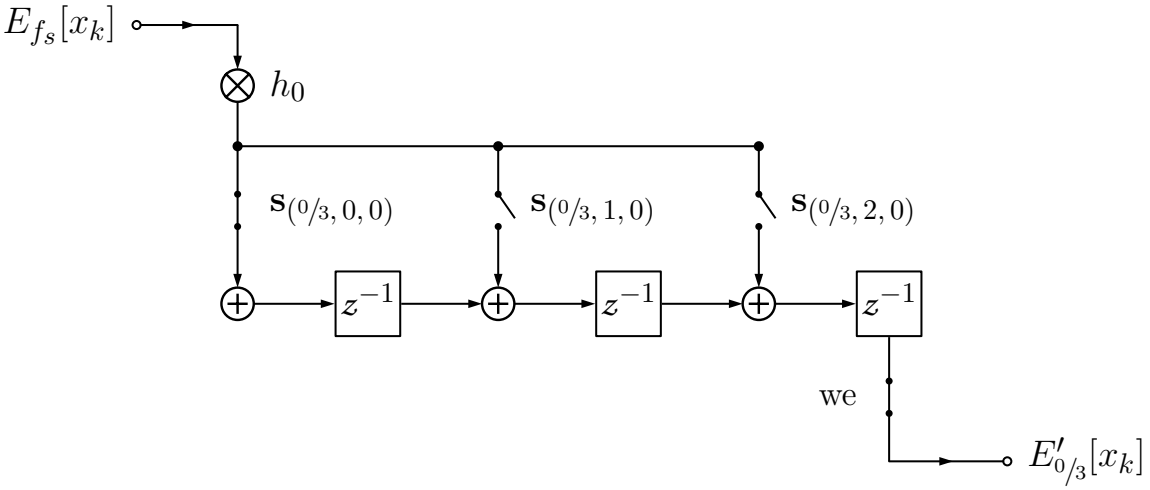


Figure 6.2: **1-D semi-systolic FIR filter for sub-pixel shift** $a = 0/3$.

In that, switch states are controlled by bits in a 2-D vector field denoted as $\mathbf{s}_{(a,w,p)}$ which is given by

$$\mathbf{s}_{(0/3,w,p)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} . \tag{6.7}$$

if $a = 0/3$. Depending on refocusing parameter a, switch state matrices $\mathbf{s}_{(a,w,p)}$ contain binary numbers with columns indexed by w for the state of each switch in the FIR filter and with rows indexed by p which loads a new row of switch states when incremented. In addition, a *write enable* (we) switch helps prevent intermediate falsified values from being streamed out.

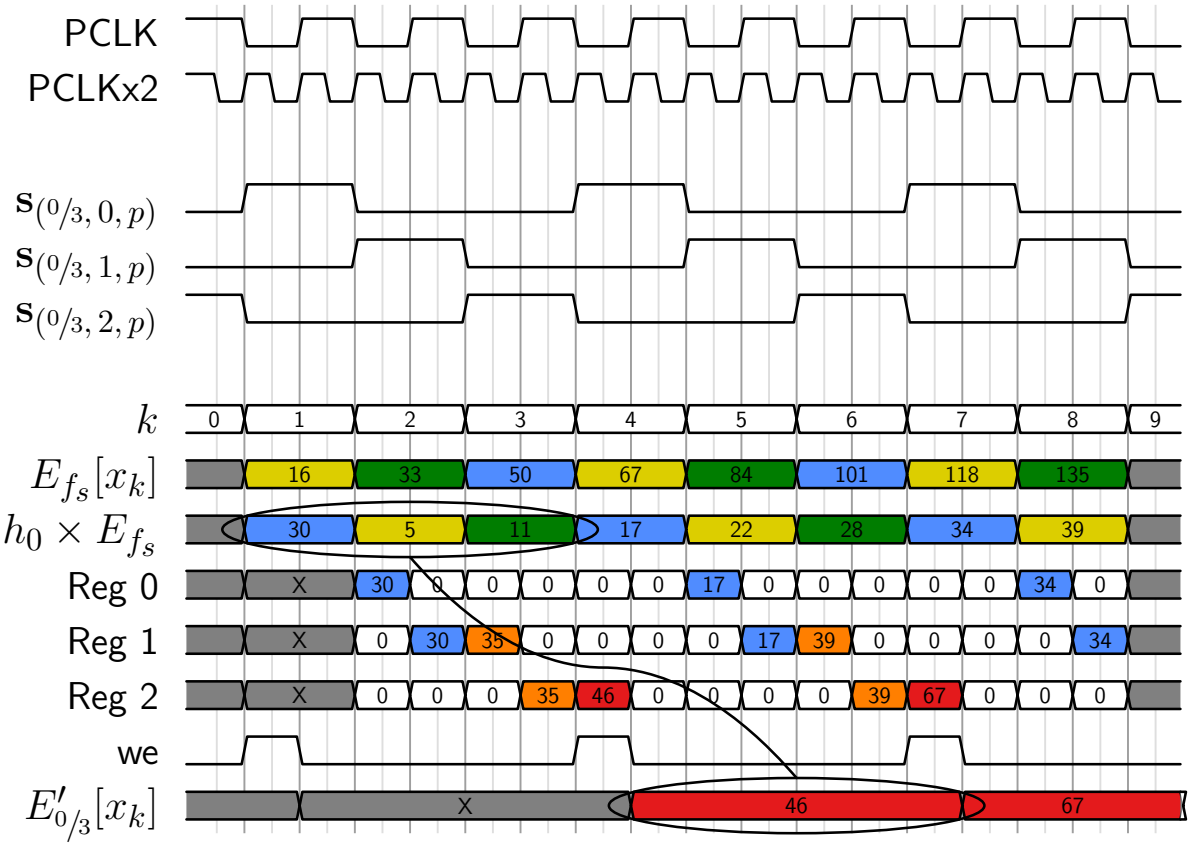


Figure 6.3: **Timing diagram of 1-D semi-systolic FIR filter for** $a = 0/3$.

For better comprehension, a timing diagram in Fig. 6.3 visualises the computational concept of the FIR design from Fig. 6.2. Here, the pixel clock signal is given as PCLK. Furthermore, the proposed architecture employs the doubled pixel clock PCLKx2 with a time period $T_{\text{PCLKx2}} = T_{\text{PCLK}}/2$ to shift and add pixel values in a single pixel clock

cycle T_{PCLK}. It is also seen that a new row of switch states is called by incrementing p every pixel clock cycle. Numbers in the data streams represent unsigned decimal 8-bit grey-scale values which are multiplied with $h_0 = 1/3$. Pixel colours match those of the SPC ray model in Chapter 3 indicating chief ray positions in micro images with $M = 3$. Orange colour highlights interim results and red signifies 1-D refocused output data. As indicated with circles, the sum of divided micro image pixels values is reflected in the output pixel $E'_{0/3}[x_k]$. The filter includes an NN interpolation which upsamples the micro image resolution by factor 3.

To refocus with $a = 1/3$, another FIR filter module is conceived based on Eq. (6.4) and depicted in Fig. 6.4.

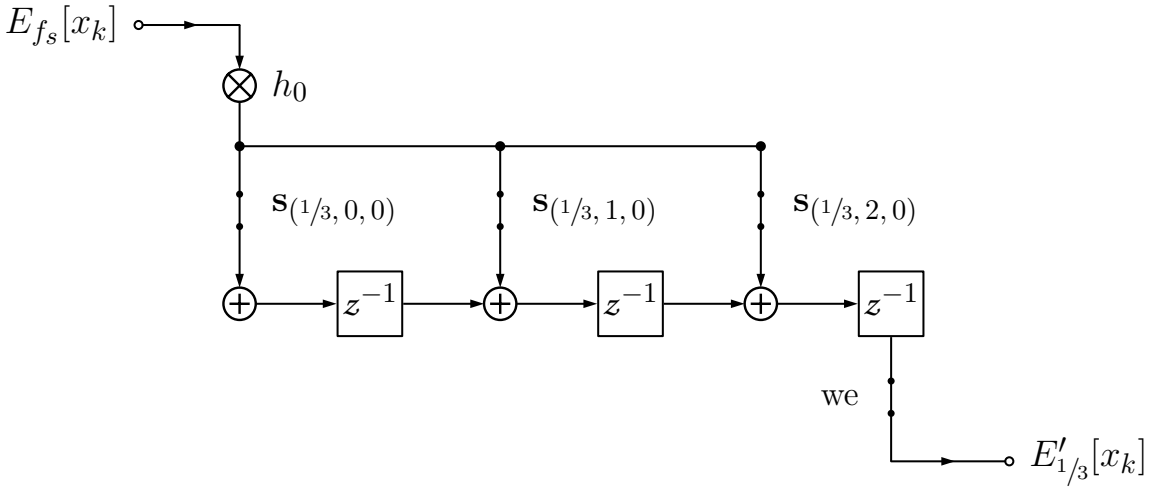


Figure 6.4: **1-D semi-systolic FIR filter for sub-pixel shift** $a = 1/3$.

In reference to the previous FIR filter with $a = 0/3$, it becomes obvious that this arrangement is identical to Fig. 6.4 except that switch states are different. The switch state matrix $\mathbf{s}_{(1/3,\,w,\,p)}$ is then given by

$$\mathbf{s}_{(1/3,\,w,\,p)} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} . \tag{6.8}$$

which means that switches remain closed at all times. A corresponding timing diagram is shown in Fig. 6.5.

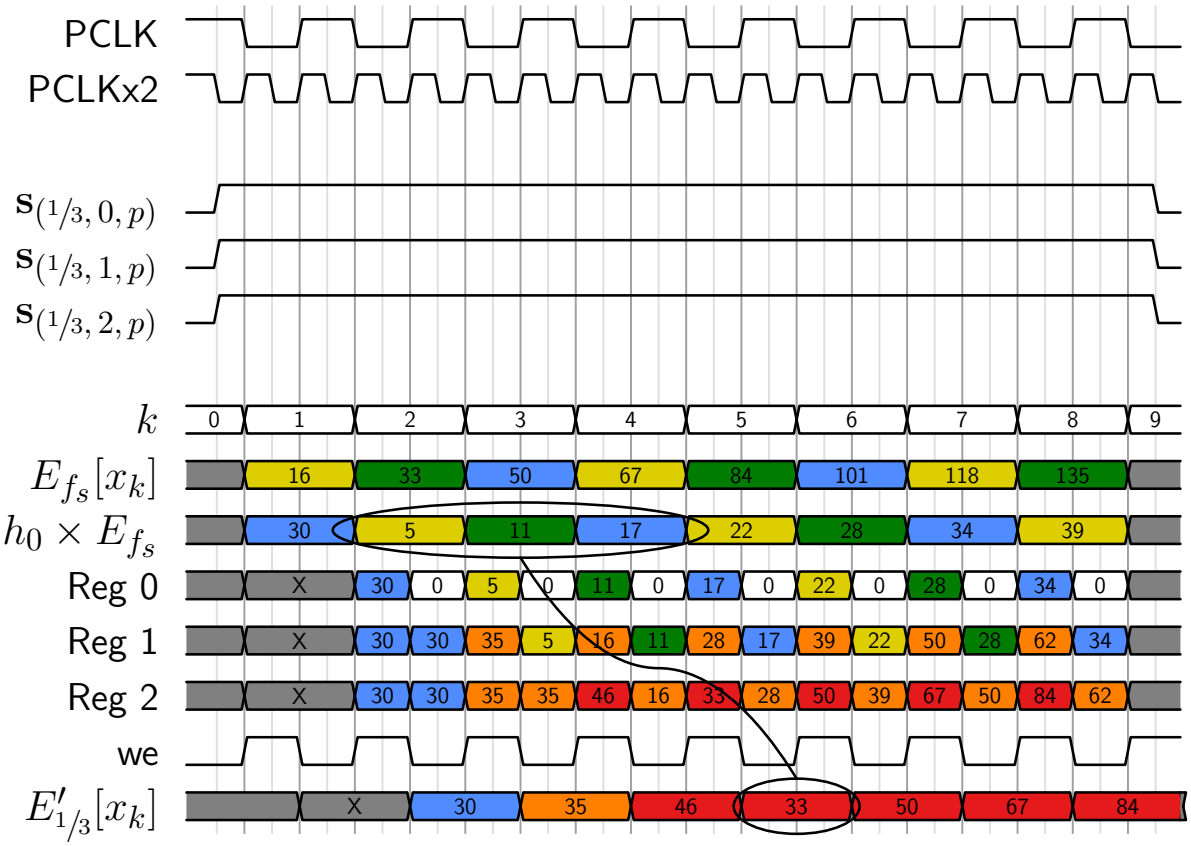


Figure 6.5: **Timing diagram of 1-D semi-systolic FIR filter for** $a = 1/3$.

Figure 6.6 depicts an FIR filter equivalent of Eq. (6.5) apparently occupying more PEs due to the fact that the distance among pixels that need to be summed up grows.

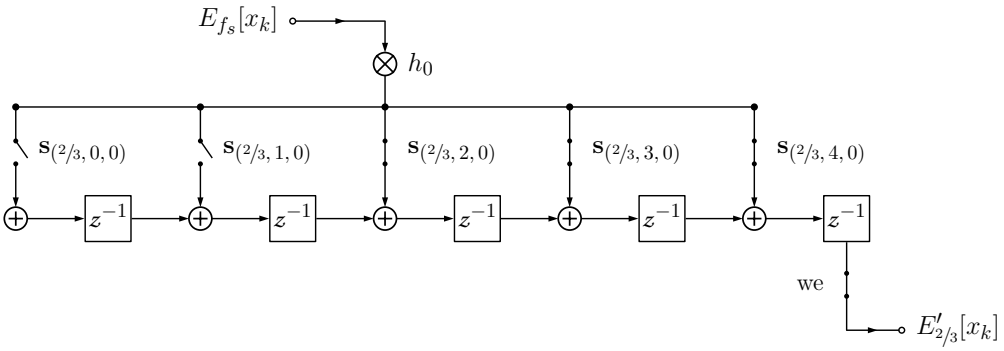


Figure 6.6: **1-D semi-systolic FIR filter for sub-pixel shift** $a = 2/3$.

The corresponding switch state matrix $\mathbf{s}_{(2/3,\,w,\,p)}$ is as follows

$$\mathbf{s}_{(2/3,\,w,\,p)} = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}, \tag{6.9}$$

and produces a filter behaviour shown in Fig. 6.7.

As can been seen from Fig. 6.6, a large 1-D *semi-systolic* filter module implies long wires when broadcasting the multiplier outputs. Long wires cause a low-pass filter behaviour in the signals which affects the readability of falling and rising edges and therefore has to be avoided. To keep wires short in the broadcasting net, incoming bit words can be distributed to several synchronised latches before being merged in adders.

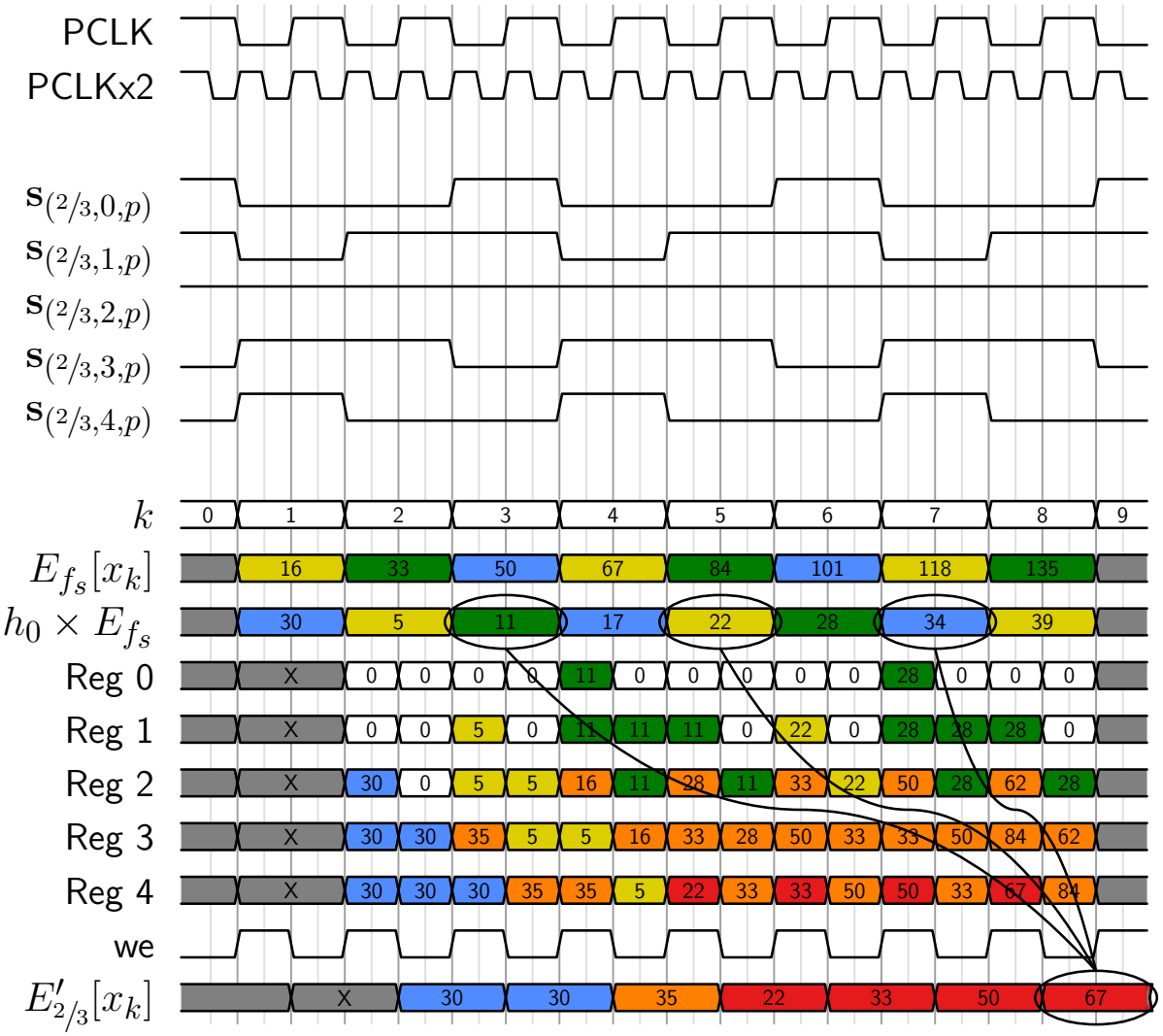


Figure 6.7: **Timing diagram of 1-D semi-systolic FIR filter for** $a = 2/3$.

6.1.2 2-D Module Array

Proposed FIR filter modules process data in 1-D and thus in horizontal or vertical direction only. Figure 6.8 shows a 2-D construct of 1-D *semi-systolic* processor modules to accomplish refocusing by processing data in both dimensions. In this example, the degree of parallelism amounts to $\iota = 3$, but could be scaled as desired until parallelism limits are reached ($\iota = L$ for image rows, $\iota = K$ for image columns).

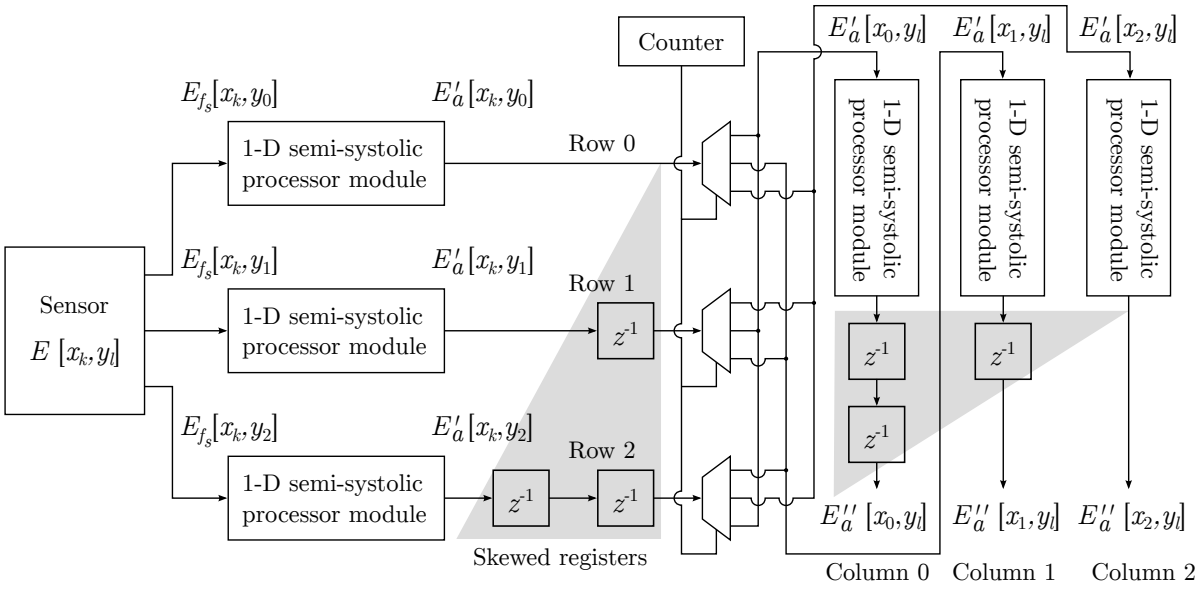


Figure 6.8: **Parallelised 2-D processing module array with $\iota = 3$.** All *semi-systolic* processor modules are identical depending on the refocusing parameter a. Skewed registers delay processed image rows $E'_a[x_k, y_l]$ with respect to each other making it possible to form an incoming image column (e.g. $E'_a[x_0, y_l]$) in consecutive order. A pixel counter controls demultiplexers to assign pixels based on their index to respective vertical processing modules. For synchronisation purposes, an additional array of skewed registers can be optionally placed behind column processor blocks.

The data flow in Fig. 6.8 is described in the following. First, pixels coming from the sensor are fed into horizontal processor blocks representing *semi-systolic* FIR filter modules. In the second stage, horizontally processed data rows are assigned to another arrangement of *semi-systolic* modules to apply the convolution in vertical direction. Here, demultiplexers are driven by a pixel counter to assist in the correct assignment of pixels values. This assures that pixels from different rows sharing index k are sent to the same vertical processing unit which produces an image column (e.g. $E''_a[x_0, y_l]$) of the final refocused image.

In order to estimate the computation time, it is assumed hereafter that the system refers to the ideal case of maximum parallelism where $\iota = L$ or $\iota = K$ for each dimension, respectively. Besides, it is supposed that colour channels are also parallelised causing no extra time delay. The shift and integration for a single output pixel refocused with $a = 1/M$ takes M pixel clock cycles in 1-D when using twice the pixel clock to process them. Taking this as an example, the overall number of steps η to compute a single image $E''_{1/3}$ with K-by-L resolution is given by

$$\eta = \text{multiplier} + M + \text{multiplier} + M + K - 1 + L - 1 + K - 1\,, \tag{6.10}$$

where 'multiplier' stands for a single step to compute the result of an incoming pixel value after multiplication. The total computation time $\mathcal{O}$ for a single image can be obtained by

$$\mathcal{O}(\eta) = \eta \times T_{\mathrm{PCLK}}\,. \tag{6.11}$$

This duration reflects the theoretical time that elapsed from the moment the first pixel $E_{f_s}[x_k, y_l]$ entered the logic gate until the final output pixel $E''_a[x_k, y_l]$ is available. When pipelining the data stream, output pixels of a subsequent image arrive directly after that letting the overall computation time for a single frame be represented by the delay time of the computational focusing system. Once the first refocused photograph is received, the number of remaining computational steps η_{sub} for every following image amounts to

$$\eta_{\mathrm{sub}} = L - 1 + K - 1\,. \tag{6.12}$$

To assess the performance limits of the presented architecture, it is worth benchmarking it with that of Pérez *et al.* [65]. Taking the example of the authors, a 3201 by 3201 pixel image ($K = L = 3201$) with 291 by 291 micro lenses was computationally refocused in 105.9334 ms at 100 MHz clock frequency. Thereby, the micro image resolution is $M = 11$ and the output image resolution amounts to 589 by 589 which is less than a sixth of the incoming image. Conversely, the proposed semi-systolic method preserves the original incoming resolution by employing an NN interpolation in $\eta = 1 + 11 + 3200 + 1 + 11 + 3200 + 3200$ steps yielding $\mathcal{O}(\eta) = 96.24$ μs computation time for a single frame when running at 100 MHz pixel clock. Each subsequent frame, however, can be processed in $\eta_{\mathrm{sub}} = 3200 + 3200$ steps which is available every $\mathcal{O}(\eta_{\mathrm{sub}}) = 64$ μs. In respect of the method proposed by Pérez *et al.* [65], this leads to an overall computation speed-up of more than three orders of magnitude reducing the processing time by 99.91 %. For comparison, an identical implementation based

on a high-level programming language takes approximately 12.1369 s per image on average.

6.2 Validation

In this section, the proposed refocusing hardware architecture is implemented on an FPGA to assess its functionality. For that purpose, the *VHSIC Hardware Description Language* (VHDL) is used to configure the FPGA where VHSIC stands for *Very High Speed Integrated Circuit*. A schematic file, generated from a VHDL compiler, is then flashed onto the FPGA chip model XC6SLX45 which belongs to the Spartan-6 device family. Figure 6.9 contains a block diagram illustrating a simplified approach to implement and validate the design proposed in the previous section.

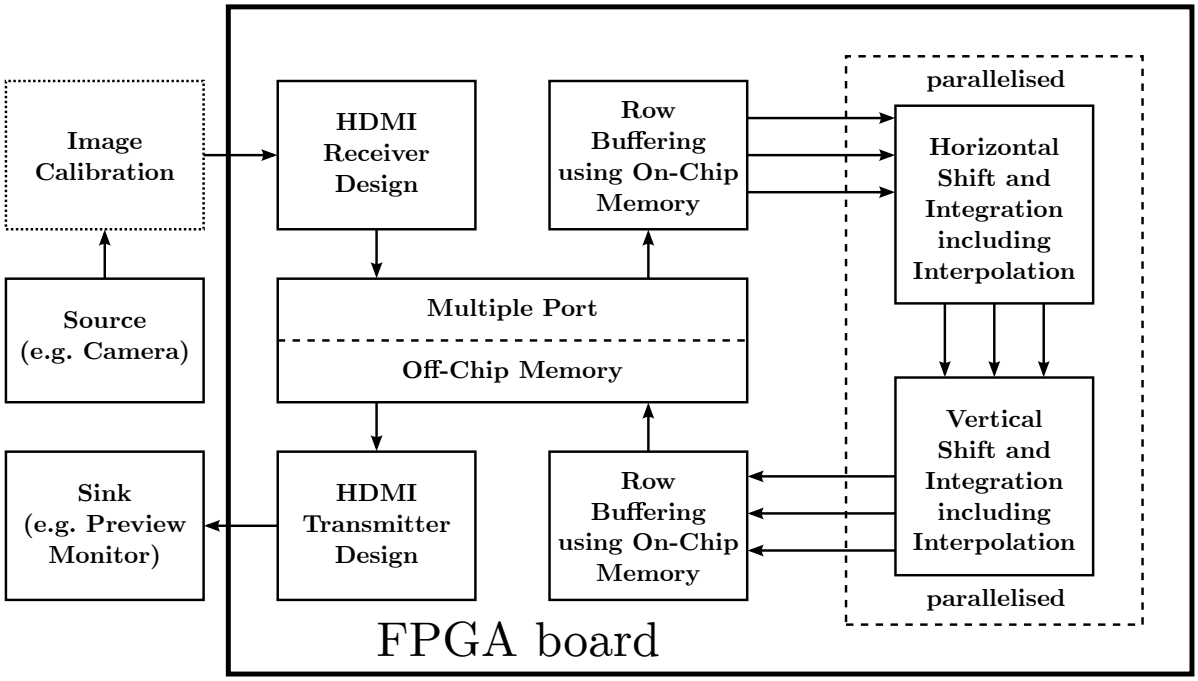


Figure 6.9: **Block diagram for signal processing**. Single arrows denote serialised whereas three arrows indicate parallelised data streams. Row buffers are employed to simulate data parallelism in the experiment.

The FPGA board features *High-Definition Multimedia Interface* (HDMI) connectors such that video frame transmission is accomplished using the *Transition Minimised Differential Signalling* (TMDS) protocol. TMDS receiver and transmitter designs have been integrated on the FPGA to fulfil deserialisation, serialisation just as de- and encoding tasks. Off-chip memory is used for buffering decoded and deserialised video frames outside the FPGA since the amount of image data exceeds internal memory storage.

Presented refocusing synthesis formulas require all micro images to be of a consistent size which is, however, not the case in raw light field photographs. As indicated with the experimental architecture in Fig. 6.9, micro image cropping remains an external process performed prior to streaming the data to the FPGA. Embedding this process on an FPGA is essential for prototyping, but left for future work. According to FIR filter designs in Section 6.1, the micro image size is reduced to $M = 3$ directional resolution.

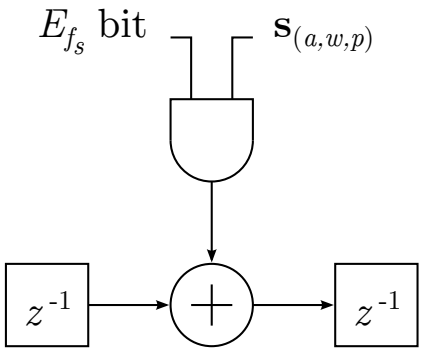


Figure 6.10: **Switch design** composed of a single incoming bit from E_{f_s}, a switch bit $\mathbf{s}_{(a,w,p)}$ bit and a logical AND gate.

Figure 6.10 schematically illustrates the switch as a circuit of logical elements for a single bit of the incoming E_{f_s} signal. In our implementation, a row of switch settings is loaded from a *Look-Up Table* (LUT) every clock cycle starting from the first row again after the last one is reached. The switch-state LUTs can be stored in *Block Random-Access Memorys* (BRAMs) which are part of the FPGA.

The integration of multiplier h_0 is also achieved using on-chip memory making it called *stored product*. In accordance with the TMDS protocol specification, a decoded pixel value is of 8 bit depth per colour channel which yields a manageable number of 256 possible results when dividing by M. Thus, quotients can be pre-calculated for a specific divisor M and stored in one BRAM per colour channel for each image row. Note that these BRAMs are read-only memories.

A screenshot from an exemplary timing diagram simulation where $a = 1/3$ and $T_{\mathrm{PCLK}} = 60$ ns is provided in Fig. 6.11. This VHDL-implemented hardware simulation shows that the filter behaves as expected justifying the conceived architecture. PCLKx2 can be obtained with a *Phase-Locked Loop* (PLL).

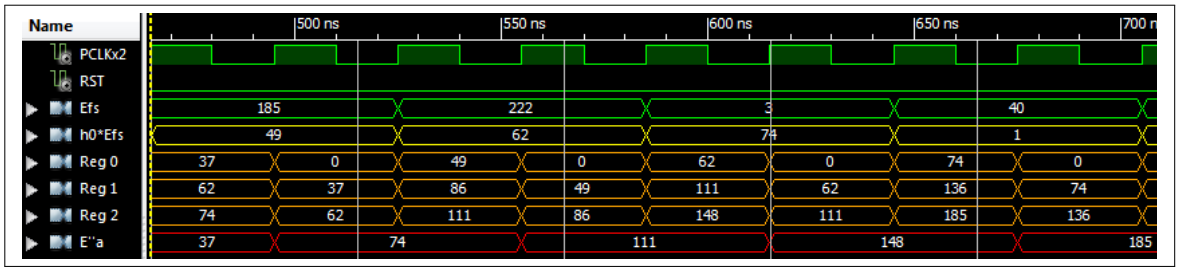
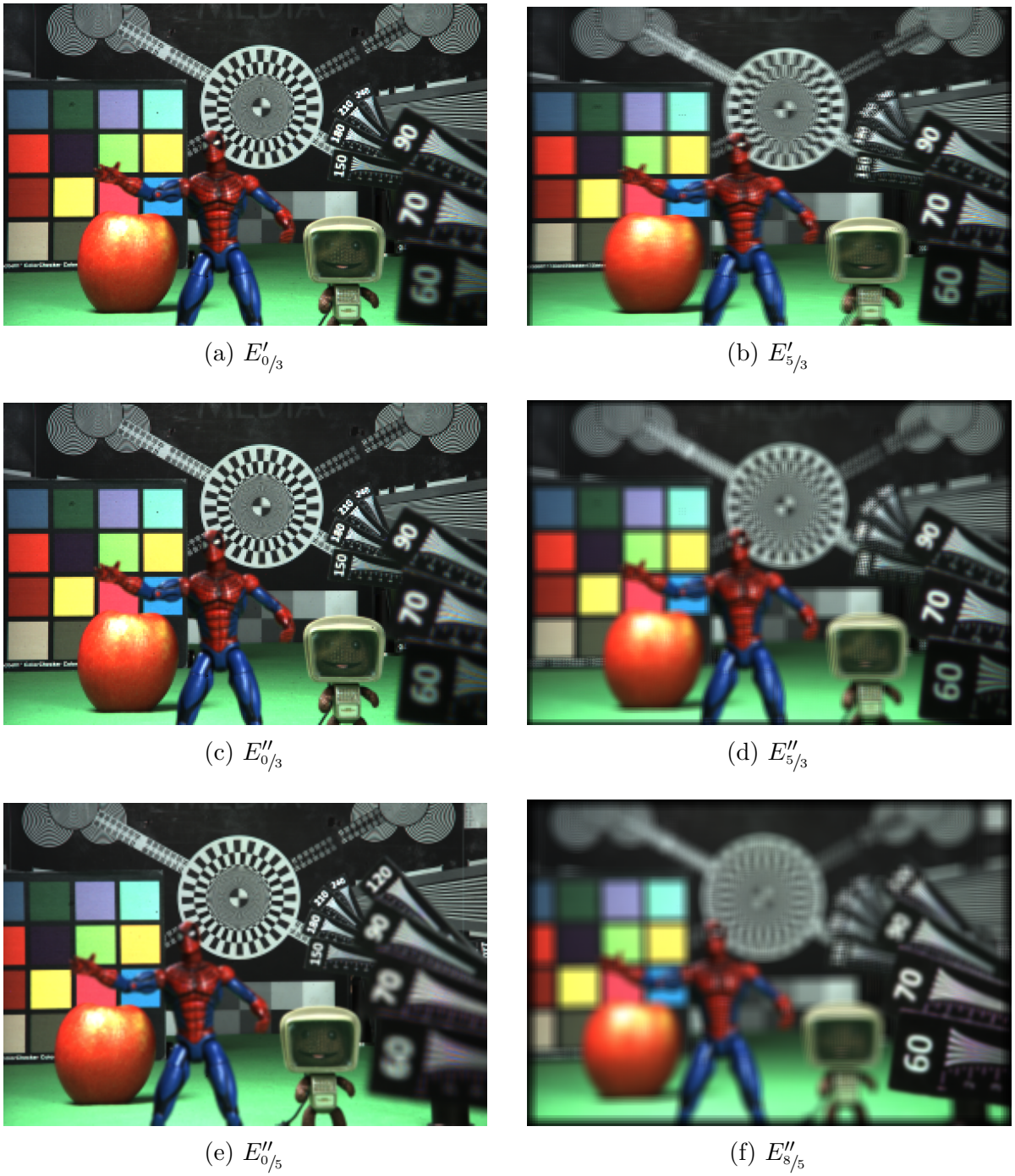


Figure 6.11: **Simulation timing diagram example**.

Figure 6.12 depicts refocused photographs computed by the proposed 2-D module array to accomplish real-time refocusing. Intermediate results after processing images in horizontal direction are seen in Figs. 6.12(a) and 6.12(b). Closer inspection of Fig. 6.12(d) indicates aliasing in blurred regions. This is due to an undersampled directional domain as there are only 3 by 3 samples per micro image ($M = 3$) in the incoming light field capture. Aliasing in synthetic image blur is an observation Ng already pointed out in his thesis [59]. To combat the aliasing problem, the author suggests to sufficiently increase the sampling rate in M. Figures 6.12(e) and 6.12(f) show refocused images obtained from a raw capture with a native micro image resolution of 5 by 5 pixels ($M = 5$) using a linear interpolation instead of nearest neighbour. There, it can be seen that aliasing artefacts are satisfyingly suppressed.



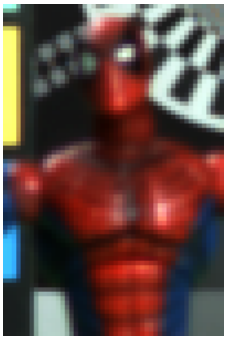
(a) $E'_{0/3}$ (b) $E'_{5/3}$

(c) $E''_{0/3}$ (d) $E''_{5/3}$

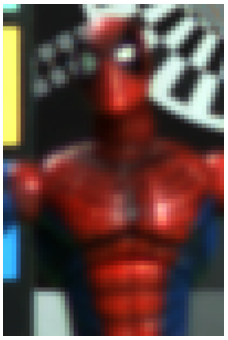
(e) $E''_{0/5}$ (f) $E''_{8/5}$

Figure 6.12: **Refocused photographs using the proposed architecture**. (a) to (d) Incoming and outgoing spatial image resolution amounts to 843 by 561 pixels with $M = 3$. (a) shows an intermediate horizontally processed image and (c) the final refocused photograph for $a = {}^{0}/_{3}$. (b) and (d) represent horizontally and entirely refocused images with $a = {}^{5}/_{3}$. For comparison, output images in (e) and (f) expose improved synthetic blur by using a linear interpolation of whole micro images with $M = 5$.

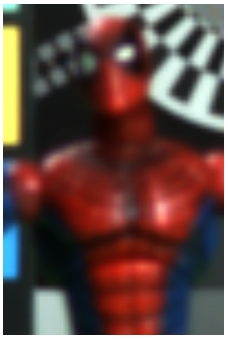
(a) Without interpolation



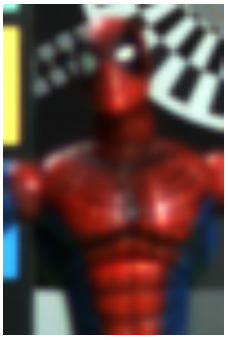
(b) NN interpolation before refocusing



(c) NN interpolation after refocusing



(d) Linear interpolation before refocusing



(e) Linear interpolation after refocusing

(f) NN interpolation with $a = 5/5$

(g) NN interpolation with $a = 6/5$

Figure 6.13: **Resolution comparison.** The images in (a) to (e) show the same region refocused with $a = 3/3$ whilst applying different interpolation techniques before shift and integration. Images in (f) and (g) are NN-interpolated versions of a different region with varying a.

A comparison of output image resolutions when interpolating whole micro images is provided in Fig. 6.13. Results in (a) to (e) suggest that interpolating a raw light field image before refocusing with $a \in \mathbb{Z}$ corresponds to a conventional interpolation of a 2-D image. On the contrary, an effective resolution enhancement can be observed when comparing (f) where $a = 5/5$ with (g) where $a = 6/5$ which are both computed from the same raw image using NN-interpolation. This is a consequence of the micro image repetition and the interleaving filter kernel of the refocusing synthesis which yields identical values for adjacent output pixels when $a \in \mathbb{Z}$ but varying intensities for the same pixels if $a \in \mathbb{R}$. This can be seen by inspecting output results of the timing diagrams in Fig. 6.3 and Fig. 6.5. Another positive side effect in repeating micro images (NN interpolation) is that it also interpolates slices a in depth domain.

6.3 Summary

In summary, it can be stated that the devised *semi-systolic* filter modules deliver a 99.91 % faster computation than Pérez *et al.* [65] whilst preserving the original resolution. This process involves an adequately short time delay of 96.24 μs for a 3201 by 3201 image resolution ensuring real-time operation. The proposed architecture can serve as a groundwork for *Application-Specific Integrated Circuit* (ASIC) chips.

A limitation of the results is that timing delays have been extrapolated and need to be verified using chip analysing tools. As the number of required PEs grows with higher image resolutions, care needs to be taken to prevent long wires in the broadcasting net. For the hardware system's reliability, it is also recommended to convert *semi-systolic* arrays into a *full-systolic* architecture.

Furthermore, to achieve consistency in M, micro image cropping has to be integrated on the FPGA chip as a preceding processing stage. Optionally, a linear or bilinear interpolation can be implemented to improve the perceived resolution although this would cause an additional time delay.

A competitive design approach may be to conceive a refocusing architecture based on the *Fourier Slice Photography Theorem* which is presented in Chapter 2. As stated at the beginning of this chapter, it is, however, expected that the *Fourier* transformation produces larger time delays. A considerable alternative to an FPGA-based implementation is the employment of a *Graphics Processing Unit* (GPU) as this takes less design effort, however, by inducing worse power efficiency. A benchmark comparison is left for future work.

CHAPTER 7

Conclusion

With light field imaging, a new way of setting the focus has found its way into photography. Various models of so-called plenoptic cameras are available for industrial and scientific applications and even for the consumer market [55, 66]. These cameras use the depth information to extract the displacement of the image plane from the intended focal plane and to refocus the image virtually after the picture or movie has been taken.

This study was set out to explore the light field geometry spanned by a *Standard Plenoptic Camera* (SPC) to determine object distances in the captured depth range. It has analysed methods to efficiently synthesise light field images, pinpoint distances to which a photograph is computationally focused just as positions of an observer's viewpoint. As the literature on this subject and specifically in the context of refocusing distance and stereoscopic baseline estimates was incomplete, this research aimed to tackle the following questions:

- Is it possible to predict the light field geometry (i.e. refocusing distance or baseline) based on the parameters of an SPC, and if so, what are the model's limitations?
- Would an efficient hardware architecture enable the SPC to accomplish refocusing in real-time if implemented on an FPGA?

The following remarks provide answers to the above questions whilst reflecting the thesis statements and identifying potential implications of the research findings.

7.1 Outcomes

This thesis introduced a paraxial light field ray model which assists to specify the geometry of light rays captured by an SPC. The model has been empirically validated for several optical setups by means of a real ray simulation tool [89] and real image captures. Experiments with our customised SPC precisely matched the model's predictions indicating no significant deviation. This may be due to the fact that the effective deviation is smaller than our SPC's depth resolution making it difficult to measure a quantitative error. When comparing the model's estimates with that of a real ray simulator [89], results suggest that the proposed light field ray model deviates by less than 0.35 %. Since optical aberrations were not taken into consideration, the error may reflect the lens system's performance in respect of its ability to suppress aberrations.

Returning to the questions posed at the beginning, it is now possible to state that the computational focusing distance of an SPC can be accurately predicted and the traditional stereo triangulation successfully applied to an SPC. Thereby, this research has found that the focal lengths of micro and main lens just as the main lens focus setting control the location of refocused object planes. The model also revealed that virtual cameras can be seen as lined up on the entrance pupil with baselines constrained by the pupil's diameter which are thus smaller than that of conventional camera arrays. This phenomenon may prove to be advantageous for close range applications. The ability to predict the baseline is also necessary if SPC content shall be displayed on autostereoscopic monitors. The proposed model can be helpful for object distance estimation when used in combination with depth from focus [6] through refocusing or disparity analysis [1] based on sub-aperture images. If used for estimating object distances, the accuracy boundaries depend on the camera's capability to resolve depth which corresponds to the number of samples per micro image.

Moreover, the presented model makes it easy to grasp the concept of computational focus variation in spatial domain. It was demonstrated that this notion facilitates to derive optimised FIR refocusing filter kernels for a time- and cost-efficient hardware implementation. Simulating the conceived architecture has proven that real-time refocusing can be accomplished with a computation time of 96.24 μs per frame reducing the time interval by 99.91 % in comparison to a previous state-of-the-art attempt. In doing so, it was shown how to retain the overall image resolution in refocused photographs by interpolating whole micro images.

In addition, methods to optically calibrate a plenoptic camera using a collimator have been presented. When attempting to verify the calibration, a new practical method was discovered by chance to find objective lens pupil locations. Besides, the sub-pixel precision in SPC image rendering proved to be essential which further supports previous studies. Parts of these outcomes have been summarised and made available on **www.plenoptic.info**, a website which includes animated figures aiming to convey the idea behind a plenoptic camera to a broader and not necessarily scientific audience.

On 9th July 2015, Lytro Inc. announced a firmware release enabling its commercially available SPC to supply refocusing distance estimates on the camera screen [55]. Yet, the so-called *Depth Assist Bar* does not provide DoF boundaries. Methods as well as quantitative results justifying the accuracy are not disclosed. Articles [40, 41] concerning with this topic were published at the time of this firmware release.

7.2 Future Work

The major contribution of the proposed ray model is that it allows any SPC to be used as an object distance estimator. A broad range of applications can benefit from this invention including:

- automotive industry for tracking objects in road traffic
- photo and video cameras to obtain the computational focusing distance and baseline
- microscopes and endoscopes to estimate distances in very close depth ranges
- robotics industry, i.e. robots in space
- machine vision systems used for quality assurance on assembly lines

What is now needed is an interdisciplinary study involving the proposed distance prediction system and blur metric in conjunction with an object segmentation to develop a competitive approach to existing depth map acquisition techniques. For instance, a further study might carry out a benchmark assessment between depth recovery from refocused images and the common disparity analysis.

It is questionable at this stage to which extent the conceived model applies to other types of plenoptic cameras. Although optical compositions of FPC and SPC are similar, their image rendering processes significantly differentiate from each other. Future studies on light ray's trajectories in an FPC or even CAC are therefore recommended.

With regards to the presented refocusing hardware architecture, a more in-depth investigation is necessary to assess its power-efficiency, reliability and examine possibilities to further enhance computation times. Besides, a final hardware implementation requires micro image cropping to be done in a preceding signal processing stage which still needs to be devised.

Because refocusing from properly exposed light field images implies averaging pixel intensities, it is expected that refocused photographs provide excellent noise suppression. It will be interesting to see applications that can benefit from this low noise behaviour which is a favourable consequence of the resolution trade-off in plenoptic cameras.

APPENDIX A

Appendix

```
% written by Christopher Hahne
% for the University of Bedfordshire
% Last update 13/01/16
% Matlab version R2015b (8.6.0.267246)
%
% File description:
% This function takes optical parameters of a standard plenoptic camera
% and computes the distance at which a synthetically focused image
% exhibits least blur. The measurement unit is millimetre.
%
% For further assistance please contact info[at]christopherhahne.de
%
% command execution example:
% refocDist(.009,2.75,0,.125,111.0324,193.2935,-65.5563,193.2935,1,0,'on')
%
% pp := pixel pitch
% fs := focal length micro lens
% hh := space between micro lens principal planes
% pm := microlens pitch / diameter
% dA := distance from  MLA to main lens exit pupil
% fU := focal length main lens
% HH := space between main lens principal planes
% bU := separation between micro lens plane and H plane of main lens
% a := refocusing parameter for respective object plane
```

```
function [output] = refocDist(pp,fs,hh,pm,dA,fU,HH,bU,a,crop,draw)

    D = 72; % entrance pupil diameter of main lens

    % (s,u) coordinates for the intersecting rays
    umax = floor(pm/pp);
    smax = 2*umax+1;
    sc = (smax-1)/2;

    if mod(umax,2) ~= 1
        warning('Given parameters yield even micro image resolution.');
        umax = umax-1;
    end

    % analyse crop value which minimizes micro image resolution
    if (~exist('crop','var'))
        crop = 0;
    elseif crop > (umax-1)/2-1 || crop < 0
        crop = (umax-1)/2-1;
        warning('Crop value is set to %s pixels.',num2str(crop));
    else
        crop = round(crop);
    end

    % reduce micro image size if cropped
    umax = umax-2*crop;

    % warning
    if fU > bU
        warning('Image distance is smaller than focal length (fU>bU).')
    elseif a >= (smax-1)/2
        warning('Refocusing slice is out of range.')
    end

    % align intersection to be as paraxial as possible
    c = (umax-1)/2;
    sjShift = -round(a*(umax-1-2*crop)/2);

    % set starting position for micro lens s and pixel u
    j0 = sjShift;
    s0 = j0*pm;
```

```
mc0 = -s0/dA;
uc0 = -mc0*fs+s0;
i0 = c;
u0 = uc0+i0*pp;
mij0 = (s0-u0)/fs;
Uij0 = mij0*bU+s0;
Fij0 = mij0*fU;
qij0 = (Fij0-Uij0)/fU;

% neighbour for shift+a slice
j2 = a*(umax-1)+sjShift;
s2 = j2*pm;
mc2 = -s2/dA;
uc2 = -mc2*fs+s2;
i2 = -c;
u2 = uc2+i2*pp;
mij2 = (s2-u2)/fs;
Uij2 = mij2*bU+s2;
Fij2 = mij2*fU;
qij2 = (Fij2-Uij2)/fU;

% intersection of ray lines from pixel centres into object space
funD = @(x) (qij0*x+Uij0) - (qij2*x+Uij2);
% intersection of ray lines behind image sensor
funDnew = @(x) (mij0*x+s0) - (mij2*x+s2);
bNew = bU - fzero(funDnew,0);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                              %
%        Depth of Field        %
%                              %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

u0U = u0+pp/2;
u0L = u0-pp/2;
m0U = (s0+pm/2-u0U)/fs;
m0L = (s0-pm/2-u0L)/fs;
mij0U = (s0-u0U)/fs;
mij0L = (s0-u0L)/fs;
Uij0U = mij0U*bU+s0+pm/2;
Uij0L = mij0L*bU+s0-pm/2;
Fij0U = mij0U*fU;
```

```
Fij0L = mij0L*fU;
qij0U = (Fij0U - Uij0U)/fU;
qij0L = (Fij0L - Uij0L)/fU;

u2U = u2+pp/2;
u2L = u2-pp/2;
m2U = (s2+pm/2-u2U)/fs;
m2L = (s2-pm/2-u2L)/fs;
mij2U = (s2-u2U)/fs;
mij2L = (s2-u2L)/fs;
Uij2U = mij2U*bU+s2+pm/2;
Uij2L = mij2L*bU+s2-pm/2;
Fij2U = mij2U*fU;
Fij2L = mij2L*fU;
qij2U = (Fij2U - Uij2U)/fU;
qij2L = (Fij2L - Uij2L)/fU;

funDm = @(x) (qij2U*x+Uij2U) - (qij0L*x+Uij0L);
funDp = @(x) (qij0U*x+Uij0U) - (qij2L*x+Uij2L);

% intersection of ray lines behind image sensor
funDmNew = @(x) (mij0L*x+s0-pm/2) - (mij2U*x+s2+pm/2);
bMinusNew =  bU - fzero(funDmNew,0);
funDpNew = @(x) (mij0U*x+s0+pm/2) - (mij2L*x+s2-pm/2);
bPlusNew = bU - fzero(funDpNew,0);

% check if plane is at infinity
if (fU == bU || fU > bU) && a <= 0
    d = inf;
    dY = qij2*(d-fs-hh-bU-HH) + Uij2;
    dPlus = inf;
    dMinus = fzero(funDm,fs+hh+bU+HH+fU) + bU + HH;
    if dMinus < 0
        dMinus = inf;
    end
    dNew = inf;
    dPlusNew = inf;
    dMinusNew = (1/fU-1/bMinusNew)^-1 + bU + HH;
else
    if fU > bNew  % check if plane goes beyond infinity
        warning('Refocused object plane goes beyond infinity.');
    end
```

```
    d = fzero(funD,fs+hh+bU+HH+fU) + bU + HH;
    if d < 0
        d = inf;
    end
    dY = qij2*(d-fs-hh-bU-HH) + Uij2;
    dPlus = fzero(funDp,fs+hh+bU+HH+fU) + bU + HH;
    if dPlus < 0
        dPlus = inf;
    end
    dMinus = fzero(funDm,fs+hh+bU+HH+fU) + bU + HH;
    if dMinus < 0
        dMinus = inf;
    end
    dNew = (1/fU-1/bNew)^-1 + bU + HH;
    dPlusNew = (1/fU-1/bPlusNew)^-1 + bU + HH;
    dMinusNew = (1/fU-1/bMinusNew)^-1 + bU + HH;
end

% Depth of field
DoF = dPlus - dMinus;

% comparison of image and object side approach
if round(d)~=round(dNew) || round(dPlus)~=round(dPlusNew) ...
        || round(dMinus)~=round(dMinusNew)
    warning('Image and object side approach yield different results.');
end

% output vector
output = [d dMinus dPlus];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                          %
%       plot diagram       %
%                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%

if (~exist('draw','var'))
    draw = 'off';
end

if strcmp(draw,'on')
```

```
% plot rays on an optical bench
plotLim = 5000; % set 50 meter as maximum plot distance
if dPlus > 0 && dPlus ~= Inf
    largeVal = dPlus;
elseif dMinus > 0 && dMinus < plotLim
    largeVal = dMinus;
else
    largeVal = plotLim;
end

xmax = round( largeVal + largeVal/10 );
sampleDensity = xmax/5000;
x = 0:sampleDensity:xmax;

% rays from sensor to micro lens h plane
y0(x<=fs) = mij0*x(x<=fs) + u0;
y2(x<=fs) = mij2*x(x<=fs) + u2;
y0Aux(x<=fs) = mc0*x(x<=fs) + uc0;
y2Aux(x<=fs) = mc2*x(x<=fs) + uc2;

y0(x>fs&x<=fs+hh) = s0;
y2(x>fs&x<=fs+hh) = s2;
y0Aux(x>fs&x<=fs+hh) = s0;
y2Aux(x>fs&x<=fs+hh) = s2;

% rays from micro lens h to H1 plane of main lens (thin lens)
y0(x>fs+hh&x<=fs+hh+bU) = ...
    mij0*x(x>fs+hh&x<=fs+hh+bU) + s0 - mij0*(hh+fs);
y2(x>fs+hh&x<=fs+hh+bU) = ...
    mij2*x(x>fs+hh&x<=fs+hh+bU) + s2 - mij2*(hh+fs);
y0Aux(x>fs+hh&x<=fs+hh+bU) = ...
    mc0*x(x>fs+hh&x<=fs+hh+bU) + s0 - mc0*(hh+fs);
y2Aux(x>fs+hh&x<=fs+hh+bU) = ...
    mc2*x(x>fs+hh&x<=fs+hh+bU) + s2 - mc2*(hh+fs);

% rays in space between principal planes of main lens
y0(x>fs+hh+bU&x<=fs+hh+bU+HH) = Uij0;
y2(x>fs+hh+bU&x<=fs+hh+bU+HH) = Uij2;
y0Aux(x>fs+hh+bU&x<=fs+hh+bU+HH) = ...
    mc0*x(x>fs+hh+bU&x<=fs+hh+bU+HH) + mc0*HH;
y2Aux(x>fs+hh+bU&x<=fs+hh+bU+HH) = ...
    mc2*x(x>fs+hh+bU&x<=fs+hh+bU+HH) + mc2*HH;
```

```
% rays from main lens plane to object space
y0(x>fs+hh+bU+HH) = ...
    qij0*x(x>fs+hh+bU+HH) + Uij0 - (qij0*(fs+hh+bU+HH));
y2(x>fs+hh+bU+HH) = ...
    qij2*x(x>fs+hh+bU+HH) + Uij2 - (qij2*(fs+hh+bU+HH));
y0Aux(x>fs+hh+bU) = 0*x(x>fs+hh+bU) + 0 - (0*(fs+hh+bU));
y2Aux(x>fs+hh+bU) = 0*x(x>fs+hh+bU) + 0 - (0*(fs+hh+bU));

% DoF rays
yUp(x<=fs) = m0U*x(x<=fs) + u0U;
yUp(x>fs&x<=fs+hh) = s0+pm/2;
yUp(x>fs+hh&x<=fs+hh+bU) = ...
    mij0U*x(x>fs+hh&x<=fs+hh+bU) + s0+pm - m0U*fs - mij0U*hh;
yUp(x>fs+hh+bU&x<=fs+hh+bU+HH) = Uij0U;
yUp(x>fs+hh+bU+HH) = ...
    qij0U*x(x>fs+hh+bU+HH) + Uij0U - (qij0U*(fs+hh+bU+HH));

yLm(x<=fs) = m0L*x(x<=fs) + u0L;
yLm(x>fs&x<=fs+hh) = s0-pm/2;
yLm(x>fs+hh&x<=fs+hh+bU) = ...
    mij0L*x(x>fs+hh&x<=fs+hh+bU) + s0-pm - m0L*fs - mij0L*hh;
yLm(x>fs+hh+bU&x<=fs+hh+bU+HH) = Uij0L;
yLm(x>fs+hh+bU+HH) = ...
    qij0L*x(x>fs+hh+bU+HH) + Uij0L - (qij0L*(fs+hh+bU+HH));

yUm(x<=fs) = m2U*x(x<=fs) + u2U;
yUm(x>fs&x<=fs+hh) = s2+pm/2;
yUm(x>fs+hh&x<=fs+hh+bU) = ...
    mij2U*x(x>fs+hh&x<=fs+hh+bU) + s2+pm - m2U*fs - mij2U*hh;
yUm(x>fs+hh+bU&x<=fs+hh+bU+HH) = Uij2U;
yUm(x>fs+hh+bU+HH) = ...
    qij2U*x(x>fs+hh+bU+HH) + Uij2U - (qij2U*(fs+hh+bU+HH));

yLp(x<=fs) = m2L*x(x<=fs) + u2L;
yLp(x>fs&x<=fs+hh) = s2-pm/2;
yLp(x>fs+hh&x<=fs+hh+bU) = ...
    mij2L*x(x>fs+hh&x<=fs+hh+bU) + s2-pm - m2L*fs - mij2L*hh;
yLp(x>fs+hh+bU&x<=fs+hh+bU+HH) = Uij2L;
yLp(x>fs+hh+bU+HH) = ...
    qij2L*x(x>fs+hh+bU+HH) + Uij2L - (qij2L*(fs+hh+bU+HH));
```

```
da=strcat('$d_{',num2str(a),'}$');
daMinus=strcat('$d_{',num2str(a),'-}$');
daPlus=strcat('$d_{',num2str(a),'+}$');

yDim = D/12;
padding = D/36;
% labels da, da-, da+
plot(d, dY,'ko',[0, round((d+20)/100)*100],[0, 0], ...
    0,uc0,'ko',0,uc2,'ko',0,u0,'go',0,u2,'bo');
text(d+DoF/16,yDim-padding+1,da,'FontSize',16,'color','c', ...
    'interpreter','latex')
text(dMinus-dMinus/100,yDim-padding+1,daMinus,'FontSize',16, ...
        'color','r','interpreter','latex', ...
        'HorizontalAlignment','right')
text(dPlus+DoF/16,yDim-padding+1,daPlus,'FontSize',16, ...
        'color','k','interpreter','latex')

% diagram title, labels, size and attributes
h = gca;
grid off, axis on
xlabel('$z_U$~[mm]','FontSize',20,'interpreter','latex')
ylabel('$(u,s)$~[mm]','FontSize',20,'interpreter','latex')
axis([0 xmax -yDim-padding+2 yDim+padding-2]);
set(h,'FontSize',18)

% DoF plots
line('XData', x, 'YData', yLm, 'LineStyle', '-', 'Color', 'r');
line('XData', x, 'YData', yUm, 'LineStyle', '-', 'Color', 'r');
line('XData', x, 'YData', yUp, 'LineStyle', '-', 'Color', 'k');
line('XData', x, 'YData', yLp, 'LineStyle', '-', 'Color', 'k');

% intersecting rays for center of corresponding refocusing plane
line('XData', x, 'YData', y0Aux, 'LineStyle', '-', 'Color', 'y');
line('XData', x, 'YData', y2Aux, 'LineStyle', '-', 'Color', 'y');
line('XData', x, 'YData', y0, 'LineStyle', '-', 'Color', 'g');
line('XData', x, 'YData', y2, 'LineStyle', '-', 'Color', 'b');

% intersection planes
line('XData', [d d], 'YData', ...
    [yDim+padding -yDim-padding], 'LineStyle', '--', 'Color', 'c')
line('XData', [dPlus dPlus], 'YData', ...
    [yDim+padding -yDim-padding], 'LineStyle', '--', 'Color', 'k')
```

```
line('XData', [dMinus dMinus], 'YData', ...
    [yDim+padding -yDim-padding], 'LineStyle', '--', 'Color', 'r')

% sensor
line('XData', [0 0], 'YData', [sc*pm -sc*pm], ...
    'LineStyle', '-', 'LineWidth', 0.25);
pixelY0 = uc0-pp/2-ceil(pm/pp)/2*pp:pp:uc0+pp/2+ceil(pm/pp)/2*pp;
pixelY2 = uc2-pp/2-ceil(pm/pp)/2*pp:pp:uc2+pp/2+ceil(pm/pp)/2*pp;
pixelX = zeros(size(pixelY0));
line('XData', pixelX, 'YData', pixelY0, ...
    'Marker', '+', 'MarkerSize', 3);
line('XData', pixelX, 'YData', pixelY2, ...
    'Marker', '+', 'MarkerSize', 3);

% micro lens grid
lensY = -sc*pm+pm/2:pm:sc*pm+pm/2;
lensf = -sc*pm:pm:sc*pm;
lensX = (fs+hh) * ones(size(lensY));
line('XData', [fs fs], 'YData', [sc*pm -sc*pm], ...
    'LineStyle', '--', 'LineWidth', 0.25);
line('XData', [fs+hh fs+hh], 'YData', [sc*pm -sc*pm], ...
    'LineStyle', '-', 'LineWidth', 0.25);
line('XData', lensX, 'YData', lensY, 'Marker', '+', ...
    'MarkerSize', 4); % lens borders
line('XData', lensX, 'YData', lensf, 'Marker', '.', ...
    'MarkerSize', 6); % optical axis

% main lens principal planes
line('XData', [fs+hh+bU+HH, fs+hh+bU+HH], 'YData', ...
    [-D/2, D/2], 'LineStyle', '--', 'Color', 'k');
line('XData', [fs+hh+bU, fs+hh+bU], 'YData', ...
    [-D/2, D/2], 'LineStyle', '--', 'Color', 'k');
text(fs+hh+bU+HH+2,yDim-padding+1,'$$H_{1U}$$', ...
    'FontSize',12,'color','k','interpreter','latex')
text(fs+hh+bU+2,yDim-padding+1,'$$H_{2U}$$', ...
    'FontSize',12,'color','k','interpreter','latex')

% main lens focal point
line('XData', [fs+hh+bU+HH+fU, fs+hh+bU+HH+fU], 'YData', ...
    [-D/20, D/20], 'LineStyle', ':', 'Color', 'k');

% optical axis zU
```

```
        line('XData', [0, xmax], 'YData', [0, 0], ...
            'LineStyle', '-', 'Color', 'k');
    end
end
```

```
% written by Christopher Hahne
% for the University of Bedfordshire
% Last update 13/01/16
% Matlab version R2015b (8.6.0.267246)
%
% File description:
% This function takes optical parameters of a standard plenoptic camera
% and computes the object distance as well as the baseline of
% synthetically reconstructed viewpoint images. The measurement unit
% is millimetre.
%
% For further assistance please contact info[at]christopherhahne.de
%
% command execution example:
% triaSPC(.009,2.75,.125,111.6268,193.2935,-65.5563,193.2935,-6,2);

% pp := pixel pitch
% fs := focal length micro lens
% pm := microlens pitch / diameter
% dA := distance from  MLA to main lens exit pupil
% fU := focal length main lens
% HH := space between main lens principal planes
% bU := separation between micro lens plane and H plane of main lens
% i := index for virtual camera
% dx := disparity value

function [ output ] = triaSPC(pp,fs,pm,dA,fU,HH,bU,i,dx)

    j=1;
    s=j*pm;
    mc = -(s/dA);
    uc = -mc*fs+s;
    u0 = uc+pp*i;
    m01 = (s-u0)/fs;
    U01 = m01*bU+s;
    m00 = -pp*i/fs;
    U00 = m00*bU;
    q01 = (m01*fU-U01)/fU;
    q00 = (m00*fU-U00)/fU;
    phi = atand(q00);
    funTri = @(x) (q01*x+U01) - (q00*x+U00);
    xInt = fzero(funTri,0);
```

```
    EP = bU+HH+xInt;
    preB = q00*xInt+U00; % Baseline at entrance pupil
    bNew = bU;
    ppNew = (-q01*bNew+preB)-(-q00*bNew+preB);

    % triangulation
    dxNew = dx'*ppNew;
    preZ = preB * bNew ./ (dxNew + bNew * -tand(phi));
    preZ = preZ + EP;

    % output vector
    output = [preB phi preZ];
end
```

APPENDIX B

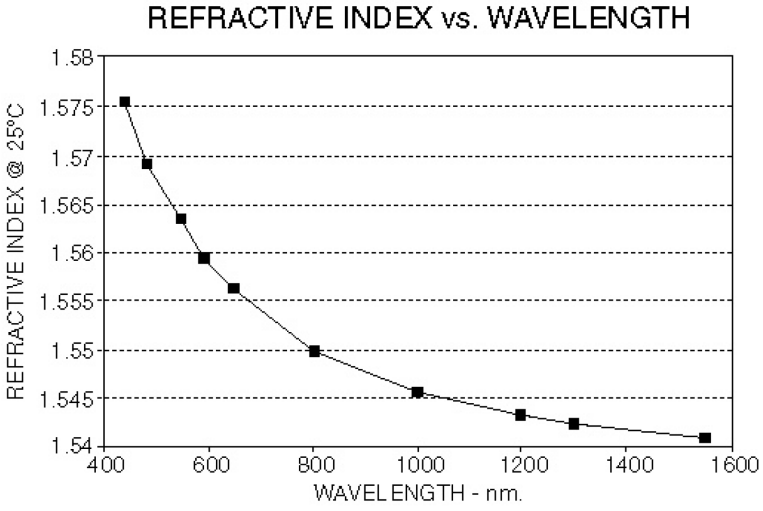
Appendix

factorX = $\bar{k}$ - floor($\bar{k}$);
factorY = $\bar{l}$ - floor($\bar{l}$);
nFactorY = 1 - factorY;
nFactorX = 1 - factorX;
if *factorY >= 0.5* **then**
    if *factorX >= 0.5* **then**
        $\gamma_{(0,0)}$ = factorY * factorX;
        $\gamma_{(1,0)}$ = nFactorX * factorY;
        $\gamma_{(0,1)}$ = nFactorY * factorX;
        $\gamma_{(1,1)}$ = nFactorY * nFactorX;
    else
        $\gamma_{(0,0)}$ = factorY * nFactorX;
        $\gamma_{(1,0)}$ = factorX * factorY;
        $\gamma_{(0,1)}$ = nFactorY * nFactorX;
        $\gamma_{(1,1)}$ = factorX * nFactorY;
    end
else
    if *factorX >= 0.5* **then**
        $\gamma_{(0,0)}$ = nFactorY * factorX;
        $\gamma_{(1,0)}$ = nFactorX * nFactorY;
        $\gamma_{(0,1)}$ = factorY * factorX;
        $\gamma_{(1,1)}$ = nFactorX * factorY;
    else
        $\gamma_{(0,0)}$ = nFactorY * nFactorX;
        $\gamma_{(1,0)}$ = factorX * nFactorY;
        $\gamma_{(0,1)}$ = factorY * nFactorX;
        $\gamma_{(1,1)}$ = factorX * factorY;
    end
end
pixel = $\gamma_{(0,0)} \times E\left[x_{\lfloor \bar{k} \rfloor + 0},\, y_{\lfloor \bar{l} \rfloor + 0}\right] + \gamma_{(1,0)} \times E\left[x_{\lfloor \bar{k} \rfloor + 1},\, y_{\lfloor \bar{l} \rfloor + 0}\right]$
$+\gamma_{(0,1)} \times E\left[x_{\lfloor \bar{k} \rfloor + 0},\, y_{\lfloor \bar{l} \rfloor + 1}\right] + \gamma_{(1,1)} \times E\left[x_{\lfloor \bar{k} \rfloor + 1},\, y_{\lfloor \bar{l} \rfloor + 1}\right]$;

Algorithm 1: Pseudo code for MIC interpolation weighting scheme.

APPENDIX C

Appendix



$$n_{25°C} = 1.5375 + \frac{8290.45}{\lambda^2} - \frac{2.11046 \times 10^8}{\lambda^4}$$

n = refractive index at wavelength
λ = wavelength in nanometers

Figure C.1: **MLA manufacturer's refractive index at $\vartheta = 25$°C.**

Bibliography

[1] Abbeloos, W. [2010], Real-time stereo vision, Master's thesis, Karel de Grote-Hogeschool University College.

[2] Abbeloos, W. [2012], 'Stereo matching', `http://www.mathworks.com/matlabcentral/fileexchange/28522-stereo-matching`.
Online; accessed: 22th April 2015.

[3] Adelson, E. H. and Bergen, J. R. [1991], The plenoptic function and the elements of early vision, *in* 'Computational Models of Visual Processing', Cambridge, MIT Press, pp. 3–20.

[4] Adelson, E. H. and Wang, J. Y. [1992], 'Single lens stereo with a plenoptic camera', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**(2), 99–106.

[5] Aggoun, A., Tsekleves, E., Swash, R., Zarpalas, D., Dimou, A., Daras, P., Nunes, P. and Soares, L. [2013], 'Immersive 3D holoscopic video system', *MultiMedia, IEEE* **20**(1), 28–37.

[6] Asada, N., Fujiwara, H. and Matsuyama, T. [1998], 'Edge and depth from focus', *International Journal of Computer Vision* **26**(2), 153–163.

[7] Bailey, D. G. [2011], *Design for Embedded Image Processing on FPGAs*, Wiley-Blackwell.

[8] Bass, M., DeCusatis, C., Enoch, J., Lakshminarayanan, V., Li, G., Macdonald, C., Mahajan, V. and Van Stryland, E. [2010], *Handbook of Optics, Third Edition Volume II: Design, Fabrication and Testing, Sources and Detectors, Radiometry and Photometry*, Vol. 2, 3 edn, McGraw-Hill, Inc., New York, NY, USA.

[9] Bishop, T. E., Zanetti, S. and Favaro, P. [2009], Light field superresolution, *in* 'IEEE International Conference on Computational Photography (ICCP)'.

[10] Bobick, A. F. and Intille, S. S. [1999], 'Large occlusion stereo', *Int. J. Comput. Vision* **33**(3), 181–200.

[11] Bodrogi, P. and Khan, T. Q. [2012], *Illumination, Color and Imaging: Evaluation and Optimization of Visual Displays*, Wiley Series in Display Technology, Wiley.

[12] Bok, Y., Jeon, H.-G. and Kweon, I. S. [2014], Geometric calibration of micro-lens-based light-field cameras using line features, *in* 'Proceedings of European Conference on Computer Vision (ECCV)'.

[13] Bracewell, R. N. [1956], 'Strip Integration in Radio Astronomy', *Australian Journal of Physics* **9**, 198–217.

[14] Brown, D. C. [1966], 'Decentering distortion of lenses', *Photometric Engineering* **32**(3), 444–462.

[15] Broxton, M., Grosenick, L., Yang, S., Cohen, N., Andalman, A., Deisseroth, K. and Levoy, M. [2013], 'Wave optics theory and 3-D deconvolution for the light field microscope', *Opt. Express* **21**(21), 25418–25439.

[16] Burger, W. and Burge, M. J. [2009], *Principles of Digital Image Processing: Core Algorithms*, 1 edn, Springer Publishing Company, Inc.

[17] Caldwell, B. [2000], 'Fast wide-range zoom for 35 mm format', *Opt. Photon. News* **11**(7), 49–51.

[18] Chaudhuri, S. and Rajagopalan, A. N. [1999], *Depth From Defocus: A Real Aperture Imaging Approach*, 1 edn, Springer Publishing Company, Inc.

[19] Cho, D., Lee, M., Kim, S. and Tai, Y.-W. [2013], Modeling the calibration pipeline of the lytro camera for high quality light-field image reconstruction, *in* 'IEEE International Conference on Computer Vision (ICCV)', pp. 3280–3287.

[20] Coffey, D. F. [1936], Apparatus for making a composite stereograph, Technical Report 2063985, U.S. Patent Comission.

[21] Dansereau, D. G. [2014], Plenoptic Signal Processing for Robust Vision in Field Robotics, PhD thesis, University of Sydney.

[22] Dansereau, D. G., Pizarro, O. and Williams, S. B. [2013], Decoding, calibration and rectification for lenselet-based plenoptic cameras, *in* 'IEEE International

Conference on Computer Vision and Pattern Recognition (CVPR)', pp. 1027–1034.

[23] Dudnikov, Y. A. [1970], 'Autostereoscopy and integral photography', *Sov. J. Opt. Technol.* **37**(7), 422–426.

[24] Dudnikov, Y. A. [1971], 'Elimination of image pseudoscopy in integral photography', *Sov. J. Opt. Technol.* **38**(3), 140–143.

[25] Dudnikov, Y. A. [1974], 'Effect of three-dimensional moiré in integral photography', *Sov. J. Opt. Technol.* **41**(5), 17–20.

[26] Dudnikov, Y. A. and Rozhkov, B. K. [1978], 'Selecting the parameters of the lens-array photographing system in integral photography', *Sov. J. Opt. Technol.* **45**(6), 349–351.

[27] Dudnikov, Y. A. and Rozhkov, B. K. [1979], 'Limiting capabilities of photographing various subjects by the integral photography method', *Sov. J. Opt. Technol.* **46**(12), 736–738.

[28] Dudnikov, Y. A., Rozhkov, B. K. and Antipova, E. N. [1980], 'Obtaining a portrait of a person by the integral photography method', *Sov. J. Opt. Technol.* **47**(3), 562–563.

[29] Ellison, I. [2014], 'What is the focal length of the first generation lytro camera?', `https://support.lytro.com/hc/en-us/articles/200863590-What-is-the-focal-length-of-the-first-generation-Lytro-camera-`.
Online; accessed: 7th October 2014.

[30] Ernst, H., Geschke, D., Kirsten, P. and Schenk, W. [2013], *Physikalisches Praktikum*, Vieweg+Teubner Verlag.

[31] Fossum, E., Mendis, S. and Kemeny, S. [1995], 'Active pixel sensor with intra-pixel charge transfer'.
US Patent 5,471,515.

[32] Garbe, C. S., Voss, B. and Stapf, J. [2012], Plenoptic particle streak velocimetry (pPSV): 3D3C fluid flow measurement from light fields with a single plenoptic camera, *in* '16th Int. Symp. on Applications of Laser Techniques to Fluid Mechanics'.

[33] Georgiev, T. and Lumsdaine, A. [2009], Superresolution with plenoptic 2.0 cameras, *in* 'Frontiers in Optics 2009/Laser Science XXV/Fall 2009 OSA Optics & Photonics Technical Digest', Optical Society of America, p. STuA6.

[34] Georgiev, T. and Lumsdaine, A. [2012], The multi-focus plenoptic camera, *in* 'Digital Photography VIII', Vol. Proc. SPIE 8299.

[35] Georgiev, T., Lumsdaine, A. and Goma, S. [2011], Plenoptic principal planes, *in* 'Computational Optical Sensing and Imaging'.

[36] Georgiev, T., Zheng, K. C., Curless, B., Salesin, D., Nayar, S. and Intwala, C. [2006], Spatio-angular resolution tradeoff in integral photography, *in* 'Proceedings of Eurographics Symposium on Rendering'.

[37] Hach, T. and Steurer, J. [2013], A novel RGB-Z camera for high-quality motion picture applications, *in* 'Proceedings of the 10th European Conference on Visual Media Production', CVMP '13, ACM, New York, NY, USA, pp. 4:1–4:10.

[38] Hahne, C. and Aggoun, A. [2014], Embedded FIR filter design for real-time refocusing using a standard plenoptic video camera, *in* 'Digital Photography X', Vol. Proc. SPIE 9023.

[39] Hahne, C., Aggoun, A., Haxha, S., Velisavljevic, V. and Fernández, J. C. J. [2014*a*], Baseline of virtual cameras acquired by a standard plenoptic camera setup, *in* '3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2014', pp. 1–3.

[40] Hahne, C., Aggoun, A., Haxha, S., Velisavljevic, V. and Fernández, J. C. J. [2014*b*], 'Light field geometry of a standard plenoptic camera', *Opt. Express* **22**(22), 26659–26673.

[41] Hahne, C., Aggoun, A. and Velisavljevic, V. [2015], The refocusing distance of a standard plenoptic photograph, *in* '3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)'.

[42] Hecht, E. [2001], *Optics*, 4 edn, Addison Wesley.

[43] Huang, F., Chen, K. and Wetzstein, G. [2015], 'The Light Field Stereoscope: Immersive Computer Graphics via Factored Near-Eye Light Field Displays with Focus Cues', *ACM Trans. Graph. (SIGGRAPH)* (4).

[44] Iocchi, L. [1998], 'Stereo vision: Triangulation', `http://www.dis.uniroma1.it/~iocchi/stereo/triang.html`.
Online; accessed: 2nd March 2015.

[45] Isaksen, A. [2000], Dynamically reparameterized light fields, Master's thesis, Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

[46] Isaksen, A., McMillan, L. and Gortler, S. J. [2000], Dynamically reparameterized light fields, *in* 'Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques', SIGGRAPH '00, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, pp. 297–306.

[47] Ives, F. E. [1903], 'Parallax stereogram and process of making same'.
US Patent No. 725,567.

[48] Ives, H. E. [1930], 'Parallax panoramagrams made with a large diameter lens', *Journal of the Optical Society of America* **20**(6), 332–340.

[49] Johannsen, O., Heinze, C., Goldluecke, B. and Perwass, C. [2013], On the calibration of focused plenoptic cameras, *in* 'Time-of-Flight and Depth Imaging', Vol. 8200 of *Lecture Notes in Computer Science*, Springer, pp. 302–317.

[50] Levoy, M. and Hanrahan, P. [1996], Light field rendering, Technical report, Stanford University.

[51] Levoy, M., Ng, R., Adams, A., Footer, M. and Horowitz, M. [2006], 'Light field microscopy', *ACM Trans. Graph.* **25**(3), 924–934.

[52] Lippmann, G. [1908], 'Épreuves réversibles donnant la sensation du relief', *Académie Des Sciences* pp. 446–451.

[53] Lipson, S. G., Lipson, H. and Tannhauser, D. S. [1995], *Optical Physics*, Optical Physics, Cambridge University Press.

[54] Lumsdaine, A. and Georgiev, T. [2008], Full resolution lightfield rendering, Technical report, Adobe Systems, Inc.

[55] Lytro Inc. [2015], 'Lytro ILLUM User Manual', `https://s3.amazonaws.com/lytro-corp-assets/manuals/english/illum_user_manual.pdf`.
2.1 edn.

[56] Marr, D. and Poggio, T. [1976], Cooperative computation of stereo disparity, Technical report, Massachusetts Institute of Technology, Cambridge, MA, USA.

[57] Mavridaki, E. and Mezaris, V. [2014], No-reference blur assessment in natural images using fourier transform and spatial pyramids, *in* '2014 IEEE International Conference on Image Processing (ICIP)', pp. 566–570.

[58] Ng, R. [2005], 'Fourier slice photography', *ACM Trans. Graph.* **24**(3), 735–744.

[59] Ng, R. [2006], Digital Light Field Photography, PhD thesis, Stanford University.

[60] Ng, R. and Hanrahan, P. [2006], Digital correction of lens aberrations in light field photography, *in* 'International Optical Design Conference', Vol. Proc. SPIE 6342, Stanford University.

[61] Ng, R., Levoy, M., Brèdif, M., Duval, G., Horowitz, M. and Hanrahan, P. [2005], Light field photography with a hand-held plenoptic camera, Technical Report CTSR 2005-02, Stanford University.

[62] Nolf, M. [2012], 'Lytro specifications: A deeper look inside', `http://lightfield-forum.com/2012/11/lytro-specifications-a-deeper-look-inside/`.
Online; accessed: 12th May 2015.

[63] Nonn, T., Kitzhofer, J., Hess, D. and Brücker, C. [2012], Measurements in an ic-engine flow using light-field volumetric velocimetry, *in* '16th Int. Symp. on Applications of Laser Techniques to Fluid Mechanics'.

[64] Otsu, N. [1979], 'A threshold selection method from gray-level histograms', *IEEE Transactions on Systems, Man, and Cybernetics* **9**(1), 62–66.

[65] Pérez, J., Magdaleno, E., Pérez, F., Rodríguez, M., Hernández, D. and Corrales, J. [2014], 'Super-resolution in plenoptic cameras using FPGAs', *Sensors* **14**(5), 8669.

[66] Perwass, C. and Wietzke, L. [2012], Single lens 3D-camera with extended depth-of-field, *in* 'Human Vision and Electronic Imaging XVII', Vol. Proc. SPIE 8291, Raytrix GmbH.

[67] Perwass, U. and Perwass, C. [2012], 'Digital imaging system, plenoptic optical device and image data processing method'.
US Patent App. 13/264,173.

[68] Radon, J. [1917], 'Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten', *Akad. Wiss.* **69**, 262–277.

[69] Rodríguez-Ramos, L. F., Marín, Y., Díaz, J. J., Piqueras, J., García-Jiménez, J. and Rodríguez-Ramos, J. M. [2010], FPGA-Based Real Time Processing of the Plenoptic Wavefront Sensor, *in* 'Adaptative Optics for Extremely Large Telescopes', p. 7007.

[70] Schubert, E. F. [2006], *Light-Emitting Diodes*, Proceedings of SPIE - The International Society for Optical Engineering, Cambridge University Press.

[71] Sokolov, P. P. [1911], 'Autostereoscopy and integral photography by prof. lippman's method', *Moscow State University Press* .
Translator: Ekaterina Avramova, Editor: Daniel Reetz.

[72] Steurer, J. H., Pesch, M. and Hahne, C. [2012], 3D holoscopic video imaging system, *in* 'Human Vision and Electronic Imaging XVII', Vol. Proc. SPIE 8291, Arnold & Richter GmbH & Co. KG.

[73] Tao, M., Hadap, S., Malik, J. and Ramamoorthi, R. [2013], Depth from combining defocus and correspondence using light field cameras, *in* 'International Conference on Computer Vision (ICCV)', pp. 673–680.

[74] Tao, M., Srinivasa, P., Malik, J., Rusinkiewicz, S. and Ramamoorthi, R. [2015], Depth from shading, defocus, and correspondence using light-field angular coherence, *in* 'Computer Vision and Pattern Recognition (CVPR)'.

[75] TRIOPTICS [n.d.], 'MTF measurement and further parameters', `http://www.trioptics.com/knowledge-base/mtf-and-image-quality/`.
Online; accessed: 3th October 2015.

[76] Vaish, V. [2007], Synthetic Aperture Imaging Using Dense Camera Arrays, PhD thesis, Stanford University.

[77] Veeraraghavan, A., Raskar, R., Agrawal, A., Mohan, A. and Tumblin, J. [2007], 'Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing', *ACM Trans. Graph.* **26**(3).

[78] Venkataraman, K., Lelescu, D., Duparré, J., McMahon, A., Molina, G., Chatterjee, P., Mullis, R. and Nayar, S. [2013], 'Picam: An ultra-thin high performance monolithic camera array', *ACM Trans. Graph.* **32**(6), 166:1–166:13.

[79] Wanner, S., Fehr, J. and Jähne, B. [2011], Generating EPI representations of 4D light fields with a single lens focused plenoptic camera, *in* 'Proceedings of the 7th International Conference on Advances in Visual Computing (ISVC)', Volume Part I, Springer-Verlag, pp. 90–101.

[80] Wei, W., Zhong, Z., Yinghe, W. and Tao, Y. [2013], 'Endoscopic method for light field video camera'.
CN Patent App. CN 201,310,331,566.

[81] Wheatstone, C. [1838], 'On some remarkable, and hitherto unobserved, phenomena of binocular vision', *Philosophical Transactions* **128**, 371–394.

[82] Wimalagunarathne, R., Madanayake, A., Dansereau, D. G. and Bruton, L. T. [2012], A systolic-array architecture for first-order 4-D IIR frequency-planar digital filters, *in* 'IEEE International Symposium on Circuits and Systems (ISCAS)', pp. 3069–3072.

[83] Xilinx, Inc. [2002], 'A 1D Systolic FIR', `http://www.iro.umontreal.ca/~aboulham/F6221/Xilinx%20A%201D%20systolic%20FIR.htm`.
Online; accessed: 12th December 2015.

[84] Xu, Z., Ke, J. and Lam, E. Y. [2012], 'High-resolution lightfield photography using two masks', *Opt. Express* **20**(10), 10971–10983.

[85] Yanagisawa, M. [1990], 'Optical system having a variable out-of-focus state'.
US Patent 4,908,639.

[86] Yang, Y., Yuille, A. L. and Lu, J. [1993], Local, global, and multilevel stereo matching, *in* 'IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)', pp. 274–279.

[87] Yu, Z., Yu, J., Lumsdaine, A. and Georgiev, T. [2012], An analysis of color demosaicing in plenoptic cameras, *in* 'IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)', pp. 901–908.

[88] Zeller, N., Quint, F. and Stilla, U. [2014], 'Calibration and accuracy analysis of a focused plenoptic camera', *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **II-3**, 205–212.

[89] Zemax LLC [2011], 'Optical design program (version 110711)'.

Notation

A	Aperture
A'	Exit pupil
A''	Entrance pupil
A''_i	Virtual camera position at the entrance pupil indexed by i
$\overline{A''H_{1U}}$	distance from entrance pupil to principal plane
$\mathbf{A}$	1^{st} Chief ray coordinate pair for intersection at refocusing plane
$\mathbf{A}\pm$	1^{st} DoF ray coordinate pair for intersection at DoF planes
a	Refocusing parameter
a_U	Main lens object distance
$a_U{}'$	Refocused object distance
$a_U{}'_{\pm}$	Refocused main lens object distance for DoF boundaries
a_s	Micro lens object distance
B	Baseline
$B_{A'}$	Baseline in SPC at exit pupil
B_G	Baseline in SPC at entrance pupil
$\mathcal{B}, \mathcal{B}_\alpha$	Change of basis operator in *Fourier Slice Photography Theorem*
$\mathbf{B}$	2^{nd} Chief ray coordinate pair for intersection at refocusing plane
$\mathbf{B}\pm$	2^{nd} DoF ray coordinate pair for intersection at DoF planes
b	Image distance of a lens
b_U	Main lens image distance
$b_U{}'$	Refocused main lens image distance
$b_U{}'_{\pm}$	Refocused main lens image distance for DoF boundaries
b_s	Micro lens image distance
c	Number of pixels separating centroid from micro image edge
$\bar{\mathbf{c}}_{j,h}$	Vector containing centroid coordinates of all MICs indexed by j, h

D	Resolvable feature size in Abbe's diffraction limit
$d_{A'}$	Distance from MLA to exit pupil A'
$d_{A''}$	Distance from MLA to entrance pupil A''
d_a	Distance from MLA to refocused plane
$d_{a\pm}$	Distance from MLA to DoF boundaries of refocused plane
d_f	Distance from MLA to the plane the main lens is focused on
d_s	Distance between sensor and back vertex of MLA
d_M	Side length of equilateral triangle in hexagonal MLA arrangement
E'_a	Illuminance of 1-D refocused photograph after horizontal process
E''_a	Illuminance of 2-D refocused photograph
E_{bin}	Illuminance after binary image conversion
E_{b_U}	Illuminance at MLA
E_{b_s}	Illuminance at sensor without presence of main lens
E_{f_s}	Illuminance at SPC's sensor plane
E_{hex}	Illuminance of intermediate sub-aperture image
E_i, $E_{(i,g)}$	Illuminance of 1-D array of sub-aperture images i and 2-D array (i, g)
E_{int}	Illuminance after interpolation to achieve sub-pixel precision
E_{norm}	Illuminance of normalised SPC image E_{f_s}
E_{rot}	Illuminance SPC image after MLA rotation
ERR	Error as percent deviation
e	Valid and fix number for horizontal micro lens index j
F	Focal plane
$F_{i,j}$	Ray intersection with focal plane
f_U	Main lens focal length
f_s	Micro lens focal length
$\widehat{f}_{i,j}(z)$	Linear function of z representing a chief ray's path in object space
$\mathring{f}_{i,j}(z)$	Linear function of z representing a chief ray's path in image space
$\mathring{f}_{i,j\pm}(z)$	Linear function of z representing DoF rays' path in image space
G_X	Signal with dimensionality X
G_Y	Signal with dimensionality Y
G	Gap between virtual cameras indicating the baseline

g	Index for vertical samples in directional domain v
H	Number of micro lenses in vertical direction
H_{1U}	Object side principal plane of main lens
H_{2U}	Image side principal plane of main lens
$\overline{H_{1U}H_{2U}}$	Spacing between principal planes H_{1U} and H_{2U}
H_{1s}	Object side principal plane of MLA
H_{2s}	Image side principal plane of MLA
$\overline{H_{1s}H_{2s}}$	Spacing between principal planes H_{1s} and H_{2s}
HE	High frequency energy in blur metric
h	Index for vertical samples in spatial domain t
h_0	Multiplier coefficient in *semi-systolic* FIR filter
h_M	Height of equilateral triangle in hexagonal MLA arrangement
I_{b_U}	Irradiance at MLA
I_{f_s}	Irradiance at SPC's sensor plane
I_U	Irradiance at main lens' aperture
i	Index for horizontal samples in directional domain u
J	Number of micro lenses in horizontal direction
j	Index for horizontal samples in spatial domain s
K	Horizontal resolution of entire SPC image
k	Index for global horizontal domain in SPC image
$\bar{k}_{j,h}$	Horizontal centre coordinate of micro lens (j, h)
L	Vertical resolution of entire SPC image
L_F	Light field
L_{b_U}	Light field captured by an SPC
$\mathcal{L}$	Number of illuminance levels in greyscale image
$(\Delta\,\ell)_{\min}$	Min. distance between image point peaks to fulfil *Rayleigh* criterion
l	Index for global vertical domain in SPC image
$\bar{l}_{j,h}$	Vertical centre coordinate of micro lens (j, h)
M	One-dimensional micro image resolution

M_a	Luminous emittance
$m_{c,j}$	Slope of MIC chief ray entering MLA
$m_{c+i,j}$	Slope of chief ray entering MLA
$m_{\{c+i,j\}\pm}$	Slopes of DoF rays entering MLA
$N_{i,j}$	Virtual sensor plane in SPC triangulation
N_s	f-number of micro lens?
$\mathbb{N}$	Natural numbers
n	Refractive index of lens material
O_L	Optical centre of left camera
O_R	Optical centre of right camera
o	Offset value for index of central micro lens
P	Plenoptic function
$\mathcal{P}$	Integral projection operator in *Fourier Slice Theorem*
p	Row index for switch state matrix $\mathbf{s}_{(a,w,p)}$
p_p	Pixel pitch
p_M	Micro lens pitch
p_N	Virtual pixel pitch
Q_H	Horizontal border separating low from high frequencies
Q_V	Vertical border separating low from high frequencies
$\mathbb{Q}$	Rational numbers
$q_{i,j}$	Chief ray slope in object space
$\mathbf{R}_z$	Rotation matrix around z
$\mathcal{R}$, $\mathcal{R}_{j,h}$	Region labelled with (j, h)
R_{s1}	Radius of curvature at front vertex of MLA
R_{s2}	Radius of curvature at back vertex of MLA
$\mathfrak{R}_\alpha$	Refocusing operator
$\mathbb{R}$	Real numbers
r	Light ray in Isaksen's ray model
r_A	Radius of an *Airy* pattern's central peak disc

S	Sharpness in blur metric
$\mathcal{S}$	Slicing operator in *Fourier Slice Theorem*
s	Horizontal spatial sampling domain at MLA
s'	Horizontal spatial sampling domain at object plane
$\mathbf{s}_{(a,w,p)}$	Switch states matrix
T	Time domain
TE	Total energy in blur metric
T_{PCLK}	Time period for pixel clock cycle
T_{PCLKx2}	Time period for half pixel clock cycle
$\mathbf{T}$	Translation matrix
t	Vertical spatial sampling domain at MLA
t_s	Physical thickness of MLA
$^{\mathsf{T}}$	Matrix transpose
U	Horizontal directional sampling domain at main lens
u	Horizontal directional sampling domain at micro image
$u_{c+i,j}$	Micro image position at pixel centre with indices
$u_{\{c+i,j\}-}$	Micro image position at pixel border towards MIC
$u_{\{c+i,j\}+}$	Micro image position at pixel border closer to the micro image edge
V	Vertical directional sampling domain at main lens
v	Vertical directional sampling domain at micro image
w	Column index for switch state matrix $\mathbf{s}_{(a,w,p)}$
X	Dimensionality of a signal
$\mathcal{X}$	*Fourier*-transformed signal of 2-D image
x	Horizontal domain of an image
$\mathbf{x}$	Matrix containing horizontal MIC coordinates
Y	Dimensionality of a signal
y	Vertical domain of an image
$\mathbf{y}$	Matrix containing vertical MIC coordinates

Z	Distance from optical centre of a lens to object point
Z_0	Object point in stereoscopic triangulation
$Z_{i,\Delta x}$	Distance from entrance pupil to object point in SPC triangulation
$\mathbb{Z}$	Integer numbers
z	Depth domain
z^{-1}	Unit-sample delay
z_U	Optical axis of main lens
z_j	Optical axis of micro lens with index j
z_i'	Optical axis of virtual camera with index i
α	Refocusing parameter in *Fourier Slice Photography Theorem*
β	Rotation angle of MLA
$\boldsymbol{\beta}$	Vector containing rotation slope and offset value
β_m	Rotation slope of MLA
β_n	Offset value for rotation slope of MLA
γ	Weighting coefficient in MIC interpolation
Γ	Coefficients in Cauchy's formula
δ	Vertical centre of micro lens columns
Δx	Disparity
ϵ	Index for summation in Cauchy formula
ε	Shift value for chequerboard image arrangement
η	Number of computation steps
θ	Angle of incident light ray entering main lens aperture
$\theta_{\max}$	Half-angle of the maximum cone of a light beam
ϑ	Temperature
ι	Degree of parallelism
κ	Imaginary unit in complex numbers
λ	Wavelength
Λ	Coefficients in Sellmeier 1 formula
ξ	Left position for cropping window in blur metric
Ξ	Right position for cropping window in blur metric
ϖ	Top position for cropping window in blur metric
π	Ratio of a circle's diameter to its circumference
Π	Bottom position for cropping window in blur metric
ρ	Vertical frequency domain

σ	Horizontal frequency domain
Σ	Summation symbol
τ	Threshold value
Φ	Tilt angle of camera
Φ_i	Tilt angle of virtual camera in SPC
Υ	Coefficients in Sellmeier 1 formula
ψ	Index in vertical frequency domain
Ψ	Vertical frequency border
ω	Index in horizontal frequency domain
Ω	Horizontal frequency border

(I.)	MLA specification with $f_s = 1.25$ mm
(II.)	MLA specification with $f_s = 2.75$ mm