



Imaging Through Disorder: Lens-less Cameras

Author

William Gasson

St Edmund Hall

August 10, 2021

Supervisors

Martin Booth

Julian Fells

Abstract

Imaging has long been defined by optics focusing light into image sensors. However, with modern computing it is possible to eliminate the necessity of having a lens, since the data from a sensor can be used to recreate an image of the scene, even when the lens has been removed. By looking at the how light is transformed on its way to the sensor, the system can be modelled. From this understanding, it is then possible to decide on an algorithm best suited to solve this reconstruction. This paper sets out two methods for modelling the transformation of light from source to sensor based on the specifics of the system. Firstly, when the system is modelled as linearly spatially invariant and, secondly, when spatial invariance cannot be assumed. The first system is solved by formulating an optimisation problem and using gradient descent to find the predicted image. For the second problem, the pseudo-inverse is used to work out how the image is related to the scene. These two methods show that the approach taken needs to be tailored to the specific situation to get the best results.

Contents

1 Acknowledgements	iii
2 Introduction	1
2.1 Overview	1
2.1.1 Aims of the Project	1
2.1.2 Structure of the Report	2
3 Background	3
3.1 Development of Imaging	3
3.2 Lens Based Imaging	4
3.3 Lens-less Imaging	4
3.3.1 X-ray Imaging	4
3.3.2 Coded Apertures & Masks	5
3.3.3 Fibre Optics	5
4 System with Spatially Invariant PSF	7
4.1 System Overview	7
4.1.1 Approximations	7
4.1.2 Hardware	8
4.2 Algorithm	10
4.2.1 Gradient Descent	10
4.2.2 Alternating Direction Method of Multipliers	12
4.2.3 Discussion	15
4.3 Characterisation	16
4.3.1 Imaging System	16
4.3.2 Field of View	18
4.3.3 Resolution	20
4.4 Conclusion	20
5 System with Spatially Variant PSF	21
5.1 System Overview	21
5.1.1 Hardware	22
5.1.2 Differences	22
5.2 Pseudo Inverse	23
5.2.1 Singular Value Decomposition	25
5.2.2 Compression	29
5.2.3 Moore-Penrose Pseudo-Inverse	29
5.2.4 Limiting Modes	30
5.2.5 Alternative Algorithm	31

5.3	Characterisation	32
5.3.1	Creating a Functioning System	32
5.3.2	Field of View	33
5.3.3	Resolution	35
5.3.4	Accuracy	36
5.4	Walsh-Hadamard Functions	36
5.4.1	Walsh Functions Definition	38
5.4.2	Walsh Functions in Optical Systems	38
5.5	Conclusion	40
6	Discussion	42
6.1	Comparison of Characteristics	42
6.1.1	Field of View	42
6.1.2	Resolution	42
6.1.3	Time	42
6.1.4	Versatility	43
6.2	Recommendations for Future Work	43
6.2.1	Imaging Real Objects	44
6.2.2	Applications	44
7	Conclusion	45
7.1	Completing Project Aims	45
7.1.1	Advantages of Lens-less Imaging	45
7.1.2	Problems to be Solved	46
Bibliography		46
A	Appendix	49
A.1	Code & Results	49
A.2	Health & Safety	49

Acknowledgements

I would like to thank my supervisors Martin Booth and Julian Fells for their support throughout this year. They helped me understand the subject area and suggested the best way of researching this project. They also had the foresight to design a project where the equipment was compact and portable enough to be transported, so I was still able to collect some physical data despite the turmoil of the past year. I would also like to thank my family for putting up with me through lockdown.

Introduction

2.1 Overview

This report discusses a project on the feasibility of lens-less imaging in different situations and from this work out its flaws, and potential uses. It explores how to build and test a system that can take images of a scene without the aid of optics focusing the light on to the sensor. A lens produces an image on a sensor with a clear relationship between the spatial distribution of intensity in the image and that in an object. However, this is not the only way a useful image can be retrieved. There is the same amount of information recorded with or without the lens. Theoretically, the only information required is the transformation of the light by the optical element between the scene and the camera. If this transformation is known, then images can be recreated despite an object scrambling the light. So, the goal is to find how the light is scattered to create the blur that is obtained without a lens. This would allow cameras to be more versatile and open up a wide range of potential applications from smaller cameras to imaging through fibre optics.

2.1.1 Aims of the Project

The initial goal of the project was to build and test a lens-less imaging system to gain an understanding of lens-less imaging and then evaluate of its advantages and disadvantages. Once this setup was functioning, the next step was to design a novel system that could reliably reconstruct the main points of a scene. Here, it was necessary to use readily available parts and relatively low computational expense so that the project could be easily replicated and show a wider range of applications. Then the capabilities of the design were to be tested to determine its properties. The aim was to present a design that could be improved upon by others to achieve better resolution and versatility in the future. Finally, the future applications of this technology were explored based on the properties that were derived and what it could

reasonably be expected to do in the future.

2.1.2 Structure of the Report

This project begins by exploring the history and literature surrounding imaging and lens-less imaging. The two main sections of the report follow, the first of which is based on work done in [1] to create a lens-less system. In this section, it was assumed that all points of light represent a linear shift of the point spread function (PSF). The second section considers how images can still be reconstructed if this no longer can be assumed to be the case. Finally, some conclusions were drawn from the research done on each method.

This report was written and researched by William Gasson under the supervision of Martin Booth and Julian Fells, who suggested possible areas which could be examined and gave assistance whenever needed. This paper builds on work done in the 4th year project completed by Saad Rana in "Diffuser Based imaging and wavefront sensing" [2] with a greater focus on the imaging system, along with the work done in the DiffuserCam paper [1], which was the starting point for this project.

Background

3.1 Development of Imaging

Imaging began with the pinhole camera created by Ibn Al-Haytham (965–1040AD). This camera was able to project an inverted image of the outside scene onto a board. However, the issue with a pinhole camera is that the larger the size of the hole, the more diffuse the image but if the hole is smaller, less light enters the system resulting in a dimmer image. When lenses were developed, they started to become an integral part of imaging. Optical instruments began to appear in 1270 with spectacles made in Florence, and high quality glass in Venice [3]. The progress that led to making spectacles was mostly due to the ability to make frames rather than understanding how the optics focused light. This has been a theme in imaging, as many advancements have come through the ability to position a lens more precisely as opposed to changing the form of the lens. In the 16th century, more optical devices began to appear, such as a telescope (developed by Galileo), microscopes, and images projected through painted glass plates. Along with these developments came equations describing how light travels and passes through a lens. Over the next few centuries, the quality of glass slowly improved and more of its properties were understood. This led up to 1865 when Maxwell developed equations describing light as electromagnetic waves. Meanwhile in 1837, the earliest form of photography was created by Daguerre, who exposed an image onto a polished silver plate. The diffraction properties of lenses were incorporated into photography to allow the light to be focused on to the sensor. This meant more light could reach the sensor so the time needed to take an image was reduced. Further improvements in how images are stored with the development of film and subsequently digital photography have occurred, but it is optics that have remained at the core of imaging in modern-day photography.

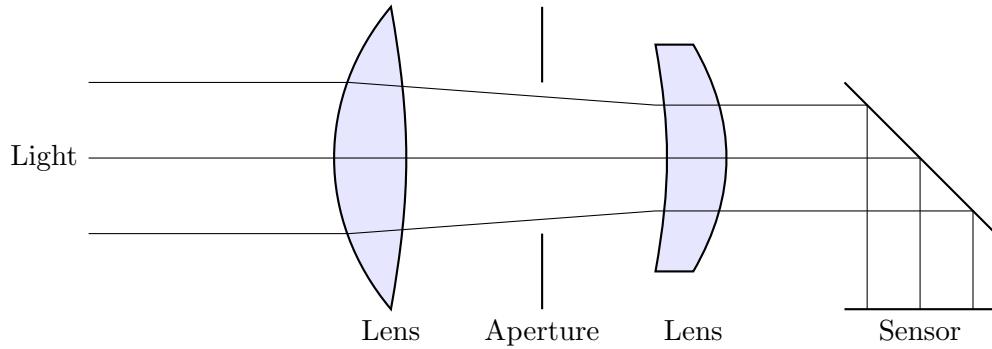


Figure 3.1: Diagram of a camera system showing how light travels to the sensor

3.2 Lens Based Imaging

Before discussing the modern lens-less imaging, it is important to understand how lenses and other optical elements are used in imaging and why they have been such an integral part of the field since they were developed. Cameras use groups of lenses to focus light through an aperture and shutter. This light is then incident on a mirror which reflects the light onto a view finder or sensor. When a photograph is taken, the shutter opens exposing the sensor to the light. The lenses are designed to make sure the light reaching the sensor is a direct representation of the scene being imaged.

3.3 Lens-less Imaging

3.3.1 X-ray Imaging

Lens-less imaging has moved on from pinhole cameras and silver plates. One of the main areas of its development has been in astronomy. Coded apertures and lens-less imaging have been necessary in astronomy for imaging certain wavelengths such as X-rays, which can pass straight through most materials. Dicke [4] describes a method using many randomly positioned pinholes to create a camera for X-rays. With this method it was possible to form images of stars whose light was in the X-ray range. This method cannot be directly replicated to image visible light as a visible light pinhole camera is affected to a much greater extent by diffraction leading to poor resolution. Also this method was designed to work for a very small number of photons. However principles of this technique can be transferred to our system.

3.3.2 Coded Apertures & Masks

Recently there has been more research into coded apertures outside astronomy. The paper [5] shows how a coded aperture can be used as a mask in front of the sensor to modulate the light in a binary fashion. This paper showed how it was possible to change the form of a camera with the mask very close to the sensor, producing a very flat camera which is able to reconstruct images and film. Another paper by Levin et al [6] shows how coded apertures can also be used for depth sensing. This paper used a camera with a lens, but showed how more information can be calculated using coded apertures. The coded aperture allows the distance from the camera to be calculated, which in turn allows for refocusing of the image. These ideas are continued in the paper [1], in which the coded aperture was replaced by a simple piece of tape acting as a diffuser. This paper first shows how the blurred images taken through the diffuser can be used to reconstruct the scene, and then moves on to take full 3D images. The quality of images taken with this set up was improved in the paper [7]. This paper highlights how lens-less imaging is an area where machine learning can be used to improve the reconstruction of the scene: a model can be trained to learn which inputs lead to certain outputs, which can then be extrapolated for unseen images. The machine learning does not replace the methods found in [1], but instead is used to supplement the output. This means there is a better reconstruction produced than would be the case if only machine learning or system based calculations were done individually. There has also been research into imaging with only an image sensor [8], which looks at how different points are combined to produce a scene. This approach is built on further in the paper [9], to show how the same method will hold true when the sensor is not pointing at the scene. In this case, the researchers are just imaging the edge of a panel of glass, which has a small amount of light from the scene reflected and refracted into the sensor. These papers show both the practical examples of lens-less systems and the feasibility of different methods. This report will hope to build on the work done.

3.3.3 Fibre Optics

There has also been some related research with fibre optics as light that has travelled through a fibre optic will be scattered. Much as it will be though our system. There is a range of techniques as there are a range fibre optics. Porat et al [10] and Orth et al [11] show methods to image through a fibre optic bundle. Paper [10] uses speckle correlation which is a similar idea to the looking at the shifts in the caustic pattern through a diffuser. While Caramazz et al [12] show it is also possible through a multimode

fibre. The reconstruction is done though calculating the inverse transformation matrix. Imaging through optical fibres opens a range of useful applications. As the papers show it is possible to image both at the microscopic level and of a wide field image.

System with Spatially Invariant PSF

4.1 System Overview

Before trying to image and find the capabilities of the system, it was important to understand the physics behind the operation. A normal imaging system can be modelled by the equation 4.1 which states the image taken on a sensor b is a function of the scene v . So, in order to work out how to reconstruct images from the sensor data, the relationship between the scene and the image needs to be discovered. To determine this relationship, one must look at the way the light reaches the sensor from the scene needs to be looked at. Several approximations on how this light travels can be assumed.

$$b = f(v) \tag{4.1}$$

4.1.1 Approximations

1. Shift invariance: This assumes that the same caustic pattern will appear on the sensor, except with a shift from a point source of light regardless of where the point source is.
2. Linearity: This assumes multiple point sources will produce a pattern which is the superposition of the individual caustic patterns from the individual point sources.

These assumptions mean that the system can be assumed to be linear shift invariant (LSI) so the system can be modelled by equation 4.2, where the image is a function of the scene transformed h . These approximations allow transformation to be determined from a single image of a point source: the point spread function (PSF).

$$f(v) = h * v \quad (4.2)$$

From Equations 4.1 and 4.2 it can be seen that to reconstruct the scene from the image this transformation function needs to be determined for the system. Then to go from the raw image to the scene the inverse function needs to be found. However, there are problems associated with this which will be discussed further in the algorithm section.

4.1.2 Hardware

The assumptions described above require a diffuser to scatter the light in a particular fashion. The diffuser is created by using a double sided piece of scotch tape. This has the properties required as it is nearly transparent so it allows most of the light to freely travel to the sensor. But due to some imperfections the light will be scattered as it travels through it. When it is very close to the sensor this makes the PSF be a caustic pattern. The rest of the equipment needed to make this project is listed below and then the reasoning behind it is explained.

- Raspberry Pi 4 Model B with 4GB RAM
- Raspberry Pi v2 8MP 1080p Camera Module
- ARDUINO MEGA 2560
- Sparkfun RGB LED Matrix Panel - 32x32
- ThorLabs breadboard and mounts
- ThorLabs variable aperture
- Double sided scotch tape
- Wires and plugs to power and connect the system together

This project built on the work done in the paper [1], which used a Raspberry Pi and its associated camera. This was a suitable device to use, as it was relatively simple to remove the lens from the Pi-camera to expose the sensor. The Pi-camera is also low cost, which is important as once the lens is removed there will be fragile components exposed. If these components got damaged either when the lens is removed or the camera is jolted the camera would break. The DiffuserCam paper used a light and piece of card to

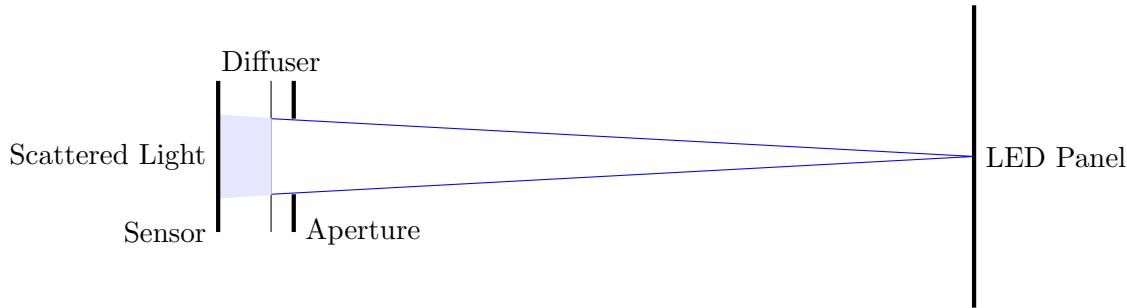


Figure 4.1: Diagram showing the diffuser system. Light travels from the LED panel and is scattered by the Diffuser before it reaches the sensor.

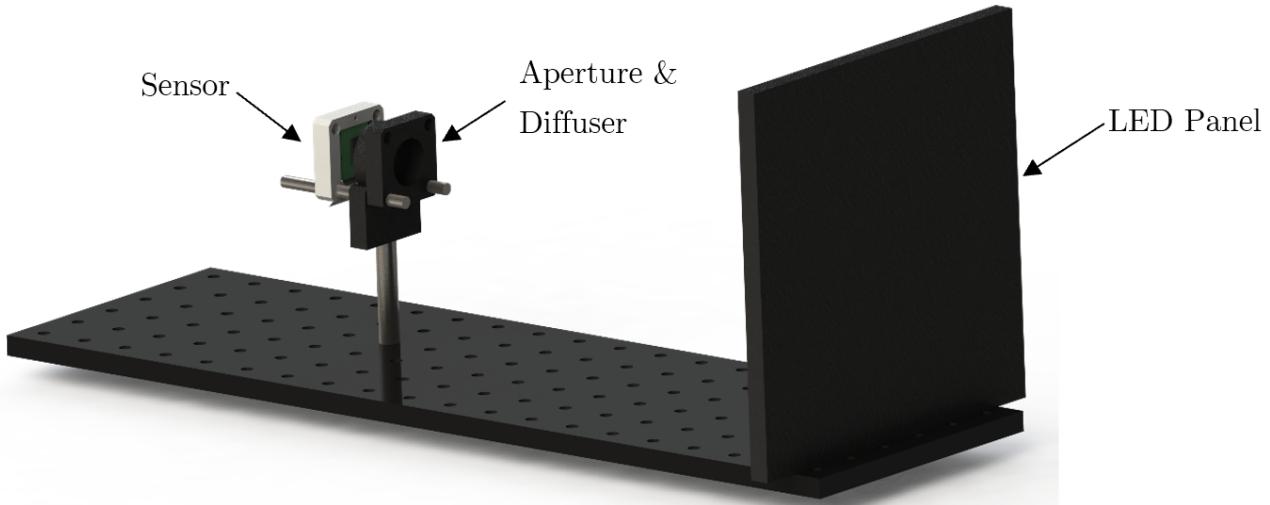


Figure 4.2: A render of the physical system. On the left is the LED panel which is facing the camera, which is held in a 3D printed mount behind a variable aperture. All the parts are attached to the ThorLabs breadboard to keep their positions constant.

create a PSF, in this project it was replaced by a 32 x 32 RGB LED Matrix Panel. This allowed the scene the camera was pointing at to be controlled. The panel was controlled by the Arduino. The Arduino Mega was needed as the LED panel requires higher processing speed and more memory than an Arduino Uno can reliably produce. All the components were held in place using a ThorLabs stainless steel breadboard, which allows the relative positions of the camera, diffuser, and LED panel to remain constant to each other when data was being collected. This equipment went along with the goal of creating a system which is easily obtainable and affordable. The Raspberry Pi and Arduino are very common devices and can be repurposed for other projects.

4.2 Algorithm

Section 4.1 describes the problem to be solved and shows that in order to reconstruct the scene from the sensor data, the transformation function needs to be inverted. This transformation h can be characterised by a matrix \mathbf{M} and a crop \mathbf{C} representing the dimensions of the sensor. However, due to the nature of the matrix this is difficult to invert. So instead it is possible to formulate the reconstruction as an optimisation problem. To do this b need to be represented by a vector where all the columns of the raw image data are stacked together. The formulation of this problem is shown in Equation 4.3, where \mathbf{CM} will therefore represents the matrices which transforms the scene to the sensor. Convex optimisation is a large topic and there are many methods available to solve these types of problem. This section will describe a couple of the available options to calculate the reconstruction of the image.

$$\mathbf{v}^* = \operatorname{argmin} \frac{1}{2} \|\mathbf{CMv} - \mathbf{b}\|_2^2 \quad (4.3)$$

To solve this problem, it needs to be converted into a version which contains elements that are known. \mathbf{CM} need to be formulated based on the PSF and image that has been taken. The transformation matrix \mathbf{M} is a convolved with the scene to produce the image. A convolution it can be represented by a multiplication in the Fourier domain. This means \mathbf{Mv} can be written as the multiplication in the Fourier domain of the PSF with the current prediction scene as shown in Equation 4.4.

$$h * v = \mathbf{CMv} = \mathbf{CF}^{-1} \operatorname{diag}(\mathbf{Fh}) \mathbf{Fv} \quad (4.4)$$

4.2.1 Gradient Descent

One way to solve the problem described in Equation 4.3 is to use an algorithm called gradient descent. This is an algorithm designed to use the gradient of a function to find the minimum value. It works by calculating the gradient at your starting point, then travelling a small step in that direction [13]. This process is then repeated at the point the you have moved to. So, each iteration of the algorithm moves downhill towards the minima. This means the gradient descent algorithm can be used to find the minimum value of a function, hence the approximation to v , the scene. The algorithm can be described as follows:

1. $\Delta x = -\nabla f(x)$
2. Choose step size α
3. Update $x_{k+1} = x_k + \alpha \Delta x_k$

Gradient descent is a very simple algorithm to implement and works reasonably reliably. A disadvantage of gradient descent is that, depending on the initial point on the function and the shape of the function, the algorithm may end up trapped in a local minimum instead of the global minimum. This is not a problem if the function is convex. As the problem that needs to be solved is a squared norm, it is a convex problem. This means there will not be any local minima where the algorithm could get stuck.

Calculating the Gradient

Gradient descent can be applied to our problem. To do this a number of functions need to be defined. The problem shown above in 4.3 will be the function that will be minimised. The first step of the algorithm is to determine the direction of the gradient Δx which is shown in Equation 4.5. This is the direction that needs to be travelled to find the minimum.

$$\Delta f(v) = (\mathbf{C}\mathbf{M})^T(\mathbf{C}\mathbf{M}\mathbf{v} - \mathbf{b}) \quad (4.5)$$

As shown in 4.5 the gradient includes the term $(\mathbf{C}\mathbf{M})^T$ so this needs to be calculated. The inverse of a crop will be a padding function. As \mathbf{M} is known \mathbf{M}^T can be found by inverting it. Equation 4.6 shows the value of $(\mathbf{C}\mathbf{M})^T$ in terms that can be known.

$$(\mathbf{C}\mathbf{M})^T = \mathbf{F}^{-1}\text{diag}(\mathbf{F}\mathbf{h}) * \mathbf{F}\mathbf{C}^T \quad (4.6)$$

Learning Rate

The second step of this algorithm is to find the learning rate α . There are several things to be considered when deciding the step size. The higher it is, the faster it will converge to the minimum. However, if it is too large it will result in the solution not being accurate as the estimate will jump from one side to the other of the optimal value and never reach it. This means that ideally the learning rate should vary, depending on how close to the solution it is. However, having a varying step size requires fine tuning so it decreases at the correct rate and doesn't slow down too fast. To pick a constant learning rate that

is large enough that it converges quickly (but not so large that it will jump), there are bounds given in Equation 4.7 [14]. This gives a bound so a value just below the upper bound can be taken ensuring it converges as fast as possible.

$$0 < \alpha < \frac{2}{\|(\mathbf{CM})^T \mathbf{CM}\|_2} \quad (4.7)$$

Implementation

Now that the gradient and learning rate have been found this problem can be formulated as a loop which can be programmed. The code needs to iterate through and at each stage calculate $\alpha * \Delta f(v)$ and then subtract it from the current point v_k . A starting point v_0 needs to be set, it can arbitrarily be set to a point where each pixel is predicted to be half on. As the system is representing pixels, negative values should be removed. This gives the formation in Equation 4.2.1 below to find the next step.

$$\mathbf{v}_0 = I/2 \quad (4.8)$$

for k=0 to iterations

$$\mathbf{v}_{k+1} = \mathbf{v}_k - \frac{1.8}{\|(\mathbf{CM})^T \mathbf{CM}\|} (\mathbf{CM})^T (\mathbf{CM} \mathbf{v}_k - \mathbf{b})$$

if $\mathbf{v} \leq 0$; $\mathbf{v} = 0$

4.2.2 Alternating Direction Method of Multipliers

An alternative to gradient descent proposed in [15] would be the alternating direction method of multipliers (ADMM). ADMM is another algorithm that can be used to solve convex optimisation problems. It works by reducing a large problem into smaller ones which can be more easily solved. To solve an optimisation problem with this method the dual problem is found. The dual problem can be explained through Lagrange multipliers.

Lagrange Multipliers

A method to solve convex optimisation problems is Lagrange multipliers. This algorithm works by finding the point where your objective function is tangential to the constraints of your problem. This will

be at the optimum point, as the objective function would be as high as possible without violating a condition. As they are tangential, the gradient of the two functions will be proportional, and the constant of proportionality is called the Lagrangian. These facts allow you to form the Lagrangian function as shown in Equation 4.10. Which can help solve problems such as the one shown in Equation 4.2.2.

$$\min_{x \in X} f(x) \quad (4.9)$$

$$S.t. \quad Gx \preceq a$$

$$Hx = b$$

$$L(x, \lambda, \nu) = f(x) + \lambda^T(Gx - a) + \nu^T(Hx - b) \quad (4.10)$$

From this the Lagrangian dual function can be created by taking the infimum as shown in Equation 4.11.

$$g(\lambda, \nu) = \inf(L(x, \lambda, \nu)) \quad (4.11)$$

This can then be formulated as a dual problem shown in Equation 4.2.2. Solving this problem if weak duality holds will provide a bound on the original problem as weak duality means the solution to the dual problem will be greater or equal to the original problems solution.

$$\max_{\lambda} \quad -\lambda^T h - \nu^T b \quad (4.12)$$

$$S.t. \quad f(x) + G^T \lambda + H^T \nu = 0$$

$$\lambda \succeq 0$$

Augmented Lagrangian

In order to calculate the ADMM, the augmented Lagrangian is needed. This is a modification to the method of Lagrange multipliers as it adds a quadratic penalty function. The augmented Lagrangian is

described in Equation 4.13

$$L(x, \lambda, \nu) = f(x) + \lambda^T(Gx - a) + \frac{c}{2}\|Gx - a\|^2 + \nu^T(Hx - b) + \frac{d}{2}\|Hx - b\|^2 \quad (4.13)$$

Implementation

As in gradient descent the problem to be solved is Equation 4.3. This can be simplified by using dummy variables and which then add constraints 4.15. A regularisation term can also be added to penalise if the gradient isn't sparse. A sparse matrix is preferred as it makes calculations quicker by setting as many elements in the matrix to zero.

$$\mathbf{v}^* = \arg \min_{w \geq 0, u, x} \frac{1}{2}\|\mathbf{C}x - \mathbf{b}\|_2^2 + \tau\|u\|_1 \quad (4.14)$$

$$S.t. \quad x = \mathbf{M}\mathbf{v}, \quad u = \Psi, \quad w = \mathbf{v} \quad (4.15)$$

Then the augmented Lagrangian can be calculated. This is done as described above including introducing the penalty term. Also, an I term is added to make use that the values are positive by penalising negative values.

$$\begin{aligned} L(u, x, w, \mathbf{v}, \zeta, \mu, \rho) = & \frac{1}{2}\|\mathbf{C}x - \mathbf{b}\|_2^2 + \tau\|u\|_1 + \frac{\lambda_1}{2}\|\mathbf{M}\mathbf{v} - x\|_2^2 + \zeta^T(\mathbf{M}\mathbf{v} - x) \\ & + \frac{\lambda_2}{2}\|\Psi\mathbf{v} - u\|_2^2 + \mu^T(\Psi\mathbf{v} - u) \\ & + \frac{\lambda_3}{2}\|\mathbf{v} - w\|_2^2 + \rho^T(\mathbf{v} - w) \\ & + I(w) \end{aligned} \quad (4.16)$$

This problem is solved by taking the maximum of the infimum of the Lagrangian. Intermediate update steps are carried out as shown in Equation 4.18. Each step of the dual problem solves the maximisation through gradient ascent.

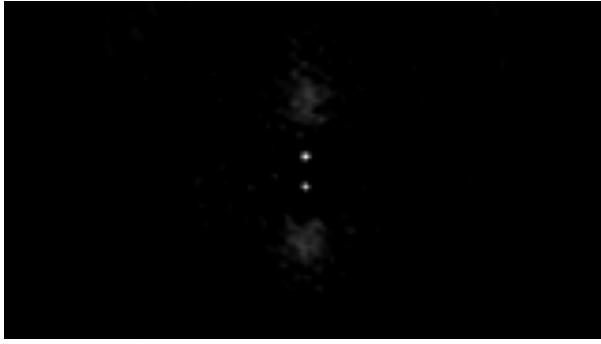


Figure 4.3: The reconstruction of two dots using gradient descent. The two dots are clearly visible in both images showing either methods work.

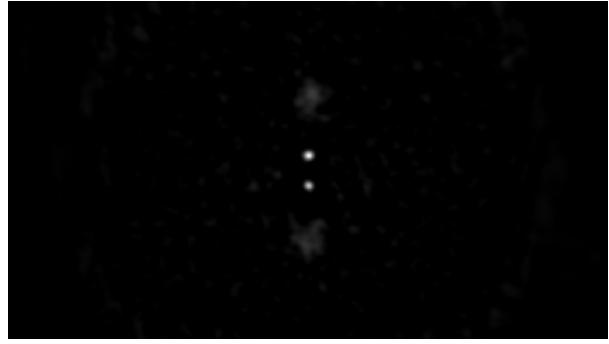


Figure 4.4: The reconstruction of two dots using ADMM. In both images there are two more diffuse blurs above and below the dots.

$$\text{Primal Updates} = \begin{cases} u_{k+1} & \arg \min_u \tau \|u\|_1 + \frac{\lambda_2}{2} \|\Psi v_k - u\|_2^2 + \mu_k^T (\Psi v_k - u) \\ u_{k+1} & \arg \min_x \frac{1}{2} \|\mathbf{C}x - \mathbf{b}\|_2^2 + \frac{\lambda_1}{2} \|Mv_k - x\|_2^2 + \zeta_k^T (\Psi v_k - x) \\ w_{k+1} & \arg \min_w \frac{\lambda_3}{2} \|v_k - w\|_2^2 + \rho^T (v_k - w) + I(w) \\ v_{k+1} & \arg \min_v \frac{\lambda_1}{2} \|Mv - x\|_2^2 + \zeta^T (\Psi v_k - x) + \frac{\lambda_3}{2} \|v - w_k + 1\|_2^2 + \rho_k^T (v_{k+1} - w_{k+1}) \end{cases} \quad (4.17)$$

$$\text{Dual Updates} = \begin{cases} \zeta_{k+1} & \zeta_k + \lambda_1 (Mv_{k+1} - x_{k+1}) \\ \mu_{k+1} & \mu_k + \lambda_2 (\Psi v_{k+1} - u_{k+1}) \\ \rho_{k+1} & \psi_k + \lambda_3 (v_{k+1} - w_{k+1}) \end{cases} \quad (4.18)$$

4.2.3 Discussion

In this project, gradient descent was used instead of ADMM, since despite the fact that it is a much faster algorithm, there are more parameters that need to be controlled to find the optimal solution. As there were no time constraints, the slowness of gradient descent was not an issue. It allowed the method to be shown to be functional and the characteristics of the system to be found. Furthermore, the results of using either algorithm produced very similar results as shown in Figures 4.3 & 4.4, which are the reconstructions using the same raw data but calculated using gradient descent and ADMM respectively. As the results are similar and gradient descent requires less fine tuning, gradient descent was chosen.

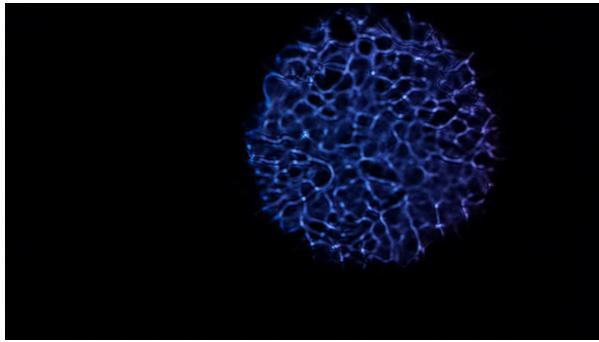


Figure 4.5: This image shows the caustic PSF of the system. It is slightly offset from the centre as the camera was not perfectly aligned with the aperture.

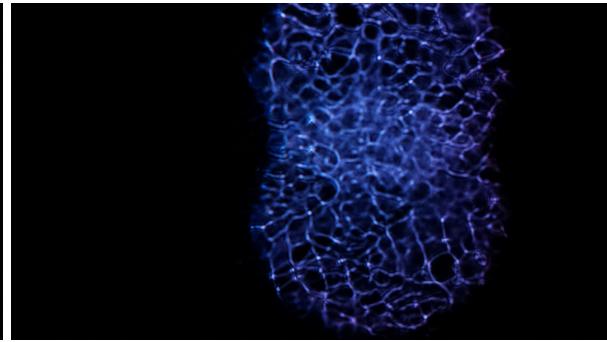


Figure 4.6: The image of the scene through the diffuser. It shows two of the caustic patterns in Figure 4.5 shifted up and down and then overlaid on top of each other.

4.3 Characterisation

4.3.1 Imaging System

The first step to characterise this system is to make a functioning system. To get it working a correctly calibrated PSF is needed. This is done by setting up a piece of tape to act as a diffuser to scatter light in a consistent fashion onto the sensor; then shining a light from a point source onto the diffuser. By adjusting the distance between the sensor and diffuser, a caustic pattern can be created. The diffuser needs to be the correct distance, as if it is too far it will just create a diffuse blur. The appropriate distance that was found to produce a caustic that was not blurred was 1-2mm. The size of the aperture also needs to be adjusted so the PSF fits on the sensor, so when it is translated, no new parts of the caustic pattern appearing in the data. Figure 4.5 shows an example of how this pattern looks. Then an image of the desired scene needs to be taken. In this example, the reconstruction is just of two dots one above the other. This is a simple example, but clearly depicts how the method works. These two images are put through the algorithm described in section 4.2.1 to produce the output in Figure 4.3. The output image clearly shows two distinct dots in the centre: one above the other, which is what the camera was imaging. There are also two more diffuse patches above and below the dots, which could be because of aliasing and represent a harmonic of the original signal.

The two main areas of the camera's performance that will be measured are its resolution and field of view. These are important metrics to examine the performance of the system, as they show the size of the image that can be taken and how detailed that image will be.

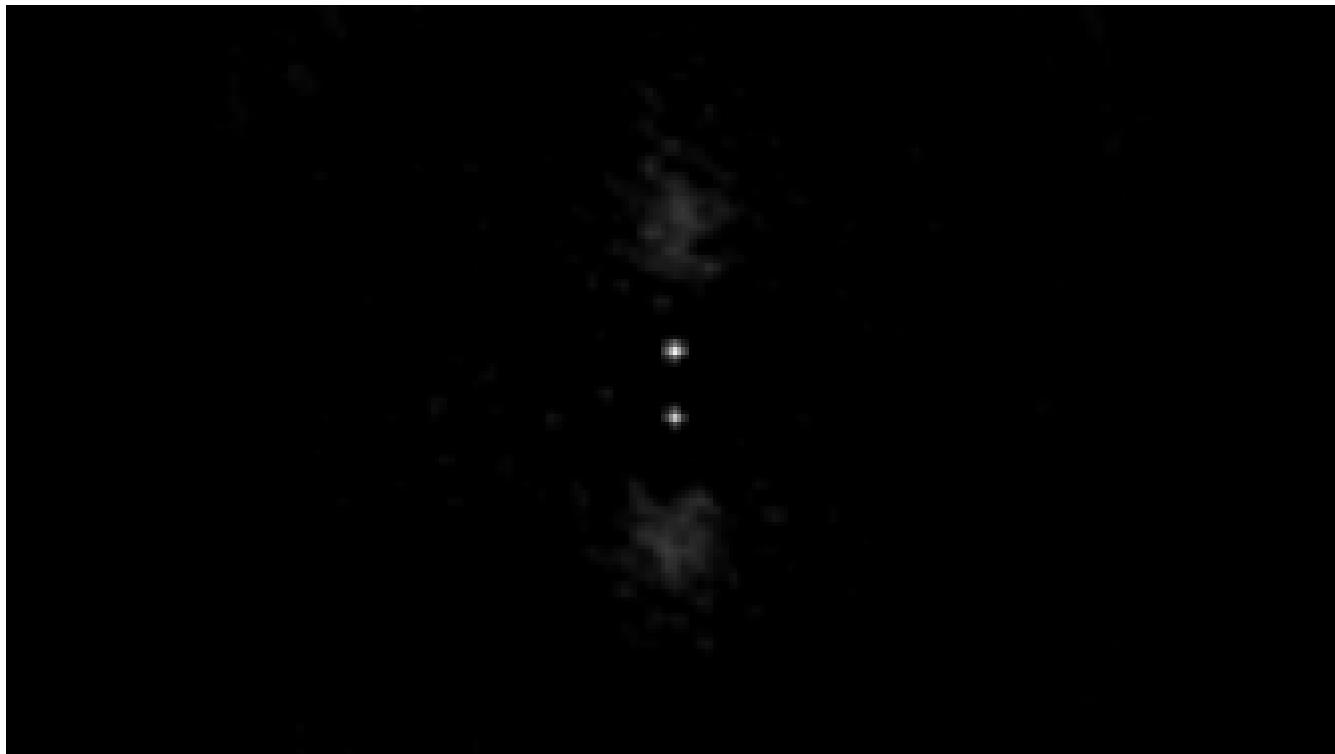


Figure 4.7: The image produced from reconstructing Figure 4.6 from using the PSF in Figure 4.5. Gradient descent was used to accurately show two dots one above the other. There is also some noise above and below the part which has accurately been reconstructed. This is impressive as the algorithm begins with each point at half intensity and was able accurately find the two points where the lights were shining.

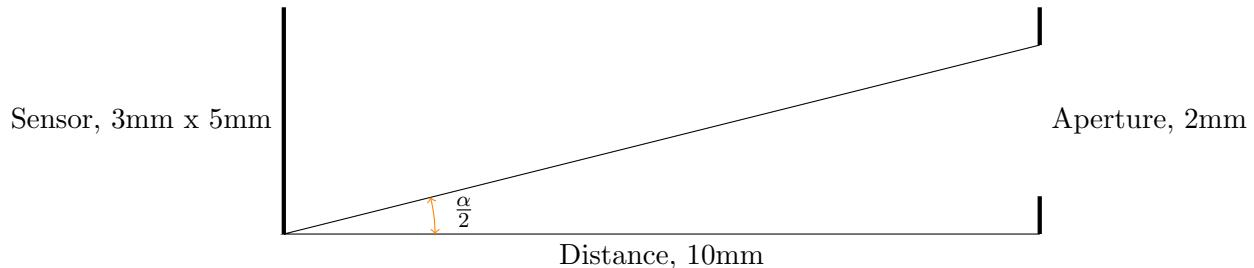


Figure 4.8: A diagram showing how the angle the geometry of the system predicts

4.3.2 Field of View

The field of view is the observable area that the camera can accurately image. This angle can be determined in this case by finding the area on the LED panel from which the camera can reconstruct an LED.

Prediction

The field of view can be predicted from the geometry of the system. This can then be compared to the experimental value to see if it is the geometry or a different limiting factor that determines the size of the field of view. The dimensions (as shown in the diagram below) of the aperture (2mm), and sensor (3mm x 5mm), and the distance between them (10mm) are used to calculate the angle. This gives an angle of $\frac{\alpha}{2} = 14^\circ$ for the vertical axis, so the predicted field of view is 28° and $\frac{\alpha}{2} = 19^\circ$ for the horizontal plane, to give a field of view of 38° .

Method

The field of view can be calculated experimentally by taking an image of a cross to determine the number of LEDs in the x and y planes that can be separated from each other. If the LED cannot be separated it can be assumed to be out of the field of view. This will give a width and height that are inside the field of view. From this and the distance from sensor to scene, the angular field of view can be calculated with simple trigonometry.

Results

From the figure 4.9 it can be seen that 15 LEDs vertically and 22 horizontally can be separated. This can be converted into a distance, as each LED has a centre 6mm from its neighbour. So the width is 9cm and height 13.2cm. The distance from the scene to the diffuser is 27cm, so by applying equation 4.19 a

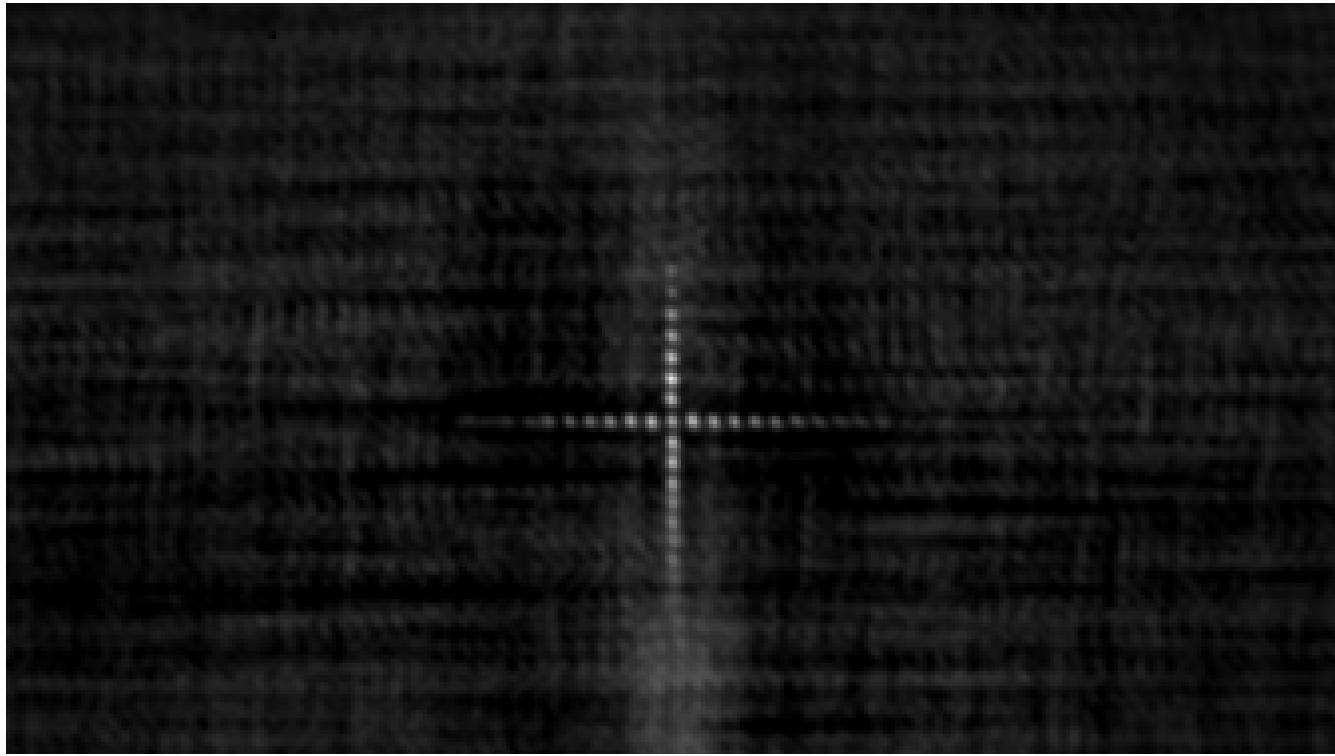
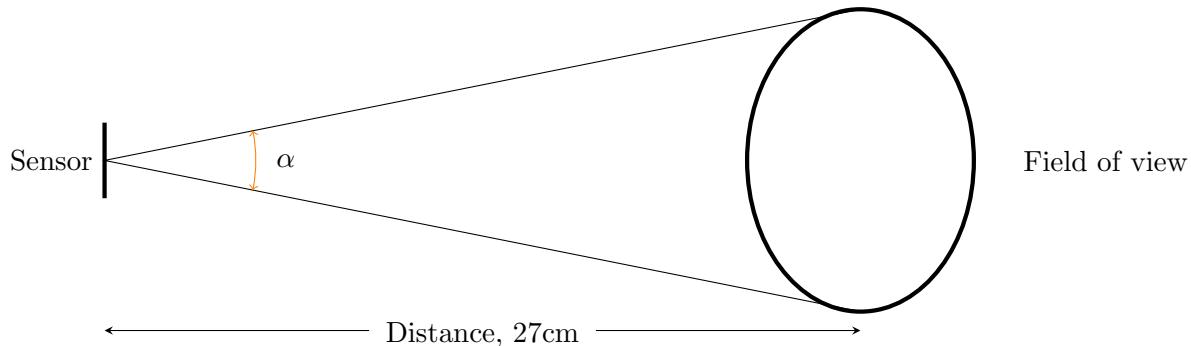


Figure 4.9: The reconstruction of two lines of LEDs in the shape of a cross. Not all the LEDs in the cross were recreated which allows the field of view to be calculated.

field of view of 19° degrees vertically and 27.5° horizontally can be calculated. This is lower than the predicted values of 28° and 38° which suggests that there is either some edge effects happening or the limiting factor is not the geometry of the system, but the linear shift invariance approximation breaking down. The fact that the dots become less bright and more diffuse as they move from the centre suggests this is the case.

$$\frac{\alpha}{2} = \arctan \left(\frac{\text{Radius of FOV}}{\text{distance}} \right) \quad (4.19)$$



4.3.3 Resolution

Method

The resolution of a camera is its ability to separate two points from each other. The closer together two points can be and still be separate, the more detail can be shown in an image. To find the resolution of the system after it has been processed, the minimum gap between two points that can be separated needs to be determined. As this project is focused on imaging an LED panel, an obvious test is to see whether the camera can differentiate between two neighbouring LEDs.

Results

The same cross in Figure 4.9 can be used to determine the resolution as the neighbouring LEDs are discernable, so the resolution is at least 6mm. However, there is also a distinct reduction in how clearly the LEDs are shown the further away from the centre the light is. This could be due to a number of factors. Firstly, the field of view is limited by the aperture (as described in the previous Section 4.3.2) so the LEDs at the edge of the field of view will not be able to get light to properly register on the sensor and therefore will not be reconstructed as clearly. Secondly, the further from the centre the weaker the spatial invariance approximations will be. When the light is coming from a point further away from the centre, the pattern created will no longer be an exact copy of the PSF shifted. This means the points on the cross will not be reconstructed as well so become smaller, more diffuse and blur into each other.

4.4 Conclusion

This section has shown how it is possible to image with a lens-less camera. The initial aim of this project, to build the system described in [1] has been successfully undertaken and testing has been completed. It is not equivalent to the performance of current cameras with lenses, as the field of view and resolution are limited. However, it demonstrates that the technique is practically viable and is a stepping stone for extending the ideas into more applications. In the next section, these ideas are expanded upon when looking at a system with when the light is transformed in a different way as it reaches the sensor.

System with Spatially Variant PSF

5.1 System Overview

In this section, the method is extended to deal with spatially variant PSFs. As shown in the previous Section 5.3, it is possible to reconstruct a scene using a camera where the lens is replaced by a diffuser. This design assumes that the image is constructed from multiple point sources, all of which produce a caustic pattern equivalent to the PSF of the system after undergoing a linear translation. This section of the paper will now examine how to reconstruct a scene, if that is not the case.

Equation 5.1 applies to the system here, as it did in the previous system, as the image b taken is still related to the scene v by some transformation h . However, to simplify the problem the new system will use the fact that the scene is an LED panel. As the LED panel is a 32 x 32 array, there are only 1024 points to predict. If each light is also limited to ON or OFF, this dramatically simplifies the problem. This is because it creates a system that can be modelled to determine the transformation from each light source to the sensor, as opposed to the previous system for which it was assumed that each light source was transformed in the same way. This, however, will require a much larger transformation matrix as it has to represent 1024 point sources as opposed to just one. This section will look at an alternative way to reconstruct the scene from this large matrix. Instead of the gradient descent or ADMM described in the previous Section 4.2, a pseudo-inverse will be used to calculate the scene.

$$b = h * v \tag{5.1}$$

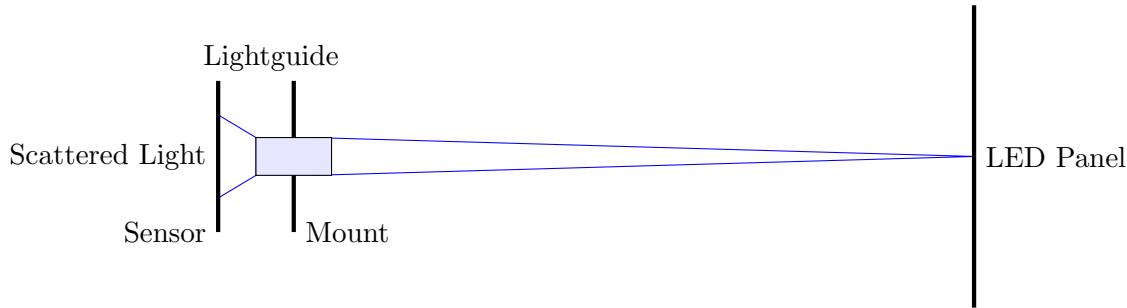


Figure 5.1: Diagram of the light guide system. Here instead of the diffuser scattering light a plastic tube scatters the light in a different fashion.

5.1.1 Hardware

In order to create a system which has the properties described, the light must be transformed from the scene to the sensor in a different way. So, instead of a thin piece of tape being used to diffuse the light close to the sensor, a light-guide (PLP2-10MM Bivar, Panel Mount LED Light Pipe, Clear Round Lens) was used. This is a thin plastic tube 1cm long and 2mm in diameter. This acts differently to the diffuser as the light enters it and is scattered and reflected from the edges for the entire length of the light-guide. As a result, the light that enters from a point source in one place is not transformed in the same way as if that point source was translated. This means it has the desired properties and is spatially variant but still linear, so the image will consist of the superposition of two separate point sources' images. These hardware changes move the project closer to interesting applications such as an optical fibre. This is because the properties of an optical fibre are much closer to a light-guide than a diffuser. This would be a very interesting avenue of research in the future.

5.1.2 Differences

In the system that is modelled in this section, the shift invariance approximation does not hold true. The linearity however is still assumed to hold. This means that it is not assumed that each point source of light will produce the same pattern just translated. Instead, each point of light will produce a different pattern depending on its location.

To create this system the diffuser is replaced by a light guide, a thin transparent plastic tube. The PSF produced with this new set up is shown in Figure 5.3. This image does not have the fine caustic pattern seen in the images taken through the diffuser. It is much more of a diffuse pattern. The original gradient descent method was tested on images taken with this the light guide instead of the diffuser. It

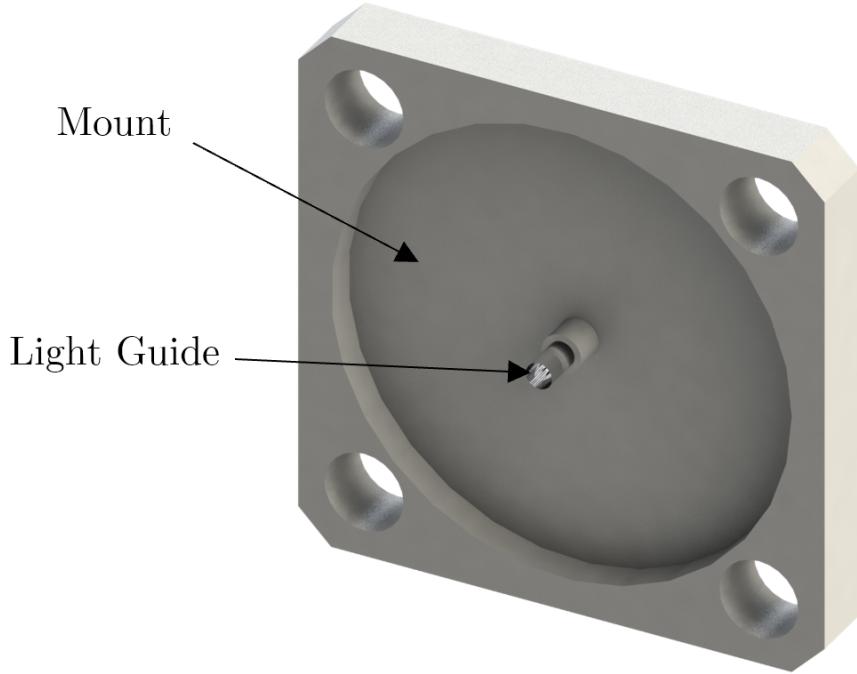


Figure 5.2: The light guide and mount designed to use the same rails as the camera, so it remains in line with the sensor.

was unable to reconstruct the scene of two dots and produced a blur as shown in Figure 5.5. This shows the system described in Section 4 was not able to recreate images as it was in the previous setup. The two dots in the scene were very close to the point source the PSF was taken from. As the reconstruction was not successful, it suggests that there isn't spatial invariance even over small distances. So, the reconstruction would deteriorate even more as points further from the point source are imaged. To improve the reconstruction a more complete calibration of the system could be taken. The further from the point the PSF is taken, the more inaccurate the reconstruction will be. This increase distance reduces the extent of correlation between the pattern and the PSF. A way to get a more complete understanding of the transformation from the scene to the sensor would be to take images of each of the lights on the panel. This would then give the relationship between each different point of the scene, and how it is transformed by the optical system onto the camera.

5.2 Pseudo Inverse

As explained in Section 5.1, the inverse needs to be calculated in order to predict which LEDs were turned on. As the transformation matrix is not square, a Moore-Penrose pseudo-inverse is calculated instead. This is done using the singular value decomposition (SVD) to break the matrix up.



Figure 5.3: The PSF of the new system with the light guide. There is no caustic pattern



Figure 5.4: This image shows a second PSF from a different point source. It is not just a shifted version of the other PSF

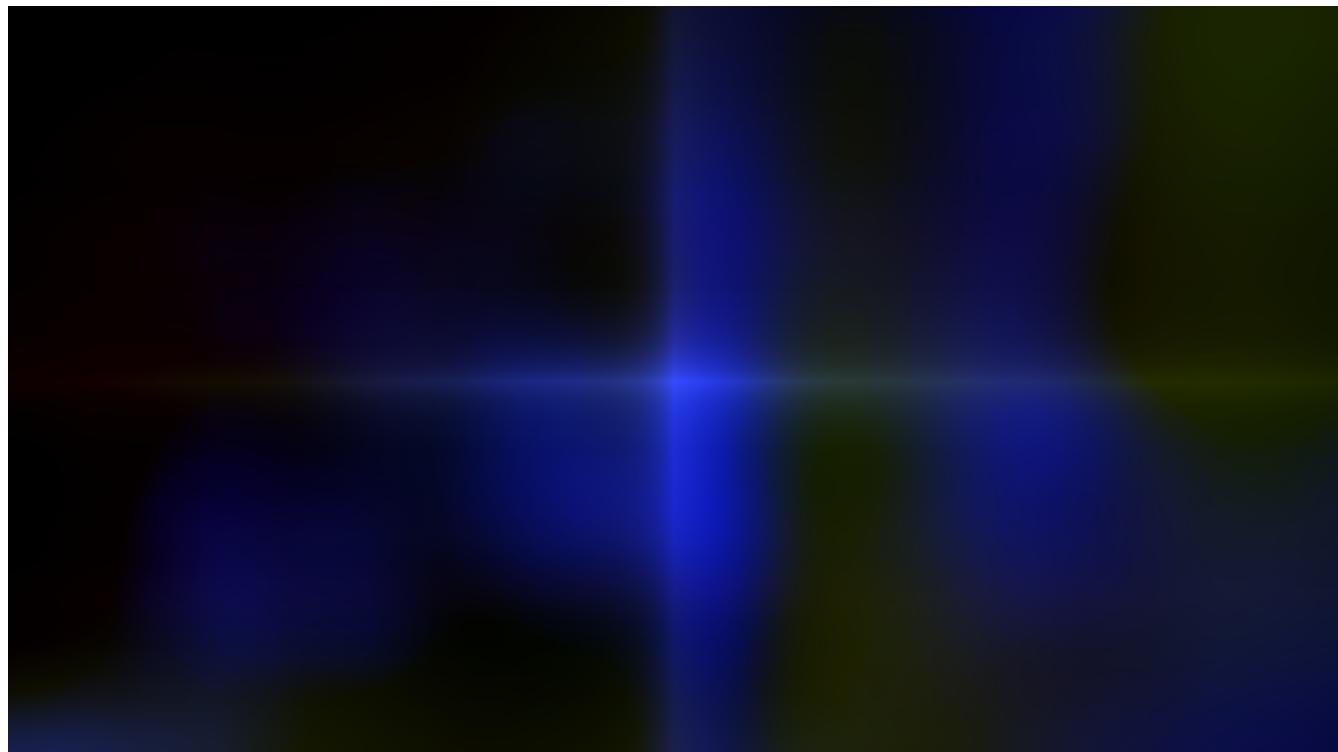


Figure 5.5: The reconstruction gradient descent created from two dots 2cm apart in a horizontal line after the diffuser had been replaced by a light guide. As the light guide scrambles the light, the algorithm is no longer able to accurately recreate the scene.

5.2.1 Singular Value Decomposition

SVD works by separating a rectangular matrix into three components: $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$ [16]. This represents translating the original data into a coordinate system with a diagonal covariance. Each of these new matrices has interesting properties, $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ and $\mathbf{V}_T\mathbf{V} = \mathbf{I}$ so they are both orthogonal. In this case, \mathbf{u}_i are the left singular vectors, σ_i are the singular values and \mathbf{v}_i are the right singular vectors. The SVD is calculated by finding the eigenvalues of $\mathbf{A}^T\mathbf{A}$ and these give the singular values. The corresponding eigenvectors then make up \mathbf{V} and the eigenvectors of $\mathbf{A}\mathbf{A}^T$ then are \mathbf{U} . These are displayed in Equation 5.2.

$$\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{pmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} \begin{pmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{pmatrix} \quad (5.2)$$

Calculating the System Modes

To learn more about the system and how the data is collected, the SVD can be used to determine the system modes. This identifies the LEDs which have the largest contribution to the image brightness on the system, and how that effect is seen on the sensor to be determined. This was done by imaging the complete array of LEDs individually, and combining this data into a matrix which represents the transformation from each LED to each pixel. This matrix was then put through SVD. Figure 5.6 shows the first 16 modes (\mathbf{U} in the SVD) of the system with the corresponding patterns (\mathbf{V}) in Figure 5.7 that would cause them to appear. The modes for the LED include negative values; as an LED cannot absorb light, these could not be produced individually in real life. Instead the first mode acts as a positive offset, so when combined with other modes the total is positive.

Randomised SVD

The method used in this project to calculate the SVD of the transformation matrix was randomised SVD. This allows the SVD to be calculated faster, particularly if not all the modes are needed. This method was introduced in paper [17], as there are now more large datasets which have high numbers of dimensions but a low intrinsic rank r . The method is described in [18] Section 1.8. The method works by randomly sampling the column space of matrix \mathbf{X} . This is done because there is a high probability that it will find the subspace that will span the dominate columns of the U matrix of the SVD. The SVD for $\mathbf{X} \in \mathbb{R}^{n \times m}$

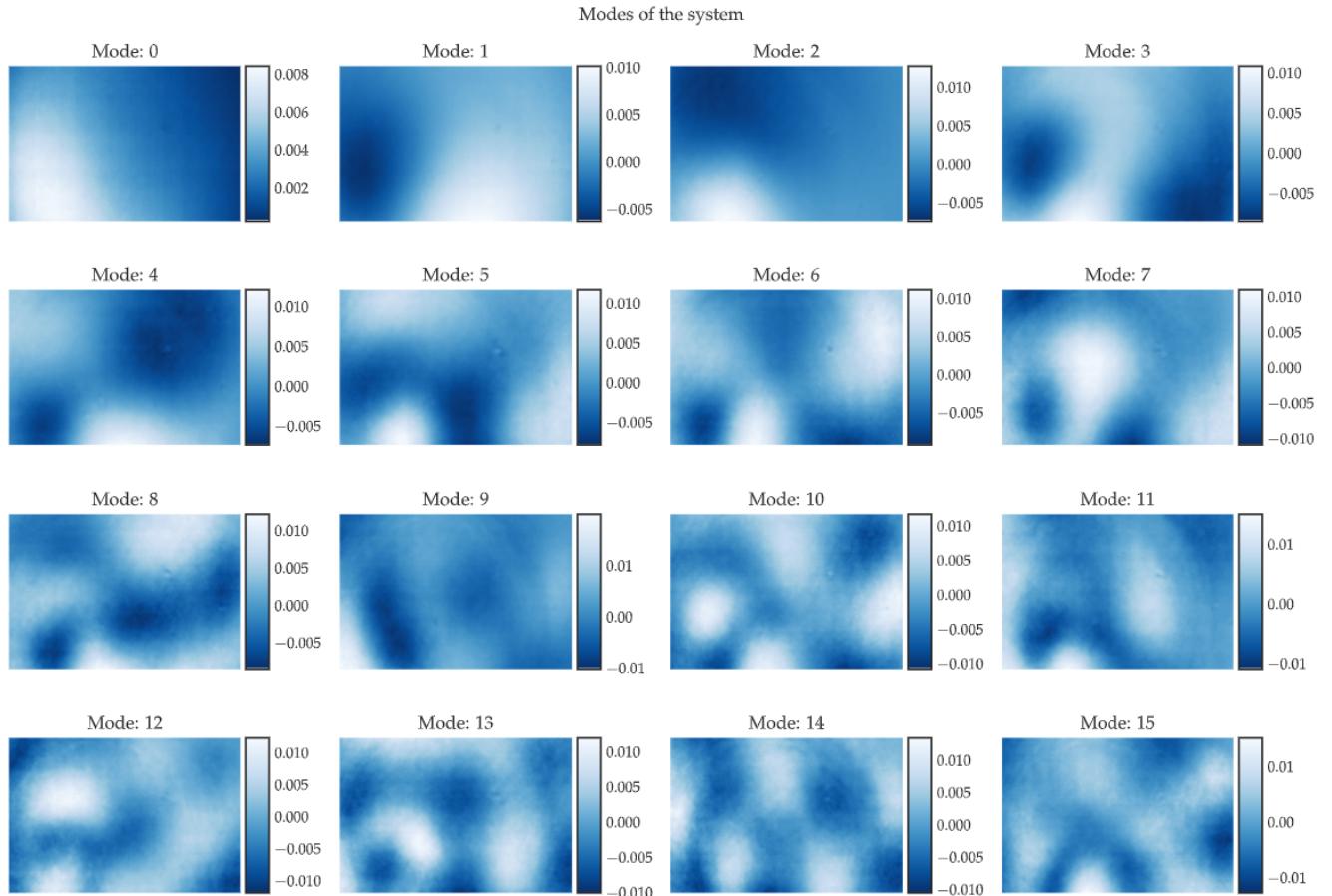


Figure 5.6: The first 16 modes of the system calculated by taking the U term of the SVD of the matrix which represents the transformation from scene to sensor.

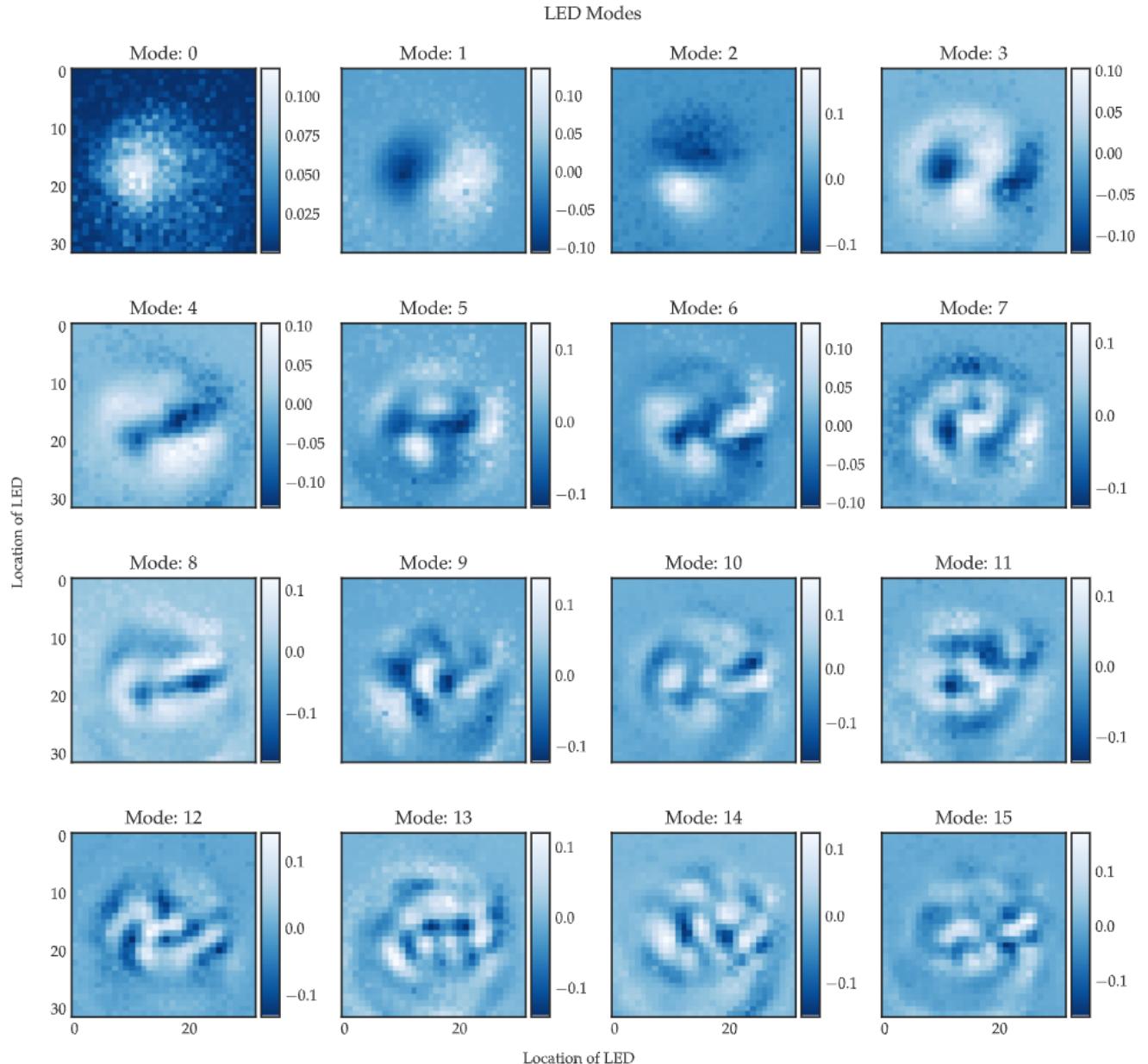


Figure 5.7: The LED values which correspond to each of the first 16 modes depicted in Figure 5.6. These are the V terms of the SVD.

is approximated by $\mathbf{U}_r, \Sigma_r, \mathbf{V}_r^t$. This is done through the algorithm below:

1. Determine the order of the rank r that is needed.
2. A random projection $\mathbf{P} \in \mathbb{R}^{m \times r}$ is created.
3. This \mathbf{P} is used to sample \mathbf{X} to create $\mathbf{Z} = \mathbf{X}\mathbf{P}$. This shrinks columns of \mathbf{X} from m to r .
4. \mathbf{Z} is then broken down through QR decomposition $\mathbf{Z} = \mathbf{Q}\mathbf{R}$ into orthonormal bases for \mathbf{Z} . These are assumed to be equivalent for \mathbf{X} due to high dimensional geometry and the random nature of \mathbf{P} .
5. With the basis \mathbf{Q} it is possible to project \mathbf{X} into a smaller space $\mathbf{Y} = \mathbf{Q}^T\mathbf{X}$.
6. The SVD of \mathbf{Y} can then be done $\mathbf{Y} = \mathbf{U}_y \Sigma \mathbf{V}^T$. As \mathbf{Q} is an orthonormal approximation of \mathbf{X} , Σ and \mathbf{V} are the same for \mathbf{X} & \mathbf{Y} .
7. It is then lifted back to give $\mathbf{U} = \mathbf{Q}\mathbf{U}_y$.

Improvements

One improvement would be to do oversampling. By taking 5 or 10 extra columns instead of just r the accuracy of the approximation can be improved as it increases the likelihood of getting the dominant subspace of \mathbf{X} . Additionally, a further improvement could be found through power iteration. Instead of calculating the SVD of the real \mathbf{X} it can be raised to the power q . This reshapes the matrix to make sure there is a drop off in the magnitude of the singular values allowing a smaller value of r to be taken to get a better result. Raising the power increases the computational time. Yet this change is not as great as actually computing the power: instead of computing \mathbf{X} to the power q the calculation can be done when you are multiplying by \mathbf{P} .

$$\mathbf{X}^{(q)} = (\mathbf{X}\mathbf{X}^T)^q \mathbf{X} \quad (5.3)$$

$$\mathbf{X}^{(q)} = \mathbf{U} \Sigma^{2q-1} \mathbf{V}^T \quad (5.4)$$

Error Bound

There are guaranteed error bounds which show the accuracy of the approximation, shown in Equation 5.5 [18]. This shows how the error bound is tunable so the difference between the real matrix and the

approximation is dependent on the desired rank. The expected error can also be calculated.

$$\|\mathbf{X} - \mathbf{QY}\|_2 \leq \sigma_r + 1(\mathbf{X}) \quad (5.5)$$

5.2.2 Compression

The process of first loading 1024 images and then calculating their pseudo inverse is computationally expensive. Therefore, in order to complete this process in an acceptable time frame with the computational power available, the images have to be compressed down from 3280×2464 to 300×200 . This is not quite equal in the horizontal axis to the vertical axis, however, it has not affected the result. The compression is done using a bicubic algorithm, which calculates the output pixel value from interpolating the original pixels that are combined when the image is compressed down to a smaller size. This allows the calculation to be done on smaller matrices and arrays without compromising the accuracy too much.

5.2.3 Moore-Penrose Pseudo-Inverse

To find the reconstruction of the scene a linear system such as the one in Equation 5.6 needs to be solved. This is normally done by finding the inverse as in Equation 5.7. However, not all matrices are invertible. This is the case for our problem as the matrix is not square so cannot be inverted. So an approximate solution can be calculated by using the pseudo inverse. As in Equation 5.8 where A^+ is the pseudo inverse.

$$\mathbf{Ax} = \mathbf{b} \quad (5.6)$$

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \quad (5.7)$$

$$\mathbf{x} \approx \mathbf{A}^+\mathbf{b} \quad (5.8)$$

The pseudo-inverse was defined as an approximation to an inverse in 1955 by Penrose [19] for any matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ as \mathbf{A}^+ . A pseudo-inverse is unique for each matrix and has the properties shown in Equation 5.11. These properties allow it to be used to find the solution of the our problem.

$$\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A} \quad (5.9)$$

$$(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+ \quad (5.10)$$

$$(5.11)$$

To find a pseudo-inverse with the properties described above the first SVD of the matrix needs to be calculated. Then the singular values need to be inverted. This process is shown in Equation 5.12.

$$\mathbf{A}^+ = \mathbf{U}\Sigma^{-1}\mathbf{V} \quad (5.12)$$

5.2.4 Limiting Modes

One of the main reasons for using SVD is to reduce the number of dimensions in the system. This is because if the underlying trends are extracted from a large dataset, the calculations can be done on a much smaller matrix and so are much faster. In this project there are two main reasons to limit the number of singular values. Firstly, it reduces the time and number of computations required to calculate the pseudo-inverse. Secondly, it focuses on the more significant modes which have a larger impact on the reconstruction, and removes the modes which introduce noise into the system. Limiting modes should be done if there are a few modes which have a much larger effect, and then the singular values rapidly drop off. This can be seen in Figure 5.8 which shows the first hundred singular values, where there is a sharp decay in the first 20 modes. Given that there are 1024 singular values, the majority of them will only have a minor effect and so can be removed without causing a significant drop in the quality of the reconstruction. Also, keeping only the higher order singular values will remove the noisy lower order values. This will have the benefit of removing the effect of LEDs from outside the field of view, whose sensor data is very low and so will have a higher signal to noise ratio and should not be included in the calculations.

Determining the Number of Modes to Use

As previously explained, it is important to reduce the number of modes used in calculating the pseudo-inverse to improve the speed and quality of the reconstruction. However, determining the number of

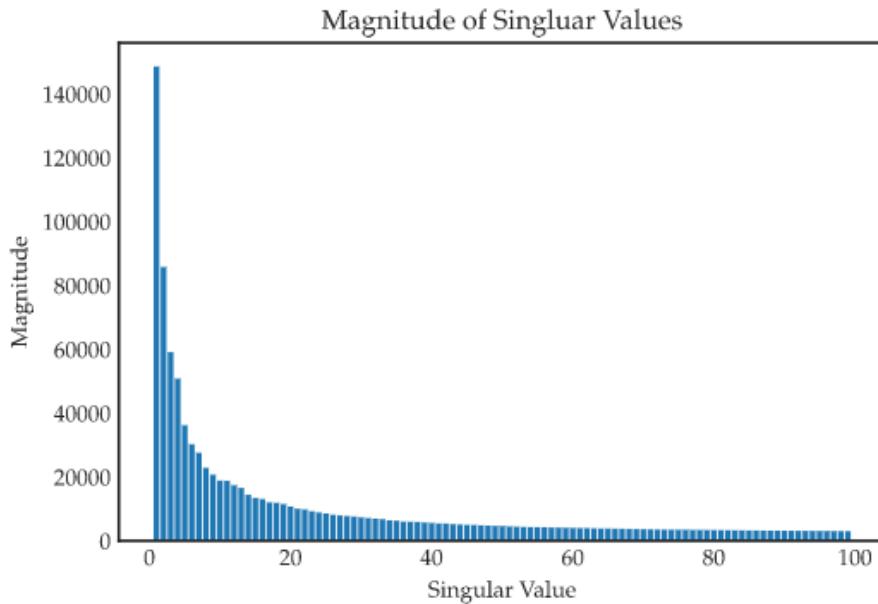


Figure 5.8: The size of the first 100 singular values. This shows there is a sharp drop off after the first few values. This means that limiting the number of modes used should not have a significant effect on the quality of the reconstruction

singular values required to optimally reconstruct the scene is challenging. To represent how using different numbers of modes changes the calculation, Figure 5.9 shows the output when different numbers of modes were used. The scene photographed was of four LEDs arranged at the corners of a square. In the reconstruction with low numbers of singular values, the first 8 and 16 modes (seen in Figure 5.7) are combining together. But there are not enough modes for there to be a combination which allows the summation to make the scene. As more modes are used, the image becomes clear. However, beyond 512 modes the result stops improving, and there becomes more noise around the corners. Different images will require differing numbers of modes to recreate them. This scene was of 4 LEDs which is unlike most of the more significant modes. This meant it required a large number modes to be recreated accurately. A different scene which was closer to a particular mode would require far fewer modes.

5.2.5 Alternative Algorithm

There is an alternative method that could be used described in the paper [8]. In this paper the reconstruction is done though solving a regularised linear inverse problem, which reconstructs the scene from the image following the equation 5.13. This is solved though an L2 minimisation. α is the regularisation parameter used to determine the noise suppression.

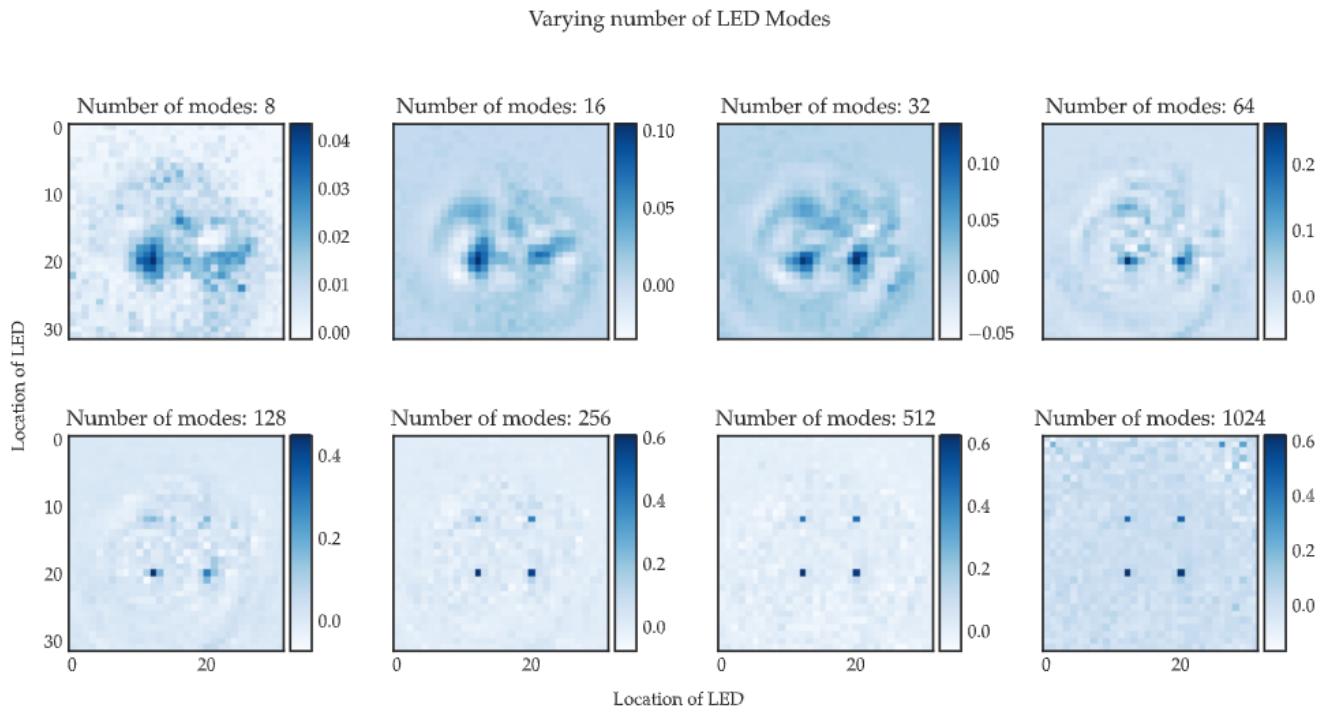


Figure 5.9: Reconstructions of four dots when a varying number of modes were used to create the pseudo-inverse. The first row with 8 to 64 modes shows the issue with having too few modes, as the scene cannot be accurately recreated. From 256 modes to 512 modes there are good reconstructions. At 1024 modes there is more noise, particularly in the corners.

$$\hat{\mathbf{x}} = \operatorname{argmin} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \alpha^2 \|\mathbf{x}\|_2^2 \quad (5.13)$$

This is a more comparable process to the method used in the first system, framing the reconstruction as an optimisation problem then finding the minimum. Which is why this approach wasn't used in this report, as the aim of this project is to show there are a range of techniques that can be used in lens-less imaging. By using an alternative method it has shown that this is the case.

5.3 Characterisation

5.3.1 Creating a Functioning System

The system just described in Section 5.2 needed to be implemented and tested. To do this, each LED was displayed on the panel and imaged. These images were then compiled into a transformation matrix. Then the pseudo-inverse was calculated by decomposing the matrix through SVD, as described in the Section 5.2. Once the inverse transformation was calculated any test scene displayed could be imaged, resulting

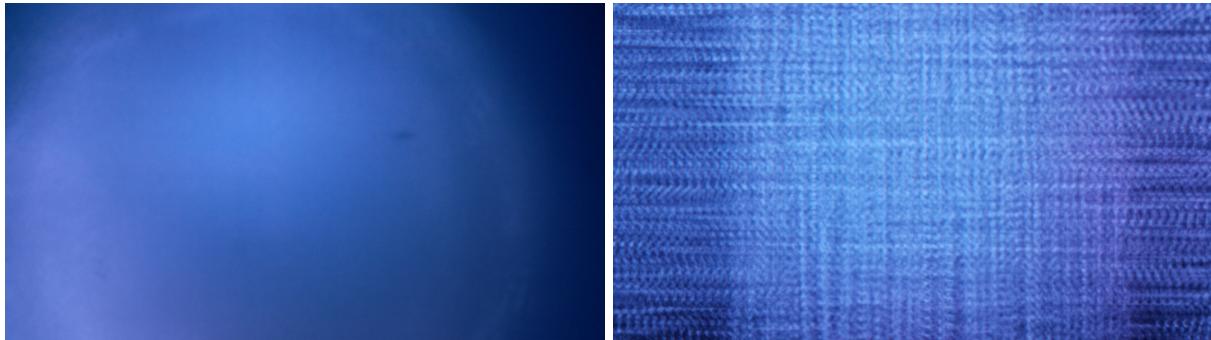


Figure 5.10: This raw sensor data through the light guide when imaging a cross. This raw data looks nothing like a cross. It has been blurred by the light guide but it can still be reconstructed to give Figure 5.12.

Figure 5.11: This image is of the raw sensor data when imaging a cross though the diffuser. It different to the image through the light guide as there is a clear repeating pattern due the spatial invariance.

in the image shown in Figure 5.10. This is then multiplied by this matrix to find the reconstruction. Figure 5.12 displays this; it is the reconstruction of a cross made from two lines of 15 LEDs positioned in the centre of the panel. In the graph there is a distinct cross in the centre, with the values rising above the low level noise.

5.3.2 Field of View

Method

As with the previous system, the angular field of view can be calculated and compared to the predicted value from the geometry of the system. This can be done in a couple of different ways. One method is to look at images taken of each LED individually, and then take the sum of all the pixel intensities in each picture. These sums can be compared to each other and if they are less than 20% of the value of the image with the highest sum, can be assumed to not be providing enough light to produce a good reconstruction. These images can be disregarded as they will just introduce noise into the system and can be taken to be outside the field of view. An alternative method would be to look at the sum of the SVD \mathbf{V} term, since pixels that are only included in low order modes will be more affected by noise and so should be taken as outside the field of view. Then with both methods, using the size of the area that can be accurately predicted together with some trigonometry, the angular field of view can be calculated. If the noise of the system was known, this could be compared to the signal to predict the signal to noise ratio (SNR). This could then be used to calculate the threshold in a more scientific fashion rather than an arbitrary cut-off.

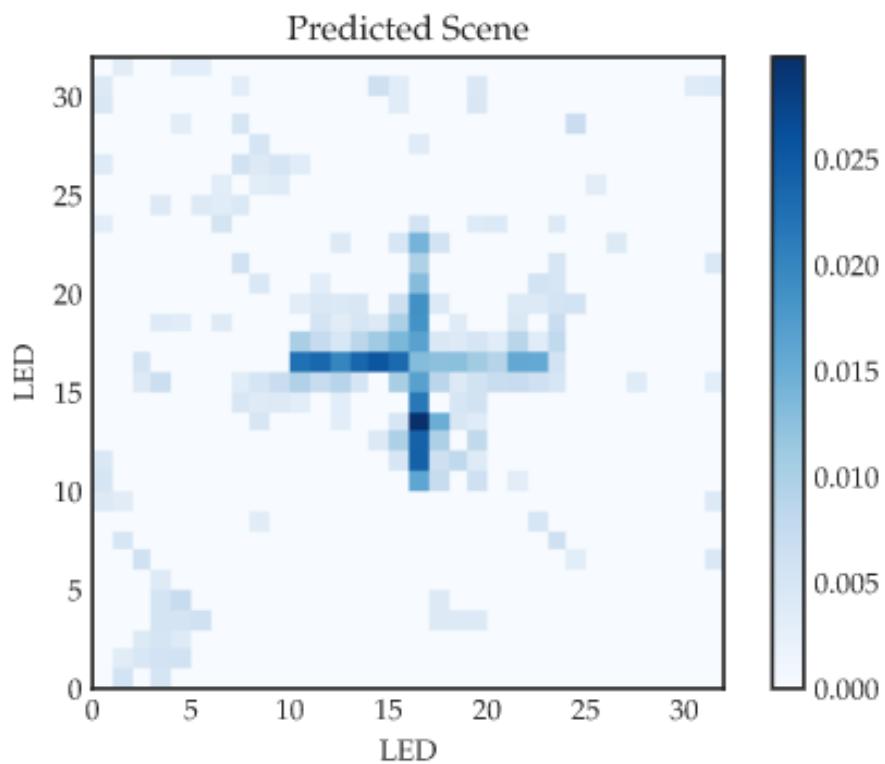


Figure 5.12: This figure shows the reconstruction from the sensor data shown in Figure 5.10. The points where the LEDs are switched on (in the shape of a cross) are much higher than the background noise showing they have been accurately predicted. However, not all the points in the cross are as far above the cut off.

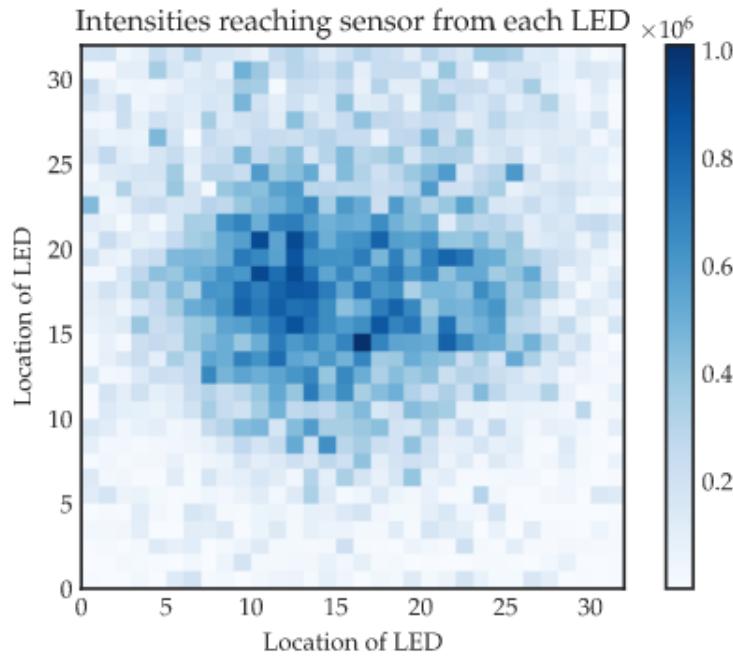


Figure 5.13: This figure depicts the total pixel values from images taken with each LED turned on individually.

Results

Figure 5.13 shows the amount of light from each LED which is reaching the sensor. Those LEDs which when imaged had a total pixel intensity of above 20% of the brightest image, produced a rough circle with a diameter of 22 LEDs. As stated before, each LED has a width of 6mm giving the diameter of the field of view to be 13cm. Again, the distance from the scene to the light guide is 27cm. This gives a field of view of 27.5° . It is possible to calculate the distance from the panel required to allow the entire screen to be inside the field of view. For this to be the case the radius of the field of view needs to be 13.4cm, as this is equal to the length from the centre to the corner of the panel. The angular field of view was just calculated to be 27° . So, the distance required to accurately image the whole panel was found to be equal to 55.6cm.

5.3.3 Resolution

Method

As stated in the previous section the resolution of a camera is its ability to separate two points from each other. To determine the resolution of the system, the minimum distance between differentiable points needs to be calculated. This can be tested by imaging two lights and varying the separation between

them, to ascertain when the system can no longer determine they are lights from different sources. Given that this imaging system is only tested on an LED panel where each point is 6mm apart, the algorithm should be able to separate each one from another. The system has a maximum possible resolution equal to the number of LEDs.

Results

The results of this test are shown in Figure 5.14 which displays the predicted total pixel value for each LED. The first column shows the results (in blue) when two LEDs are switched on next to each other. The central column shows the results when two LEDs are switched on with a spacing of one LED between them, and the third column when there is a gap of two LEDs between them. Overlaid over all of them are the results when each LED is switched on individually. From this graph it can be seen that the correct entries are higher than the baseline for all entries. When the lights are at the edge of the panel, the peaks are sharp and there is more noise at the edge of its field of view. This shows that in the centre of the field of view, the resolution is equal to the maximum it could be for this system.

5.3.4 Accuracy

If the aim were not to reconstruct an image but to predict which LEDs are switched on, it would be possible to test the accuracy of the system. To get an output for which LEDs are predicted to be ON or OFF, a simple threshold can be put in place. This can then be compared to the scene to determine the accuracy of the reconstruction.

5.4 Walsh-Hadamard Functions

Originally the calibration matrix was calculated from imaging each LED individually and combining the data from each image. An alternative method would be to image the LEDs arranged in patterns then combine these images to form a matrix. This matrix would then need to be multiplied by another matrix to tell the system which LEDs are illuminated, which would then produce a matrix equivalent to the original method.

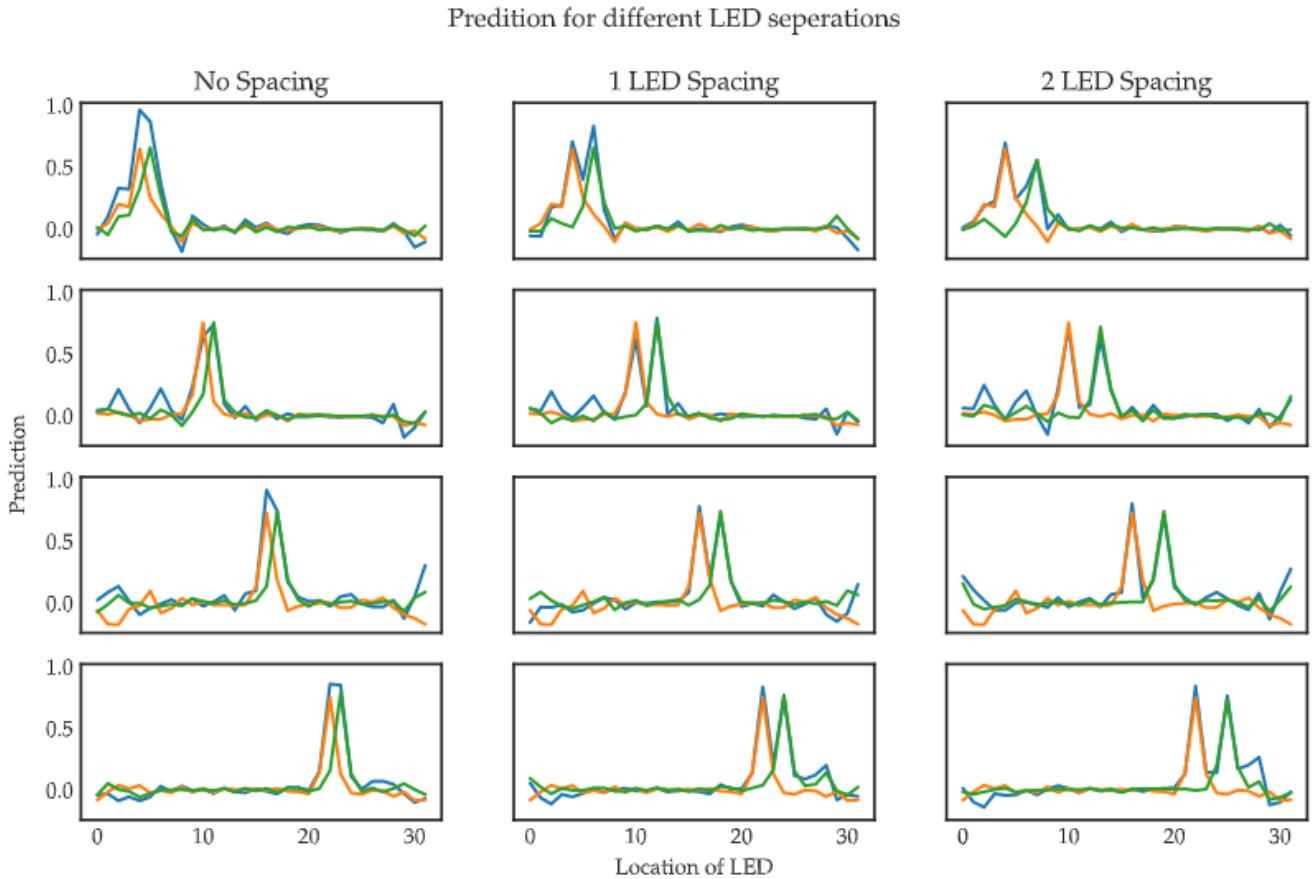


Figure 5.14: Calculating the minimum separation between two LEDs which is still discernable. In the first column the two LEDs are neighbouring, in the second column there is one LED between them and in the third there is a gap of two LEDs. The blue lines depict the value predicted of each LED in the line when two LEDs are on. This is compared to the green and orange lines which show the value when each LED is on individually. The orange and green lines overlap the blue line substantially, which means that each LED is identified individually and is separable from its neighbour.

5.4.1 Walsh Functions Definition

A pattern that could be used to replace imaging each light individually is one formed from the Walsh-Hadamard function. In the paper [20] the idea for Walsh functions is set out. These are a series of normal orthogonal basis functions which can be used to represent any discrete function. They split the range into binary values with half being 1 and half -1. As these are setting the values of LEDs, -1 represents OFF and 1 represents ON. The equation for a Walsh function in 1D is shown in Equation 5.17. To go from this to 2D Walsh functions that can image:

1. Take the 1D Walsh functions to represent the x direction
2. Take another 1D Walsh function to represent the y direction.
3. Pass them both through an AND gate to produce each of the matrices. This is shown in Figure 5.15.

$$\phi_0 = 1, 0 \leq x \leq 1 \quad (5.14)$$

$$\phi_1 = \begin{cases} 1 & \text{if } 0 \leq x < 1/2 \\ -1 & \text{if } 1/2 < x \leq 1 \end{cases} \quad (5.15)$$

$$\phi_{n+1}^{(2k+1)}(x) = \begin{cases} \phi_n^{(k)}(2x) & \text{if } 0 \leq x < 1/2 \\ (-1)^{k+1} \phi_n^{(k)}(2x - 1) & \text{if } 1/2 < x \leq 1 \end{cases} \quad (5.16)$$

$$\phi_{n+1}^{(2k)}(x) = \begin{cases} \phi_n^{(k)}(2x) & \text{if } 0 < x \leq 1/2 \\ (-1)^k \phi_n^{(k)}(2x - 1) & \text{if } 1/2 < x < 1 \end{cases} \quad (5.17)$$

5.4.2 Walsh Functions in Optical Systems

Walsh functions are useful in optical systems as they have various properties (as described in [21]). There are benefits to including this type of function in this project. In particular only turn on one LED at a time to calculate the transformation matrix caused the signal to be low so the SNR is higher especially at increased distances. Walsh functions could be used to calculate the transformation matrix instead. The Walsh function for every element of order two is self-inverse. This means that to convert the images taken of each Walsh function into a matrix that is of the same form as in the previous method, the images of

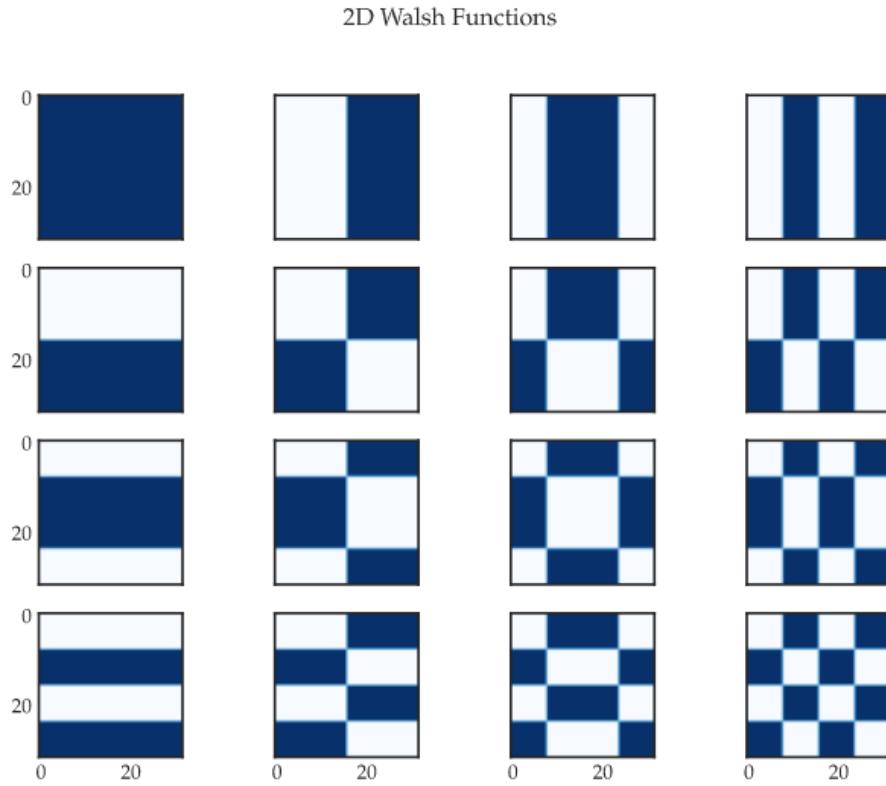


Figure 5.15: 2D Walsh functions which were projected onto the LED panel and photographed. These pictures were then used to determine the transformation matrix

Walsh functions only need to be multiplied by the Walsh functions that were used to generate it.

Distance

One of the main reasons for changing to using the Walsh functions is because of increasing the distance from sensor to the LED panel. Varying the distance changes the resolution and field of view, because moving the sensor further from the scene means a larger radius circle will be in the field of view. As the distance increases, the light intensity reduces by the square of the distance, following equation 5.18. This means that as the sensor is moved away from the light the signal it will pick up is lower but the noise will remain the same. The signal to noise ratio will increase so there becomes less usable information to reconstruct the scene, which would result in an inferior reconstruction.

$$I \propto \frac{1}{d^2} \quad (5.18)$$

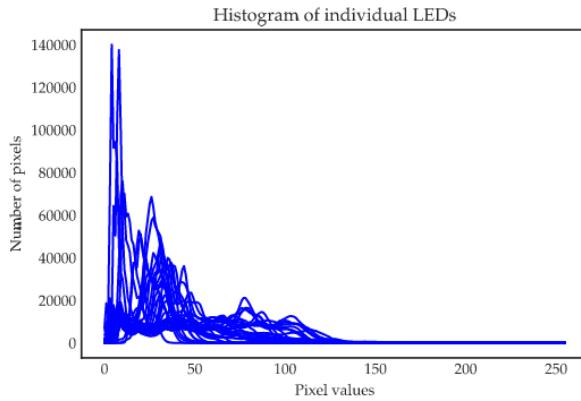


Figure 5.16: The pixel values imaging each LED individually. There is a large range in intensity between different images as the closer to the centre the LED is, the more light will reach the sensor despite all the LEDs being the same brightness.

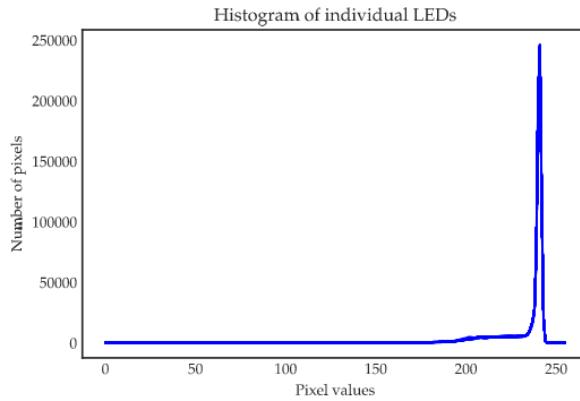


Figure 5.17: The pixel values when imaging Walsh functions. As the functions have half the LEDs switched on and evenly spread out, all the images have very similar intensities at much higher values compared to Figure 5.16.

Over-saturation

When using Walsh functions, the potential issue of over-saturation could arise. As the pixels output a value between 0 and 255, if there is too much light reaching the sensor the pixels will start saturating. This will result in the contrast between pixels reducing as they are limited, and in the extreme the contrast will be zero. Such saturation is a problem as it would mean the data could not be used to reconstruct an image. This is not an issue in this case as the pixel values when imaging the Walsh function do not go above 250 as shown in Figure 5.17. If the shutter speed was increased or the sensor moved closer it could become one.

Using Walsh functions is a similar idea to using SVD, as it is building the real image by combining together various modes. Much like with the SVD, it would also be possible to use a limited number of modes. Reducing the number of Walsh basis functions used would decrease the time required to take the images for the transformation matrix.

5.5 Conclusion

In this section an alternative method of lens-less imaging has been described and tested. The goal here was to design a system which could reliably reconstruct the main points of a scene. Issues such as field of view, peripheral noise and resolution have been explored and resolved above. The visual images in

this chapter clearly demonstrate the delicate balance in finding the optimal number of modes to use to achieve a representation of the main points of the image. In spite of the comparative simplicity of the equipment used, the fundamental task has been achieved. The next step would be refining and improving the rendition of the image. The concluding section of this report will compare these developments to the first tested system from Section 4.

Discussion

6.1 Comparison of Characteristics

In this section, we compare the results of the systems described in Sections 4 and 5 with each other. As both systems have their own advantages and disadvantages due to the nature of the problem they were solving.

6.1.1 Field of View

The field of view of each system is relatively similar at around 27.5° . This appears to be due to both systems having a similar sized aperture at the same distance from the sensor. This is quite low compared to modern cameras which is often around 58° . However, with a different set up where the apertures could be closer to the sensor the field of view could be increased.

6.1.2 Resolution

The resolution of the two systems is different, as the spatially variant system uses the knowledge that the panel is a 32×32 board. So the output is just a 32×32 image but it can separate an LED from its neighbouring one. The linear system can also separate neighbouring LEDs from each other, but it does become less clear as you move away from the centre. The invariant system, however, is not limited to the 32×32 image and can find more detail if it is there.

6.1.3 Time

Another metric that needs to be considered is how long it takes to image with each system. The time can be split into two parts: firstly time to collect data and calibrate the system, and secondly time to

perform the reconstruction on a test image. The diffuser system only requires two images to be taken, the calibration PSF and the test image. The light guide system required 1025 images. This means it takes around two hours to take all the images in the second system. With the Walsh method this could be cut down, but it is still far greater. The computational time can vary based on the number of factors for both systems; the extent of the compression of the image is, if the image is taken in colour or the the number of iterations used in the first system, for the second system the number of modes used in the pseudo-inverse will also affect the time taken.

6.1.4 Versatility

While both systems are designed to solve similar problems, they have different limitations. The first system requires an object that has a consistent effect on the light and is very close to the sensor to make sure the LSI property holds. This means that it can have a much smaller calibration and a resolution much closer to that of a camera with a lens. However, the invariant system could work for a wider range of transparent objects in front of the sensor, or nothing at all. This means it can be expanded to image through a wide range of designs, but this also means it requires a much larger calibration process and it is currently limited to imaging the 32 x 32 LED panel.

6.2 Recommendations for Future Work

There is a wide range of topics to continue researching in this area, which are unfortunate beyond the scope of this report. If this project was continued, more work could be done on researching how well this system can be diversified. Looking at the optical properties of different objects and how they scatter light into the sensor from liquids to more reflective materials.

A third area which this report could be expanded to include, would be to look more into how the results could be improved though machine learning. This is discussed in the paper [7]. This takes the interesting approach of using machine learning to supplement, instead of replace, the mathematics of the system. This would require a large dataset to be taken, which then could be used for training and validation. Machine learning could greatly increase the versatility of the system as the training data can be taken at a range of light intensities and distances to create a robust dataset.

6.2.1 Imaging Real Objects

This report has mainly focused on the technical capabilities of this system but it would be remiss not to include a section on its effectiveness imaging real scenes. There is a large difference between the controlled environment created using a dark room and an LED panel to imaging real objects. Real objects may not be well lit and most of light on them will be reflected, as opposed to the illuminated single points of the LEDs that were used in this research. This means real objects are dimmer and exist as a gradient of brightness as opposed to a simple ON or OFF. Additionally, imaging in a real environment involves constantly changing ambient brightness and light from sources other than the object you are taking a photograph of. Moving to imaging real objects could be done in stages; imaging the LED panel is just the first step. From this it is possible to move to a computer LCD screen, and then to real objects which are well lit, before using the system as a normal camera.

6.2.2 Applications

Fibre Optics

More research could be done into how this methodology could be applied to optical fibres. This would allow a simple method which was able to image through an optical fibre. There has already been some research in this, however, this system could also give interesting results if applied to this problem. Improvements in imaging through fibre optics, which would have numerous applications. Imaging objects remotely is very important either to monitor areas that would be hazardous for humans (such as inside nuclear reactors) or inaccessible areas where creating a hole larger than a few millimetres could damage the object being imaged.

Medical Imaging

Another important application of lens-less imaging would be medical imaging, since by removing the lens you can reduce the dimensions of the camera. A lens was originally designed to increase the light received by the sensor; however, it can have the opposite effect when the camera is shrunk. Developments based on this research could allow much smaller cameras which can be used for medical imaging and provide better quality imaging by letting more light reach the sensor.

Conclusion

7.1 Completing Project Aims

The aims for this project set out in the Introduction, were to build and test a version of the DiffuserCam and then to create an alternative model for a different set of specifications. These goals were achieved and both systems were analysed, with the systems performing reliably. This paper has researched a range of algorithms and showed that they can be used to solve this problem. But each algorithm needs to be implemented with an understanding of the design of the system. These methods have produced good quality reconstructions of test images. One of the practical goals was that the design should be possible with readily available parts, to increase the accessibility for others to develop the work in the future. The imaging data was captured on a Raspberry Pi and the computing was done on a laptop, without having to resort to using a higher performance computer such as Oxford’s ARC to do the calculations. The clarity of using these simple elements has enabled clear testing of the concepts, the results can be improved in the future on more sophisticated equipment. One area that would have benefited from a longer project is testing the boundaries of the system further. This could include testing more complex scenes, and looking at how changing the dimensions of the light guide affects the reconstruction.

7.1.1 Advantages of Lens-less Imaging

Lens-less imaging allows for a more complete understanding of the imaging system and allows the design of cameras to be radically changed. When the requirement for large heavy lenses is removed, a camera can be reimaged. The fact that the system was still able to recreate the scene even after a dramatic compression of the image is promising. As computing power increases it will be possible to use images that are less compressed, and do the calculations in a shorter period of time. This will allow more nuance

in the image to appear and a higher resolution panel to be imaged.

7.1.2 Problems to be Solved

There are still many issues with lens-less imaging. Firstly, the resolution was far below that of standard cameras currently available. Secondly, the processing required to reconstruct these lower resolution images is still too computationally costly for most useful applications. Finally, the second system required a large calibration process where the system had to stay in the same place in a dark room for several hours to before it would work on any test data. This calibration also required large matrices to be saved. With the Walsh functions it is possible to limit the number of modes imaged to reduce this issue, but this would have a negative effect on resolution. Hopefully in the future this problems will be examined and solved.

Bibliography

- [1] N. Antipa, G. Kuo, R. Heckel, B. Mildenhall, E. Bostan, R. Ng, and L. Waller, “Diffusercam: lensless single-exposure 3d imaging,” *Optica*, vol. 5, pp. 1–9, Jan 2018.
- [2] S. Rana, “Diffuser based imaging and wavefronts,” 2020.
- [3] J. N. Mait, “A history of imaging: Revisiting the past to chart the future,” *Opt. Photon. News*, vol. 17, pp. 22–27, Feb 2006.
- [4] R. Dicke, “Scatter-Hole Cameras for X-Rays and Gamma Rays,” *The astrophysical journal*, vol. 153, p. L101, Aug. 1968.
- [5] M. S. Asif, A. Ayremlou, A. Sankaranarayanan, A. Veeraraghavan, and R. G. Baraniuk, “Flatcam: Thin, lensless cameras using coded aperture and computation,” *IEEE Transactions on Computational Imaging*, vol. 3, no. 3, pp. 384–397, 2017.
- [6] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, “Image and depth from a conventional camera with a coded aperture,” *ACM Trans. Graph.*, vol. 26, p. 70–es, July 2007.
- [7] K. Monakhova, J. Yurtsever, G. Kuo, N. Antipa, K. Yanny, and L. Waller, “Learned reconstructions for practical mask-based lensless imaging,” *Opt. Express*, vol. 27, pp. 28075–28090, Sep 2019.
- [8] G. Kim, K. Isaacson, R. Palmer, and R. Menon, “Lensless photography with only an image sensor,” *Appl. Opt.*, vol. 56, pp. 6450–6456, Aug 2017.
- [9] G. Kim and R. Menon, “Computational imaging enables a see-through lens-less camera,” *Opt. Express*, vol. 26, pp. 22826–22836, Sep 2018.
- [10] A. Porat, E. R. Andresen, H. Rigneault, D. Oron, S. Gigan, and O. Katz, “Widefield lensless imaging through a fiber bundle via speckle correlations,” *Opt. Express*, vol. 24, pp. 16835–16855, Jul 2016.

- [11] A. Orth, M. Ploschner, E. R. Wilson, I. S. Maksymov, and B. C. Gibson, “Optical fiber bundles: Ultra-slim light field imaging probes,” *Science Advances*, vol. 5, no. 4, 2019.
- [12] P. Caramazza, O. Moran, R. Murray-Smith, and D. Faccio, “Transmission of natural scene images through a multimode fibre,” *Nature Communications*.
- [13] S. Boyd and L. Vandenberghe, *Convex Optimization*. USA: Cambridge University Press, 2004.
- [14] “Gradient descent with constant learning rate for a quadratic function of multiple variables.”
- [15]
- [16] G. Strang, *Introduction to Linear Algebra*. Wellesley, MA: Wellesley-Cambridge Press, fifth ed., 2016.
- [17] N. Halko, P.-G. Martinsson, and J. A. Tropp, “Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions,” 2010.
- [18] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2019.
- [19] R. Penrose, “A generalized inverse for matrices,” *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 51, no. 3, p. 406–413, 1955.
- [20] J. L. Walsh, “A closed set of normal orthogonal functions,” *American Journal of Mathematics*, vol. 45, no. 1, pp. 5–24, 1923.
- [21] M. de and L. Hazra, “Walsh functions in problems of optical imagery,” *Optica Acta*, vol. 24, pp. 221–234, 02 1977.

Appendix

A.1 Code & Results

The code used for all the data collection and calculations has been uploaded to GitHub. There is more data and results that weren't included in this report that have also been uploaded to the same location. All the coding was done using Python and Jupyter Notebooks, except for the code to illuminate the LED panel which was done in Arduino. There is a list of the required packages needed to run the code. To access it go to GitHub <https://github.com/WilliamGasson/LenslessImaging>. The code for the DiffuserCam project that the invariant system was derived from can be also be found on GitHub here <https://github.com/Waller-Lab/DiffuserCam-Tutorial>

A.2 Health & Safety

The risk involved in a project needs to be examined before its undertaking. Attached is the risk assessment for this project, which describes the predicted hazards that might arise along with the measures in place to mitigate the risk that these hazards present.

Department of Engineering Science

4YP Risk Assessment

Description of 4YP task or aspect being risk assessed here: <i>(Read the Guidance Notes before completing this form)</i> Building and Testing a lens-less imaging setup		4YP Project Number: 12280
Site, Building & Room Number: 17 Norham Gardens	Approx size of equipment/apparatus used or built (in metres): Height: 20cm Width: 20cm Length: 50cm	Photo provided? YES/NO
Assessment undertaken by: William Gasson	Signed: 	Date: 03/11/2020
Assessment Supervisor: Martin Booth	Signed: 	Date: 3/11/2020

Assessing the Risk*		RISK MATRIX		LIKELIHOOD (or probability)			
CONSEQUENCES	High	Medium	Low	Remote			
	Severe	High	High	Medium	Low		
	Moderate	High	Medium	Medium/Low	Effectively Zero		
	Insignificant	Medium/Low	Low	Low	Effectively Zero		
Consequences	Negligible	Effectively Zero	Effectively Zero	Effectively Zero	Effectively Zero		

You can do this for each hazard as follows:

- Consequences:** Decide how severe the outcome for each hazard would be if something went wrong (i.e. what are the Consequences?) Death would be "Severe", a minor cut to a finger could be regarded as "Insignificant".
- Likelihood:** How likely are these Consequences to actually happen? Highly likely? Remotely likely, or somewhere in between?
- Risk Rating:** Start at the left of the coloured Matrix. On your chosen Consequences row, read across until you are in the correct Likelihood column for the hazard in question. For example, an outcome with Severe consequences but with a Low probability of actually happening equates to a Medium risk overall. In this case "Medium" is what should be written in the Risk.

Hazard (potential for harm)	Persons at Risk	Risk Controls In Place (existing safety precautions)	Risk*	Future Actions identified to Reduce Risks (but not in place yet)
Damage to eyesight from LEDs	William Gasson	The LEDs are low intensity consumer items so there is minimal risk of exposure. In normal operation the lights will be covered when in use by black card.	Low	There are no lasers involved so a "Introduction to Light Sources" Safety Training is not required
Electricity running through microcontroller/ LED panel	William Gasson	The equipment is all operating at low voltage, so there is minimal hazard. Only work on the wiring when the parts are not connected to mains power	Low	

Hazard (potential for harm)	Persons at Risk	Risk Controls In Place (existing safety precautions)	Risk*	Future Actions identified to Reduce Risks (but not in place yet)
Using computers for extended periods of time	William Gasson	A computer usage risk assessment form has been completed and the recommendations from that will be followed	Low	
Transport of equipment between department and home	William Gasson	Equipment consists of small, low risk items, so no special arrangements are required.	Low	

Department of Engineering Science



Additional guidance for 4th Year Project students
using laptop computers

Laptops are designed for short term portable use. They have smaller keyboards and screens, compact pointing devices and very little adjustability. Prolonged use therefore can lead to discomfort.

Wherever possible, a separate keyboard, mouse and laptop stand should be used to improve the ergonomic arrangement in line with the general guidance on good ergonomics.

Laptop use, a summary:

- Use an external mouse and keyboard wherever possible.
- Use a proper desk rather than a bench or lap.
- Keep the mouse and keyboard at the edge of your desk.
- Adjust your chair so your desk allows your arms to work comfortably at the keyboard.
- Raise the laptop so the top of the screen is just below your eye level. Use a laptop stand, stack of books, or similar.
- Sit back in your chair and keep your head directly over your shoulders.
- Don't lean on your desk while typing and keep your elbows in line with your shoulders, hanging loosely at your sides.
- When using the laptop keyboard, try not to drop your wrists onto the wrist rest. Instead, move your hands freely across the keyboard and keep your hands in line with your elbows.
- Avoid prolonged carrying of laptops. If the laptop must be carried regularly over distance e.g. to another building on a weekly basis, a backpack carrier rather than a briefcase design.
- Security of individuals and the data held on laptop computers is also an issue of concern. If possible, tasks should be avoided that involve lone users carrying or storing portable equipment in any area where theft is a possibility.

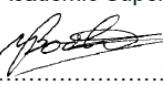


Photographs reproduced from Health & Safety Executive's publication "Work with display screen equipment" L26

Department of Engineering Science
Supplementary Questions for 4th Year Project Students



Factor	Answer	Things to Consider	Record details here
Has the checklist covered all the problems that may arise from working with the VDU?	<input type="checkbox"/> <input checked="" type="checkbox"/> Yes No		I am completeing my project on a laptop so I will follow the guidance on page 11
Are you free from experiencing any fatigue, stress, discomfort or other symptoms which you attribute to working with the VDU or work environment?	<input type="checkbox"/> <input checked="" type="checkbox"/> Yes No	Any aches, pains or sensory loss (tingling or pins and needles) in your neck, back shoulders or upper limbs. Do you experience restricted joint movement, impaired finger movements, grip or other disability, temporary or permanently	I am free from any fatigue, stress and other symptoms
Do you take adequate breaks when working at the VDU?	<input type="checkbox"/> <input checked="" type="checkbox"/> Yes No	Periods of two minutes looking away from the screen taken every 20 minutes and longer periods every 2 hours Natural breaks for taking a drink and moving around the office answering the phone etc.	I will take breaks as described. My project gives me natural breaks to edit the physical devices
How many hours per day do you spend working with this computer?	<input type="checkbox"/> <input checked="" type="checkbox"/> 1-2 3-4 <input type="checkbox"/> <input checked="" type="checkbox"/> 5-7 8 or more	3-4	
How many days per week do you spend working with this computer?	<input type="checkbox"/> <input checked="" type="checkbox"/> 1-2 3-5 <input type="checkbox"/> 6-7	3-5	
Please describe your typical computer usage pattern		I naturally work on the computer in the mornings and evening and work from print outs in the afternoon, this means I am not working on the computer for too many consecutive hours.	

<u>Student Declaration and Academic Approval</u>		
<u>Student Declaration:</u>		<u>Academic Approval</u>
I have completed the DSE Workstation Checklist and the Supplementary Questions for my computer-related risk assessment for 4YP Project Number indicated below:		I confirm my approval of this 4YP DSE Risk Assessment. Academic Supervisor's Name: (<i>please print</i>)
4YP Project Number:	12280
4YP Student's Name (<i>please print</i>)	William Gasson	Academic Supervisor's Signature:  J
4YP Student's Signature:	william gasson	