



Zhang, Yiwei (2017) *Investigations into applications of photometric stereo and single-pixel imaging*. PhD thesis.

<http://theses.gla.ac.uk/8554/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten:Theses
<http://theses.gla.ac.uk/>
theses@gla.ac.uk

Investigations into Applications of Photometric Stereo and Single-pixel Imaging

Yiwei Zhang

A thesis submitted in fulfillment of
the requirements for the degree of
Doctor of Philosophy

Optics Group
School of Physics and Astronomy
College of Science and Engineering
University of Glasgow

October 2017

Declaration of Authorship

I, YIWEI ZHANG, declare that this thesis titled, ‘ Investigations into Applications of Photometric Stereo and Single-pixel Imaging ’ and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Yiwei Zhang

*“ The reasonable man adapts himself to the world.
The unreasonable one persists in trying to adapt the world to himself.
Therefore, all progress depends on the unreasonable man.”*

George Bernard Shaw

*“ If you don't have any shadows you're not in the light.”
“ So there's nothing more provocative than taking a genre that
everybody who's cool hates - and then making it cool.”*

Lady Gaga

Abstract

Computational image reconstruction is generally an inverse procedure which helps to recover the original information in a scene. Various imaging techniques have been developed to extract certain kinds of information for applications in different fields. The focus of this thesis is to improve two elegant and powerful methods among those approaches, namely, photometric stereo and single-pixel imaging, into a more practical and applicable phase.

With the advances in modern imaging technology, 3D information is playing an increasingly significant role in real-world applications, from robotic vision, manufacturing, entertainment, and biology to security. While an immense amount of research has been conducted over the last few decades, the requirement of generating a rapid and accurate estimation of scene depth information with a cost-efficient system remains challenging. In the first work, we developed an inexpensive computational camera system allowing fast 3D reconstruction of objects based on the principle of photometric stereo. By analysing the estimated 3D data of various objects, we noticed good quantitative agreement with the known reference object with a wide viewing angle. With a low-cost accessory, our system provides a simplified reconstruction routine alongside a high efficiency, which extends its portability and capability for practical applications.

Single-pixel imaging is an emerging paradigm which utilises spatial correlation of light with a single-pixel detector to form an image. It provides an alternative strategy to conventional imaging techniques which rely on a pixelated sensor for spatial resolution. In the second work, we combined photometric stereo with

single-pixel imaging to evolve a new 3D imaging system with an efficient real-time sampling scheme. By utilising a high-speed structured illumination and four single-pixel detectors, multiple images of a scene with different shading profiles were able to be reconstructed with perfect pixel registration for depth estimation, empowering 3D imaging of dynamic scene. A compressive strategy, known as evolutionary compressed sensing, was further employed to improve the frame rate of 3D single-pixel video at an expense of only a modest reduction in image quality. This system represents a step-forward towards real-time 3D single-pixel imaging.

By using single-pixel imaging technique, it offers a feasible solution for situations that are costly or constrained with conventional pixelated camera sensor, for instance, near-infrared (NIR) imaging and fluorescence imaging through multimode fibres. However, the signal-to-noise ratio (SNR) scales poorly when increasing the single-pixel imaging resolution. In the last work, we developed a NIR single-pixel imaging system with micro-scanning, an optimisation approach that generates a higher-resolution image while maintaining the SNR of the lower-resolution images where it is derived from. With the use of sunlight and an infrared heat lamp as the illumination sources and a set of NIR bandpass filters, our system indicated a well capability of revealing the water absorption underneath the surfaces of plant leaves and fruits compared to an expensive pixelated InGaAs camera. Additional efforts were devoted to further improve the image quality of a modified single-pixel imaging system that allows visible and NIR dual-band detection simultaneously.

Acknowledgements

This thesis would not have been possible without the support and guidance of many wonderful people. First and foremost, I would like to give a massive thank you to my physics supervisor, Miles Padgett, for taking me on when my initial lead supervisor Rob Jenkins in psychology moved away, and guiding me through the major period of my PhD with much patience and numerous vital pieces of advice. My sincere gratitude also goes to Rob Jenkins for mentoring me in the first year and Monika Harvey in psychology for joining in as a valuable second supervisor.

A big thanks to the University of Glasgow for funding this research with the Lord Kelvin Adam Smith (LKAS) PhD Scholarship.

I am incredibly grateful to all my colleagues in the optics group, for cake, drink, chat, support through hard times and wonderful times in between. I would especially like to thank Matt Edgar and Graham Gibson who have patiently addressed my countless queries and assisted in building many of the experimental set-ups employed during my PhD. Both have taught me a great deal in the lab. Thanks to Baoqing Sun, Ermes Toninelli, Mingjie Sun, Ruifeng Liu, Ruiqing He, Lixiang Chen, Shouqian Chen and many other folks for all the enjoyable discussion. I would also like to thank Niamh Friel in psychology for those genuine suggestions.

In addition, my heartfelt thanks goes to my wonderful friends in England, Yang Yang, Rui Shang, Songhui Sun, Ying Guo and Lewis Andrews, and my bosom friends in Scotland, Marcos Orellana, Lisa van Gent, Craig Purdie, Nan Qi, Weiwei Chen and Alan Gallacher. I would also avail this opportunity to place on record my gratitude to all my unnamed friends for the invaluable support and encouragement.

Your love and affection has left me with sumless wonderful memories along the way of my PhD.

Last, but certainly not the least, I would like to express my deep thanks to my parents and my sister for their dedicated support and unwavering belief in me. They are my most priceless treasure for life.

Dedicated to my beloved parents

Wanhua Zhang and Xiaolan Li

&

my inspirational sister

Yitao Zhang.

Contents

Declaration of Authorship	i
Abstract	iii
Acknowledgements	v
Contents	viii
List of Figures	xi
List of Tables	xiii
Abbreviations	xiv
List of Publications	xvi
1 Introduction	1
1.1 Preamble	1
1.2 Conventional 3D optical imaging techniques	4
1.2.1 Interferometry	4
1.2.2 Structured Illumination	5
1.2.3 Stereo vision	6
1.2.4 Photometric stereo	7
1.3 Single-pixel Imaging	8
1.3.1 Single-pixel imaging theory	9
1.3.2 Computational imaging systems with single-pixel detectors	10
1.4 Conclusion	11
2 A Fast 3D Camera System using Photometric Stereo	13
2.1 Introduction	13
2.1.1 Contributions	14
2.2 Analysis of surface reflection property	14
2.2.1 Specular and diffuse reflection	14

2.2.2	Lambertian reflectance model	16
2.2.3	Shape from shading	18
2.3	Photometric stereo model	19
2.3.1	Classic mathematical calculation model	20
2.3.2	Fast surface gradient calculation model	22
2.4	Experimental set-up	24
2.5	Image acquisition	27
2.6	Quantitative and qualitative analysis	30
2.7	Conclusion	38
3	Real-time 3D Imaging with Single-Pixel Detectors	40
3.1	Introduction	40
3.1.1	Contributions	41
3.2	Computational single-pixel imaging	42
3.2.1	Light projection with a digital micro-mirror device	43
3.2.2	Structured illumination with Hadamard patterns	45
3.2.3	Differential signal acquisition approach	51
3.3	3D imaging system with single-pixel detectors	53
3.3.1	Experimental setup	53
3.3.2	Results	57
3.4	Evolutionary compressed sensing	61
3.5	Real-time video from a 3D single-pixel computational imaging system	65
3.6	Conclusion	69
4	NIR Single-Pixel Imaging with Micro-Scanning	71
4.1	Introduction	71
4.1.1	Contributions	72
4.2	NIR single-pixel detection	72
4.2.1	Experimental setup of NIR imaging system	73
4.2.2	NIR comparison with and without using polarisers	75
4.2.3	NIR imaging with bandpass filters	76
4.3	Micro-scanning imaging	78
4.4	NIR micro-scanning imaging with solar radiation	82
4.5	InGaAs NIR imaging comparison	86
4.6	Conclusion	88
5	Additional Improvements of Single-Pixel System with VIS and NIR Dual-Band Detection	90
5.1	Introduction	90
5.1.1	Contributions	91
5.2	Dual-band imaging modification	91
5.3	Single-pixel dual-band detection with fruits	93
5.4	Additional methods for improving the SNR of single-pixel images	94

5.4.1	Signal averaging	94
5.4.2	Randomisation of pattern sequence	96
5.4.3	Regularisation	97
5.5	Conclusion	99
6	Conclusion and Future Work	100
A	Extra NIR Single-Pixel Imaging Data of Different Objects	104
B	Extra Data of Single-Pixel Dual-Band Detection	108
	Bibliography	112

List of Figures

1.1	Schematic representation of a single-pixel imaging system	9
2.1	Specular and diffuse reflection	15
2.2	An example of different reflections	15
2.3	Lambertian Reflectance Geometry	18
2.4	Principle of photometric stereo	20
2.5	Alternative method of gradient calculation	22
2.6	3D camera system set-up	25
2.7	Program back control interface	26
2.8	Program frontal interface	27
2.9	3D reconstruction pipeline	29
2.10	3D height map	29
2.11	Four objects with different shapes	31
2.12	3D reconstruction of different shape objects	31
2.13	3D Reconstruction process	32
2.14	3D reconstruction results	33
2.15	Mannequin head height map comparison	34
2.16	Human face 3D reconstruction	36
2.17	An image pair acquired from the textured 3D height map of an object with slightly different perspectives.	37
3.1	Computational single-pixel imaging approaches	42
3.2	DMD operating principle	44
3.3	Raster scanning	46
3.4	Multi-pixel scanning with random patterns	47
3.5	Hadamard matrices	51
3.6	Illustration of a structured illumination pattern and its inverse	52
3.7	3D single-pixel experimental setup	54
3.8	3D imaging system components	55
3.9	Signal acquisition procedure	56
3.10	Structured illumination	58
3.11	3D reconstructions comparison using structured illumination at dif- ferent image resolutions	59
3.12	Video sample of a 3D reconstructed moving object	60
3.13	Pattern rearrangement with compressive sensing algorithm	61
3.14	Comparison of 3D imaging using evolutionary compressed sensing	63

3.15	A one second video clip of a moving object at 64x64 pixel resolution with compressive sensing	64
3.16	A 10-second video clip of a moving skull at 128x128 pixel resolution with compressive sensing	66
3.17	A 10-second video clip of a moving Santa at 128x128 pixel resolution with compressive sensing	67
3.18	Elapsed-time comparison	68
4.1	2D single-pixel NIR imaging system setup	73
4.2	The spectral power distribution of infrared heat lamp R125	74
4.3	PDA20CS InGaAs detector spectral response curve	74
4.4	NIR outcomes comparison with and without polariser	75
4.5	NIR imaging of an apple with a set of bandpass filters	77
4.6	The trade-off between image resolution and signal-to-noise(SNR) in single-pixel imaging	78
4.7	The micro-scanning approach	79
4.8	Micro-scanning HR image pixel reconstruction	80
4.9	SNR comparison of images with and without using micro-scanning	81
4.10	NIR micro-scanning imaging with solar radiation	83
4.11	Solar Radiation Spectrum	84
4.12	Effect of atmospheric absorption of radiation on sea-level solar NIR imaging	85
4.13	Water detection with a 1400 ± 35 nm NIR bandpass filter	86
4.14	The spectral sensitivity range of Goldeye G-008 InGaAs camera	87
4.15	InGaAs NIR imaging comparison	88
5.1	Dual-band imaging modification	91
5.2	Experimental images of a piece of painting with different layers of acrylic colour	92
5.3	Single-pixel dual-band imaging results of a ripe avocado without averaging	93
5.4	NIR imaging results of an apple using different averaging values	95
5.5	NIR imaging results with and without randomisation	96
5.6	NIR imaging results before and after regularisation	98
A.1	Sea-level solar NIR imaging outcomes 01	105
A.2	Sea-level solar NIR imaging outcomes 02	106
A.3	Sea-level solar NIR imaging outcomes 03	107
B.1	NIR imaging results of avocados using different averaging values	109
B.2	NIR imaging results of bananas using different averaging values	110
B.3	NIR imaging results of a pomegranate using different averaging values	111

List of Tables

2.1	Deviations between measured values and true values	35
3.1	Relative RMS error comparison at 128x128 pixel resolution with five different compression ratios.	62

Abbreviations

1D	One dimensional
2D	Two dimensional
3D	Three dimensional
ADC	Analog-to-digital converter
ALP	Accessory light-modulator package
BW	Bandwidth
CC	Colour camera
CCD	Charge-coupled device
CMOS	Complementary metal-oxide-semiconductor
CS	Compressed sensing
DAQ	Data acquisition
DCT	Discrete cosine transform
DMD	Digital micromirror device
DLP	Digital light processing
DSLR	Digital single-lens reflex
EA-SLM	Electrically activated spatial light modulator
GPU	Graphics processing unit
HR	High-resolution
ISO	International standards organization
InGaAs	Indium-Gallium-Arsenide
IR	Infrared
LCOS	Liquid crystal on silicon
LED	Light-emitting diode
LR	Lower-resolution

MEMS	Micro-electro-mechanical-system
NIR	Near-infrared
NRMSE	Normalized root mean square error
OA-SLM	Optically activated spatial light modulator
OSC	Optical semiconductor chip
PD	Photodetector
PDA	Photodiode array
PMT	Photomultiplier tube
QVGA	Quarter video graphics array
RGB	Red green blue
RMS	Root mean square
RMSE	Root mean square error
SFS	Shape from shading
SLM	Spatial light modulator
SNR	Signal-to-noise ratio
SP	Short-pass
STL	Stereolithography
SWIR	Shortwave infrared
TV	Television
USB	Universal serial bus
VI	Virtual instrument
VIS	Visible

List of Publications

- [1] Y. Zhang, G. M. Gibson, R. Hay, R. W. Bowman, M. J. Padgett and M. P. Edgar. A fast 3D reconstruction system with a low-cost camera accessory. *Scientific reports* 5, 10909; doi: 10.1038/srep10909, 2015.
- [2] Y. Zhang, M. P. Edgar, B. Sun, N. Radwell, G. M. Gibson and M. J. Padgett. 3D single-pixel video. *Journal of Optics*, 18(3):035203, 2016.
- [3] Y. Zhang, M. P. Edgar, N. Radwell, G. M. Gibson and M. J. Padgett. Fast Compressive 3D Single-pixel Imaging. *3D Image Acquisition and Display: Technology, Perception and Applications*, TW3A-3, Optical Society of America, 2016.

Chapter 1

Introduction

1.1 Preamble

The human eye is a complex optical system which detects and interprets information from visible light in the surrounding environment to provide visual perception. A tremendous amount of effort in the past centuries has been aimed to replicate its functionality of perceiving colours and depth in intricate detail. Based on the structure characteristics of our eyes, a typical two-dimensional imaging system generally consists of a cornea-like lens to form an image by focusing the rays of light, a pupil-like aperture to restrict the amount of light that gets through, and a retina-like device which contains the photosensitive material.

Since the earliest known attempt of imaging began in 1021 with use of a small hole in one side of a dark room in Iraq [1], imaging system technology has gone through some momentous phases. The first phase of the technology began with camera obscura in 1550, which used a simple lens to focus light from an external scene onto a wall or a drawing board inside a light-proof box, tent or room so the image can be traced[2]. This technology migrated to chemical-based imaging in 1826 when the first successful sun-picture was made in France [3]. Silver chemistry, with which light is focused through a lens to reduce silver ions to silver metal to create an

image [4], is the most common form of chemical imaging technology. In 1884, film was developed to replace the photographic glass plate [5]. The next phase in the evolution of imaging system technology was digital imaging, when electronics had started moving away from vacuum tubes to small transistors since 1920s [6]. In 1969, the charge-coupled device (CCD) was conceived by Boyle and Smith to generate an image by counting the number of light photons falling onto a silicon chip [7]. This allowed optical image data to be transformed into an electronic format for digital processing. Complementary metal-oxide-semiconductor (CMOS) technology was later developed in 1970s as a competitive digital imaging solution to CCD [8]. Both technologies are based on arrays of pixel sensors to gather photons of light and convert them into a visible image, but with different schemes of capturing the electrical data [9]. Thanks to tremendous advances in semiconductor and electronic component technology, the processing power of computers has been growing at a blazing rapid rate. This has led imaging system technology to step into the current phase, computational imaging, in which various image processing approaches have been implemented to further enhance the capabilities of digital imaging systems and overcome their limitations.

In particular, three dimensional (3D) computational imaging, a technique to retrieve the depth information of a 3D object, has become available and popular in a wide range of disciplines, for instance, prototyping, object recognition, robot navigation, visualisation and animation [10]. However, the requirements of many 3D imaging systems, which involve high accuracy, fast estimation, low cost, portability and flexibility, are still quite challenging.

To respond to the challenge, my colleagues and I firstly developed a cost-efficient and portable 3D imaging system that allows fast and accurate estimation of object's depth information in a scene, as presented in chapter 2. Photometric stereo, a technique for measuring the surface normals of objects from a set of 2D images taken with a fixed viewpoint and different lighting conditions [11], was implemented in our system by utilising a consumer digital camera and a low-cost

accessory to provide a simply and efficient 3D reconstruction routine. A quantitative analysis was made on the acquired 3D data by comparing to the reference 3D information of objects, showing a good agreement.

Comparing to conventional digital imaging techniques, single-pixel imaging, which obtains spatial information of an object by sampling the scene through a set of microstructured illumination masks [12, 13], is becoming a remarkable alternative with the advent of single-photon-sensitive detectors, high-speed micro-electro-mechanical-systems(MEMS) devices, super-fast data acquisition(DAQ) equipments and high-performance computers. By combining photometric stereo and single-pixel imaging, we then derived a modified 3D imaging system that provides a perfect pixel-registration among 2D images captured with different shadows features, as shown in chapter 3. A compressive sensing algorithm was applied in the system to speed up the processing rate of continuous 3D reconstruction of a scene exhibiting dynamic behaviours.

One significant advantage of single-pixel imaging is that it allows one to use ultra-high sensitive light sensors to explore spectral bands beyond visible wavelength for imaging by avoiding conventional silicon-based sensor arrays in digital cameras which exhibit no sensitivity at wavelengths exceeding 1000nm [14]. In chapter 4, a near-infrared(NIR) single-pixel imaging system is demonstrated with a micro-scanning approach employed to improve the signal-to-noise ratio(SNR) when increasing the imaging resolution. Results of plant leaves and fruits imaged under the sea-level sun and a heat lamp at different NIR spectral bands with both our system and an expensive commercial InGaAs infrared camera are compared visually with both showing good capability for revealing water absorption underneath the surface. Chapter 5 interprets a fine-adjusted dual-band single-pixel imaging system with some other attempts carried out to improve its SNR performance.

In this chapter, I shall first give a brief review to some of the widely used conventional 3D optical imaging approaches, in which particular attention has been

placed on photometric stereo. Then the single-pixel imaging technique that can acquire images without the need of a pixelated sensor will be introduced, along with some previous research which inspired the imaging systems that are presented in this thesis.

1.2 Conventional 3D optical imaging techniques

The depth information of a three dimensional scene always get lost when capturing from a conventional camera perspective due to its two dimensional receptive field. Retrieval of this lost information has been particularly crucial in recent decades on account of a rapidly growing demand for 3D content in the field of robotics, medicine, entertainment, etc. Based on this background, various methods have been developed for 3D shape capturing. Here we will review some of the commonly used conventional 3D optical techniques with non-contact measurements.

1.2.1 Interferometry

On the smallest scale, interferometry is an extensively utilized optical approach for surface depth measurement on 3D structures. The principle of interferometry is based on the interference phenomenon that two waves (usually light, radio or sound waves) with the same intensity and frequency traversing the same space will be superimposed to form a resultant wave with an amplitude equalling to the sum of the amplitudes that are produced by these two waves separately [15, 16]. To put it concretely, if the two merging waves are in phase, the amplitude of the new wave will increase significantly (constructive interference), whilst if they are out of phase, the amplitude will be decreased (destructive interference).

When the two waves are spatially extended, we can observe the variations over a surface area with alternating regions of bright and dark bands, known as interference fringes. Phase differences between two waves occur when travelling different optical path lengths caused by the shape differences between a test surface and a reference one. By analysing the interference patterns, the path length differences can be determined, by which the surface relief of test object can be computationally extracted. Based on the wavelength of the incident waves, interferometry can provide a height measurement of 3D structures with surface profiles varying from a few nanometres to several centimetres [17, 18]. As a non-contact measuring method, it avoids the risk of damage or deformation of the test surface. With the employment of lasers and optical fibres [19], interferometry is able to provide dynamic surface measurements with high resolution and outstanding accuracy on the order of a small fraction of the wavelength of light used.

1.2.2 Structured Illumination

Structured illumination is a reliable contactless technique, initially developed for microscopy applications, which enables 3D feature extraction of an object using full-field illumination optics. Having a calibrated projector-camera pair with a known distance, a set of mathematically constructed light patterns are sequentially illuminated onto an object and a set of images are synchronously captured by the camera (or cameras). If the object surface is planar without any topographic variation, the pattern shown in the corresponding image will be the same as the projected light pattern. However, if the surface of the object is non-planar, the pattern in the captured image will appear to be distorted with locally varying displacements caused by the geometry of the object [20]. Based on the displacement of projected patterns, it enables an exact retrieval of the 3D coordinates of each point on the object surface following the principle of geometric triangulation.

The structured light patterns can be generated by either a laser with the appropriate optics or a digital projector manipulated by a computer. As the intensity value of each light pattern is generally not required and only its geometric displacement is used in this approach, the surface measurement is largely independent of the object's optical properties, such as the colour and the reflectance. There is another 3D measurement approach with structured illumination based on the detected light intensities, which will be covered later on in section 1.3. Depending on the optical resolution of projected patterns, structured illumination can provide 3D surface profile measurements with an error of millimetres [21–23].

1.2.3 Stereo vision

In addition to the aforementioned methods, stereo vision, also known as stereopsis or stereoscopic vision, is another widely performed 3D approach that deduces the spatial shape and position of an object through parallax between the corresponding pixels from two (or more) images of the object as observed from slightly different viewpoints [24, 25]. Due to noise and image structure, the most difficult and time-consuming procedure in a stereo vision system is the registration of the images, or more precisely, the identification of corresponding pixels. Once the correspondence for each pixel is determined, the object's 3D information can be calculated directly using the principle of triangulation.

To simplify the correspondence search problem, a geometric limitation known as the epipolar constraint is generally applied in most stereo systems [26, 27], which narrows the search space for each pixel in one image from two-dimensions into a one-dimensional line (epipolar line) in the other one by ensuring the optical axes of the cameras are parallel and the images are captured at the same height of focal points with the same focal length. With this standard stereo geometry, those lines follow along the horizontal scan lines of the images, where the matching process can be handled efficiently. For stereo systems with an arbitrary camera

setup, a transformation process known as image rectification can be employed to perfect camera alignment [28–30]. There have been numerous algorithms with further constraints utilised for tackling the correspondence problem in stereo vision [31–33], the outcomes of which are different in terms of accuracy, robustness and computational requirements.

1.2.4 Photometric stereo

Among all kinds of cues that can be used to reconstruct 3D geometry, it is possible to infer shape from the shading on a surface. This is based on the fact that following the surface normal changes across an object, the apparent brightness varies on the basis of the angle between the local surface orientation and the incident illumination. The technique that extracts 3D information of an object from the intensity variation is known as shape from shading [34]. Most algorithms for this technique [35–37] assume that the object surface is of a uniform albedo and reflectance and rely on large data training to recover the 3D shape from a single 2D image.

To make this approach more powerful and reliable, multiple light sources are used separately to obtain different images of an object, with the same position in each image corresponding to the same surface point. This approach, pioneered by Woodham in 1980 [11], is known as photometric stereo, in which the role of light sources analogous to the cameras placed at different locations in traditional stereo [38]. It requires no hypothesis of the surface smoothness and enables satisfactory accuracy with reasonable computational cost.

There has been a lot of research effort put forth in improving the performance of 3D reconstruction using photometric stereo since it was proposed. Technically it is adequate to use two light sources for photometric stereo with known surface albedo, but using more than three can minimise the shadow effect in sampled images [39]. For general diffuse reflectance, Okatani and Deguchi [40] pointed out that there

exists a set of the surface normals for which the relation between the surface normal and the orientation of the 3-vector formed by the image brightness triplet is guaranteed to be one-to-one. With images of an object obtained under unknown lighting conditions, Basri, Jacobs and Kemelmacher [41] recovered the surface normals based on the finding that general lighting conditions for Lambertian objects can be represented using low order spherical harmonics. To enhance the resolution of photometric stereo, Tan et al. [42] demonstrated a method to recover the sub-pixel surface geometry by acquiring the distribution of surface normals along with the surface convexity of each pixel and then spatially arranging the normals among pixels in accordance with the consistency and simplicity constraints on the surface structure. A self-calibrating approach for photometric stereo was introduced by Shi et al. [43] which analyses colour and intensity profiles in the RGB and irradiance-time domains to automatically resolve both the camera's radiometric response function and the unknown lighting directions and intensities. Hansen et al. [44] presented an photometric stereo algorithm that reduces the effects of shadows in the 3D reconstruction without foreknowing the precise shadow boundaries in images by making use of both visible and near-infrared light sources. Wu et al. [45] demonstrated a robust photometric stereo method which employs advanced convex optimization techniques to handle shadows and specularities in the images and provide highly accurate estimates of surface orientations. By making use of a reference model to estimate the illumination parameter of each image, Sun et al. [46] proposed a fast 3D face reconstruction algorithm with a combination of classical photometric stereo and lighting calibration.

1.3 Single-pixel Imaging

In conventional imaging systems, a sensor array is generally required to capture a 2D image of a scene by collecting light focused through an optical lens. An alternative approach, known as single-pixel imaging [12, 47], replaces the multi-pixel

sensor with a single photosensitive detector as the imaging device by utilising a spatial light modulator (SLM) to provide either time-varying, structured detection of the scene or time-varying, structured illumination onto the scene. The final image is retrieved by correlating the recorded intensity signal with the corresponding structured illumination. This technique is closely related to ghost imaging [48, 49], which yields an image of an object by combining information from two detectors: a conventional detector with a pixlated sensor that records a light field that has never interacted with the object to be imaged, and a single-pixel detector that collects a field that has interacted with the object.

1.3.1 Single-pixel imaging theory

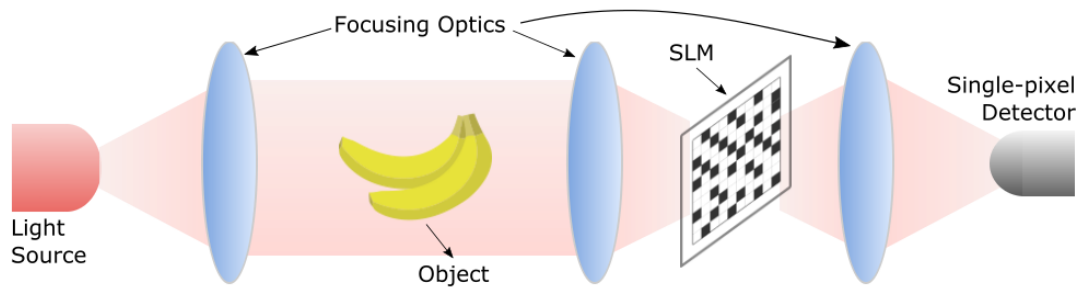


FIGURE 1.1: Schematic representation of a single-pixel imaging system. By using focusing optics, light is collimated to illuminate an object which is focused onto a SLM. By changing the masks or coded apertures on a electronically controlled SLM, the intensity of transmitted light will be different each time. The partially transmitted light was then focused onto a single-pixel detector which measures the aggregate power of the light. By analysing the correlation between measured signals and the corresponding masks, a 2D image of the object is then reconstructed.

A typical single-pixel imaging system consists of a light source, focusing optics, a SLM and a single-pixel detector as illustrated in Fig. 1.1. Light scattered from an illuminated object is focused by a lens onto a SLM, which selectively passes light onto the single-pixel detector. With a known set of binary-coded patterns displayed on the SLM and the measured signal from the detector for each pattern, an image of the object can be reconstructed. For a N -pixel image, single-pixel

measurements for a sequence of M N -element light patterns can be presented in a matrix form as follow:

$$I = \phi X \quad (1.1)$$

where I is a M -element column vector, ϕ is a $M \times N$ matrix representing the pattern set, and X is a N -element column vector representing the pixels in the image. The size of ϕ varies based on the type of patterns displayed on the SLM [12]. For a well-defined measurement matrix ϕ [50], image X can be recovered using the inverse matrix equation:

$$X = \phi^{-1} I \quad (1.2)$$

1.3.2 Computational imaging systems with single-pixel detectors

Compared to conventional imaging techniques using sensor arrays, single-pixel imaging is a remarkable alternative with several advantages. The high sensitivity of single-pixel detectors enables this technique to be employed in low-light-level imaging applications [51, 52]. It also provides a simple and reliable way of measuring the spatial distribution of multiple physical dimensions of a scene. Ref.[53] proposed an optical system that performs polarimetric imaging with a single-pixel detector. Ref.[54] demonstrated a full-colour imaging system using three spectrally-filtered single-pixel detectors. Ref.[55] introduced the design of a time-of-flight range system for acquiring 2D depth maps of piecewise-planar scenes using a single time-resolved photodetector. Ref.[56] employed several single-pixel detectors in different locations to retrieve the 3D information of an object based on shape from shading. Ref.[57] presented a framework for digital holography by combining phase-shifting interferometry with single-pixel imaging. Furthermore,

single-pixel detectors can offer a significantly broader spectral range than conventional silicon-based cameras, extending the imaging technique to different spectral regions [58].

Despite those advantages, large amounts of illumination patterns are required to reconstruct high-quality images with single-pixel imaging using traditional sampling methods. These methods fall under the classical Nyquist sampling theorem which states that in order to accurately reproduce a signal it should be periodically sampled at least twice the highest frequency contained in the signal [59]. The full measurement under this sampling scheme is memory demanding and requires long acquisition time [60]. One way to address this issue is to increase the pattern switching speed on the SLM. Significant amount of effort has been invested in the area of SLMs, including ones based on micro-electromechanical systems (MEMS) [61], digital micromirror devices (DMD) [62], and liquid crystal on silicon (LCOS) [63]. Single-pixel imaging speed can also be increased by applying a novel sampling theorem called the compressed sensing (CS), which exploits the fact that natural images tend to be sparse in the discrete cosine transform (DCT) domain, in other words, many images can be well-approximated by only a small number of non-zero expansion coefficients in terms of a suitable basis [12, 64–66]. By performing single-pixel imaging with compressive sensing, images can be retrieved with far fewer measurements than that established by the Nyquist limit.

1.4 Conclusion

In this chapter I have presented a brief historical overview of the development of imaging system technology and the principle of some widely used conventional 3D optical imaging technologies including interferometry, structured illumination, stereo vision and photometric stereo. Among those imaging technologies, photometric stereo was applied in our fast 3D reconstruction system due to the advantage that it can provide a simple and efficient reconstruction routine with a

low-cost hardware. Chapter 2 will summarise the experimental setup and methodology. I will explain the mathematical model and the computational program that we have developed for building up this 3D camera system.

Besides those conventional imaging methods, I have also reviewed an alternative imaging approach, known as single-pixel imaging, which uses a single photosensitive detector as the imaging device instead of a multi-pixel sensor in a conventional camera. In chapter 3 I will demonstrate a modified 3D imaging system by combining photometric stereo with single-pixel imaging to provide perfect pixel-registration and a capability of 3D imaging of scenes exhibiting dynamic behaviour in real-time. A big advantage of using single-pixel imaging technique is that it can be readily extended to other wavelengths with the use of exotic detectors that would be impossible or prohibitively expensive in a conventional digital camera. Chapter 4 will present a near-infrared (NIR) single-pixel imaging system with an indium-gallium-arsenide (InGaAs) amplified photodetector. A micro-scanning method and Hadamard matrices for improving the image signal-to-noise ratio (SNR) will be introduced. The results of imaging different objects with a set of NIR bandpass filters using our system will be given to indicate its ability of detecting water absorption underneath the surface of plant leaves and fruits. In chapter 5 I will demonstrate a visible (VIS) and NIR dual-band single-pixel imaging system with three optimisation approaches that have been applied or tested for further improving the NIR image quality. Chapter 6 will summarise the results of this thesis and give a brief description of the future work that could be undertaken.

Chapter 2

A Fast 3D Camera System using Photometric Stereo

2.1 Introduction

Three-dimensional (3D) imaging is a heavily explored research field that supports applications in a wide range of disciplines, such as face recognition and robot navigation. Various approaches have been developed for performing 3D imaging, each with different strengths and drawbacks depending on the specific application. This chapter will focus on the use of photometric stereo, an imaging technique that allows depth and surface orientation to be estimated by using multiple 2D images of a static object obtained from a fixed viewing point, under different illumination directions. Compared to other 3D approaches such as geometry modelling and 3D scanning, a big advantage of using photometric stereo is that it has a simple and efficient reconstruction routine. In this work, I will demonstrate a 3D camera system with a low-cost accessory allowing fast 3D reconstruction of static objects using photometric stereo. The results of imaging a selection of objects with varying geometric complexity using our camera system will be presented.

2.1.1 Contributions

The idea of utilising multiple light sources as a camera accessory to recover 3D images was the result of discussions among myself, Miles Padgett, Matt Edgar and Graham Gibson. The design and development of the lighting system was undertaken by myself with the help of Graham Gibson and Matt Edgar. The design of the fast surface gradient calculation method and subsequent 3D image reconstruction was fulfilled by myself. Many of the test objects used for this experiment were designed and 3D printed by myself. Qualitative and quantitative assessment of the 3D imaging system was undertaken by myself with support from Matt Edgar.

2.2 Analysis of surface reflection property

Given a light source, an object, and a detector, a reflectance model can be used to describe the intensity and spectral components of the reflected light from the object surface to the detector. Both the intensity and spectral composition of the reflected light are determined by the attributes of the light source and the properties of the object surface. By utilising multiple light sources with corresponding reflection models, photometric stereo enables the estimation of the three-dimensional shape of an object from 2D pictures. In this chapter, I shall present a fast 3D reconstruction system using a commercial digital single-lens reflex (DSLR) camera and a customised low-cost lighting accessory.

2.2.1 Specular and diffuse reflection

For a non-luminous object being imaged, the detected intensity can be the result of different types of surface reflection: specular and diffuse. Specular reflection is the type of reflection that occurs on smooth surfaces such as mirrors or a calm

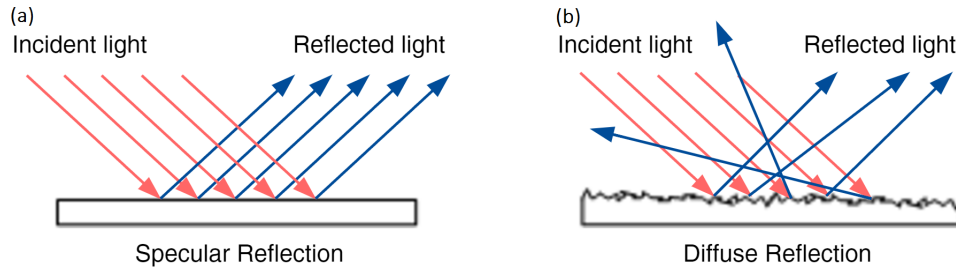


FIGURE 2.1: Specular and diffuse reflection from an object surface. (a) Light reflects from a smooth and flat surface will have the same angle as they arrive. (b) Light reflects from a rough surface will scatter in all directions.

body of water, whereas diffuse reflection is caused by rough surfaces, such as paper or fabric. Microscopically the smoothness or roughness of an surface has an enormous influence upon the subsequent reflection of a beam of light. Figure 2.1 exhibits the resultant reflection of two beams of light incident upon a smooth and a microscopically rough surface. A beam of light can be described as a bundle of individual light rays travelling parallel to each other. Figure 2.2 gives an example of specular and diffuse reflectance for a spherical object under identical lighting conditions.

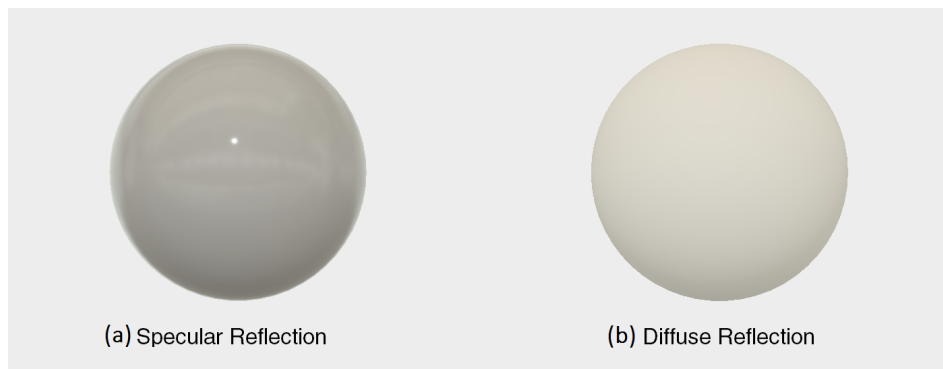


FIGURE 2.2: An example of specular and diffuse reflection for a spherical object under identical lighting conditions. (a) It shows the image result of a smooth surface with specular reflection. (b) It shows the image result of a rough surface with diffuse reflection.

In particular, the diffuse component provides intensity information about every surface on the object, whereas the specular component occurs on comparatively

fewer surface where the normal subtends an equal angle between the illumination and the detection. For this application, we want to reconstruct 3D shape from as much of the object as possible, therefore we are generally interested in the diffuse component of the measured reflected intensity. This can be achieved experimentally by choosing objects that exhibit diffuse reflection properties, or by isolating the specular component, for example by using crossed polarisers in the illumination and detection.

2.2.2 Lambertian reflectance model

A light source reflected from an object surface in the direction of a detector results in an image of the object surface. The intensity of each pixel in the image is strongly related to the reflectance properties of the corresponding point on the surface. In the field of computer vision, various reflectance models have been developed to enable the rendering of a 3D surface in a 2D space. In general, the appearance of a diffuse object with a specular varying reflection may be modelled as [67]:

$$I_P = \sum_{i=1}^k \rho_{i,P}^t [f_i(n_P, L_P, v)], \quad (2.1)$$

where I_P is the pixel intensity at point p ;

k is a fixed value of a linear combination of k basis materials;

ρ^t is a reflection coefficient that varies on the surface;

f_i is any reflectance map as a function of the viewing direction v ;

n_P is the surface normal at that point;

L_P is the incident illumination field.

One of the most basic reflection models is the Lambertian model, which simplifies the reflectance property of an object surface and is often used as a first order approximation. Other non-Lambertian models are significantly more complicated and far harder to cope with.

A Lambertian surface is defined as an ideal diffusely reflecting surface, the concept of which was first introduced by Lambert [68]. The appearance of a Lambertian surface on a detector will look equally bright from any viewing direction. In other words, the luminance of such a surface is isotropic, and the reflected intensity is independent of the viewing direction but dependent on the light source orientation. A linear Lambertian reflectance model can be represented as follows:

$$I(\vec{i}) = \lambda(\vec{i}) \vec{n}(\vec{i}) \cdot \vec{l}, \quad (2.2)$$

where $I(\vec{i})$ is the image;

$\lambda(\vec{i})$ is the albedo on the object surface;

$\vec{n}(\vec{i})$ is the object's surface normal;

\vec{l} is the light source and $\vec{l}/|\vec{l}|$ is the light source direction.

This model is also known as Lambert's cosine law because the image depends on the cosine of the angle θ between the direction of the incident light and the surface normal (see Figure 2.3). To use this model it assumes the following:

- the object surface is a Lambertian surface and its reflectance function is uniform;
- based on the object size, the distance between the detector and the object is long enough to consider the image system as an orthographic projection;
- both attached and cast shadows are ignored in the model;
- only the incident light on the front of the object surface is considered here, the incident light on the back of the surface is blocked by the object and is hence not considered;
- the illumination from the light source is assumed as incessant and uniform over the front of the surface.

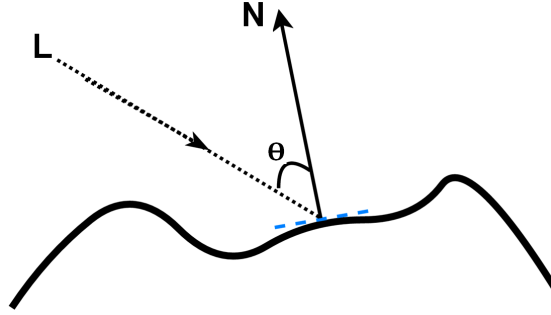


FIGURE 2.3: Lambertian Reflectance Geometry. θ is the angle between the incident light source \vec{L} and the surface normal \vec{N} .

The Lambertian reflectance model has been widely used to describe diffuse reflection over various object surfaces, for instance, cotton cloth, plastic toys, matte painting and paper, even human faces [69]. This model has also been employed in our 3D camera system.

2.2.3 Shape from shading

Giving that we are assuming a Lambertian reflectance model, the intensity of each point in a 2D image can be determined by geometrical features of the object surface in the scene and the known incident light source. The approach that retrieves 3D information from this 2D intensity image is called shape from shading (SFS), first introduced by Horn [70–75]. To use this approach with a single image, the reflectance properties of the object surface is assumed. The brightness of the object surface is associated with the specified orientation at each point (x, y, z) in gradient space. A reflectance map, $R(p, q)$, is used to describe this relationship, where the surface gradient is given conveniently by the first partial derivatives of z with respect to the corresponding x and y :

$$(p, q) = \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y} \right), \quad (2.3)$$

The reflectance map, defined in gradient space, provides an uncomplicated representation of the inherent constraint in single image brightness measurement. With a known brightness map $I(x, y)$ on the surface of an object, It restricts the possibility of what the surface orientation at certain point of the object might be. The connection between the reflectance map $R(p, q)$ and the brightness map $I(x, y)$ can be represented by the image-irradiance equation:

$$R(p, q) = I(x, y), \quad (2.4)$$

This equation has two unknown variables (p, q) , which results in the non-uniqueness of the solution. In other words, a contour of $R(p, q)$ can have the same brightness map. More constraints are needed. One method to solve this problem is to start at a single point in the image where the surface orientation is known, then grow a solution by moving to its nearest neighbour point along the contour. An extension of SFS approach is to employ multiple images rather than one image.

2.3 Photometric stereo model

The appearance of an object in an image results from the effects of illumination, object orientation, object shape and its reflectance. With a static object, the corresponding surface orientation can be calculated by analysing the object images under different illumination directions, as shown in Figure 2.4 [76]. Photometric stereo, first introduced by Woodham [11], allows depth and surface orientation to be estimated from multiple images of a static object taken from the same viewpoint, but under different illumination directions. Fundamentally, photometric stereo, which is simple and concise for Lambertian surfaces, enables 3D reconstruction of an object by analysing differences of the intensity distribution in images that are captured from at least three different illumination directions [77].

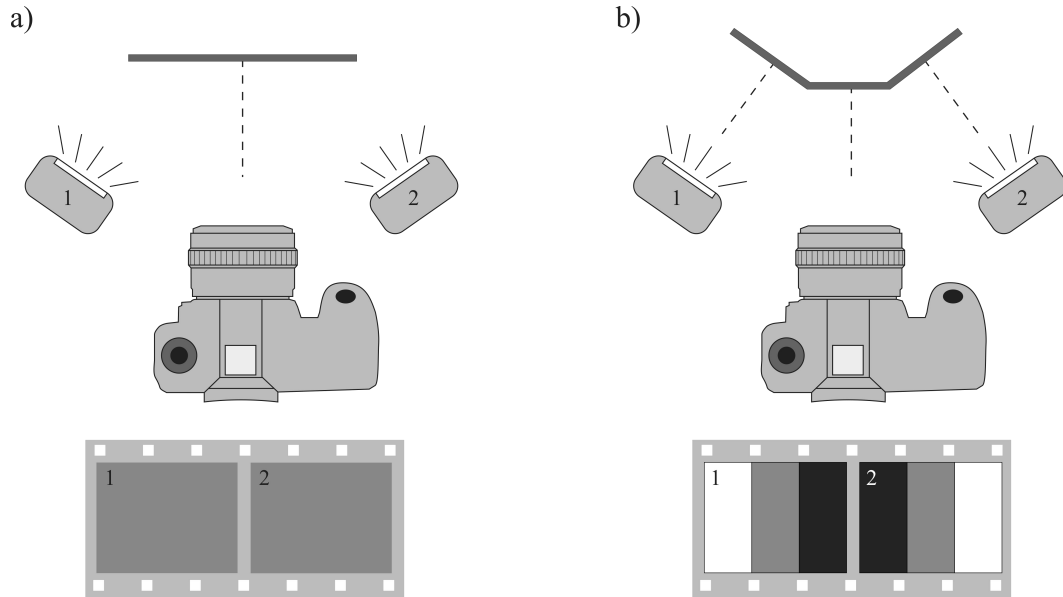


FIGURE 2.4: Principle of photometric stereo. a) If the surface of an object is flat and normal to the camera lens, images captured under axis lighting directions 1 and 2 will show the same light intensity of the object. b) If the surface of an object is uneven, images captured under axis lighting directions 1 and 2 will show different light intensities.

This approach requires some control of the lighting environment, without position changes of either object nor camera. Compared to SFS with single image, photometric stereo does not require any assumption of the surface smoothness. Each image defines a unique set of possible surface orientations along with its own reflectance map. Theoretically, three illumination directions are sufficient to obtain the surface normals, however to ensure that at least three intensity values are measured at any pixel in all acquired images and to minimise the shadow effect at certain regions, our 3D imaging system has utilized four light sources at different illumination directions.

2.3.1 Classic mathematical calculation model

According to Eq.2.2, we can transfer the Lambertian surface model into a matrix format. The intensity of a pixel in an image can be expressed as follow,

$$I_n(x, y) = \lambda (\mathbf{L}_n \cdot \mathbf{N}), \quad (2.5)$$

where $I_n(x, y)$ is the image intensity at the point (x, y) ;

$\lambda(\vec{i})$ is the surface albedo at the given point (x, y) ;

\mathbf{L}_n is the unit illumination vector pointing from the surface to the light

source, $\mathbf{L}_n = (l_x, l_y, l_z)_n$;

\mathbf{N} is the surface normal unit vector of the object at the given point (x, y) .

In our system four light sources with illumination vectors $\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3, \mathbf{L}_4$ are utilised to determine the surface orientation. Thus the equation 2.5 can be written in matrix notation as

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \lambda \cdot \begin{bmatrix} l_{x1} & l_{y1} & l_{z1} \\ l_{x2} & l_{y2} & l_{z2} \\ l_{x3} & l_{y3} & l_{z3} \\ l_{x4} & l_{y4} & l_{z4} \end{bmatrix} \cdot \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix}, \quad (2.6)$$

which equals to $\mathbf{I} = \lambda \cdot \mathbf{L} \cdot \mathbf{N}$. Given that the illumination vectors $\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3, \mathbf{L}_4$ are non-coplanar, \mathbf{L} is non-singular thus its inverse matrix \mathbf{L}^{-1} exists, and for any point (x, y) the unit surface normal can be calculated via

$$\mathbf{N} = (\mathbf{1}/\lambda) \mathbf{L}^{-1} \cdot \mathbf{I}, \quad (2.7)$$

$$\mathbf{S} = \mathbf{L}^{-1} \cdot \mathbf{I} = \lambda \cdot \mathbf{N}, \quad (2.8)$$

where \mathbf{S} can be used to calculate the surface gradient components (p, q) . From the surface normals calculated at each pixel, it is possible to determine the gradient between adjacent pixels from which we obtain the surface geometry by integration.

2.3.2 Fast surface gradient calculation model

Assuming that the detector and the light sources were relatively far away from the object surface comparing to the size of the object, I proposed an improved model to calculate the surface gradient (see Figure 2.5). The detector(camera) is set as the origin of coordinates and four same-type light sources are fixed surrounding it with the same distance, two of which are on x-axis and the other two are on y-axis.

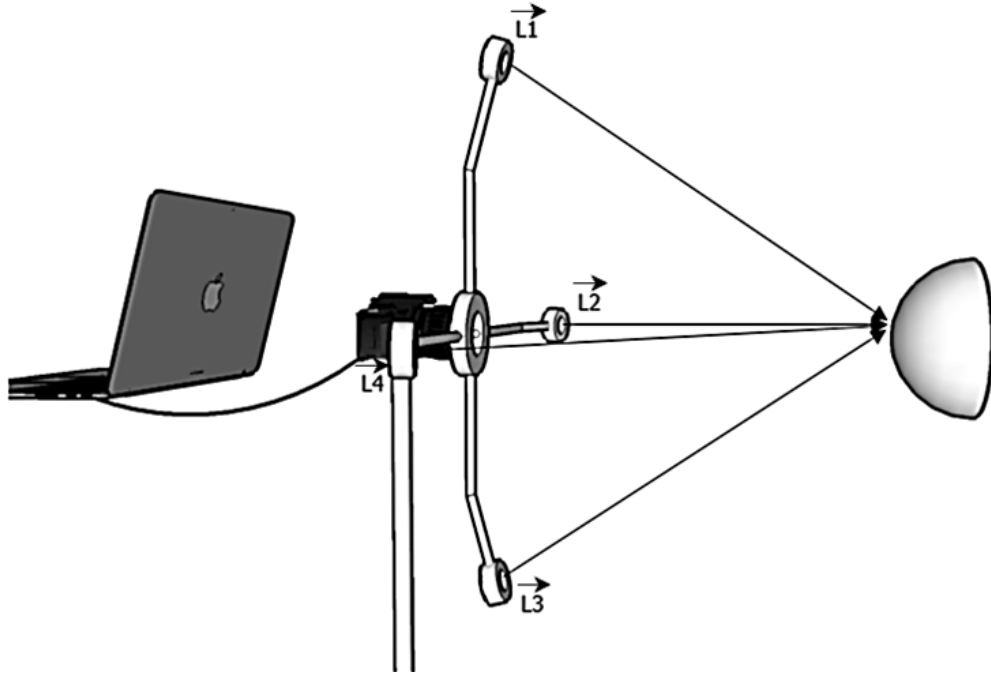


FIGURE 2.5: Alternative method of gradient calculation. Four light sources $\vec{L}_1, \vec{L}_2, \vec{L}_3, \vec{L}_4$ are used to provide illumination consecutively while the object images are collected via the detector(camera). The surface normal of the object can be calculated quickly by comparing the image intensities between \vec{L}_1 and \vec{L}_3, \vec{L}_2 and \vec{L}_4 .

For our lighting system we have four sources located as follows:

$$\vec{L}_1(0, y_L, z_L),$$

$$\vec{L}_2(-x_L, 0, z_L),$$

$$\vec{L}_3(0, -y_L, z_L),$$

$$\vec{L}_4(x_L, 0, z_L),$$

and the surface normal at any point is described by: $\vec{N}(dx, dy, dz)$. Therefore

we can write the intensity measured by the fixed camera for any surface point, resulting from each lighting vector as:

$$\begin{cases} I_{L_1} = \vec{L}_1 \cdot \vec{N} = 0 * dx + y_L * dy + z_L * dz \\ I_{L_2} = \vec{L}_2 \cdot \vec{N} = (-x_L) * dx + 0 * dy + z_L * dz \\ I_{L_3} = \vec{L}_3 \cdot \vec{N} = 0 * dx + (-y_L) * dy + z_L * dz \\ I_{L_4} = \vec{L}_4 \cdot \vec{N} = x_L * dx + 0 * dy + z_L * dz \end{cases} . \quad (2.9)$$

If we compare the image intensities between \vec{L}_1 (top) and \vec{L}_3 (bottom), \vec{L}_2 (left) and \vec{L}_4 (right),

$$\begin{cases} I_{L_1} - I_{L_3} = (0 * dx + y_L * dy + z_L * dz) - (0 * dx + (-y_L) * dy + z_L * dz) \\ \quad = 2 y_L * dy \\ I_{L_1} + I_{L_3} = (0 * dx + y_L * dy + z_L * dz) + (0 * dx + (-y_L) * dy + z_L * dz) \\ \quad = 2 z_L * dz \\ I_{L_4} - I_{L_2} = (x_L * dx + 0 * dy + z_L * dz) - ((-x_L) * dx + 0 * dy + z_L * dz) \\ \quad = 2 x_L * dx \\ I_{L_4} + I_{L_2} = (x_L * dx + 0 * dy + z_L * dz) + ((-x_L) * dx + 0 * dy + z_L * dz) \\ \quad = 2 z_L * dz \end{cases} ,$$

and then

$$\begin{cases} \frac{I_{L_1} - I_{L_3}}{I_{L_1} + I_{L_3}} = \frac{y_L * dy}{z_L * dz} = \frac{dy}{dz} * \frac{y_L}{z_L} \\ \frac{I_{L_4} - I_{L_2}}{I_{L_4} + I_{L_2}} = \frac{x_L * dx}{z_L * dz} = \frac{dx}{dz} * \frac{x_L}{z_L} \end{cases} .$$

Since x_L , y_L , z_L and I_{L_1} , I_{L_2} , I_{L_3} , I_{L_4} are known,

$$\begin{cases} \frac{dx}{dz} = \frac{I_{L_4} - I_{L_2}}{I_{L_4} + I_{L_2}} * \frac{z_L}{x_L} \\ \frac{dy}{dz} = \frac{I_{L_1} - I_{L_3}}{I_{L_1} + I_{L_3}} * \frac{z_L}{y_L} \end{cases} .$$

When $\vec{L}_1, \vec{L}_2, \vec{L}_3, \vec{L}_4$ are the unit normal vectors,

$$\vec{L}_1 (0, 1, 1),$$

$$\vec{L}_2 (-1, 0, 1),$$

$$\vec{L}_3 (0, -1, 1),$$

$$\vec{L}_4 (1, 0, 1),$$

and then

$$\begin{cases} \frac{dz}{dx} = \frac{I_{L_4} + I_{L_2}}{I_{L_4} - I_{L_2}} \\ \frac{dz}{dy} = \frac{I_{L_1} + I_{L_3}}{I_{L_1} - I_{L_3}} \end{cases} . \quad (2.10)$$

This simplifies the complexity of calculation for the surface gradient of the object. In this model, instead of calculating the surface gradient of the object at each point in the images one by one with Eq.2.6, we can quickly get a full surface gradient map of the object by comparing the intensities of all the points in the images at the same time with Eq.2.10. Both the classical and this fast calculation model have been applied on our camera system separately, and they have worked as well as each other when dealing with relatively-low resolution images due to our high-performance system hardware. When analysing with super-high resolution images, this new gradient calculation model has showed an improvement of the image process efficiency.

2.4 Experimental set-up

The 3D camera system was composed of a commercial DSLR camera (Canon EOS 5D MarkII), four white LEDs (Luxeon Rebel) surrounding a 50mm focal camera lens, fixed at a distance of 330 mm through aluminium spokes, a controller board (Arduino Uno) to enable USB control of the illumination time and direction, and a laptop running our program (on LabVIEW 2010) to communicate with the controller board and obtain real-time 2D images captured by the camera (see Fig. 2.6).

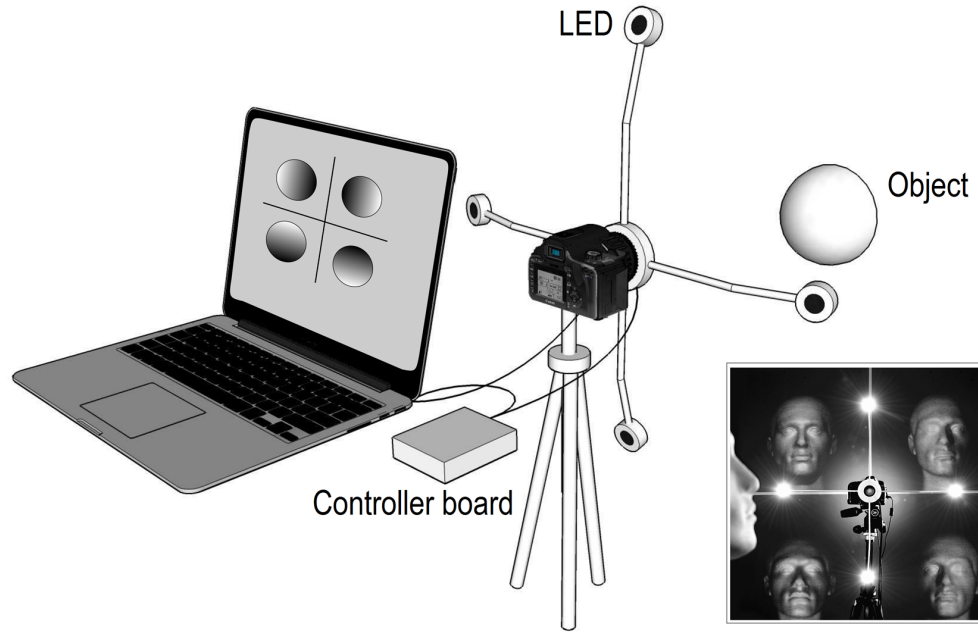


FIGURE 2.6: 3D camera system set-up. Schematic diagram of system consisting of four LEDs fixed around the camera lens and linked to a electronic controller board. Each object was placed normal to the camera lens. Both the camera and controller board were wired to a laptop controlled by our program.

Based on the camera type and software version, certain drivers and virtual instruments (VIs) were installed to connect and control the camera through the program. The camera was set with manual focus mode and controllable shutter speed, aperture value and film speed(ISO) to adapt different scenes. To make sure the camera captures each image under the corresponding light in rapid sequence, the Arduino board was set to keep each light on for 100ms in the sequence: top, left, bottom, right. The lights were set all on at the start for the system to target the object in the scene. Once the image capture signal was triggered, three of these four lights will be turned off based on the direction setting and the images were collected. Fig. 2.7 shows the control back panel of the 2D image collection section. Starting from the top left in this figure, the Arduino resource VI is the controller of the four LEDs to make sure each time only one light is switched on for a short term. The camera components are set to grab an image after each light switch and send it to the Images3D Info VI for calculating the gradient map.

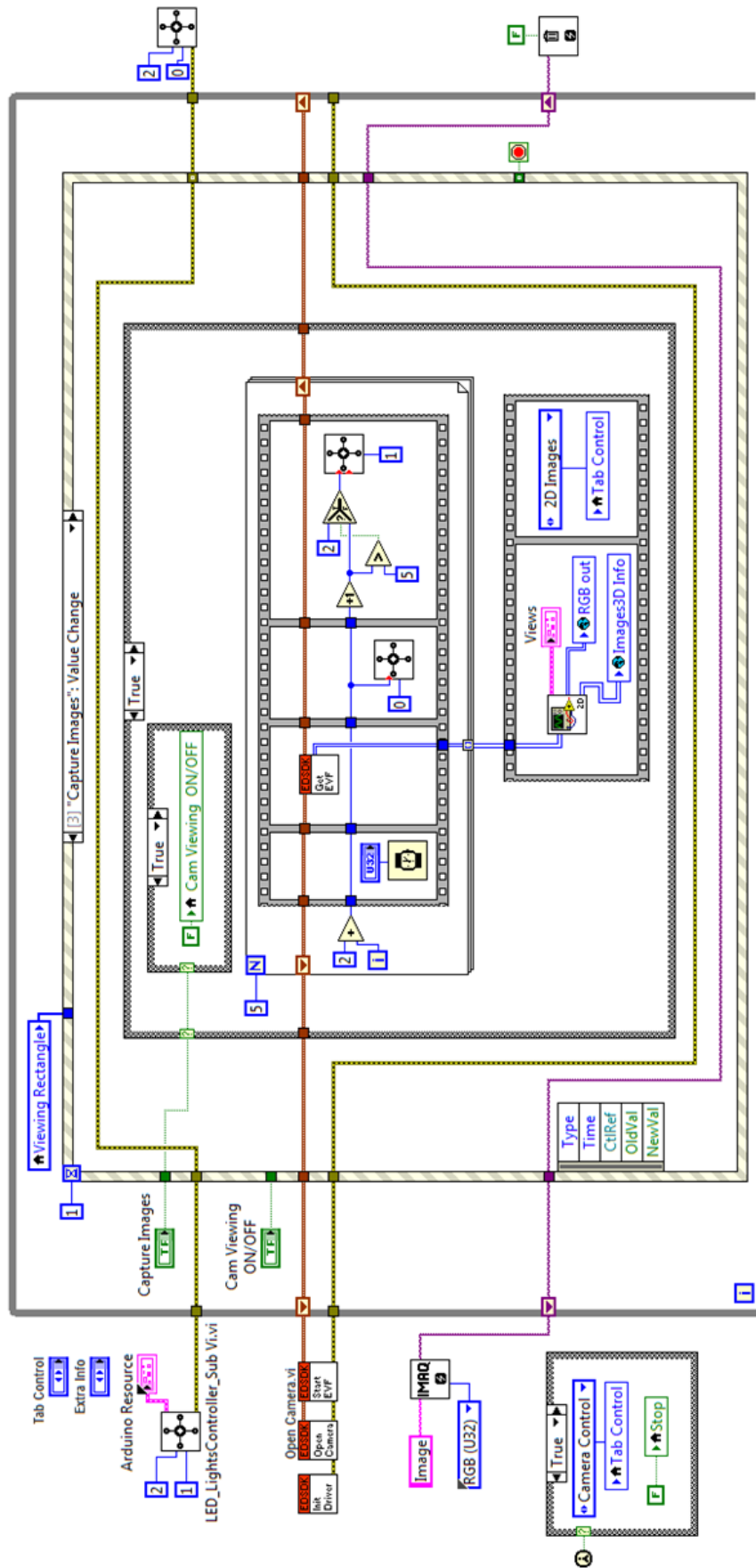


FIGURE 2.7: The control back panel of 2D image collection section for the 3D imaging program. An image is captured successively by the camera when each light is toggled on through the Arduino resource. These images are collected in the Image3D Info VI for calculating the gradient map.

2.5 Image acquisition

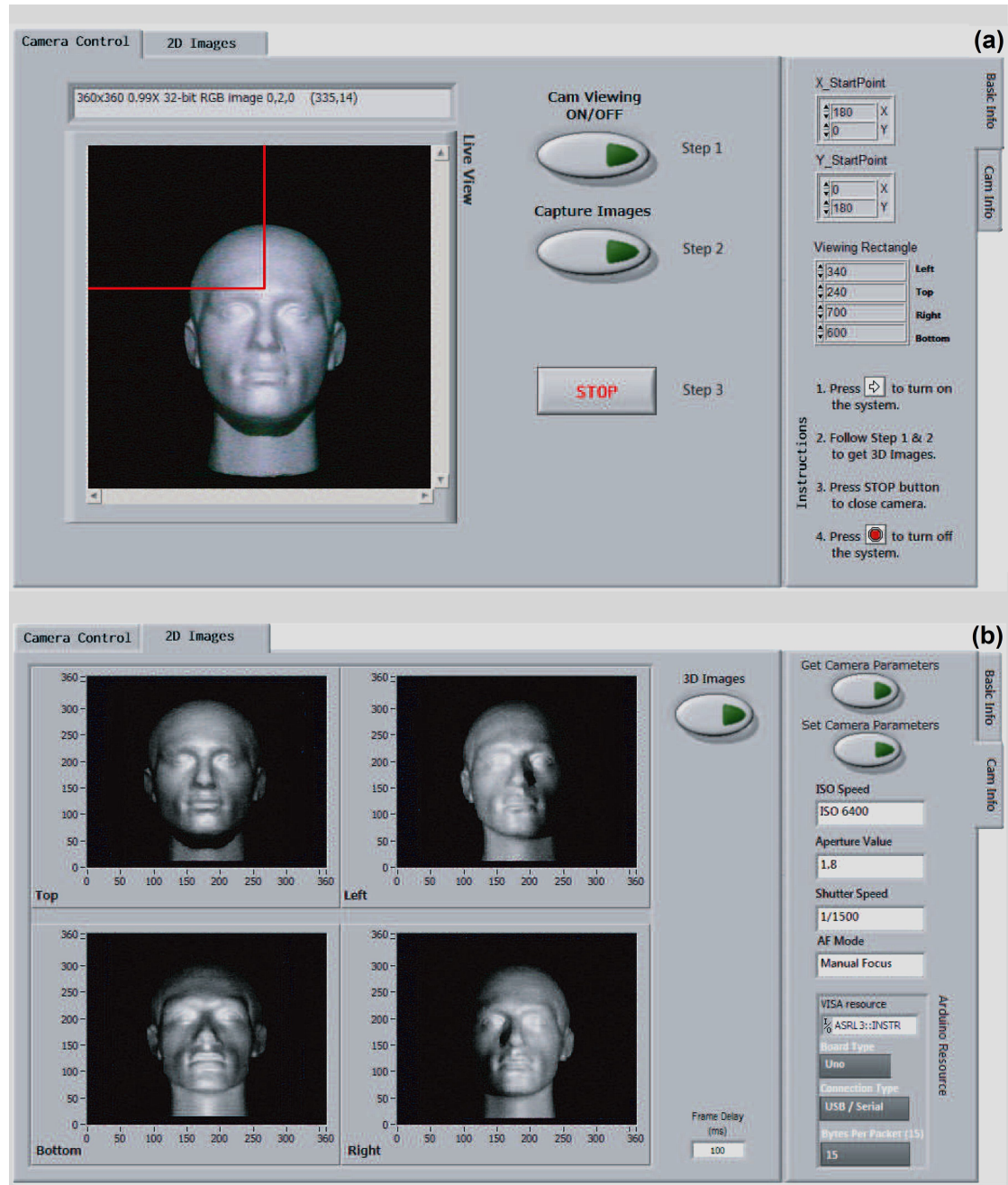


FIGURE 2.8: The control front panel of the program. (a) The Camera Control tab displays the live view from the camera with a controllable size of the region in the scene. The red cross point in the viewing area is set as the start point of reconstruction. (b) The 2D Images tab shows the four images captured under four different illumination directions. The Cam Info tab contains the manual setting information of the camera parameter used for calibration.

During the experiment, the background of each scene was set to be black, which helped to extract the object from it. A further threshold intensity value was used to reduce the black background noise. As the captured area from the camera was much larger than our need and to minimize computer memory requirements, the images were subsequently resized and down-sampled to 360 x 360 pixels before 3D image reconstruction was performed. The object position, the camera perspective, and the light positions for all images were static and known in advance. To calculate the gradient map of the object, a start point was chosen from the image that has the highest intensity value when all four LEDs were turned on at the start. As shown in Fig. 2.8 (a), the cross point of two red lines in the live view area indicates the start point for the calculation of gradient map. Once the object and camera were aligned, image acquisition was then triggered at the front control panel of the program, initiating the four LEDs to turn on successively, based on the setting and synchronized with a short camera exposure (in total less than 1 second). This procedure generated four similar 2D images of the object (see Fig. 2.8 (b)), each with different shading properties determined by the lighting environment of the time.

The software reconstruction pipeline is shown in Fig. 2.9. The intensities from those 2D images were used as the initial input, the value at each pixel in these images was compared and the maximum intensity was obtained to provide the red, green and blue (RGB) intensity values for each pixel. This provided the colour information of the scene, which was later used for mapping the 3D object texture (see Fig. 2.10). Areas that had a lower intensity value than the threshold value were identified as the background with a black colour reset. With the known coordinates, L_n , of four LEDs and the corresponding object intensity value on each pixel I_n , the object's surface normal N was then calculated via Eq. 2.6 or Eq. 2.10.

Based on the surface normal of each pixel, the gradients between adjacent pixels were calculated to obtain the surface geometry by integration, starting from the

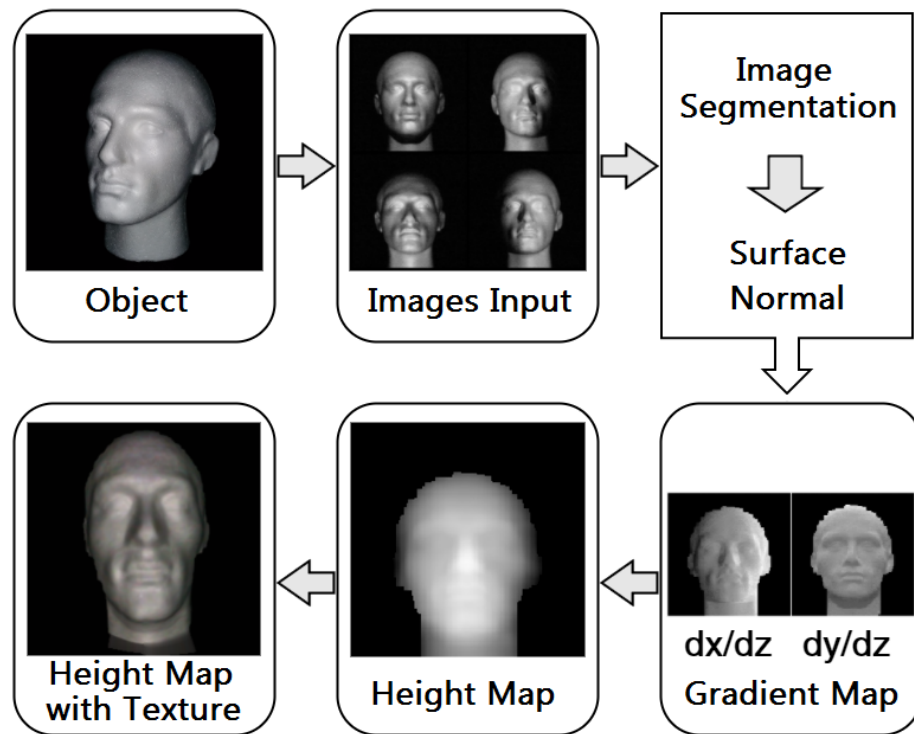


FIGURE 2.9: 3D Reconstruction pipeline. The reconstruction program executes following the sequence: image input, image segmentation, extraction of surface normal, gradient map, height map, and height map with texture.

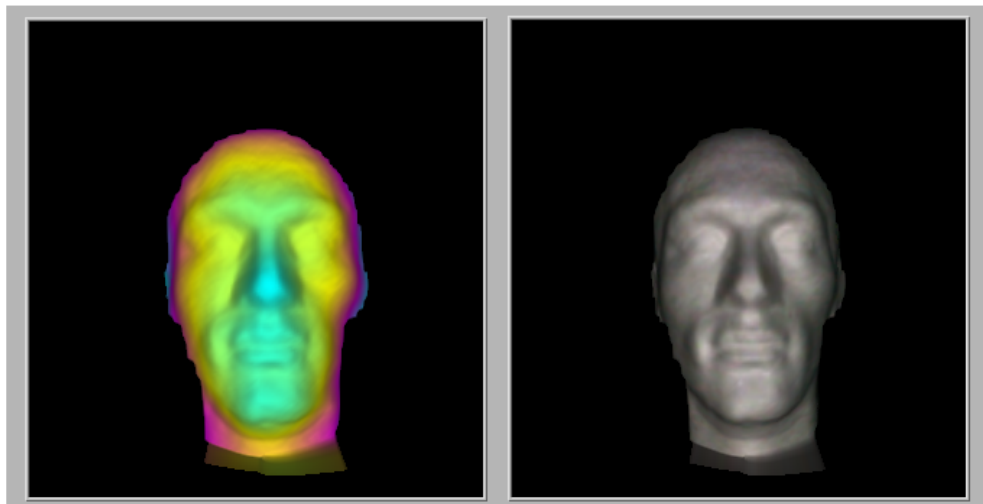


FIGURE 2.10: 3D height map. The height map on the left is untextured, in which the colour at each pixel represents its height value. The height map on the right is a textured height map with the real RGB value being mapped on top of it.

start point that was chosen from previous procedure to the outermost pixels of the object's surface. The surface height at a pixel point could be approximated with the gradient of the surface and the height of its nearest-neighbour point. Since each pixel point corresponded to the measured gradient data and had more than one nearest-neighbour point, the gradient of the surface used was the mean value of the gradients at every two contiguous points, and the surface height at each pixel point was replaced by the mean of the values counted from all its nearest-neighbour pixel points. The relative height value at each pixel was estimated one by one, after the gradient map had been calculated. The pixel points of an object could be categorized into certain two types: the internal pixel point and the boundary pixel point. When a pixel was an internal point, the reconstructed height value was set to the mean calculated from all its nearest neighbours; when a pixel was a boundary point, the measured gradient data at the point was then assumed to be accurate [76].

2.6 Quantitative and qualitative analysis

Four different-shape objects were designed (with SketchUp) and 3D printed (with Cura) as the initial trial set for 3D reconstruction with the camera system: an arc, a cube, a flat ramp and a hemisphere (see Fig. 2.11). The height of each object was set as 50 mm. The length of the arc and the ramp, along with the diameter of the hemisphere was set as 100 mm. They were all painted with a matt white spray to create a Lambertian surface. Fig. 2.12 exhibits the 3D reconstruction results of these objects. Owing to the constraints of the position between the camera and the object, it was unable to estimate the height value of the cube as its side faces were hidden from the perspective of the camera and each point at the front face has the same intensity value.

To test the accuracy and robustness of our 3D imaging system we then selected three different objects with varying geometric complexity: the hemisphere, the arc,

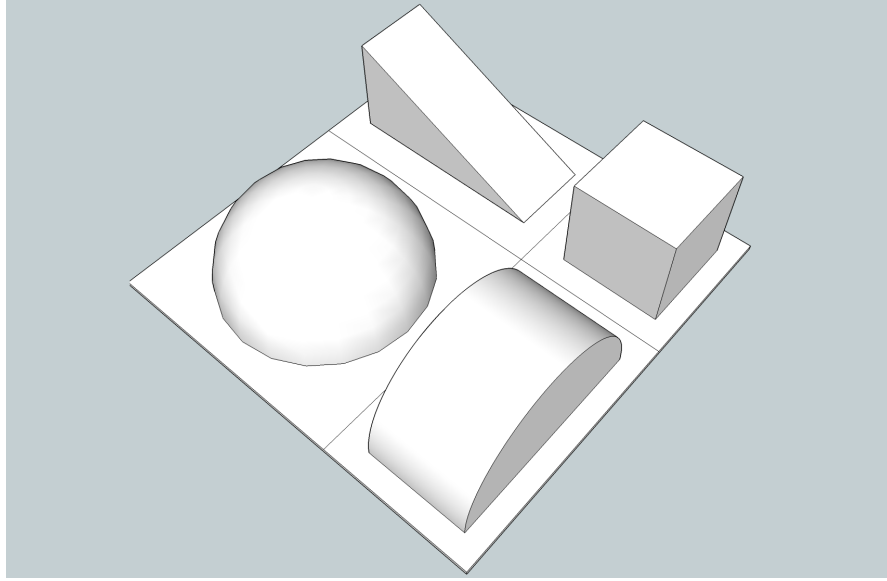


FIGURE 2.11: Four objects designed with the same height of 50 mm using software SketchUp: an arc, a cube, a flat ramp (with a length of 100 mm) and a hemisphere.

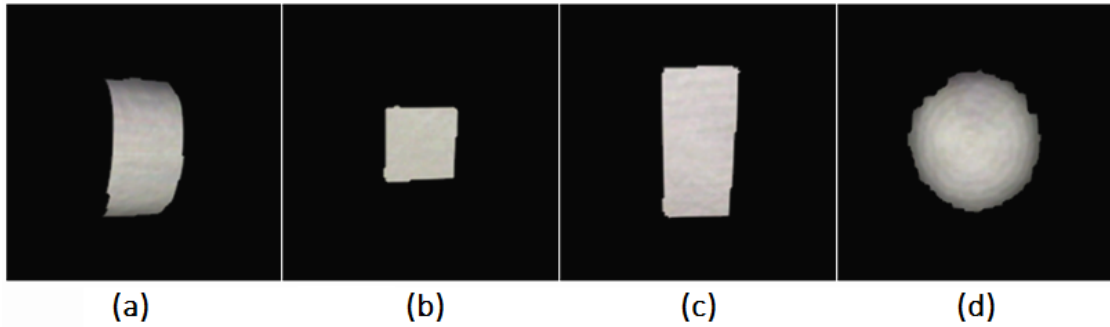


FIGURE 2.12: 3D reconstruction results of four different shape objects with a Lambertian surface using our camera system: (a) an arc, (b) a cube, (c) a flat ramp, (d) a hemisphere. It was unable to estimate the height value of the cube due to the constraints of the position between the camera and the object.

and a plastic mannequin head. The mannequin head is made from polystyrene with a smooth surface which can be considered as a Lambertian surface. It was measured to have a height (from ear to nose tip) of 160 mm.

In our experiment, the distance between the top front point of each object and the camera lens was set as 900 mm. Four white LEDs (positioned on the top, left, bottom, and right of the camera) were fixed to surround the lens, maintaining at a distance of 330 mm from the its center. For each of those objects tested in

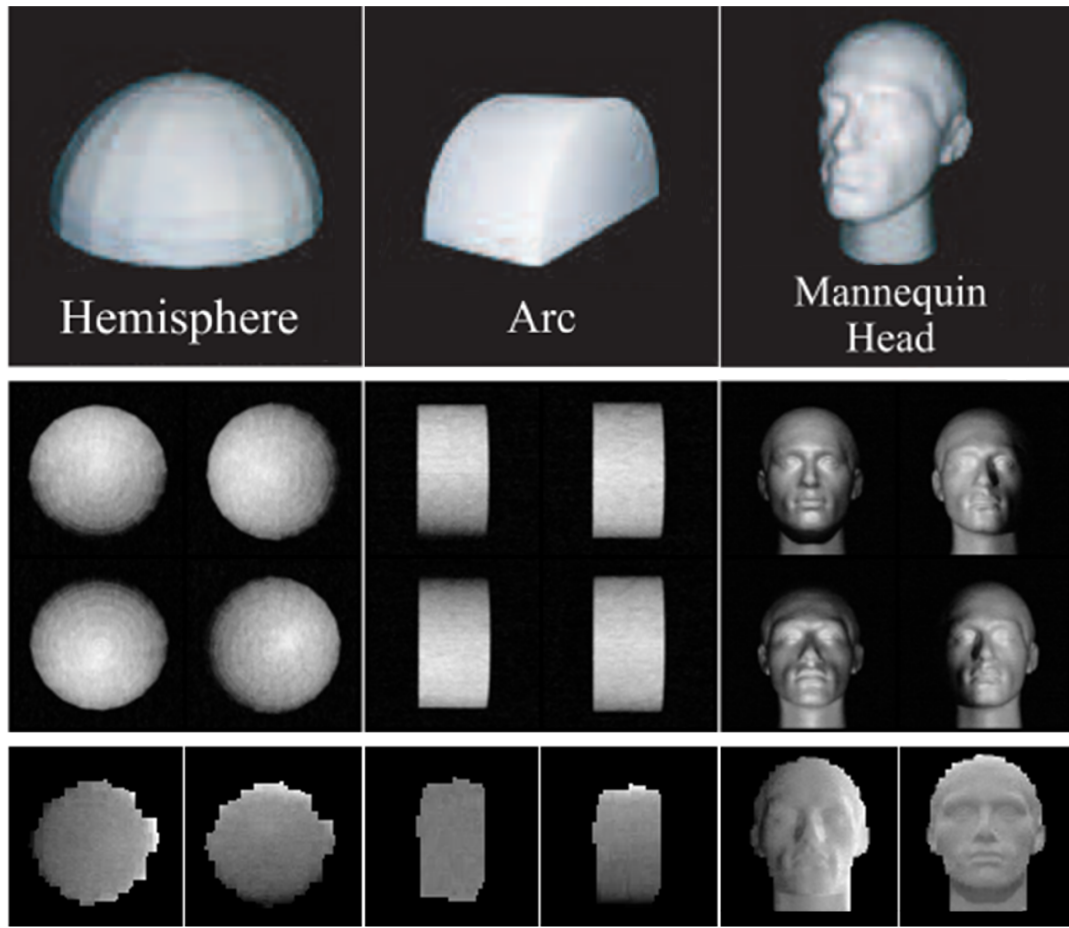


FIGURE 2.13: 3D Reconstruction process of objects: a hemisphere, an arc, and a mannequin head. Images aligned from top to bottom are: object images under uniform light conditions, images captured with four different illumination directions, object gradient maps dx/dz and dy/dz .

a dark black background environment, four images were acquired with different reflected intensity distributions, corresponding to the distinct lighting conditions. The viewing scene were cropped into a 360×360 pixel image (corresponding to $320 \text{ mm} \times 320 \text{ mm}$ virtual size). Those images were applied to calculate the gradients of each pixel which were then used to estimate the height map of the object surface. The colour value was later mapped onto the height map using the aforementioned approach (see Fig. 2.13 and Fig. 2.14).

The reconstructed height map of each object was compared with a reference height map to provide a general analysis of the 3D reconstruction quality of our camera

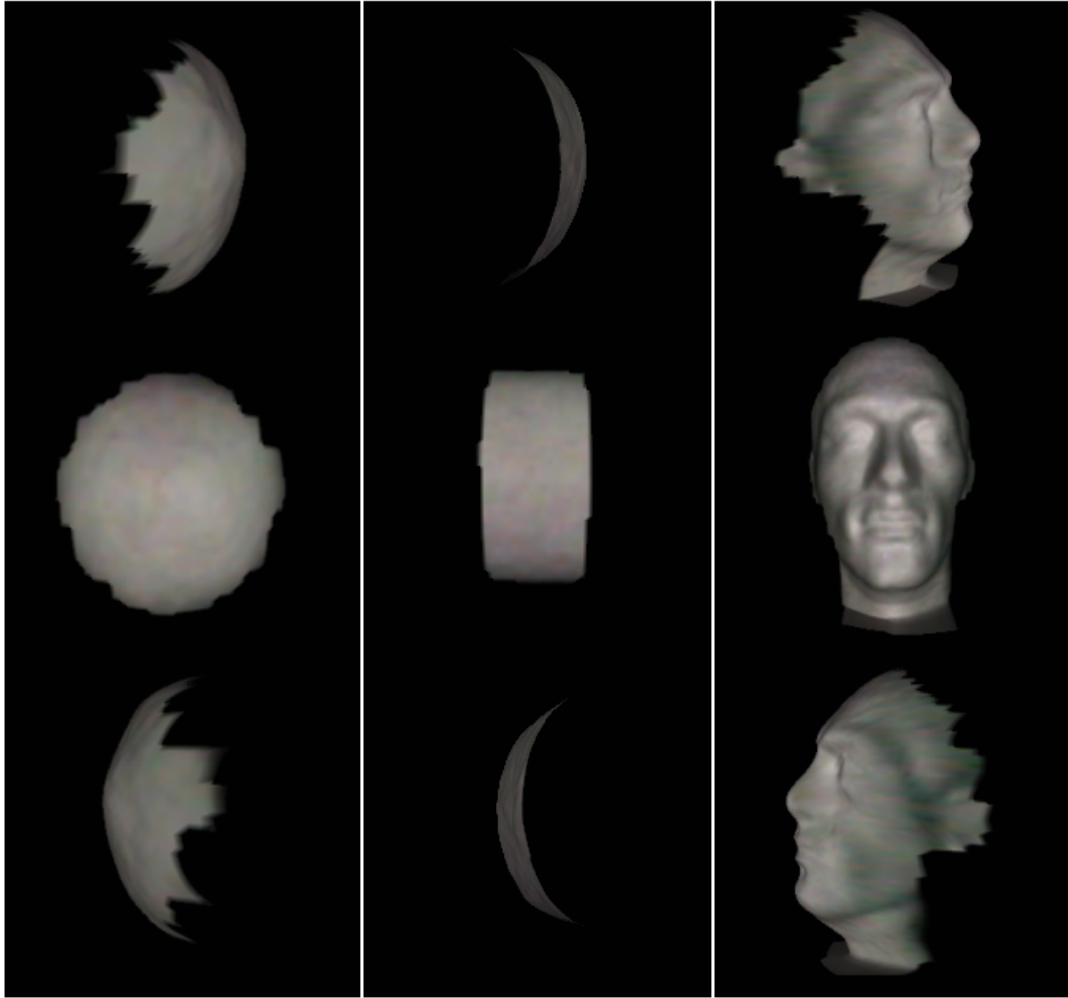


FIGURE 2.14: 3D reconstruction results of three objects: a hemisphere, an arc, and a mannequin head. Images from top to bottom are the 3D shape of objects captured from the left, center, and right directions.

system. For the arc and the hemisphere, the reference height map was obtained from the stereolithography (STL) files that I used to design them, whereas the mannequin head reference data was obtained from an expensive separately calibrated commercial stereo-photogrammetric system employing 4 high-resolution cameras (see Fig. 2.15) which represented state of the art and the 3D reconstruction process was time-consuming. For comparison, the size of measured data was scaled appropriately to map the length and width with the reference data, and the measured height map was normalised and scaled to map the maximum height

value of the reference data, which was 160mm.

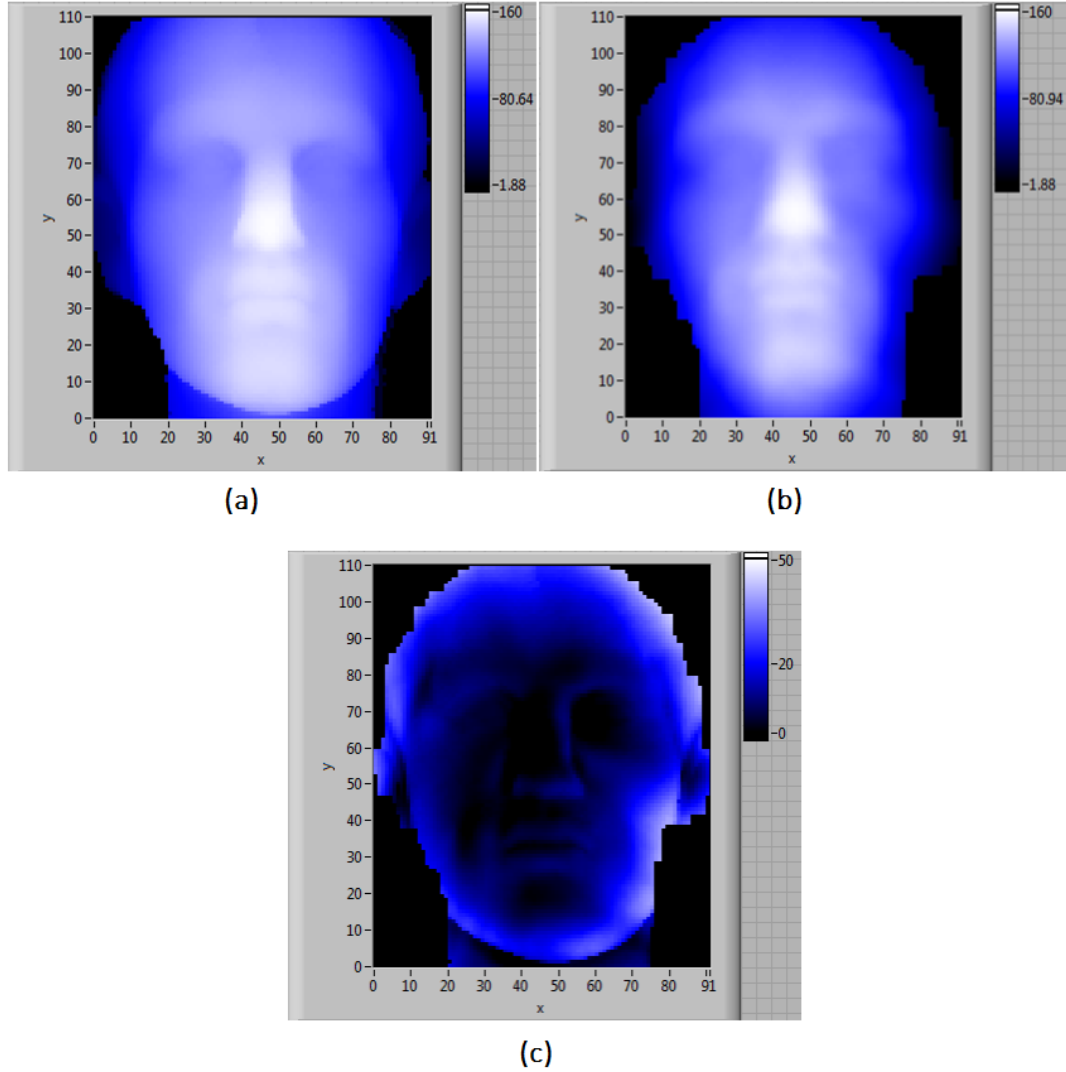


FIGURE 2.15: Mannequin head height map comparison: (a) The realigned reference height map with a known maximum value of 160mm. (b) The measured height map normalised and scaled to match the maximum height value of the reference data. (c) The difference map of the height value at each pixel between (a) and (b). The unit of the scale bar is millimetre.

As shown in Table 2.1, we presented the standard deviation of the differences between our measured height data and the reference height data by using the root mean square error (RMSE) and the normalized root mean square error (NRMSE),

referring to the variation of the RMSE [78]. The RMSE and NRMSE, correspondingly are defined as [79]:

$$RMSE = \sqrt{\left(\frac{1}{n} \sum_{i=1}^n d_i^2\right)}, \quad (2.11)$$

$$NRMSE = \frac{RMSE}{x_{max} - x_{min}}, \quad (2.12)$$

where n is the number of data pairs;

d_i is the difference between measured values and reference values;

$(x_{max} - x_{min})$ is the range of measured values.

TABLE 2.1: Deviations between measured values and true values

Object	Height(mm)	Scale Factor	RMSE (mm)	NRMSE
Hemisphere	50	2.66	2.76	5.53%
Arc	50	3.14	2.65	5.32%
Head	160	2.02	15.60	9.76%

The RMSE is one of the most widely used statistical indicator for assessing reconstructed image quality by measuring the differences between the reconstructed images with the ground truth, and the NRMSE is used to compare the accuracy between those images which may have different units and ranges [80, 81]. The value of NRMSE is commonly expressed as a percentage, where a lower value indicates less variance and hence higher accuracy. We noted close agreement for the two 3D printed objects with relatively low geometric complexity, the hemisphere and the arc, with the NRMSE of 5.53% and 5.32%, respectively. A higher NRMSE of 9.76% was observed for the mannequin head. We noticed that the regions of the object contributing most to the overall RMS error were areas of sharp edges or where the surface normal was in a direction perpendicular to the camera perspective.

Moreover, we did a simple test on our 3D camera system with different human faces as the imaging subjects, to roughly validate its capabilities for potential applications of practical affairs, such as human facial recognition. Before any data

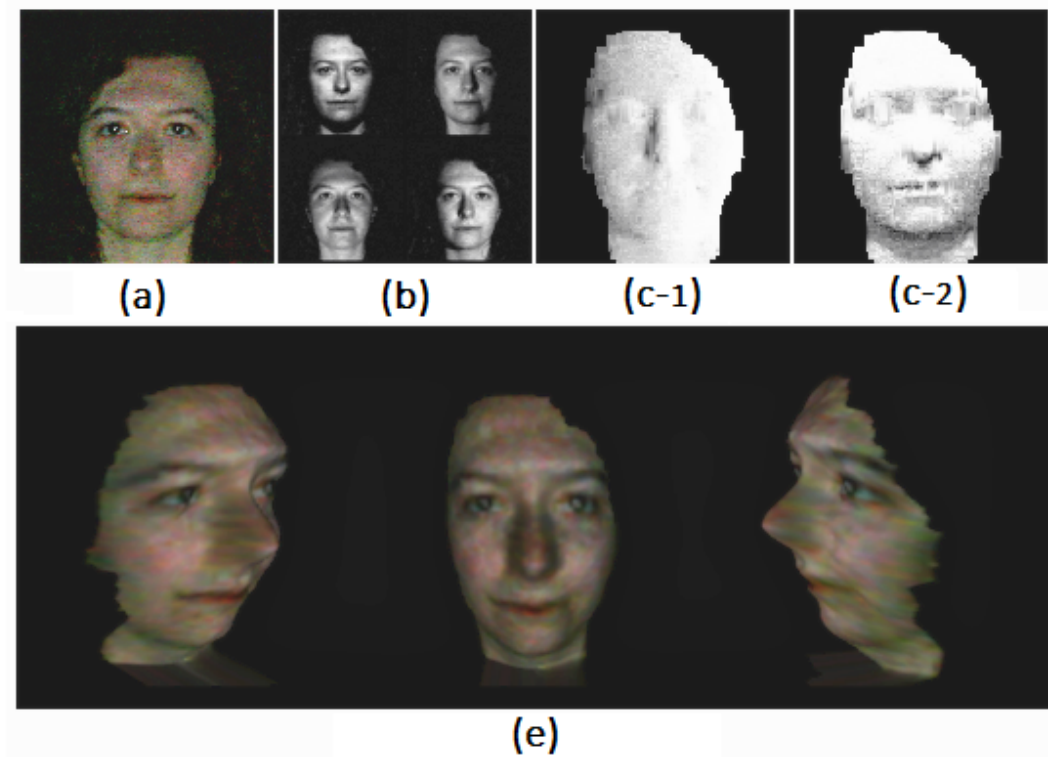


FIGURE 2.16: 3D reconstruction of a human face using our camera system. (a) Object. (b) 2D Image input. (c-1) dx/dz . (c-2) dy/dz . (d) Images of 3D reconstructed result.

was collected an ethics approval was gained from the Research Ethics Committee at University of Glasgow and a consent form for publication of reconstructed images was signed by each participant. Fig. 2.16 presented one of 3D reconstruction results of a human face. As there was no reference data for each individual human face, the quality of 3D reconstructed result was simply assessed visually and the images were only used for demonstrating the possibility of creating the 3D model of a real human face with our camera system.

In addition to reconstructing 3D images with a low-cost camera accessory and photometric stereo software, it was feasible to simultaneously construct pairs of 2D images representing slightly different perspectives of an object that would be viewed by our eyes, whose typical separation is around 60mm, since we calculated the height map of the object (see Fig. 2.17). By using those paired 2D images fetched from the textured height map, a movie was then generated by displaying

each pair simultaneously and could be displayed on a 3D-enabled TV screen for viewing. The 3D-enabled TV operates by superimposing the paired images in the movie with each having a different polarisation, and the viewer wears polarising 3D glasses to separate the left and right image for each eye. It is quite remarkable that our relatively low-cost 3D camera system is capable of providing such realistic 3D images of objects and faces for a wide range of viewing angles, despite the camera and the subject remaining fixed during the acquisition.

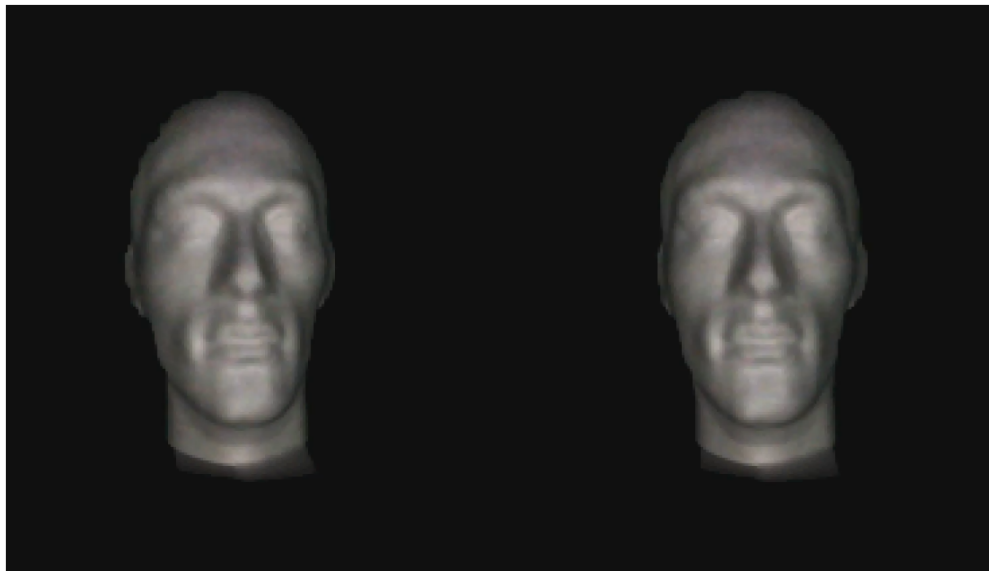


FIGURE 2.17: An image-pair acquired from textured 3D height map of an object with slightly different perspectives. The pair was superimposed and displayed on 3D-enabled TV screen with each having a different polarisation. Viewers wearing polarising 3D glasses were able to observe the displayed object in a 3D form.

As well as being used for scientific research, the 3D photometric stereo system was further used at several events for science outreach, to demonstrate the technology and help engage the public in the field of optics and imaging. A customised 3D photo-booth was built by Matt Edgar and Graham Gibson to house the equipment and attract member of the public. This formed the basis of the Creative Cameras exhibit at the Royal Society Summer Science Exhibition, which myself and colleagues helped to deliver. The associated short-film about our exhibit has become the most watched video with over 100,000 views, thus highlighting the impact

of the research. The results shown at those exhibitions confirmed the capability of our camera system to 3D reconstruct human faces under realistic workplace conditions.

2.7 Conclusion

To summarise, we have developed a fast camera system with a low-cost accessory, which utilises an Arduino controller board to associate with four white LEDs and laptop, to enable an efficient computer algorithm to reconstruct 3D images. A photometric stereo method uses multiple 2D images that are captured from a single viewing perspective with different illumination directions to estimate the depth and surface orientation of an object.

Compared to other 3D imaging methods such as 3D-scanning, it comes with a number of benefits, such as simplifying the reconstruction routine with a high efficiency. We analysed our reconstructed 3D results of different objects with various geometric complexity, noticing good quantitative agreement with the known reference object with a wide viewing angle. We also observed an increased error at regions of large gradients, where the surface normals were approaching a direction perpendicular to the camera perspective, such as shape edges, showing some limitation of our 3D camera system. Further improvements could be made by optimising the reconstruction algorithm in order to provide better height estimates at those regions and finding an efficient solution to eliminate the inner shadow effect.

We tested the 3D camera system at various science exhibitions which proved its potential capabilities for applications to practical affairs. Our inexpensive 3D camera system could be used at places that require fast 3D information collection, such as the security check at airports, or perhaps at retail clothing outlets for

personalising the browsing experience whilst shopping. It could also be applied to schools for education purpose due to the low-cost components.

Chapter 3

Real-time 3D Imaging with Single-Pixel Detectors

3.1 Introduction

As described in the previous chapter, photometric stereo is a well-established, and in some cases a convenient 3D imaging technique. Importantly, traditional 3D imaging systems using this approach demand that the scene remains completely static whilst the lighting condition changes during acquisition in order to prevent pixel matching errors, which narrows its extendibility to real-time applications. Whilst a variety of strategies [82–85] were devised to improve the 3D reconstruction accuracy of different 3D shapes with photometric stereo, it seems that there have been relatively less research on eliminating the underlying problems associated with sequential acquisitions.

The use of spectral multiplexing is one state-of-the-art method to reduce pixel matching errors, which captures the scene with an imaging system configured to measure multiple spectral channels. With this approach, the system utilises two cameras aligned co-axially with a beam splitter and spectrally filtered using two different bespoke dichroic filters, in conjunction with three bright LED light

sources with unique spectral profiles fixed at separated locations. The two filters separate the visible spectrum into six non-overlapping bands. By analysing those six-channel photographs, it captures per-pixel photometric normals and full colour reflectance synchronously, without requiring time-varying illumination. More details can be found in [86]. However, this approach still needs to match images from those two cameras, and reconstruction bias appears as a result of spectral variations for scenes with distinct materials, for instance, human faces.

In this chapter, we present a revised 3D imaging system that combines photometric stereo with a single-pixel imaging technique to tackle the matter of pixel registration among images. Furthermore, one of the compressive strategies, known as the evolutionary compressed sensing algorithm, was employed in our system to enable 3D reconstruction of scenes exhibiting dynamic behaviour in real-time. Samples of video frames with different moving objects obtained at various image resolutions and compression ratios are presented to demonstrate our system's ability to perform 3D imaging in real time.

3.1.1 Contributions

The work in this chapter was performed by myself and Matt Edgar with help from Graham Gibson, Baoqing Sun, Neal Radwell and Miles Padgett. The 3D single-pixel system was devised, partially 3D printed and assembled by myself and Graham Gibson. The evolutionary compressed sensing algorithm was conceived by Neal Radwell, optimised and deployed into our system by myself and Matt Edgar. The discussion on real-time video using 3D single-pixel imaging was a result of collaboration between myself, Matt Edgar and Miles Padgett.

3.2 Computational single-pixel imaging

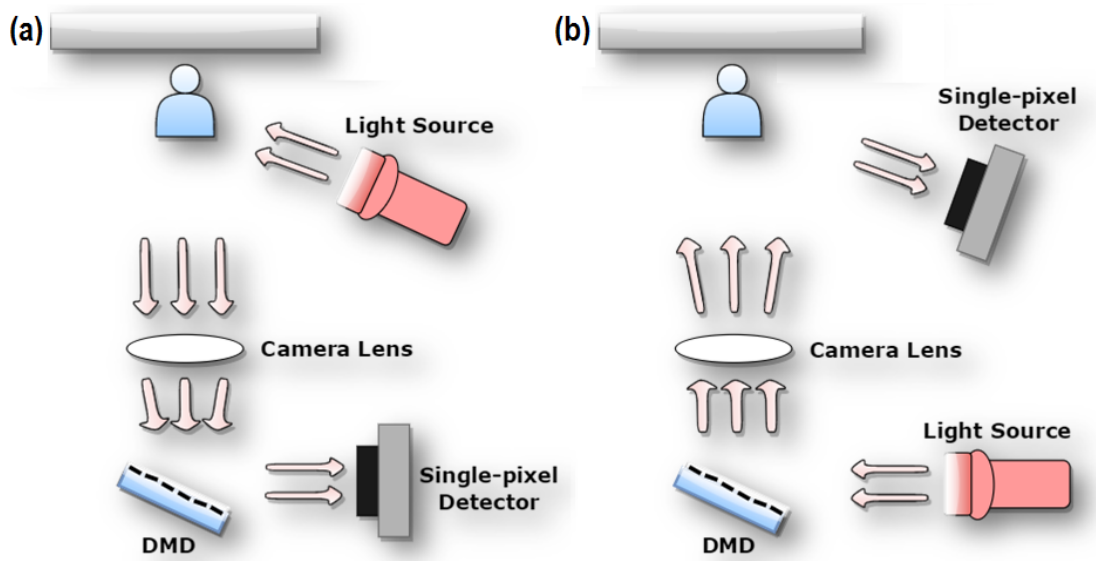


FIGURE 3.1: (a) Computational single-pixel imaging with structured detection: light reflected from an object surface is focused through a camera lens onto a digital micromirror device (DMD) which displays a set of different patterns. Different amounts of projected light are collected by a single-pixel detector and used to deduce an image of the object. (b) Computational single-pixel imaging with structured illumination: a set of structured light patterns are projected from a DMD and focused through a camera lens onto a scene. A single-pixel detector collects light from the illuminated scene to deduce an image.

As an alternative to conventional multi-pixel imaging approaches, a single-pixel imaging technique [12, 87, 88] enables a scene to be captured, illuminated by a uniform light field, with a single-pixel detector as the imaging device that measures the correlations between the collected light and a set of patterns. An optical single-pixel imaging system generally consists of an illumination source, a light detection unit and a spatial light modulator (SLM) that provides either time-varying, structured detection of an image, or by providing time-varying, structured illumination onto a scene (see Figure 3.1). To combine photometric stereo with single-pixel imaging, the SLM provides the spatial resolution and hence perspective to the scene, whereas the position of the detector has the same effect as the illumination source in a conventional photometric stereo system. This provides

a novel scheme for recovering multiple illuminations using only one physical light source and several single-pixel detectors. Thus, a structured illumination approach was chosen for this photometric stereo computational imaging system.

In general, the total amount of time for reconstructing an image with single-pixel imaging is determined by the speed of signal acquisition together with sampling size required. To improve the imaging efficiency in our system, certain strategies were employed to increase the data collection speed as well as to reduce the measurement times.

3.2.1 Light projection with a digital micro-mirror device

To obtain an image of a scene with a single-pixel detector typically requires at least the same amount of sampling as the number of desired pixels in the image. With a fixed image resolution, the reconstruction time depends on the speed of data collection. Scanning rate is an important aspect that effects the data collection efficiency. Different scanning rates in a single-pixel imaging system can be achieved by using different types of SLM with varying structures and working principles. To maximise the scanning rate, a fast-pattern switching speed of SLM was employed in our single-pixel system.

By definition, an SLM is any sort of electro-optic unit that can be used to modulate certain properties of light waves in space and time, for instance, the amplitude, phase, or polarization [89]. It's a critical structural component for most applications in optical signal processing and optical computing. Based on the information-bearing element, either an electrical or optical signal, SLM devices can be classified into two types: namely, the electrically activated SLM (EA-SLM) and the optically activated SLM (OA-SLM) [90]. In this system, we employed an EA-SLM device due to its capability of interfacing with electronic components in an electro-optic system with fast response. More specifically, we used a Texas Instruments Digital Light Processing (DLP) unit [91], a display device that consists of electronic logic

circuits, digital memory, and an optical semiconductor chip (OSC) known as a digital micro-mirror device (DMD), to perform the structured light projection.

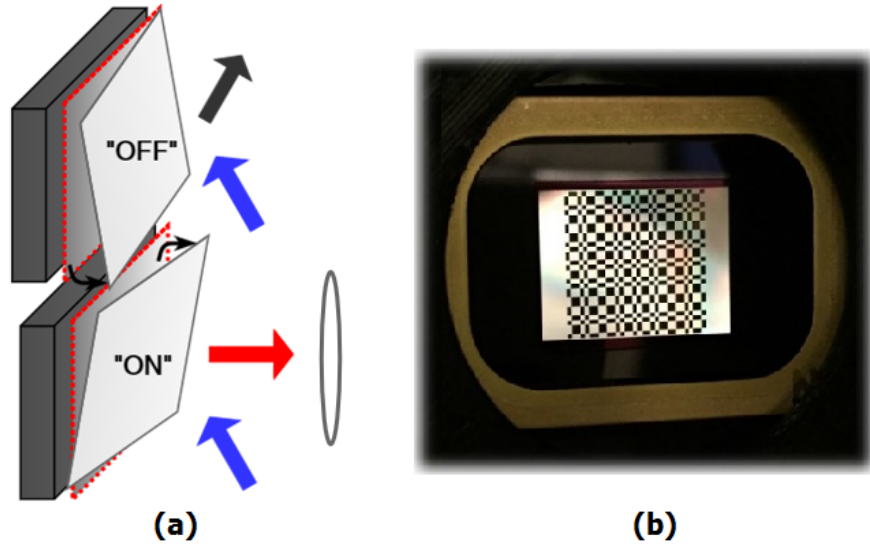


FIGURE 3.2: (a) An illustration of DMD mirrors with on and off states: when a micro-mirror reflects light towards the lens onto the scene, it corresponds to an "on" state. When a micro-mirror reflects light towards a different direction rather than the lens, it corresponds to an "off" state. (b) An illustration of a modulated light pattern on a DMD. The bright regions on the DMD represent the "on" state, and the dark ones represent the "off" state.

A DMD is a reflective SLM with its surface made up of thousands of microscopic metal mirrors in a form of a 2D rectangular array [92, 93]. It was originally developed in 1987, then diffusely applied in traditional display applications and optical modulation systems [94–96]. The mirrors on a DMD are controlled individually, through a '1' or '0' signal stored in its underlying memory cell, to rotate at a certain angle (+12 or -12 degrees), corresponding to an 'on' or 'off' state. As shown in Fig. 3.2, when a mirror is in an 'on' state, light coming from the projector light source will be reflected towards the lens to make a corresponding area appear bright on the screen. While in an 'off' state, light reaching this mirror will be reflected towards reverse angle, resulting in the corresponding area to be black on the screen. A prime attribute of DMD is that it can provide structured illumination at a high speed rate of up to 22.727 kHz. Furthermore, the operational

bandwidth of DMD can reach from 300nm to $2\mu m$, which enables its potential extension outside the scope of visible wavelength, for instance, the near-infrared wavelength, in which area most traditional imaging techniques are much more expensive and difficult to employ.

A computational imaging system using a single-pixel detector and a DMD can be regarded as having three main processes: scanning, sampling and processing. The imaging efficiency changes depending on the arrangement and optimisation between these processes. Besides a high projection speed and a wide operational bandwidth, another important advantage of using DLP is that it provides a trigger signal in reference to the structured illumination to enable synchronisation measurements between scanning and sampling, which provides stable parallel procedures to enable a faster collection rate. More specifically, when each pattern on a DMD gets projected, there will be a trigger signal released from the DLP control circuit, being sent to the data acquisition (DAQ) device. With this signal input, the DAQ device will start to acquire light intensity data from the single-pixel detector(s). A certain number of data acquisitions is acquired after each trigger signal, which is determined by the total number of detectors, the projection time of each pattern, and the data acquisition rate of the DAQ device.

3.2.2 Structured illumination with Hadamard patterns

Besides scanning rate, the scanning method is another aspect that effects the total image acquisition time. In single-pixel imaging, scanning efficiency varies between different methodologies based on the amount of signal collection that is required to generate an image in each method. One basic scanning approach is raster scanning[97, 98], which typically generates a frame of the scene by scanning pixel by pixel from the top left to the bottom right (see Fig. 3.3).

A significant advantage of using raster scanning is that it is possible to selectively erase and modify certain image area as each pixel is scanned independently.

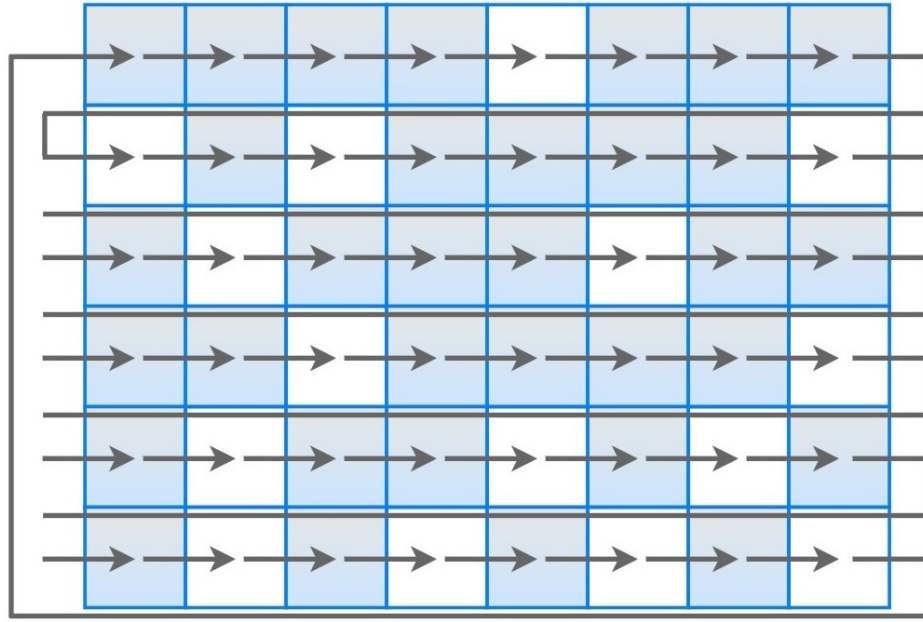


FIGURE 3.3: One basic illumination method: Raster scanning. It scans the scene pixel by pixel typically from left to right, top to bottom. Each pixel is scanned independently as only one pixel is illuminated at a time.

However, the illumination efficiency with this approach decreases when the image resolution increases and the sampling data of each pixel is spatially discrete which increases the noise level in the final image. Ref. [99] experimentally demonstrated a single-pixel imaging system with raster scanning which suffers a high noise level comparing to other methods.

Instead of raster scanning, an alternative approach, also known as pattern illumination, is to scan multiple pixels simultaneously instead of one pixel each time, which can help to improve the signal-to-noise ratio of each measurement, if the system is detector noise limited. Fig. 3.4 illustrates a single measurement of multi-pixel scanning with a random illumination pattern. As there is more light illuminates onto the scene, the ratio of useful signal collected by the detector increases.

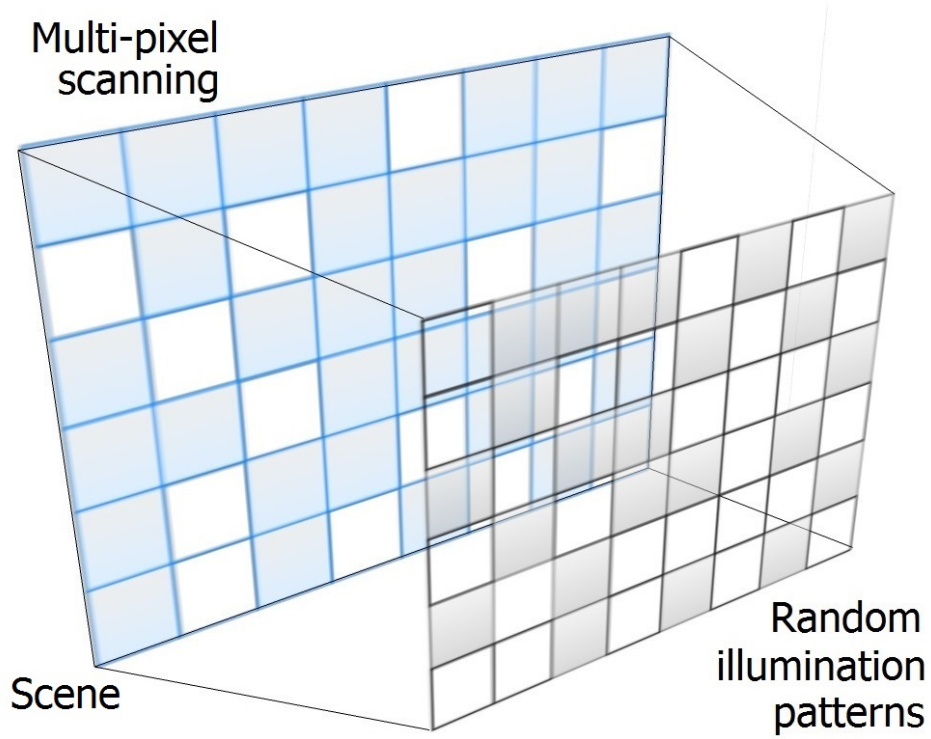


FIGURE 3.4: Illustration of a single measurement of multi-pixel scanning with a random illumination pattern. More than one region/pixel area of the scene is illuminated each time based on the light pattern, in which case the signal collected by the single-pixel detector will have a higher ratio of useful signal than the noise.

To obtain an image S of $N \times N$ pixel resolution with pattern scanning follows the equation:

$$\begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_{(m-1)} \\ P_m \end{bmatrix} \boxtimes S_{(N \times N)} = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_{(m-1)} \\ B_m \end{bmatrix}, \quad (3.1)$$

where P_m is a row vector of the m^{th} pattern that is used to project on the scene, and each pattern contains the same amount of pixels as the image, which means

$$P_m = [p_{m1}, p_{m2}, \dots, p_{m(N \times N)}];$$

$S_{(N \times N)}$ is a matrix of image S with $N \times N$ pixels $\begin{bmatrix} s_{11} & \dots & s_{1n} \\ : & \dots & : \\ s_{n1} & \dots & s_{nn} \end{bmatrix}$ and here

$$\text{we set } A \boxdot S_{(N \times N)} = AS = A \begin{bmatrix} s_{11} \\ s_{12} \\ : \\ s_{21} \\ : \\ s_{n(n-1)} \\ s_{nn} \end{bmatrix};$$

B_m is a scalar that represents the detected light intensity value with the m^{th} illumination pattern;

\boxdot is an operational symbol that we define to implement the operation:

$$P_m \boxdot S_{(N \times N)} = p_{m1}s_{11} + p_{m1}s_{11} + \dots + p_{m(N \times N)}s_{nn} = B_m.$$

Since the illumination patterns that we used only have black and white blocks representing the off and on states of mirrors (equal to 0 and 1), this equation will have the form:

$$\begin{bmatrix} 1 & 0 & \dots & 1 & 1 \\ 1 & 0 & \dots & 0 & 1 \\ : & : & \dots & : & : \\ 1 & 0 & \dots & 0 & 1 \\ 0 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} s_{11} \\ s_{12} \\ : \\ s_{21} \\ : \\ s_{n(n-1)} \\ s_{nn} \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \\ : \\ B_{(m-1)} \\ B_m \end{bmatrix}, \quad (3.2)$$

$$PS = B \quad (3.3)$$

This approach could be time-consuming if the illumination patterns are randomly selected, as there will be overlapping information among those patterns which would increase the amount of sampling that is required to form an image. Since there are $N \times N$ unknown pixels in image S , the minimum amount of patterns m that are required to solve this equation is $N \times N$, which makes P a square matrix.

According to Eq.3.3, if P is an invertible matrix (which means its inverse matrix P^{-1} exists), then we can get

$$S = P^{-1}B \quad (3.4)$$

To minimise the computational complexity of this equation and reduce measurement redundancy, the set of patterns P should be orthogonal, in which case

$$P^T P = P P^T = I \quad (3.5)$$

where P^T is the transpose of P , and $P^T = P^{-1}$;

$$I \text{ is the identity matrix and } I_n = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ : & : & \dots & : & : \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}_{n \times n}.$$

In our system, we chose to use Hadamard matrices [100] to construct the patterns to be applied as structured illumination, which simplifies the complexity of calculation and reduces the measurement redundancy.

In mathematics, a Hadamard matrix is a type of square matrix with its rows are mutually orthogonal [100–103]. The elements of a Hadamard matrix are either -1 or +1, and its equivalent definition can be given by

$$H_n H_n^T = H_n^T H_n = n I_n,$$

where H_n is a Hadamard matrix of order n . Typically the order of a Hadamard matrix n should be 1, 2, or $4i$ where i is a positive integer (see Fig. 3.5). The transpose of a Hadamard matrix, H_n^T , is still a Hadamard matrix. In a Hadamard matrix H , any two columns, the same as any two rows, are orthogonal, and its inverse matrix $H^{-1} = \frac{1}{n} H^T$.

The Hadamard matrices that we used in our system are constructed following the method that was proposed by Sylvester [104], also known as Walsh matrices [105], where

$$\begin{aligned} H_1 &= \begin{bmatrix} 1 \end{bmatrix} \\ H_2 &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ H_{2^n} &= \begin{bmatrix} H_{2^{n-1}} & H_{2^{n-1}} \\ H_{2^{n-1}} & -H_{2^{n-1}} \end{bmatrix} \end{aligned}$$

The Hadamard matrices constructed from this approach have some particular properties. Each single matrix is symmetric, in which case $H = H^T$. The elements in the first row and the first column of each Hadamard matrix are all positive (+1), while the ones in the rest rows and the rest columns are half positive (+1) and half negative (-1).

To make sure the patterns derived from each row of a Hadamard matrix are enough to cover all image pixels, we only chose Hadamard matrices of order 2^{2k} , then reshaped each row (or column) into a $2^k \times 2^k$ 2D pattern array based on the image resolution. Therefore a Hadamard matrix in the order of 2^{2k} was transferred into a complete set of 2^{2k} ($2^k \times 2^k$) structured illumination patterns in 2D, which were

displayed one after another. In our experiment, a 1024×1024 Hadamard matrix was used to generate 1024 of 32×32 Hadamard derived illumination patterns to modulate the light source for acquiring an image with 32×32 pixel resolution. By analogy, we also generated a set of 4096 64×64 Hadamard derived patterns and a set of 16384 128×128 Hadamard derived patterns for obtaining images with corresponding pixel resolutions.

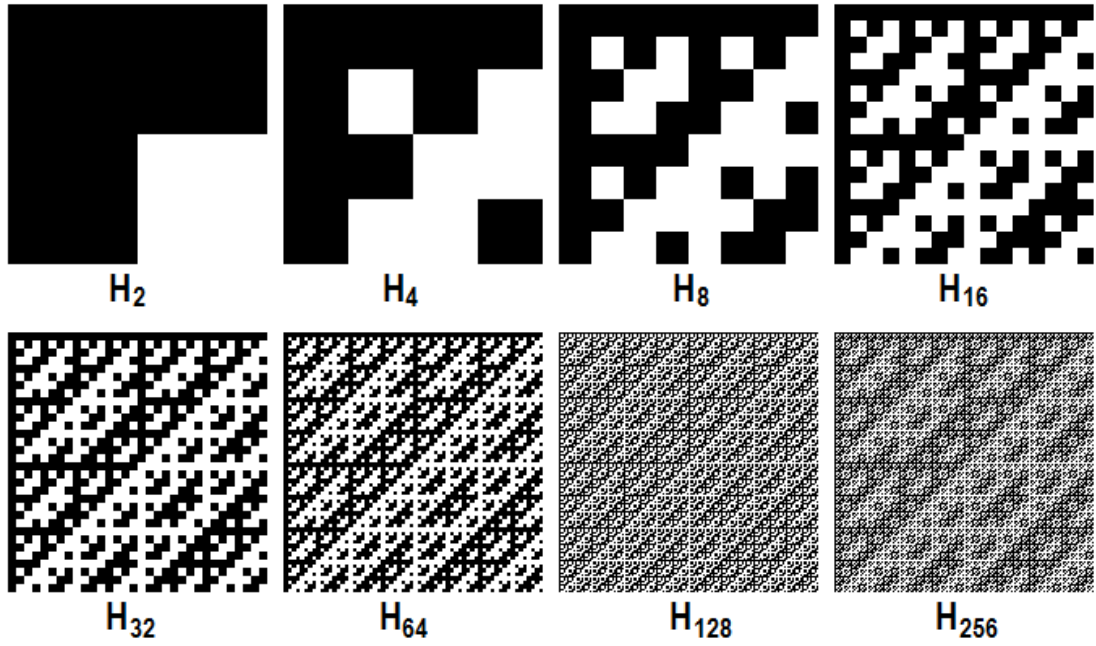


FIGURE 3.5: A Sylvester's Hadamard matrix set with orders of 2^n ($n \geq 1$).

The pattern coding with Hadamard matrix is based on the theory of combining weighing. For an image with m pixels, the theoretical SNR improvement of using a Hadamard matrix of order $n = m$ called optimum chemical balance weighing deisign is \sqrt{n} , as each pixel will be weighed n times [106–108].

3.2.3 Differential signal acquisition approach

One issue of employing Hadamard patterns on DLP is that the elements in those pattern arrays are either "+1" or "-1", while the DLP can only generate black

or white illumination blocks (equals to 0 and 1) on the scene. To solve this, a differential signal acquisition approach was proposed.

Instead of using the absolute intensity value of signal from each pattern to reconstruct the 2D image, we projected each structured pattern followed by its inverse where the black pixels became white and the white ones turned into black, and adopted the differential intensity value from those pattern pairs to replace the signal value for each corresponding Hadamard derived pattern (see Fig. 3.6). By employing this differential signal acquisition approach, the effective elements of each illumination pattern are either +1 or -1, rather than 0 or 1. In addition to this, another benefit of using this approach is that it helped to remove the background signal noise compared to the method of single-pattern projection. However, this improvement of signal to noise ratio is at the cost of overall illumination efficiency since it doubled the amount of projection patterns.

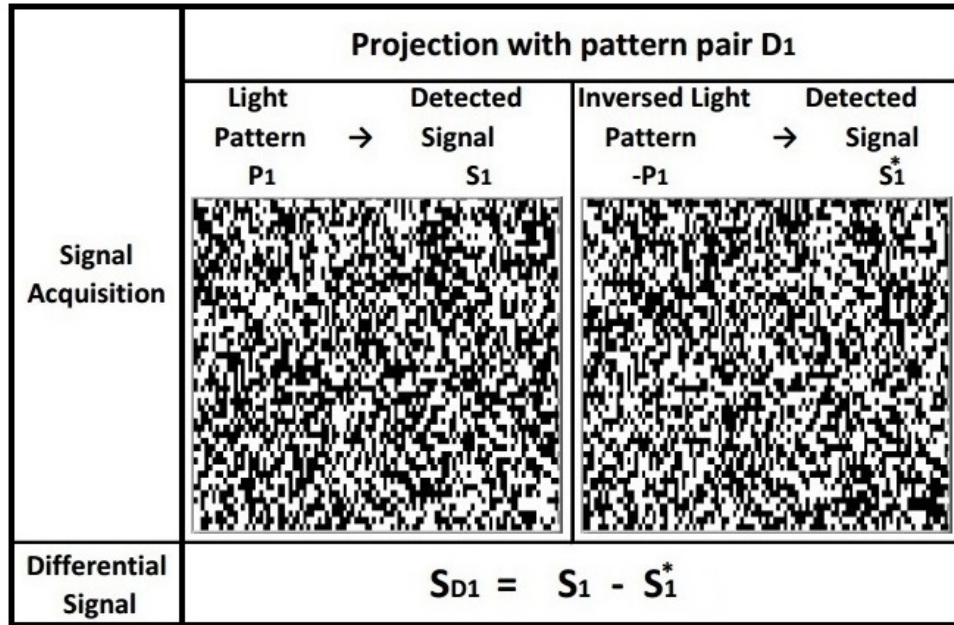


FIGURE 3.6: Illustration of a structured illumination pattern and its inverse. Pattern P_1 and $-P_1$ are a pair of inverse illumination patterns. S_1 and S_1^* are the light intensity values received from the single-pixel detector based on pattern P_1 and $-P_1$. The differential signal value S_{D1} , equalling to the subtracting of S_1 and S_1^* , is used to represent the signal value with pattern P_1 .

3.3 3D imaging system with single-pixel detectors

In addition to 2D reconstruction, 3D image reconstruction can also be produced by using single-pixel imaging technique. Rather than using only one single-pixel detector, four detectors are employed in the system at different locations, each producing a 2D image of the same scene simultaneously. It is realised that all these images have the same shape but containing different intensity distributions which are determined by the surface normal of the object and the detecting vectors. With the knowledge of the intensity distributions in these four images together with the relevant detector's location, a 3D profile of the object can be derived by following the photometric stereo approach that we discussed in the previous chapter.

3.3.1 Experimental setup

Before assessing the 3D video quality of our 3D single-pixel imaging system, we first tried to 3D reconstruct a static object. The system, as illustrated in Fig. 3.7, contained a high-speed DLP device, a camera lens with 24mm focal length attached in front of the DLP, and four spatially separated photodetectors (PDs) in fixed locations around the lens. A 3W white LED was used to illuminate the DMD chip and encoded into binary light fields. A silver-coated mirror was adjusted manually inside a 3D-printed mount so that the LED light was reflected at an angle of 24 degrees inclination to the normal and 45 degrees around the normal towards the DMD chip, in which case the illumination patterns on the DMD passed through the lens and were projected right onto the scene. As light intensity scattered from the 3D object is relatively low, a camera lens was placed in front of each detector to increase its numerical aperture.

The DLP that we used in the system is a Hi-Speed V-7000 DLP from Vialux, which contains a micro-mirror array of 1024×768 elements and an operation circuit programmed through the Accessory Light modulator Package(ALP) 4.2 controller tool with LabVIEW and connected to the computer via a USB interface. Its usable spectral range covers all wavelengths from 350nm to 2500nm, and the maximum array switching speed can go up to 22,727 Hz (we defined it at 22 kHz in our system). In our experiment, only the central 768×768 region of the DMD was used. Each pattern set was pre-loaded on the ALP in sequence.

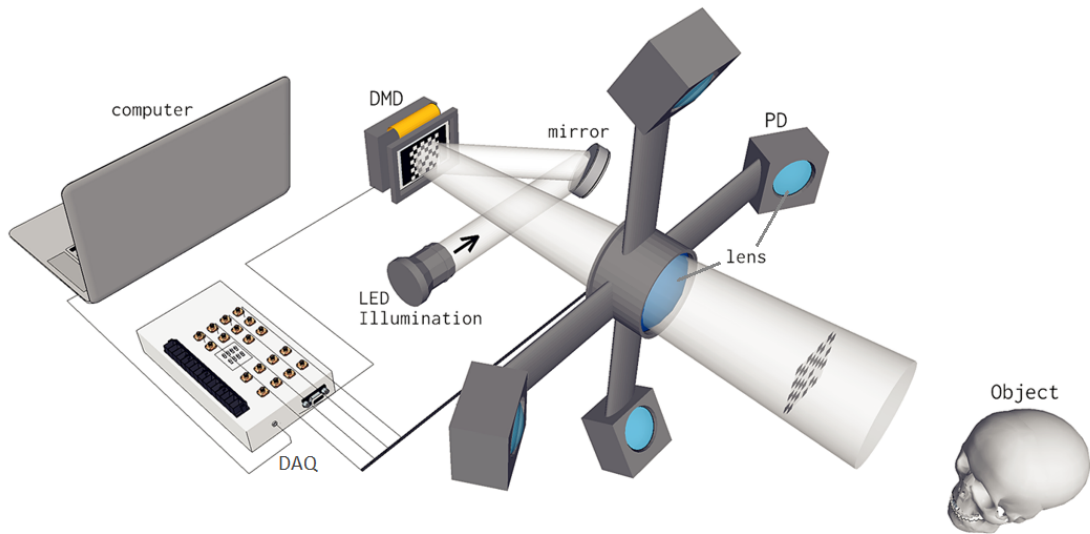


FIGURE 3.7: Illustration of 3D single-pixel experimental setup. A white LED, a fast DMD and a camera lens were combined to deliver structured illumination on the scene. Each photomultiplier tube transformed light scattered by the object to an intensity value of the scene corresponding to the patterns, which was processed by a computer to generate a 2D image via a structured scanning approach. Then a 3D image was reconstructed by combining those 2D images based on photometric stereo.

The first set of detectors we used were Thorlabs PDA100A-EC silicon amplified photodetectors (PDAs). The intensity signals we obtained from them were not strong enough to generate 2D images with a good quality, which consequently lowered the 3D reconstruction qualities.

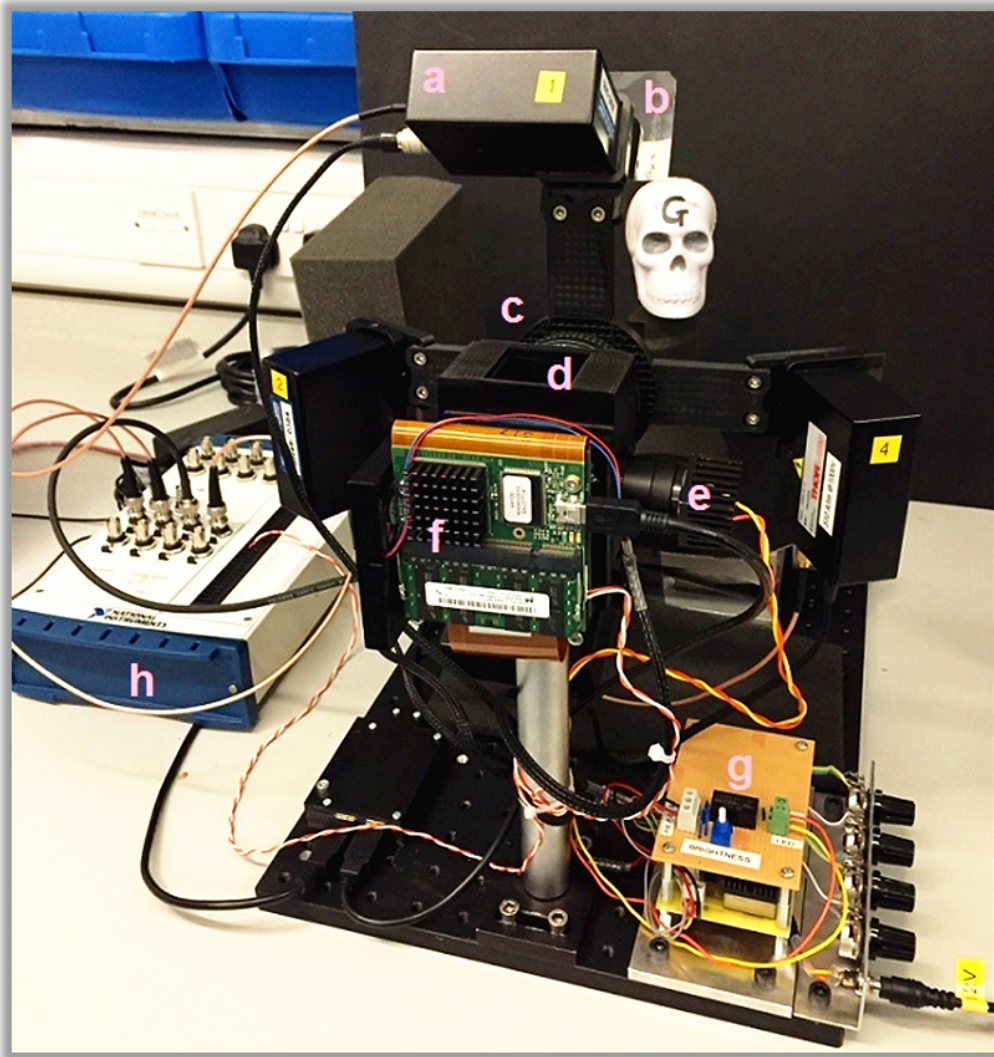


FIGURE 3.8: 3D imaging system components: a. Photomultiplier tube; b. Linear polariser sheet; c. Camera lens; d. 3D printed holder and mount with a silver mirror adjusted inside; e. White LED; f. DLP device; g. Customised electric controlling board; h. DAQ device.

To enhance the signals, we replaced them with Thorlabs PMM02 photomultiplier tubes(PMTs), which are able to detect faint optical signals from weakly emitting sources. The response spectrum of this type of PMT is between 280nm and 850nm with a peak value at 420 nm, and its response speed is significantly faster than

the PDA's 2.4 MHz. The parameters of both PMTs and LED were set through a customised electric controlling board beforehand (see Fig. 3.8).

The object we tested in our experiment was a polystyrene skull model and the surface was nearly Lambertian but also produces a component of specular reflected intensity which can easily saturate the detected signals and lead to poor reconstruction quality. It is well known that light that has been specularly reflected retains its polarization, while light that has been multiply scattered loses its polarization [109]. Therefore, to get rid of the specular reflection from the scene, some cross polarisers were applied in our system. A linear polariser sheet was attached in front of each PMT at a particular orientation, e.g. transmitting vertical polarisation, with another polariser placed on the projection lens and orientated orthogonally, e.g. transmitting horizontal. The use of crossed-polarisers isolates specular reflections at the expense of detected intensity.

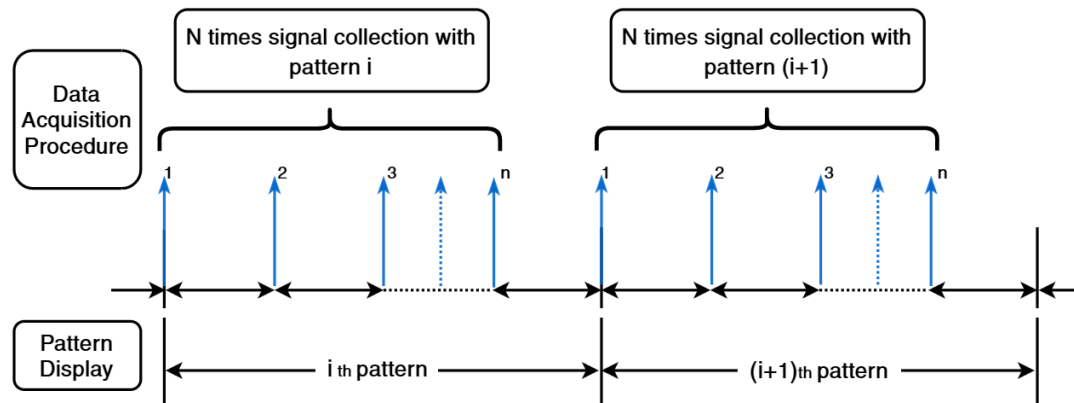


FIGURE 3.9: Signal acquisition procedure. For each projected pattern, there are N times signal collection, where N is related to DAQ speed, the number of detectors, and the exposure time of each pattern.

Once a pattern was on display, an order from the computer was sent to a data acquisition board to collect signals from PMTs. The device used in this experiment was a National Instrument portable USB DAQ (NI USB-6221) with a maximum

acquisition rate of 250 kHz for all channels. Depending on the signal acquisition rate, the number of channels used, and the exposure time required for each pattern, the number of measurements obtained for each pattern could be more than one, which allows averaging to be performed to reduce noise (see Fig. 3.9).

In our system, there were four PMT detectors connected to the DAQ device with each having a sample speed of $(250/4)=62.5$ kHz. Each illumination pattern was set to display for $50 \mu\text{s}$ (5×10^{-5} s). In which case, there were $(62.5 \times 10^3 \times 5 \times 10^{-5}) \approx 3$ times of signal collection for every single pattern and the mean was used as its signal intensity value.

To obtain a 32×32 pixel resolution 2D image with our single-pixel system following the differential signal acquisition approach, 1024 (32×32) structured Hadamard patterns plus the same amount of the inverse patterns were required, which added up to a total amount of 2048 patterns. With the DMD we used, it took the system at least $2048/(22 \times 10^3) \approx 0.09\text{s}$ for deducing such an image. By analogy, it took $\sim 0.37\text{s}$ ($8192/(22 \times 10^3)$) and 1.49s ($32768/(22 \times 10^3)$) for our single-pixel system to gain a 64×64 2D image and a 128×128 2D image. After generating those 2D images, some additional computational time was needed for our single-pixel system to perform 3D reconstruction.

3.3.2 Results

In one experiment a static object was used, in this case a polystyrene toy skull (see Fig. 3.10), and imaged with the system at 32×32 , 64×64 , and 128×128 pixels respectively, as shown in Fig. 3.11. The total frame rates, including data acquisition and reconstruction are also provided.

As employed in other single-pixel imaging work, Hadamard matrices were chosen as the sampling basis to provide structured illumination. For this type of scanning approach, where the illumination power is spread across the entire DMD, the

detection SNR for Hadamard scanning compares better than raster scanning, since the available intensity is $(N/2 - 1)$ times greater [110].

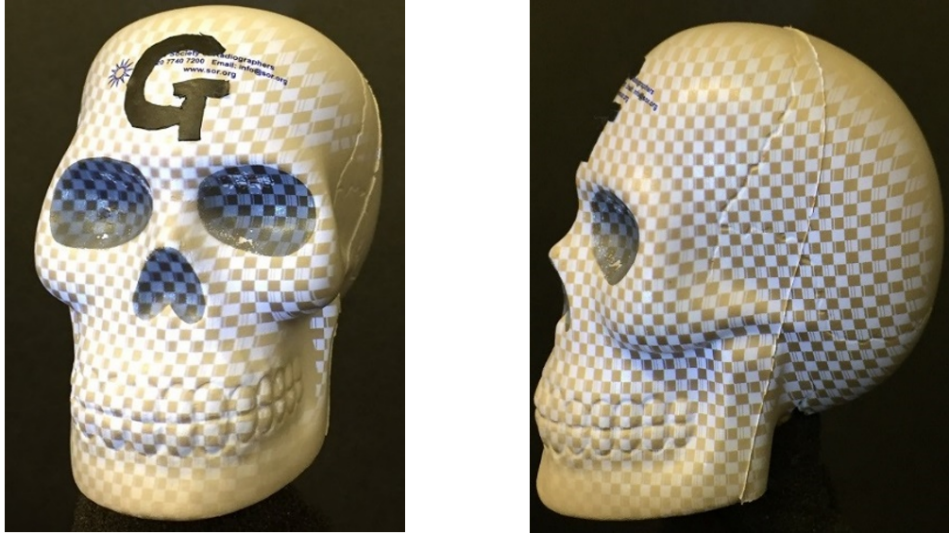


FIGURE 3.10: A structured light pattern projected onto the front and side of a skull through the camera lens. The size of skull was 80mm in length, 60mm in width and 80mm in depth

For 3D reconstruction at those three resolutions with the differential signal acquisition approach, it took our single-pixel system respectively 0.11s (32 x 32), 0.42s (64 x 64), 2s (128 x 128) in total each time. The 3D reconstruction frame-rate, as expected, was reduced when increasing the image resolution which improved the image quality.

After the first investigation, we then tested a moving object with real-time 3D reconstruction. In Fig. 3.12, a sample of video frames are shown, containing a moving skull, 3D reconstructed with our single-pixel system, at 64 x 64 pixel resolution with a frame rate of ~ 2.4 Hz. It employed a full set of 4096 illumination pattern pairs derived from a Hadamard matrix. As shown in the figure, the skull was only physically rotated in a horizontal direction. Since the 3D shape of the skull was calculated, the rendered 3D model could be orientated arbitrarily and in

this demonstration was rotated in the vertical direction at an angle range of -45 to 45 degrees in the meantime. Compared to the previous research in Ref. [56], our system enabled 3D single-pixel imaging in a video rate.

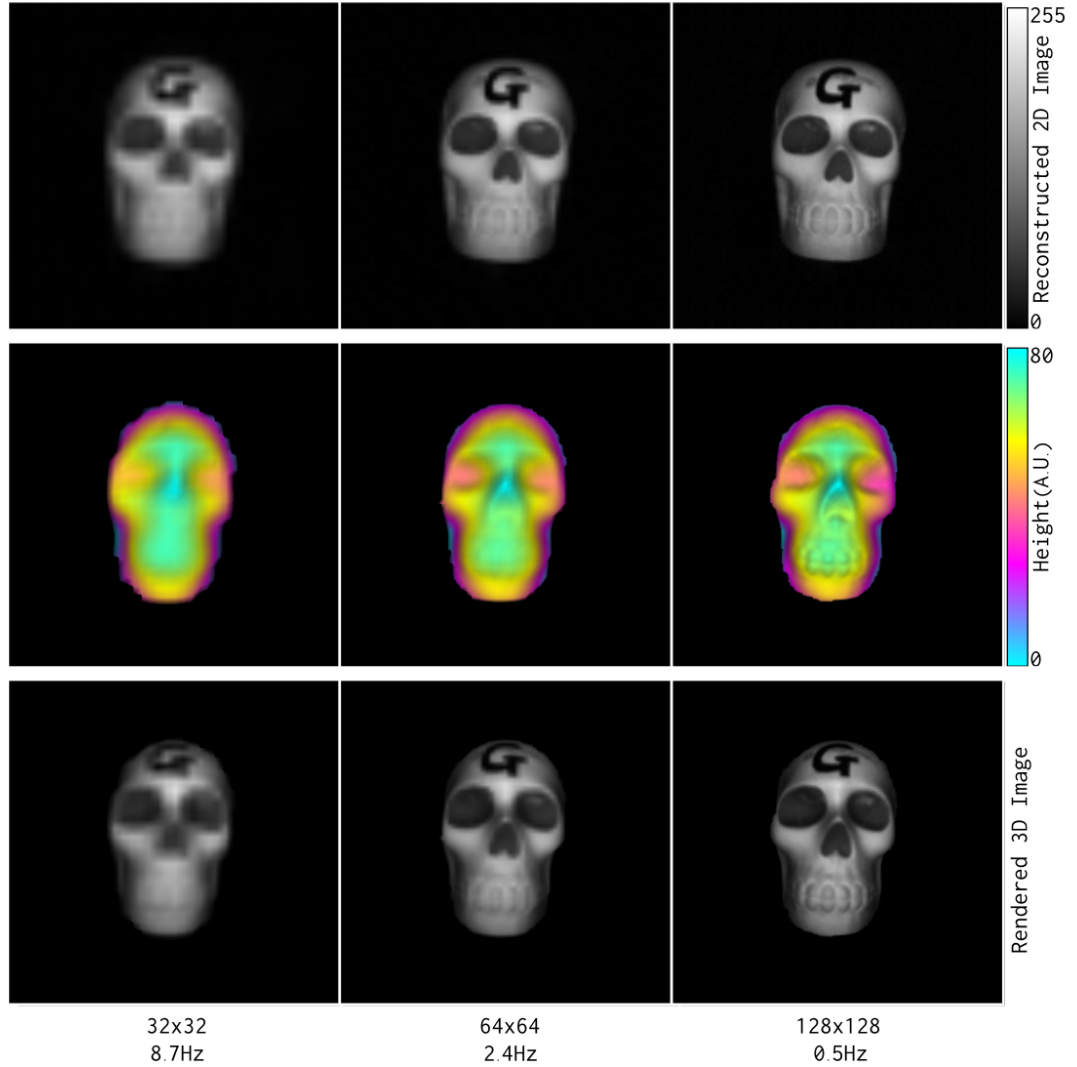


FIGURE 3.11: Comparison of 3D reconstructions of a static object using structured illumination at different image resolutions. With illumination resolutions of 32x32, 64x64, and 128x128 generated from Hadamard structured pattern pairs, the corresponding video sample rate of the 3D reconstructed skull with our system are: 8.7Hz, 2.4Hz, and 0.5Hz.

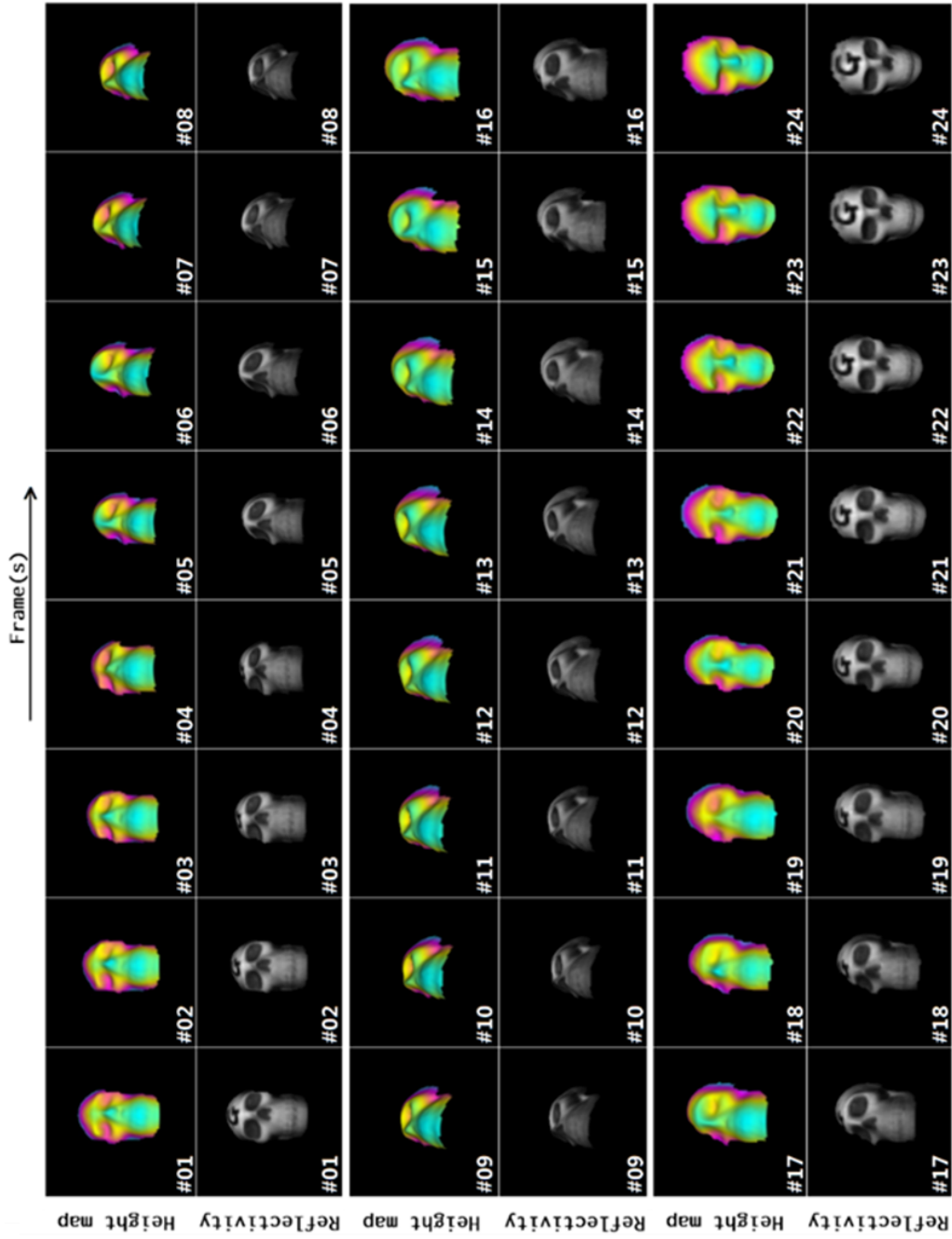


FIGURE 3.12: Video sample of a 3D reconstructed moving object. Each 64x64 pixel resolution frame was acquired at approximately 0.42s with a skull physically rotating in a horizontal direction along with its 3D reconstructed model rotating in the vertical direction, using a full set of 4096 Hadamard pattern pairs.

3.4 Evolutionary compressed sensing

To improve the video sample rate of 3D reconstructed moving objects while remaining a high image quality, we then experimented to apply compressive sensing algorithms into our system.

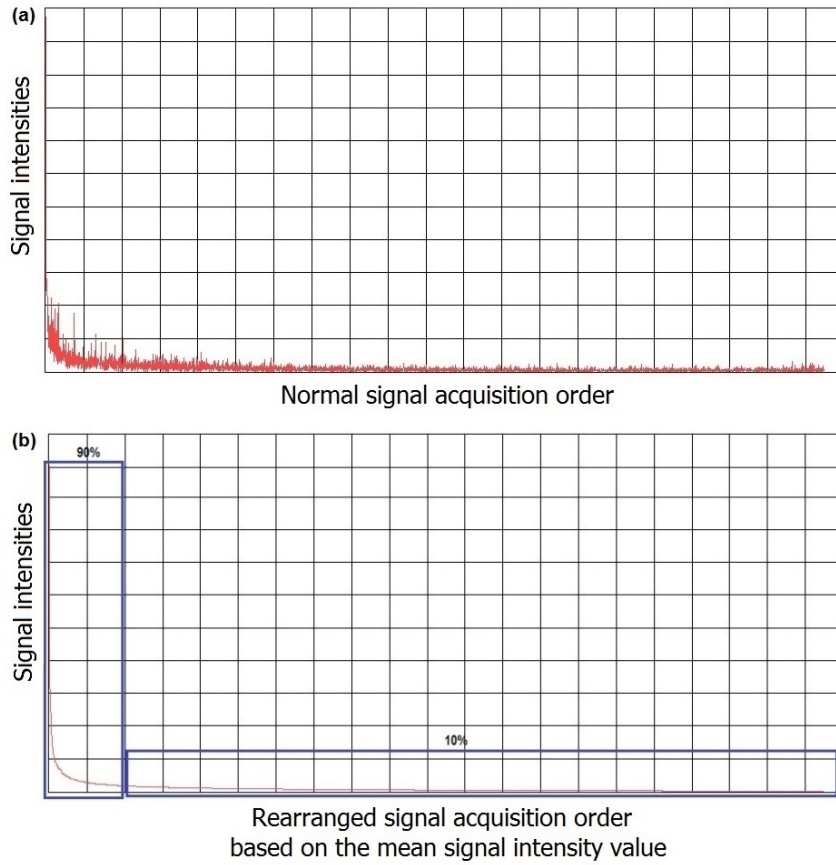


FIGURE 3.13: Pattern rearrangement with compressive sensing algorithm: (a) before reordering the illumination patterns; (b) rearranged illumination patterns based on its corresponding signal intensity.

In choosing the optimal strategy we noted that typical images could be represented by a subset of Hadamard patterns instead of a complete pattern set, and continuing adjacent frames were practically coincident to each other with only slight variations. Therefore, we rearranged the Hadamard pattern pairs based on their corresponding mean signal intensities from those four detectors, and utilized only

a certain amount of the top-ranking patterns to form the 2D images instead of using all of them. By using this approach, 90% of the subset of illumination patterns were chosen from the top-ranking rearranged pattern set, while the other 10% were randomly selected from the remainder of the complete pattern to adjust the slight variations between adjacent frames, as shown in Fig. 3.13. Following that, images generated with those pattern subsets were combined with photometric stereo to obtain 3D images. This strategy with selection of a subset of Hadamard patterns is a compromised approach to maintain a high frame-rate without decreasing the spatial resolution.

TABLE 3.1: Relative RMS error comparison at 128x128 pixel resolution with five different compression ratios.

128x128 3D single-pixel imaging with compressive sensing					
Patterns Used	12.50%	25%	50%	75%	100% (Ground truth)
Relative RMS Error	3.586	3.443	2.229	1.571	0

With implementation of this compressive sensing algorithm, we reconstructed an object at 128 x 128 pixel resolution by using five different amounts of structured pattern pairs: 16384 pattern pairs, 12288 pattern pairs, 8192 pattern pairs, 4096 pattern pairs, and 2048 pattern pairs, which were equivalent to 100% (zero-compression), 75%, 50% 25% and 12.5% compression ratio (see Fig. 3.14), and we compared the relative root-mean square (RMS) errors of the height values in those 3D reconstructed objects with the one using zero-compression (see Table 3.1). The 3D reconstructed object with zero-compression in this figure is the same as the 3D reconstruction of 128x128 pixel resolution in Fig. 3.13. The result showed that the RMS error of the object height value, as expected, increased when using less pattern pairs (higher compression), while the video sample rate was correspondingly improved.

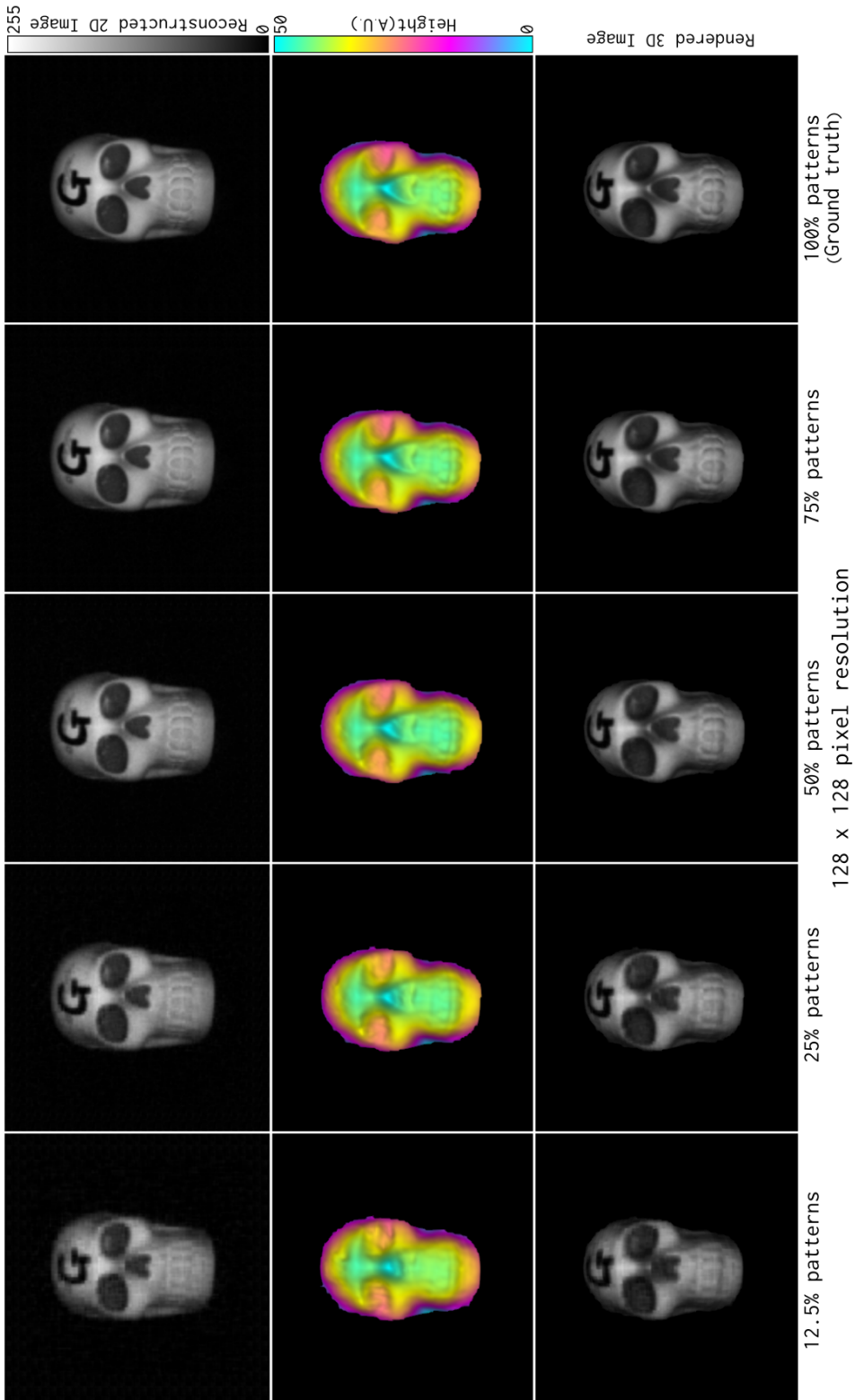


FIGURE 3.14: Comparison of 3D imaging using evolutionary compressed sensing. The skull was reconstructed at 128x128 pixel resolution with five different compression ratios: 12.5%, 25%, 50%, 75% and 100%.

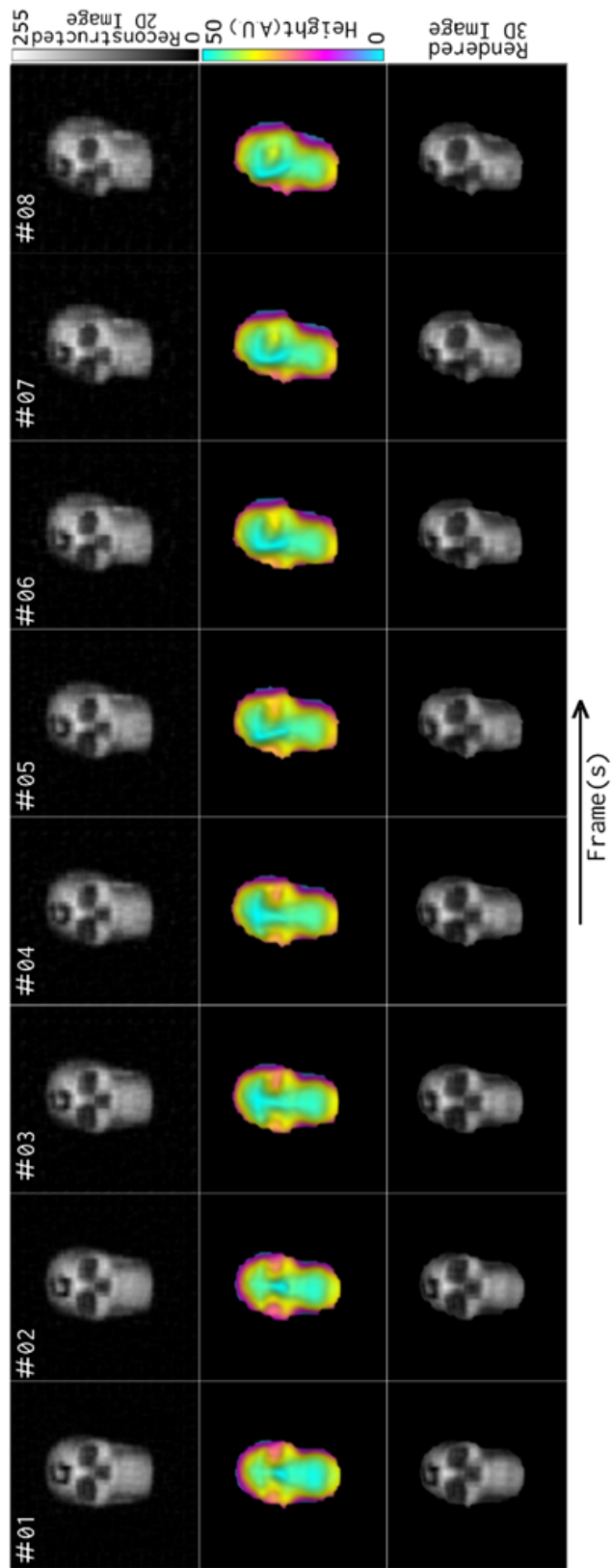


FIGURE 3.15: A one second video clip of video frames containing a moving skull 3D reconstructed at a 64x64 pixel resolution by using the compressive sensing algorithm. Each 3D frame was produced based on 1024 patterns (25% compression ratio) acquired at a frame rate of $\sim 8\text{Hz}$ for an object physically rotated in a horizontal direction.

3.5 Real-time video from a 3D single-pixel computational imaging system

Fig. 3.15 shows a video clip of a moving skull at 64x64 pixel resolution with compressive sensing in one second. Each frame was produced based on 1024 illumination patterns (25% compression ratio), which was equal to the same amount of patterns for a zero-compression 32 x 32 pixel resolution 3D image reconstruction. The frame rate of this 3D video was 7.6Hz, approximately 4 times faster than the zero-compression 64 x 64 pixel resolution one.

We also experimented 3D reconstructed a moving skull at 128x128 illumination resolution by using 4096 Hadamard patterns (25% compression ratio). The video frame rate was increased to 0.9Hz, as seen in Fig. 3.16. The object was physically rotated in a horizontal direction whilst the 3D reconstructed model was set to rotate in the vertical direction at an angle range of $[-30, 30]$, since we were fully aware of the height map of the skull. We noticed that the frame rate in this video was restrained due to the fact that the 3D reconstruction process at 128 x 128 pixel resolution started to play an important role in overall time performance when the 2D reconstruction time decreased with the compressive sensing algorithm.

Furthermore, to have a quick test on our system robustness, we replaced the skull with a smaller object which were made with a different material. In Fig. 3.17, it demonstrates a 10-second real-time video clip of a rotating Santa toy which was reconstructed at 128x128 illumination resolution at a frame rate of 0.9Hz by using 4096 Hadamard patterns (25% compression ratio). The result showed that our system had a characteristic of good robustness.

During the experiment, we noted that with the displayed patterns being reduced exponentially the corresponding imaging rates were not increased proportionally. We then recorded the elapsed time of full pattern display, 2D reconstruction, and

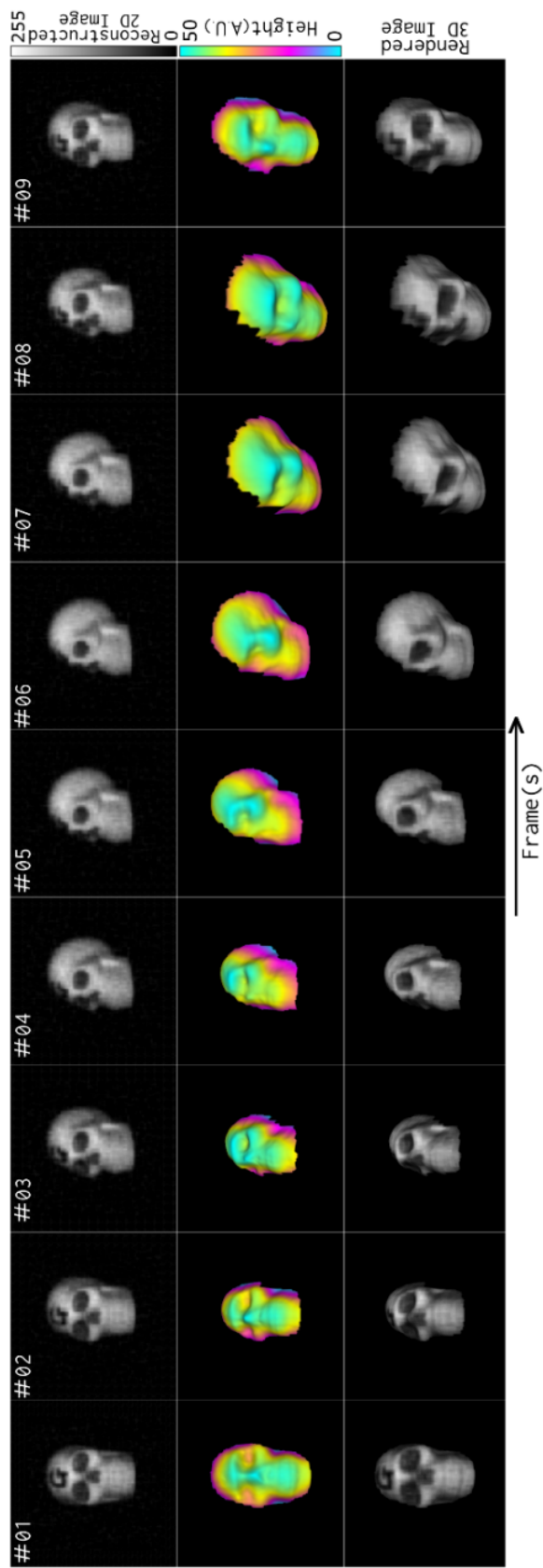


FIGURE 3.16: a 10-second video clip of a moving skull at 128x128 pixel resolution with compressive sensing. Each frame was produced based on 4096 patterns (25% compression ratio) acquired at approximately a frame rate of $\sim 1\text{Hz}$ for an object physically rotated in a horizontal direction and the 3D reconstructed model rotated simultaneously in the vertical direction at an angle range of $[-30, 30]$ degrees.

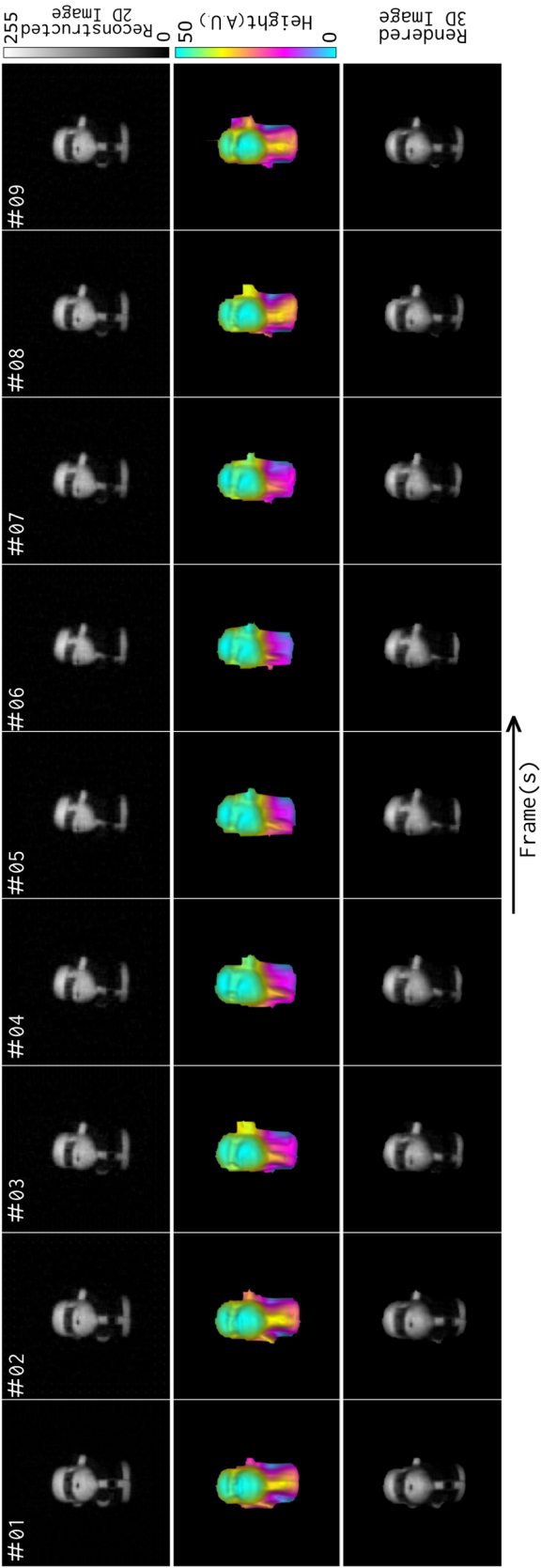


FIGURE 3.17: A 10-second video clip of a moving Santa at 128x128 pixel resolution with a 25% compression ratio (4096 patterns).
The video frame rate was $\sim 1\text{Hz}$ with the Santa object is physically rotated a horizontal direction.

3D reconstruction respectively at those three different resolutions, with results shown in Fig. 3.18.

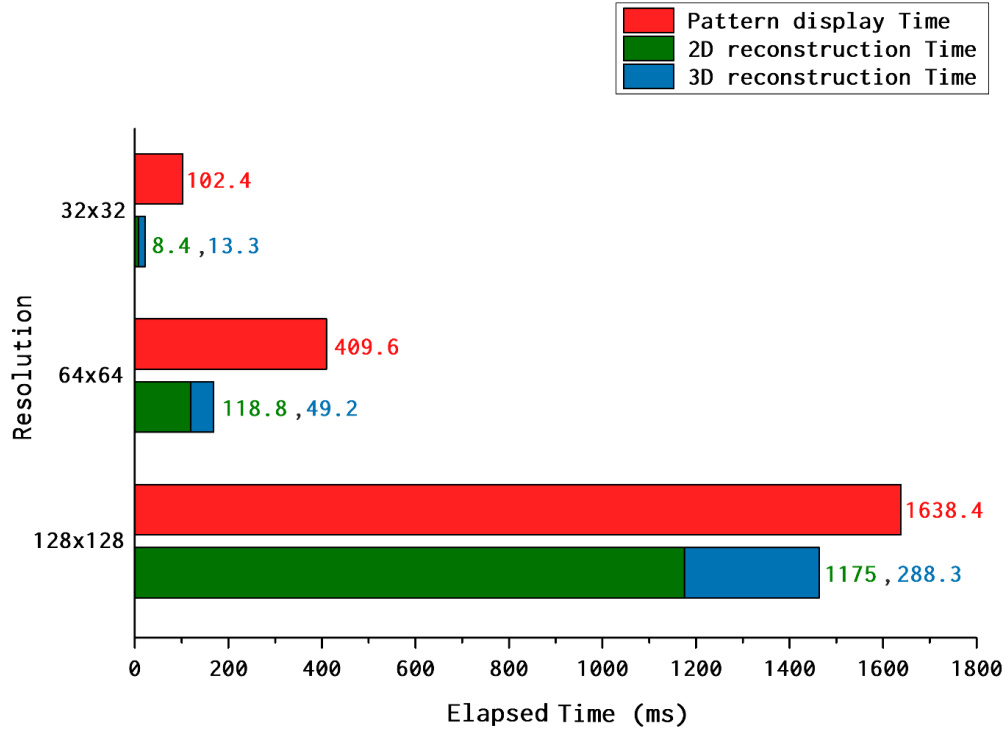


FIGURE 3.18: Elapsed-time comparison without compressive sensing algorithm at three different resolutions: 32x32, 64x64, 128x128. Each pattern was displayed for 50ms following by displaying its inverse as well. The corresponding total pattern display time, its 2D reconstruction time, and 3D reconstruction time were recorded respectively.

The result indicated that the total elapsed time of 2D and 3D reconstruction become noticeable when increasing the image resolution. By employing the compressive sensing algorithm, the time required for both pattern display and 2D reconstruction was decreased due to the smaller amount of illumination patterns. However, the time for 3D reconstruction remained unchanged as the illumination resolution was still the same, which limited the total compression rate. One possible method to improve the 3D reconstruction efficiency of our system in the future is to perform the 3D reconstruction operation with a dedicated high-performance GPU instead of the computer's CPU.

3.6 Conclusion

In conclusion, we have experimentally demonstrated a video-rate 3D imaging system based on photometric stereo which exhibited perfect pixel registrations by utilising four single-pixel detectors with a high-speed structured illumination. Instead of using raster scanning or light projection with random patterns, we employed structured illumination patterns derived from Hadamard matrices which improved the image SNR and the system sampling efficiency. A differential signal acquisition approach was introduced to eliminate the illumination background noise. As an extension of the previous work in Chapter 2, we have now shown continuous operation for 3D image reconstruction at illumination resolutions of 32x32, 64x64 and 128x128 pixels, with video frame rates of 8.7 Hz, 2.4 Hz, and 0.5 Hz, respectively.

In addition, we made use of a compressive sensing algorithm to speed-up the system sampling rate. In place of using a full illumination pattern set, we rearranged the Hadamard pattern pairs based on the corresponding signal intensities and only chose a certain amount of top-ranking pairs to generate the 2D images. For dynamic scenes the noise introduced randomly affects each image used for photometric stereo, which in some circumstances could be advantageous. This approach increased the frame rate at the expense of only a modest reduction in image quality, as evidenced by a quantitative analysis. For 3D reconstruction with 64x64 and 128x128 illumination resolution, a 25% compression ratio provided an increment of the video frame rates to ~ 8 Hz and ~ 1 Hz respectively, comparing to ~ 2.4 Hz and ~ 0.5 Hz with full pattern set sampling. We noted that with 128x128 illumination resolution, the total computational time for 2D and 3D image reconstruction, as performed on an octa-core processor, placed a limit on the achievable frame rate. Further improvement could be achieved by operating the 3D reconstruction on a dedicated high-performance GPU.

Since our 3D single-pixel imaging system does not rely on a pixelated camera sensor and the operational bandwidth of the DMD device extends beyond the visible (VIS) range, it can be readily extended to other wavelengths where traditional camera technology is prohibitively expensive, such as the near-infrared (NIR) wavelength.

Chapter 4

NIR Single-Pixel Imaging with Micro-Scanning

4.1 Introduction

Following up with the 3D single-pixel imaging experiment in chapter 3, the single-pixel detector was replaced with an Indium-Gallium-Arsenide (InGaAs) detector, sensitive to shortwave infrared wavelengths (800nm-1800nm). Importantly however, the increased detector noise compared to silicon counterparts, in addition with scarcity of high-power sources at non-visible wavelengths, limits the performance of the computational imaging system in application. This has led to the development of novel scanning strategies and reconstruction algorithms to improve the resulting image quality, the topic of which forms the basis of this chapter. Micro-scanning, sometimes termed super-sampling, is a well established method to recover higher-resolution images from a set of lower-resolution acquisitions [111]. In this work, we employed this optimisation approach to sample the spatial properties of an image at a moderate resolution, from which a higher-resolution image can be recovered using an iterative or inversion algorithm. We imaged different

objects by using various NIR bandpass filters to detect hidden information underneath the surface for potential practical applications. We also compared our imaging system with an expensive commercial InGaAs NIR camera, which indicates a reasonably well capability of revealing the water absorption underneath the surface of plant leaves and fruits.

4.1.1 Contributions

The work carried out in this chapter stems from a collaboration between Miles Padgett, Graham Gibson, Matt Edgar and myself. The design and development of the experimental hardware used to perform this experiment was the result of myself, with support from Graham Gibson and Matthew Edgar. The micro-scanning strategy was first experimentally demonstrated by Mingjie Sun, and subsequently incorporated within this system by myself and Matthew Edgar. I performed the experiments and analysed the results.

4.2 NIR single-pixel detection

As we described in Chapter 3, the wide operational bandwidth of DMD and single-pixel detectors enable the application of computational imaging schemes beyond visible wavelengths. In this chapter, we are focused on 2D single-pixel imaging at shortwave infrared wavelengths.

Instead of using four VIS single-pixel detectors and a white LED in the 3D imaging system, we displaced with one InGaAs NIR detector and an infrared light source. To protect the DMD from being over-heated by the infrared light source, we chose to perform single-pixel imaging with structure detection over structured illumination, in which case the light were irradiated to the scene first and then reflected to the DMD. The DAQ sample speed for the NIR detector was 250 kHz

and the DMD display time for each pattern was still 50 μ s. Consequently, there were $(250 \times 10^3 \times 5 \times 10^{-5}) \approx 12$ times of signal collection for every single pattern and the mean was used as its signal intensity value. We applied a set of NIR bandpass filters before the detector to examine if certain information underneath the object surface were easier to be observed at some NIR wavelength range, for instance, the overripe or rotten portions of fruits. A micro-scanning algorithm was also employed to improve the image quality.

4.2.1 Experimental setup of NIR imaging system

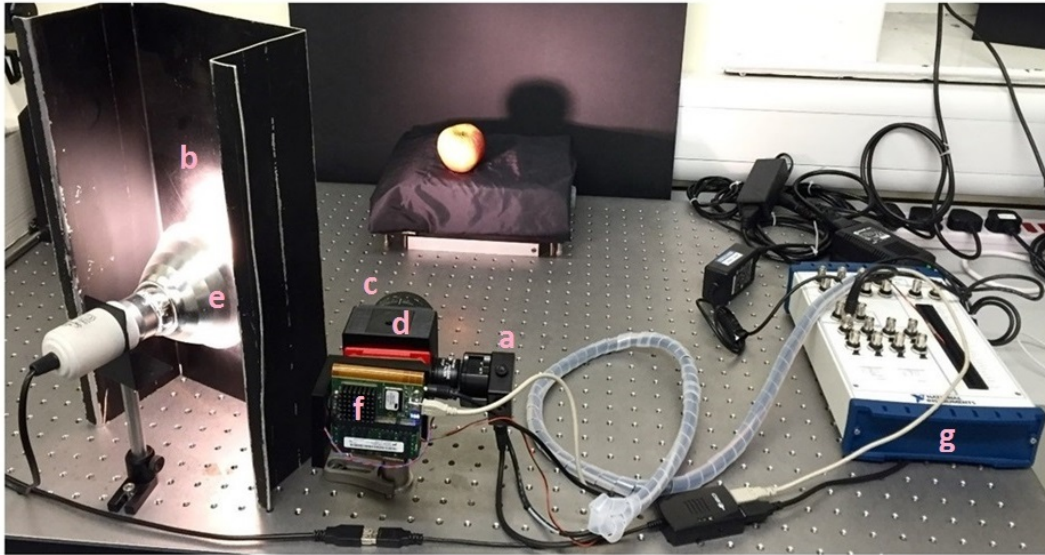


FIGURE 4.1: 2D single-pixel NIR imaging system setup: a. InGaAs NIR detector; b. Wire grid polariser on glass substrate fixed on a metal board; c. Camera lens; d. 3D printed holder and mount with a silver mirror adjusted inside; e. IR heat lamp; f. DLP device; g. DAQ device.

The system setup is shown in Fig. 4.1. A 50 mm x 50 mm wire grid polariser on glass substrate (Thorlabs) with a spectral range of 250 nm - 4000 nm and a high radiation threshold was fixed on a black metal board placed in front of an infrared (IR) light source together with a polariser sheet attached behind a 50mm

focal camera lens at a relatively perpendicular direction to remove the specular reflection.

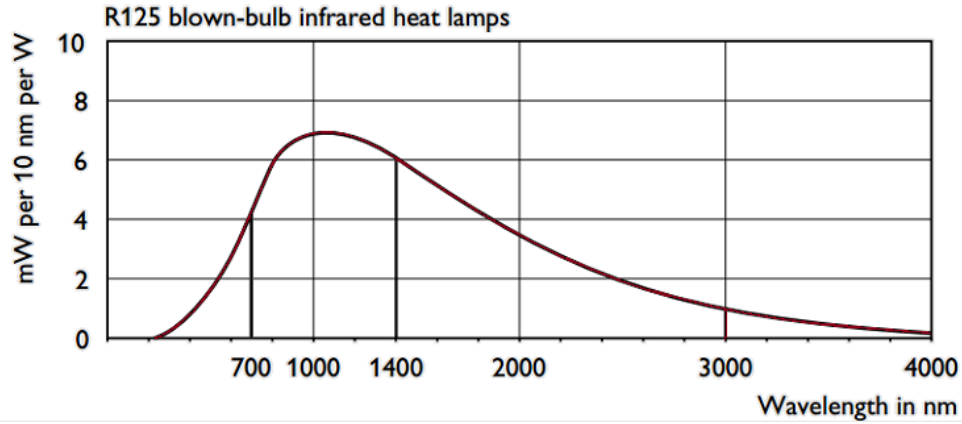


FIGURE 4.2: The spectral power distribution of infrared heat lamp R125 [112].

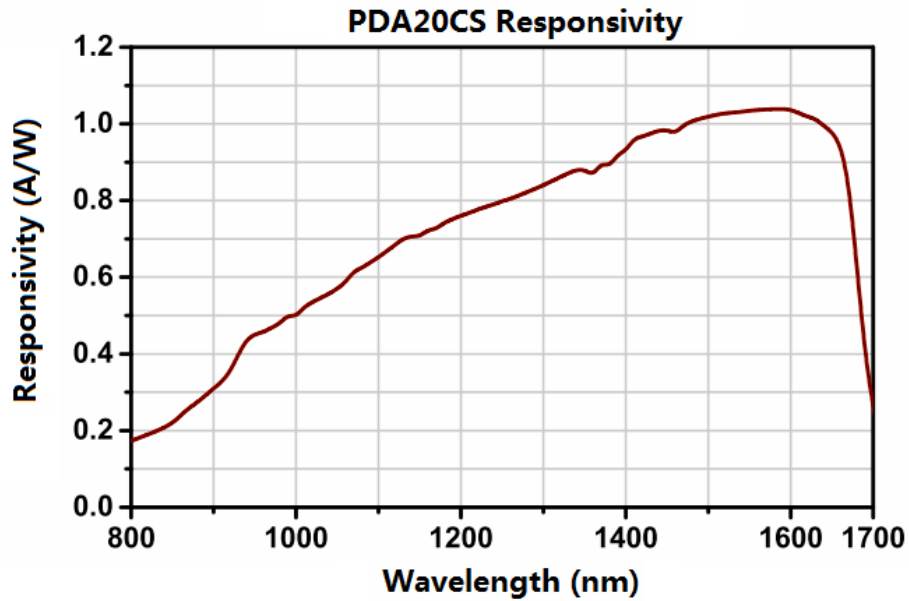


FIGURE 4.3: PDA20CS InGaAs detector spectral response curve [113].

A lens tube filter holder was mounted in front of the detector allowing different bandpass filters to get changed easily. The illumination source we adopted in the system was an IR heat lamp (Philips R125), 175W, with the spectral power distribution shown in Fig. 4.2 [112]. And the NIR detector we chose was

a (Thorlabs) PDA20CS - InGaAs Switchable Gain Amplified Detector, 10 MHz BW(bandwidth), with a spectral response curve shown in Fig. 4.3 [113].

4.2.2 NIR comparison with and without using polarisers

In one experiment an apple was imaged with and without the use of crossed polarisers to investigate the role of specular reflections and the influence on the signal to noise ratio in the image reconstruction.

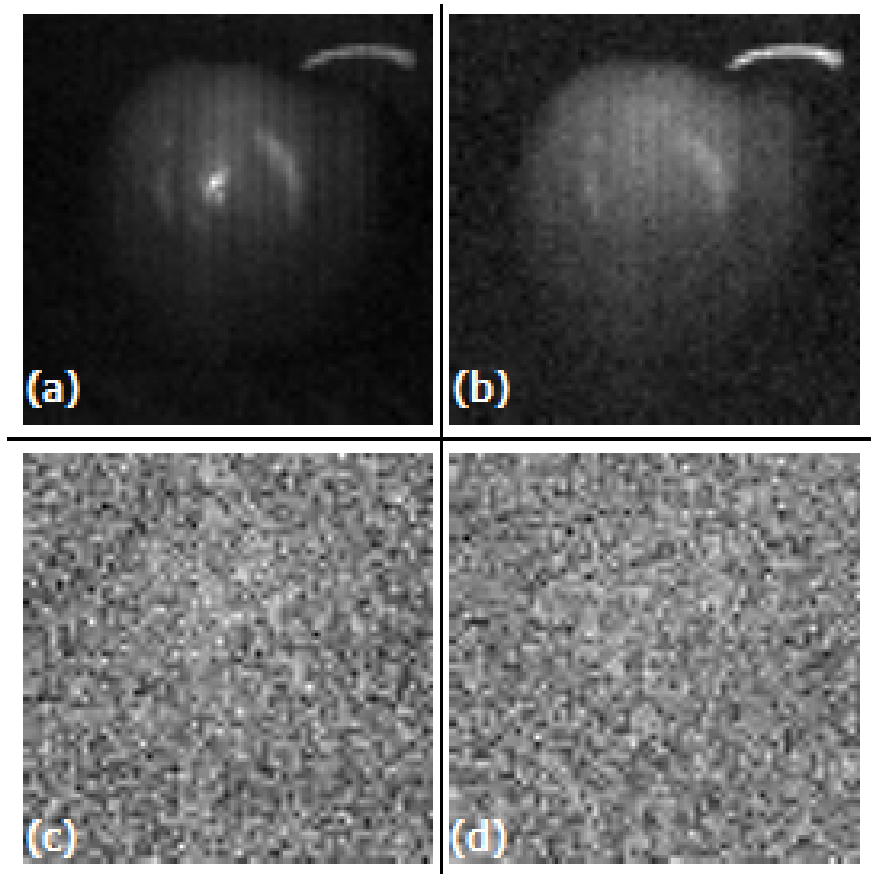


FIGURE 4.4: NIR outcomes comparison with and without polariser at a detection resolution of 64×64 . Image (a) without polariser and (b) with polariser were captured with no bandpass filter and no averaging. Image (c) without polariser and (d) with polariser were obtained with an NIR bandpass filter of 1000 ± 5 nm and 10-frame averaging.

In this experiment the Hadamard basis was used to scan the image, as described in last chapter. In addition, an IR bandpass filter of 1000 ± 5 nm (a wavelength range with the highest spectral power from the lamp) was attached in front of the detector to obtain images with 10-frame averaging. The outcomes were compared and shown in Fig. 4.4. It is worth to mention that, in all the NIR single-pixel imaging experiments, we used full Hadamard patterns for the image reconstruction rather than selecting the top-ranking subset as we did in the last chapter. This is because the purpose of the NIR single-pixel imaging experiments was to get a good quality image rather than increasing the frame rate, and by using only the top-ranking subset it would lower the image quality.

From the results, we noticed that although using the polarisers removed the specular reflection, it also reduced the light that projected onto the scene (see Fig. 4.4(a), (b)). With an extra bandpass filter (e.g. 1000 ± 5 nm), the amount of light reached to the detector was too small to form a good image, even after image averaging (see Fig. 4.4(c), (d)). For this reason, it was determined that the use of crossed linear polarisers to isolate specular reflections led to a dramatic reduction in the detected intensity and thus could not be used in this system.

4.2.3 NIR imaging with bandpass filters

We then tested some objects by employing a set of different NIR bandpass filters without the polarisers at the same resolution of 64×64 . The filter set (Thorlabs FKB-IR-10) included 10 different bandpasses: 850 ± 5 nm, 900 ± 5 nm, 1000 ± 5 nm, 1100 ± 5 nm, 1200 ± 5 nm, 1300 ± 6 nm, 1400 ± 6 nm, 1500 ± 6 nm, 1550 ± 6 nm, and 1600 ± 6 nm. Two infrared heat lamps were set on the left and right side of the DMD to provide the illumination. Images acquired after each bandpass filter was averaged 100 times to generate a better image with reduced noise. The outcomes of imaging an apple with the filter set are shown in Fig. 4.5. In this case it was necessary to average 100 frames in order to improve the image SNR. This

led to the development of alternative sampling strategies to enhance the system performance.

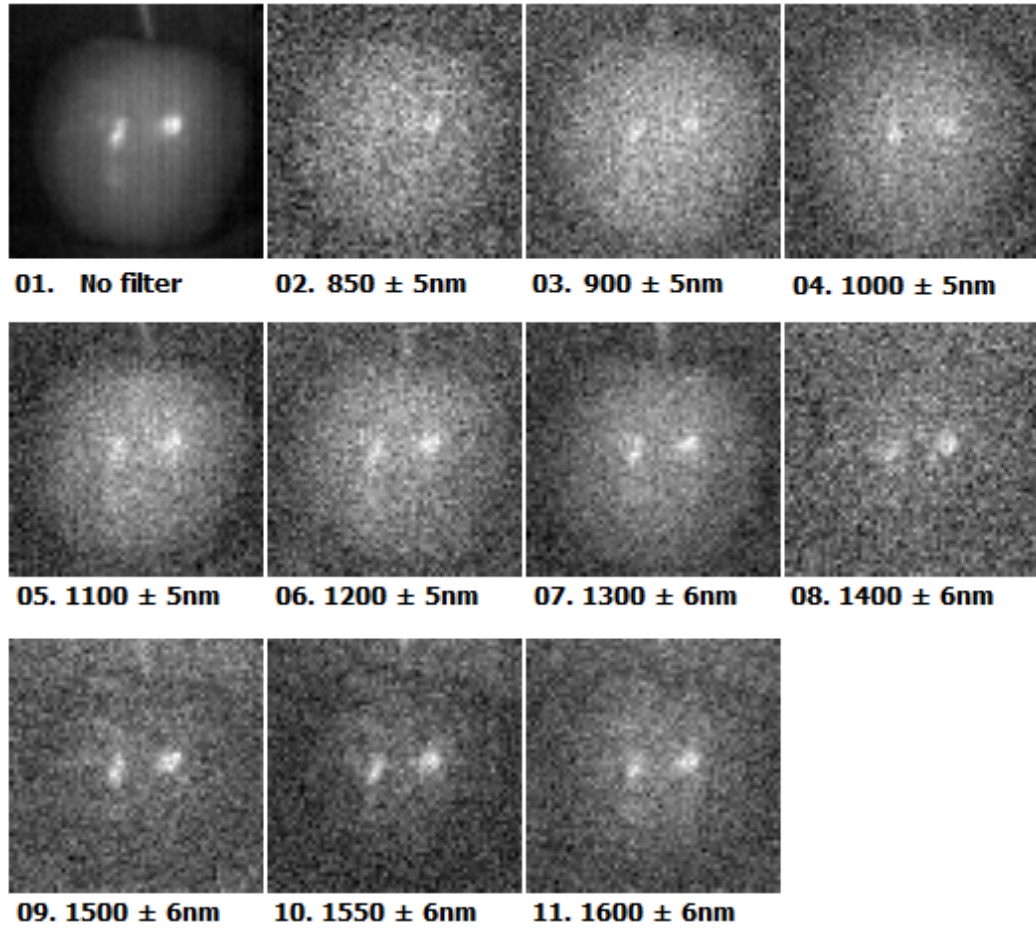


FIGURE 4.5: NIR imaging of an apple with a set of 10 bandpass filters at a detection resolution of 64×64 . No polariser was assembled here to increase the intensity of the light received by the detector. Image (01) was captured without using any bandpass filter. From image (02) to image (11), each one was obtained with the use of a different NIR bandpass filter as shown in the figure and 100-frame averaging.

4.3 Micro-scanning imaging

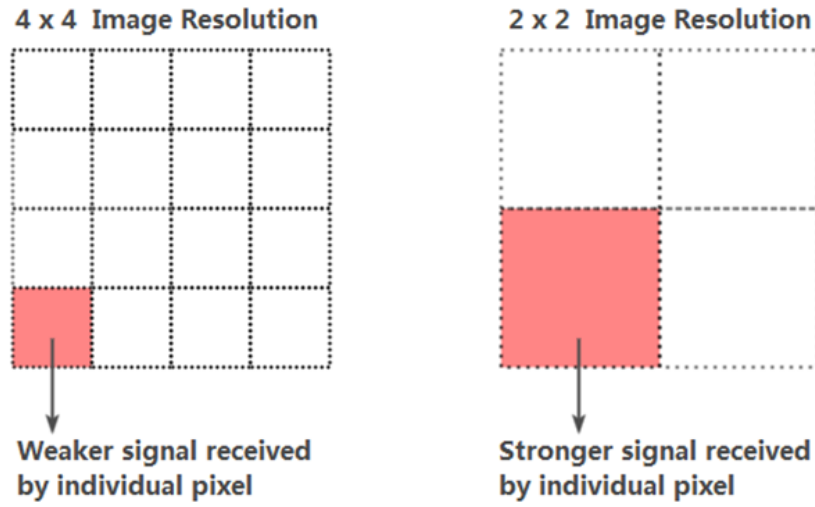


FIGURE 4.6: The trade-off between image resolution and signal-to-noise (SNR) in single-pixel imaging.

For any single-pixel imaging system employing a DMD as the spatial light modulator, the number of micro-mirrors used to scan a 'pixel' in the image, determines the light intensity variation that the detector must measure. Consequently, for a detector having fixed amplitude noise, there is an inherent relationship between the maximum image resolution and the reconstruction SNR [114–116]. In other words, images of lower resolution have higher SNR than images of higher resolutions. Fig. 4.6 illustrates the change in the reflected intensity for two different scanning resolutions. The trade-off between the image resolution and its SNR put a restriction in further single-pixel applications. One approach to improve this is the adoption of differential signal acquisition as demonstrated in 3.2.3. In this system, we also employed an optimization approach called micro-scanning [111, 117–121], capable of retrieving high-resolution (HR) images from images acquired at a lower-resolution (LR) with favourable SNR. The principle of this approach is similar to the one of super-resolution structured illumination microscopy [122].

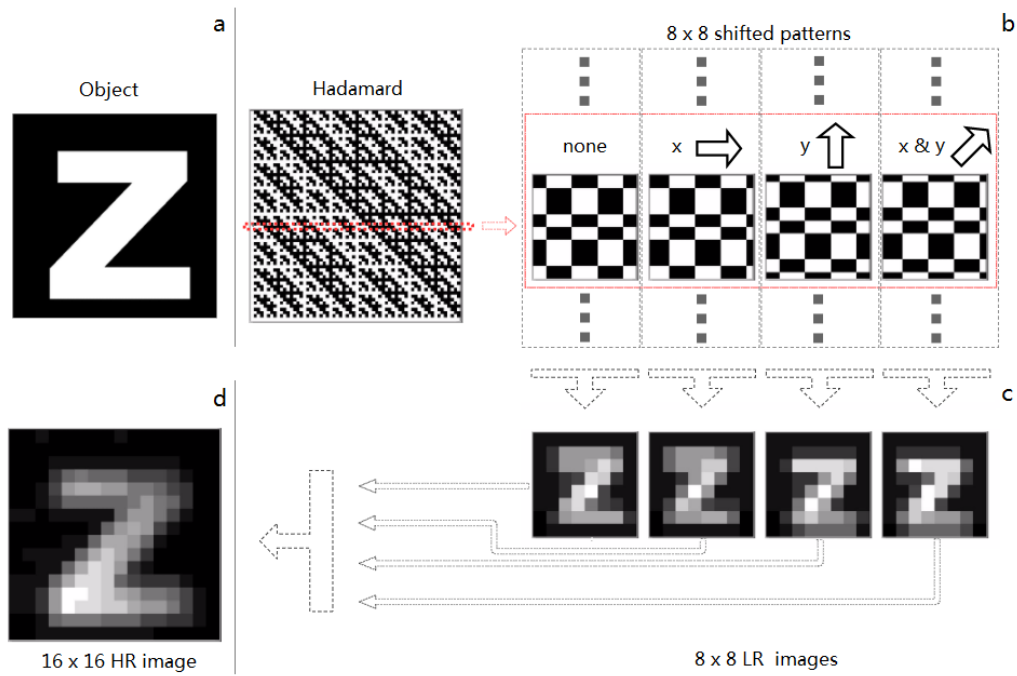


FIGURE 4.7: The micro-scanning approach. (a) the object. (b) the modification of Hadamard patterns. (c) the LR images generated through the pattern sets. (d) the NIR HR image yielded via LR images.

Figure 4.7 demonstrates the reconstruction of a 16 x 16 HR image using micro-scanning from four 8 x 8 LR images acquired from the NIR detector. In the system, we utilised the central square region of the DMD, consisting of 768 x 768 micro-mirrors, to display the patterns. Each detection block in the pattern (equal to an individual pixel) was represented by a group of micro-mirrors. For a 8 x 8 image, each pixel comprised 96 x 96 micro-mirrors in a square region. To modify the pattern set, each detection block in the patterns was then shifted half a pixel in the central square region of the DMD three times, respectively on the x-axis, the y-axis, and the x & y axes (see Fig. 4.7(b)). More to the point, all the patterns were shifted 48 micro-mirrors right on the x-axis, top on the y-axis, and top-right on the x & y axes by turns. This generated four 8 x 8 NIR LR images including the original one, with each containing different spatial information (see Fig. 4.7(c)). By re-mapping those LR images on a 16 x 16 pixel grid and averaging

the overlapped values at each grid, a 16 x 16 HR image was then reconstructed (see Fig. 4.7(d)).

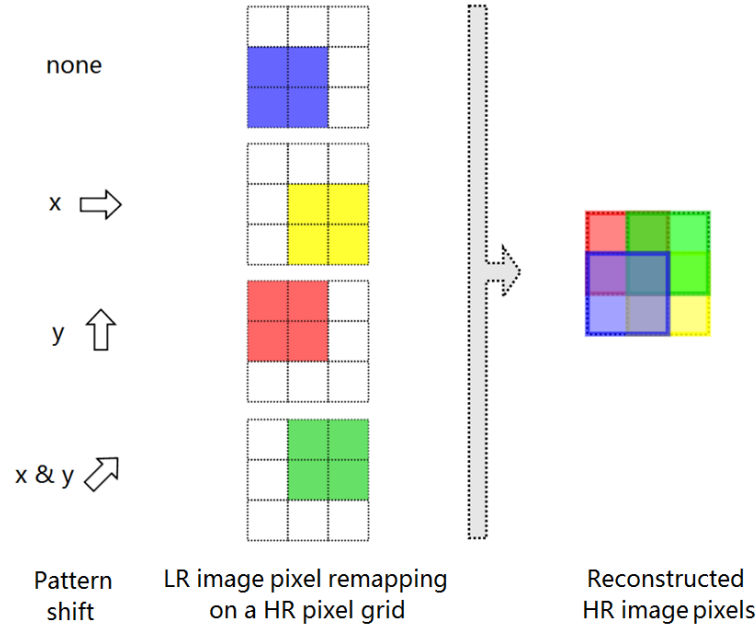


FIGURE 4.8: Micro-scanning HR image pixel reconstruction. The re-mapped LR image pixel values at the HR pixel grid were averaged to represent the HR image pixels.

Figure 4.8 illustrates the computing of a HR image with four LR images. Each coloured region corresponds to an individual LR image pixel. I_1 , I_2 , I_3 , and I_4 represent the measured pixel intensities in those LR images. The intensity value at the central pixel grid on the high-resolution grid I_c is calculated by

$$I_c = \frac{1}{4}(I_1 + I_2 + I_3 + I_4) \quad (4.1)$$

Each pixel in the HR image is generated from a specific combination of the LR pixels. Taking no account of signal noises, the reconstructed HR image of using this micro-scanning approach is mathematically equivalent to the convolution of the standard HR image that obtained with a normal Hadamard pattern set (the image resolution equals to the size of the pattern) with a kernel κ_{er} , where

$$\kappa_{er} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (4.2)$$

As this image reconstruction approach involves is analogous to smoothing adjacent pixels, the resulting HR image suffers from a moderate reduction in resolution than had the scene been scanned at a higher resolution. However, it provides a significantly improvement of the image SNR comparing to the standard method as the HR image with micro-scanning inherits the higher SNR from its lower-resolution images. Figure 4.9 demonstrates a comparison of SNR between two sets of 64 x 64 NIR images reconstructed with and without using micro-scanning. Here we calculated the SNR using [123, 124] :

$$SNR = \frac{\mu}{\sigma} \quad (4.3)$$

where μ is the signal mean and σ is the standard deviation of the noise.

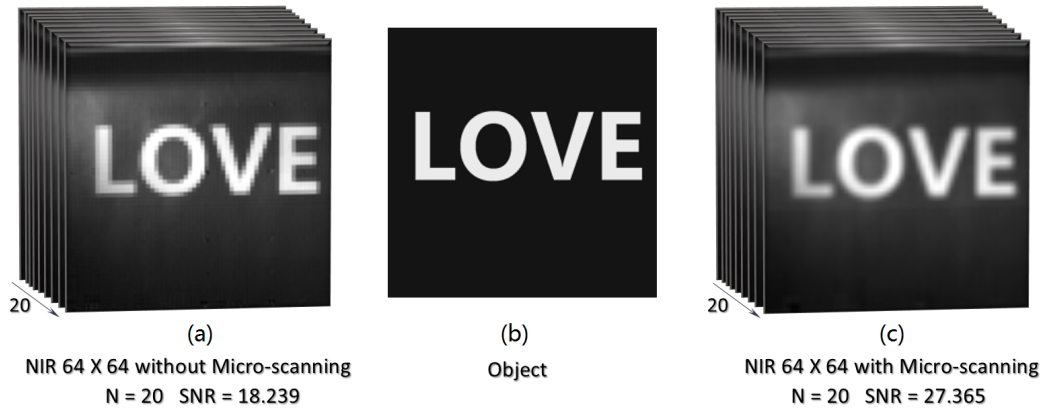


FIGURE 4.9: SNR comparison of images with and without using micro-scanning. (a) The SNR of 20 NIR 64 x 64 images reconstructed without using micro-scanning equals to 18.239. (b) The photographed object. (c) The SNR of 20 NIR 64 x 64 images reconstructed using micro-scanning improved to 27.365.

For both methods in Fig. 4.9 (a) and (c), we each generated 20 NIR 64 x 64 images [125], and then calculated the average of the mean value at each pixel in those images as μ and the average of the standard deviation of each pixel among those images as σ . The results showed an improvement of SNR with micro-scanning when reconstructing the same resolution images.

4.4 NIR micro-scanning imaging with solar radiation

In another experiment, on a clear day, solar radiation was used to illuminate a scene with more power at SWIR wavelengths than the incandescent bulb used in the laboratory. The same set of NIR bandpass filters, as mentioned in 4.2.3, was employed in front of the detector separately. Figure 4.10 exhibits the imaging outcomes of four different fruits obtained under the sunlight at a resolution of 128 x 128 using micro-scanning with 50-frame averaging. The reason for employing 50 times frame averaging in this experiment was that the intensity of light passing through each NIR bandpass filter was too low compared to the environment noise. By doing the averaging increased the SNR, which was explained in 5.4.1.

From those images, we noticed that with the 1400 ± 6 nm bandpass filter all four objects were unable to be detected. The spectral power distribution of the sun arriving on the surface of the earth was then explored [126–129].

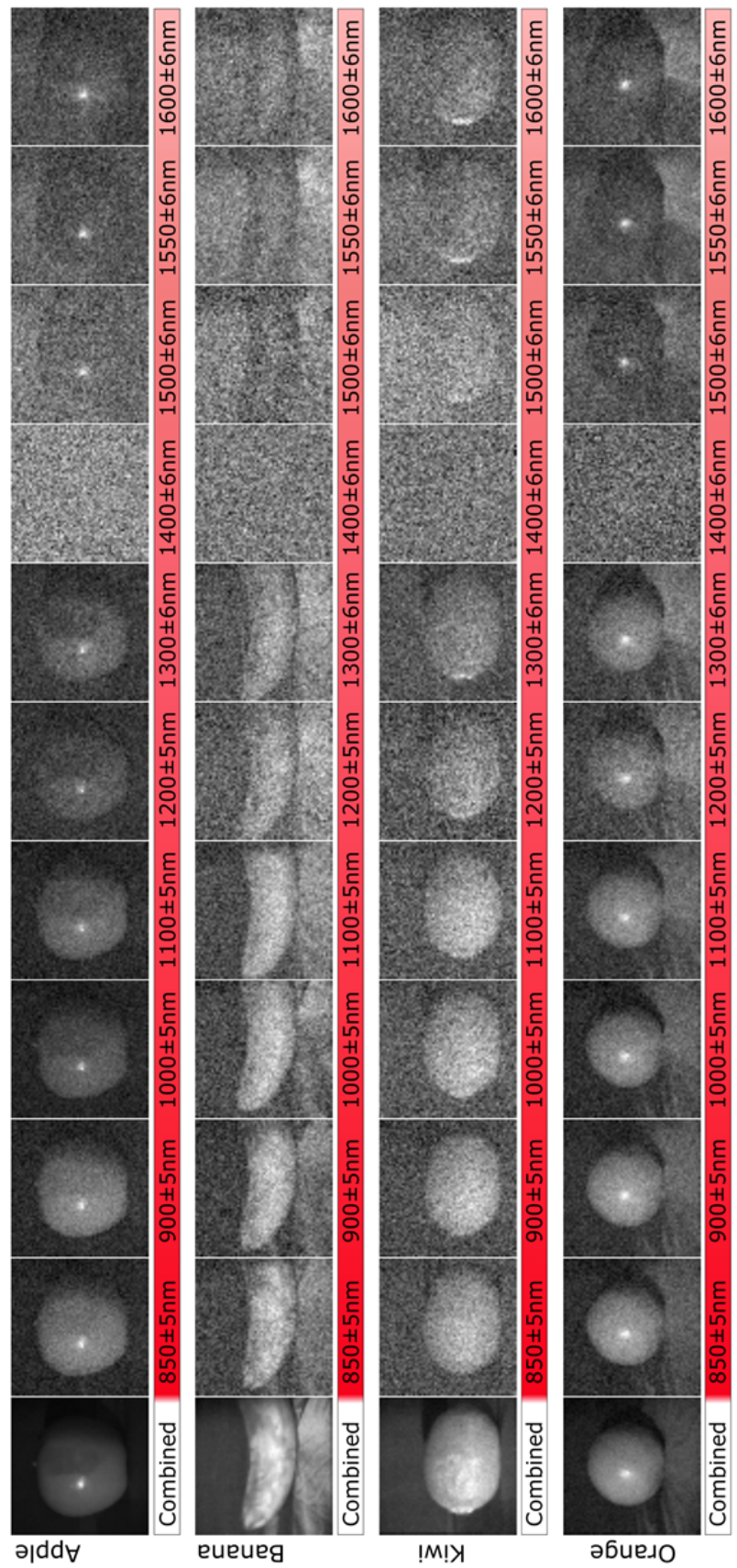


FIGURE 4.10: NIR micro-scanning imaging with solar radiation. Four different fruits, an apple, a banana, a kiwi and an orange, were imaged at a resolution of 128 x 128 with 50-frame averaging after employing the bandpass filter set. The combined images were reconstructed without utilising any filter or averaging.

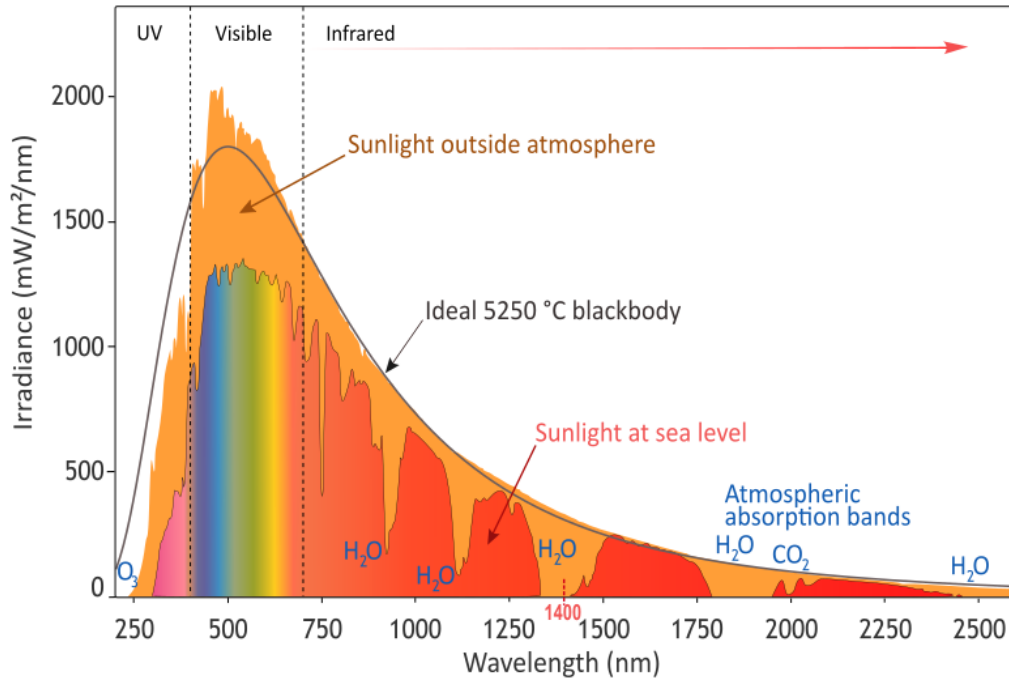


FIGURE 4.11: The spectral irradiance emitted by the sun [129]. As labelled, the brown spectrum is what hits at the top of the Earth’s atmosphere, the grey curve is the spectrum distribution from an ideal 5250 °C blackbody, and the multicoloured spectrum is the radiation at sea level.

Figure 4.11 shows the solar irradiance spectrum above the Earth’s atmosphere and at surface(/sea level) [129]. The distribution of sunlight outside atmosphere is close to an ideal 5250 °C blackbody spectrum, where a blackbody is an ideal emitter that absorbs all incident electromagnetic radiation and has a spectrum based solely on its temperature, regardless of the body’s shape or composition [130, 131]. However, gases contained in the atmosphere can absorb sunlight at certain wavelength bands, which results in a filtered solar irradiance spectrum at sea level. Among those gases, water vapour is the most effective gas that absorbs a lot of IR light at specific wavelengths [129, 132, 133].

Based on that, we then mapped those reconstructed fruits images with the sea-level solar radiation distribution, which explained the outcomes with the 1400 ± 6 nm bandpass filter (see Fig. 4.12 and Appendix A). Furthermore, the wave

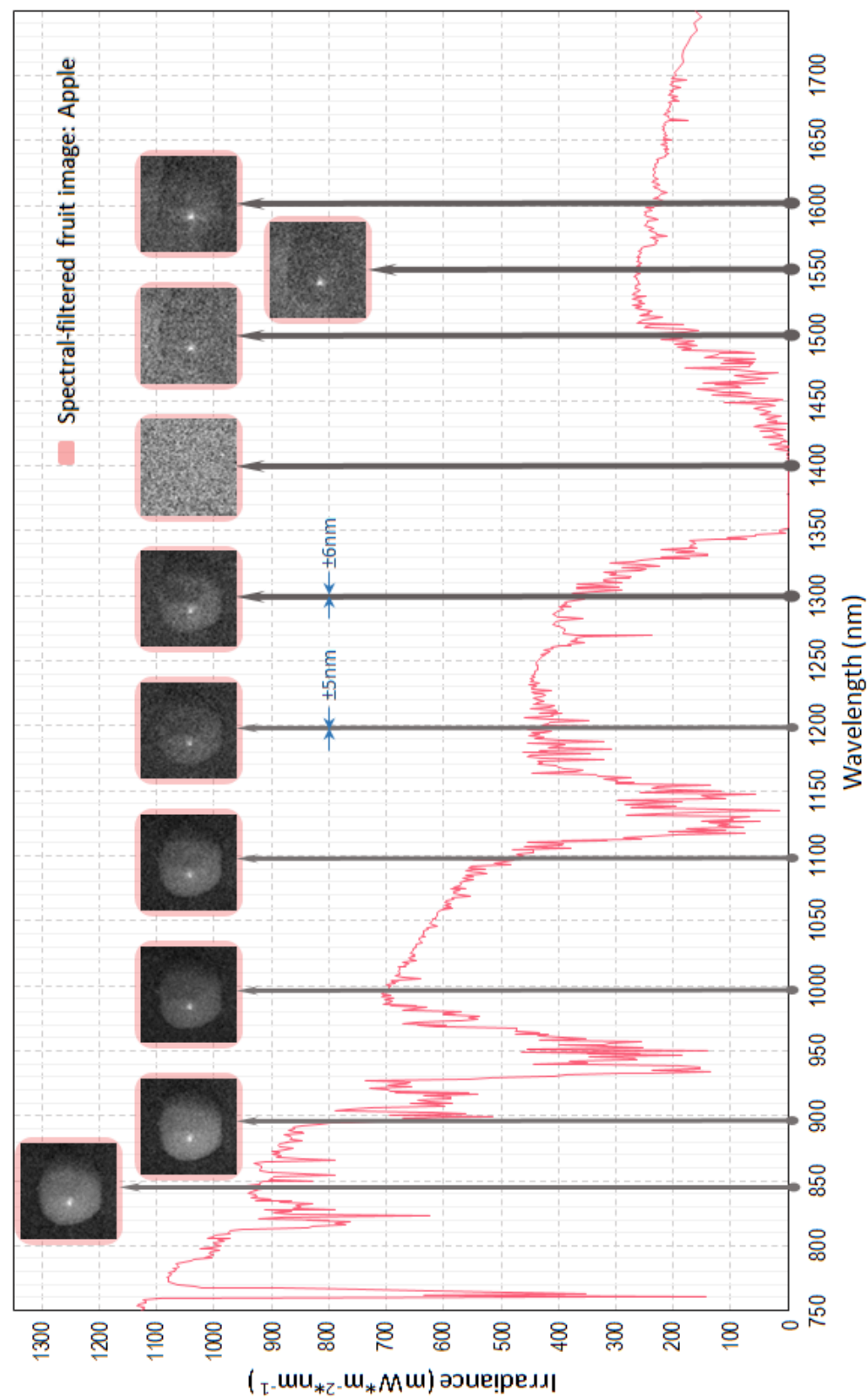


FIGURE 4.12: Effect of atmospheric absorption of radiation on sea-level solar NIR imaging. Light near the 1400nm wavelength was absorbed by water vapour mainly, which resulted in images with no signal detected at this specific waveband.

range near 1400 nm was chosen to detect the water uptake in plant leaves and the overripe or rotten portions of fruits together with the IR heat lamp as mentioned in 4.2.1. Figure 4.13 presents two 128 x 128 NIR micro-scanning images using the heat lamp and a wider NIR bandpass filter of 1400 ± 35 nm for water detection in plant leaves. By comparing those NIR images, the water absorption regions could be detected.

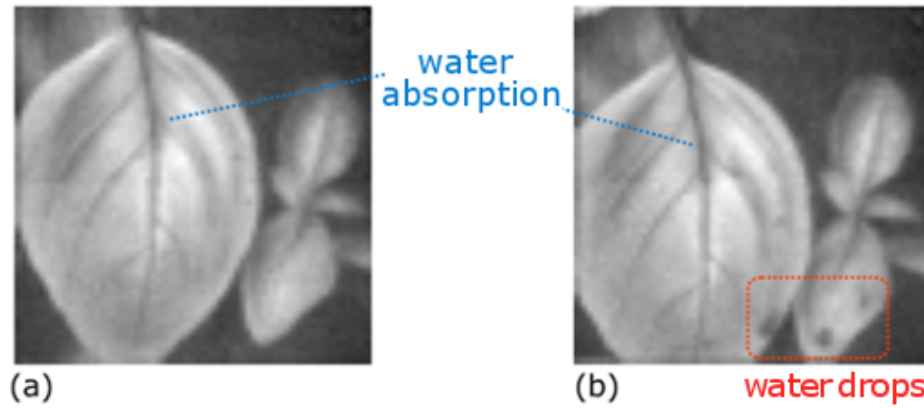


FIGURE 4.13: Water detection with a 1400 ± 35 nm NIR bandpass filter and 50 times averaging at 128 x 128 resolution. The dark areas of the plant leaves in (a) and (b) reveal the internal water absorption distribution. The region marked with a red colour in (b) shows water drops near the edge of leaves.

4.5 InGaAs NIR imaging comparison

To roughly evaluate the NIR imaging quality of our InGaAs single-pixel camera system, we compared the outcomes of imaging several objects at a resolution of 128 x 128 with the ones captured from an expensive commercial InGaAs camera, a Goldeye G-008, which has a maximum frame rate of 344 fps. The factor of frame rate and exposure time was not a big consideration in this comparison.

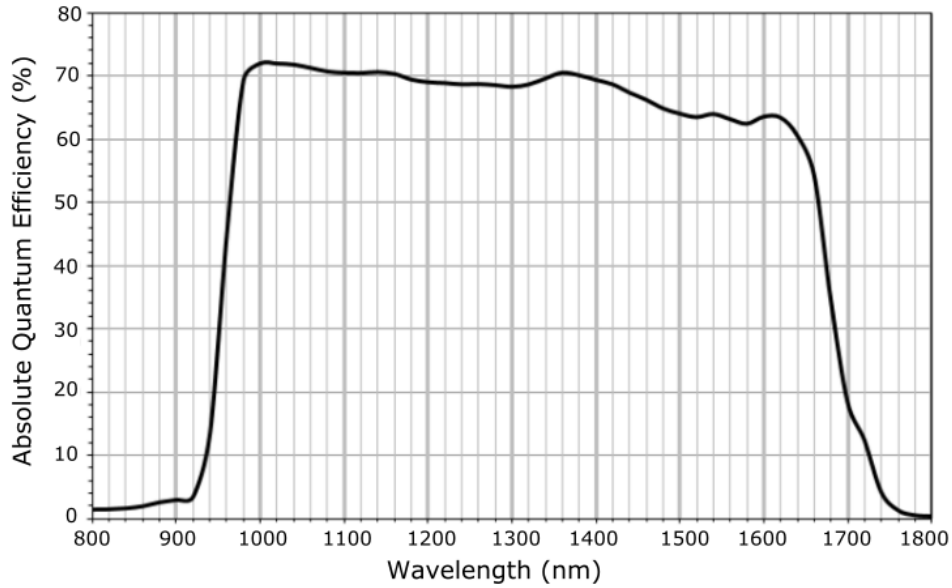


FIGURE 4.14: The spectral sensitivity range of GoldEye G-008 InGaAs camera: 900 nm to 1700 nm [134].

The short wave infrared spectrum (SWIR) of this Goldeye camera is fitted with a QVGA (Quarter Video Graphics Array) InGaAs sensor of 320 x 256 pixels, 30 μm per pixel size, making its spectral sensitivity reach from 900 to 1700 nm as shown in Fig. 4.14 [134].

For comparison, we imaged each object with and without the 1400 ± 35 nm band-pass filter by using both cameras at the same time. With the single-pixel camera, we captured the scene at a resolution of 128 x 128 by using micro-scanning. With the Goldeye camera, a larger area of the scene was photographed at a resolution of 320 x 256 and a region of 128 x 128 was later segmented to match the view from the single-pixel camera. One of the outcomes is shown in Fig. 4.15, in which both cameras were capable of detecting the water distribution underneath object's surface and the image quality with Goldeye G-008 was unsurprisingly better with a specialized adjustable focus lens.

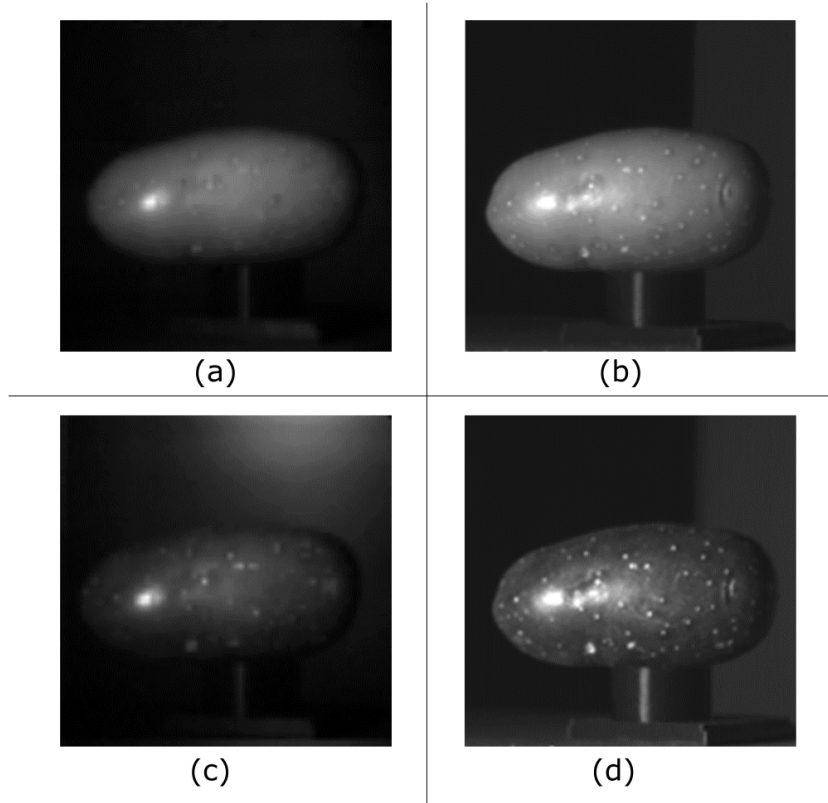


FIGURE 4.15: InGaAs NIR imaging comparison between the single-pixel camera and Goldeye G-008. Image (a) and (c) were captured by our signal-pixel camera at a resolution of 128×128 using micro-scanning. There was no image averaging applied on image (a), but 100 averaging on image (c). The frame rates for image (a) and (c) were 0.67 Hz and 0.0067 Hz. Image (b) and (d) were captured with the Goldeye G-008 at a region of 128×128 which matched the view from the single-pixel camera. There was no image averaging applied on image (b) and (d) and both of them had a frame rate of 10 Hz. Image (a) and (b) were captured without using any bandpass filter, and image (c) and (d) were captured with a 1400 ± 35 nm bandpass filter.

4.6 Conclusion

In this chapter, we have demonstrated a NIR single-pixel imaging system by utilising an InGaAs amplified photodetector. To overcome one of the main limitations in single-pixel imaging, where the SNR of reconstructed image decreases with increasing resolution, an optimisation method known as micro-scanning was employed in our system. This approach combines multiple lower resolution images acquired

with a modified set of patterns to generate a higher-resolution image, while maintaining the SNR of the lower resolution ones. By comparing images reconstructed using micro-scanning with the ones obtained from conventional sampling at a same resolution, it showed a significant improvement of the SNR of the micro-scanning images with a modest reduction in the contrast of high spatial frequencies.

We then imaged different objects with a set of NIR bandpass filters by making use of sunlight as the illumination source and figured out that light at specific IR wavelengths would normally be absorbed completely by water vapour in the atmosphere, for instance, light near the 1400 nm wavelength. Based on that, we chose a 1400 ± 35 nm bandpass filter to detect the water distribution underneath the surface of plant leaves and fruits with a heat lamp, which led to potential applications for detecting the overripe or rotten portions of fruits. We also roughly compared our NIR single-pixel imaging system with a much more expensive commercial InGaAs camera Goldeye G-008, in which both cameras demonstrated the capability of revealing the water distribution underneath object's surface and the Goldeye camera showed an unsurprisingly better image quality with a specialized adjustable focus lens. Further improvement of the imaging system was made and presented in the next chapter.

Chapter 5

Additional Improvements of Single-Pixel System with VIS and NIR Dual-Band Detection

5.1 Introduction

In this chapter, the single-pixel imaging system was modified to perform visible (VIS) and near-infrared (NIR) dual-band detection simultaneously with double light paths by making use of the DMD more efficiently. The NIR images of various fruits obtained at different IR bands were compared with the ones captured from the VIS detector and a reference colour camera to gain a better understanding of the information beneath the surfaces. Besides differential signal acquisition and micro-scanning, some other approaches have also been applied or tested in our system to improve the SNR of single-pixel images. Among those methods, single averaging and randomization of pattern sequence, as used in our system, will be briefly explained and exemplified here. A regularisation approach based on minimising the spatial curvature of the reconstructed image will be presented with further improvement planned in the future work.

5.1.1 Contributions

The work that follows was undertaken by myself with help from Matt Edgar. The dual-band detection system was partially 3D printed and structured by myself. The data collection of different fruits was performed by myself. The idea of utilising a regularisation approach for optimising the reconstructed image was the result of discussions among Miles Padgett, Matt Edgar and myself.

5.2 Dual-band imaging modification

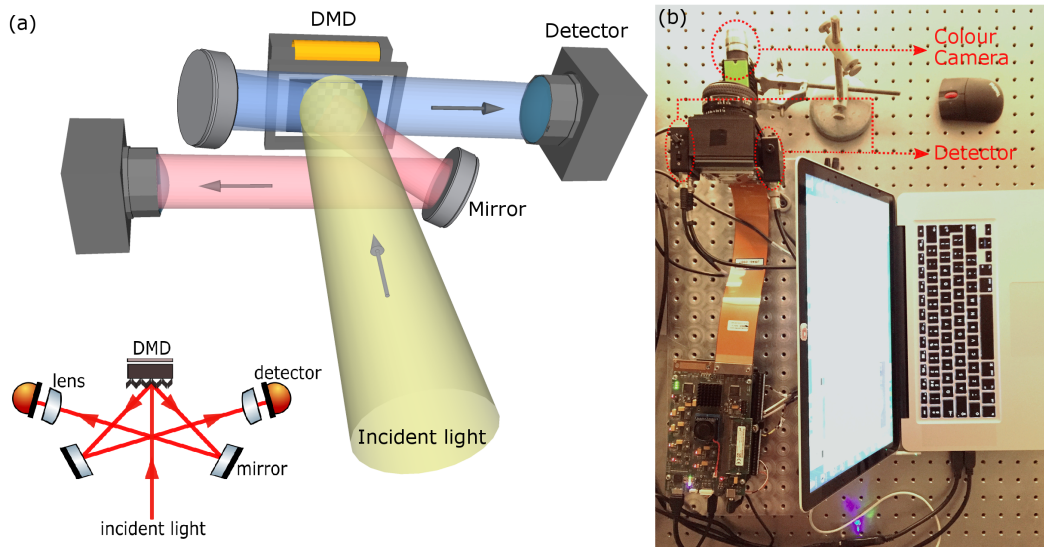


FIGURE 5.1: Dual-band imaging modification. (a) The propagation paths of light reaching from DMD to two detectors. (b) Modified single-pixel imaging system setup with VIS and NIR dual-band detection. An industrial colour camera was placed under the lens for the DMD to provide a reference image.

In most single-pixel imaging systems, the micro-mirrors on an energized DMD rotate at either $+12$ or -12 degrees to represent an 'on' or 'off' state, as mentioned in section 3.2.1. In this circumstance, only half of the light('on') reflected from the DMD is used, the other half which is reflected towards the reverse angle ('off') is perceived as the unwanted one and wasted, even though both parts contain the

same amount of information as each other with the differential signal acquisition approach (see section 3.2.3). To take advantage of both parts of the light, we modified the system with two adjusted mirrors facing towards two detectors by which it created double light paths for enabling dual-band detection simultaneously as shown in Figure 5.1(a).

Figure 5.1(b) demonstrates the modified system setup with VIS and IR detection. The VIS detector we chose was a silicon switchable gain detector, Thorlabs PDA100A-EC, for light signals over 400 to 1100nm wavelength range. A short-pass (SP) filter, Comar 716GK25, was fitted in front of the VIS detector to limit its detectable signal range to 400-716nm. The NIR detector we selected was an InGaAs switchable gain detector, Thorlabs PDA20CS, for light signals over 800 to 1800nm wavelength range. With each detector, a micro-scanning approach was applied to generate an image with a resolution of 128 x 128. An industrial colour camera (CC), JAI GO-5000-USB with 5MP and 62fps was placed under the lens that focuses light onto the DMD. It was used to provide a colour image of the scene as the reference for the reconstructed VIS and NIR single-pixel images.

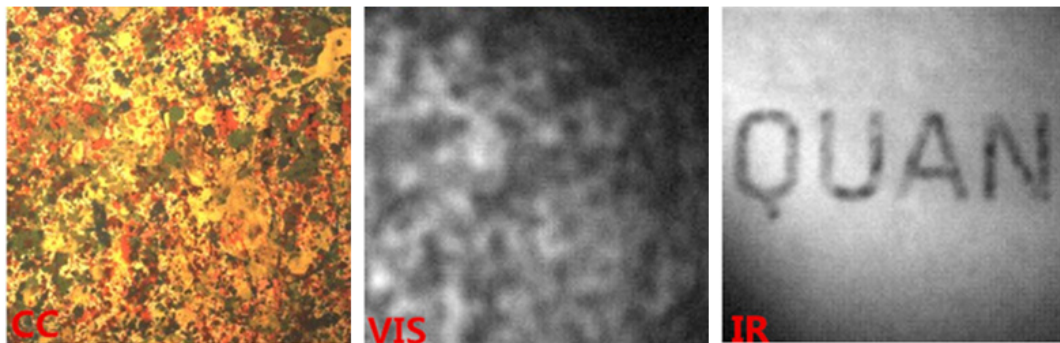


FIGURE 5.2: Experimental images of a piece of painting with different layers of acrylic colour. The CC produced a colour image of the painting and the VIS detector generated a 128 x 128 monochrome image of the painting. The NIR detector obtained a 128 x 128 monochrome image of hidden information underneath the surface of the painting.

Figure 5.2 gives a simple demonstration of different imaging results from the system on a piece of painting with various layers of acrylic colour. Both the industrial colour camera and the visible light detector captured an surface image of the painting. The reconstructed image from the NIR light detector showed the hidden words printed under the surface painting, which revealed the penetration capability of IR light with certain painting material.

5.3 Single-pixel dual-band detection with fruits

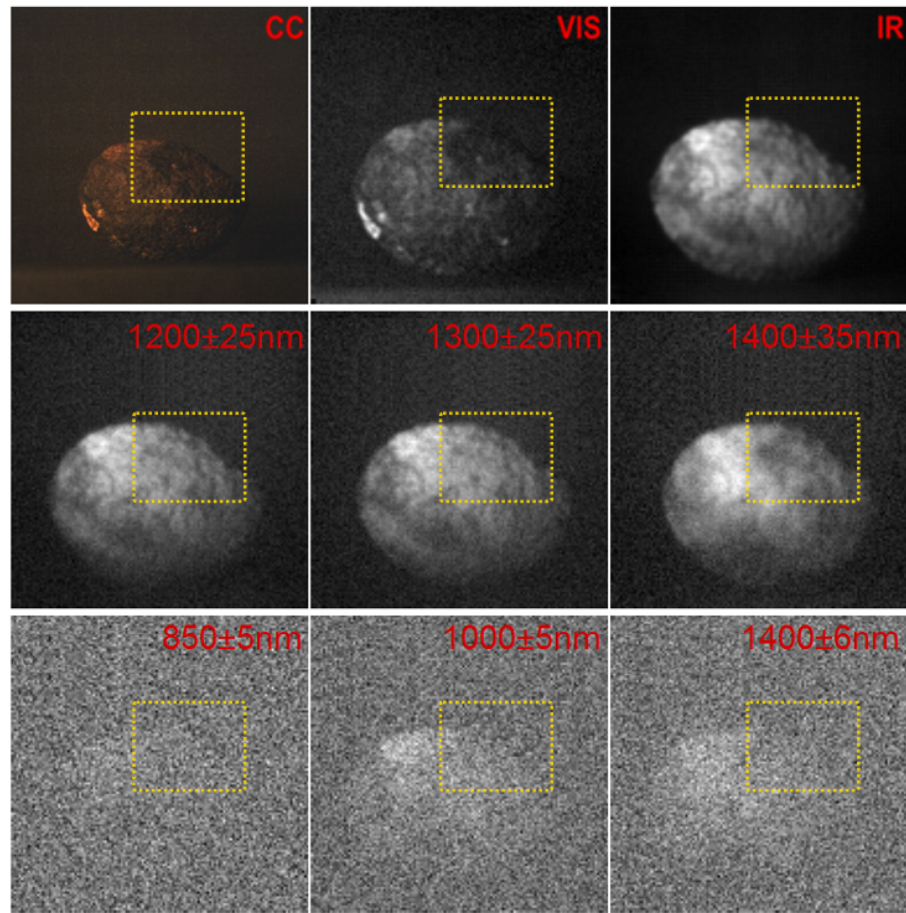


FIGURE 5.3: Single-pixel VIS and NIR imaging results of a ripe avocado with no signal averaging. By comparing the degree of shade in those images, the yellow region of the avocado was likely to be a rotten area with water.

Given an object, the degree of penetration revealed in a reconstructed single-pixel NIR image usually depends on the wavelength range of IR signals. To explore the hidden structure underneath the surface, we generated the NIR images with various fruits by utilising different bandpass filters, and compared it with the VIS and CC images. Figure 5.3 exhibits the imaging results of a ripe avocado with VIS and NIR detection at a resolution of 128 x 128 with no averaging. Based on the different degrees of shade in the yellow region among those images (due to water absorption), it revealed a potential rotten area of the avocado. This area was later proven to be a rotten area after peeling off its skin.

5.4 Additional methods for improving the SNR of single-pixel images

With single-pixel imaging, the cost of SNR when increasing the image resolution restricts its application in various fields. Diverse methods have been developed to improve it. Besides differential signal acquisition and mirco-scanning, here we will briefly demonstrate three other approaches that have been adopted or tested in our system to reduce the noise level.

5.4.1 Signal averaging

One major element from the single-pixel system hardware that contribute to the total noise is the circuit noise within the detector and the analog-to-digital converter (ADC). With a high-resolution detector and ADC, the effects of noise in the system will be shown, which leads to uncertainty in every measurement result. To increase the strength of a measured signal relative to this type of noise, signal averaging have been used in our system. Given N successive sets of measured signals S_i with noise, the averaged value S_{avg} can be calculated as follows:

$$S_{avg} = \frac{1}{N} \sum_{i=1}^N S_i \quad (5.1)$$

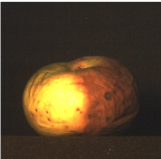
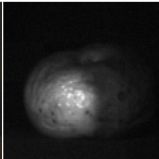
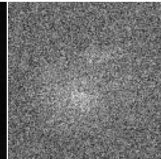
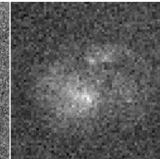
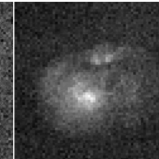
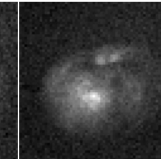

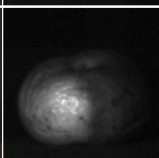
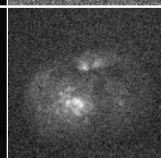
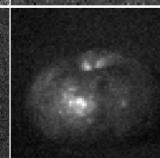
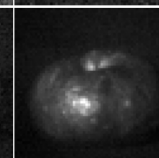
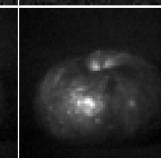
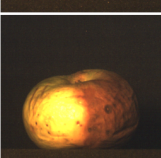
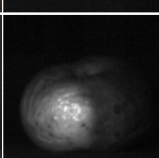
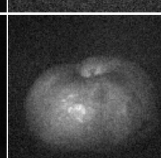
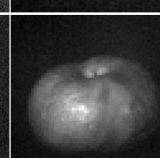
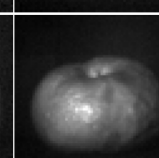
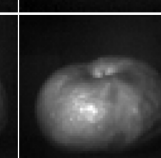
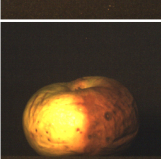
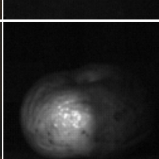
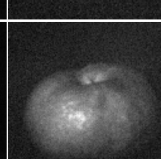
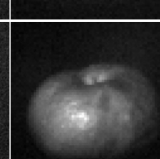
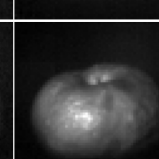
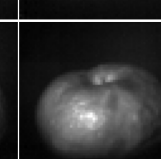
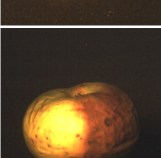
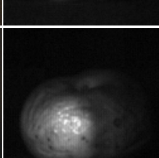
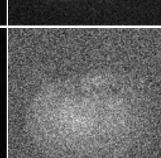
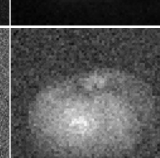
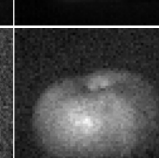
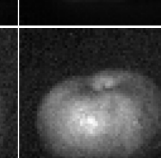
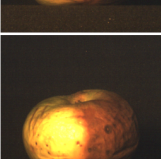
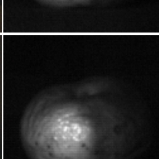
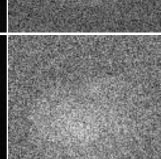
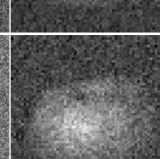
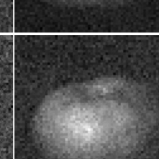
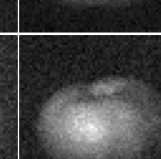
CC	VIS	IR				Bandpass Filter
		No averaging	10 averaging	50 averaging	100 averaging	
						1600 ± 6 nm
						1400 ± 35 nm
						1300 ± 25 nm
						1200 ± 25 nm
						1000 ± 5 nm
						850 ± 5 nm

FIGURE 5.4: NIR imaging results of an apple using different averaging values. The CC images were adjusted as the reference to compare with the VIS & IR single-pixel images at a resolution of 128x128 with micro-scanning. By increasing the number of averaging with the NIR signals, the effects of noise were deduced in the imaging results.

The more successive measurements included in the signal averaging will result in more reduction of the noise with longer settling time. Figure 5.4 demonstrates the improvement of IR imaging results by increasing the amount of frame averaging, where the NIR light level went down after passing through each bandpass filter

and the hardware noise became obvious. More imaging results of different fruits are shown in Appendix B. By visually comparing those images, it is obvious that the SNR in a high-noise reconstructed image can be improved by employing frame-averaging in our single-pixel system.

5.4.2 Randomisation of pattern sequence

A potential noise of using signal averaging in single-pixel imaging can be produced by the fluctuation intensities of the light source with a typical frequency of 50 Hz or 60 Hz, which could result in fixed discrepancies between individual patterns in a set.

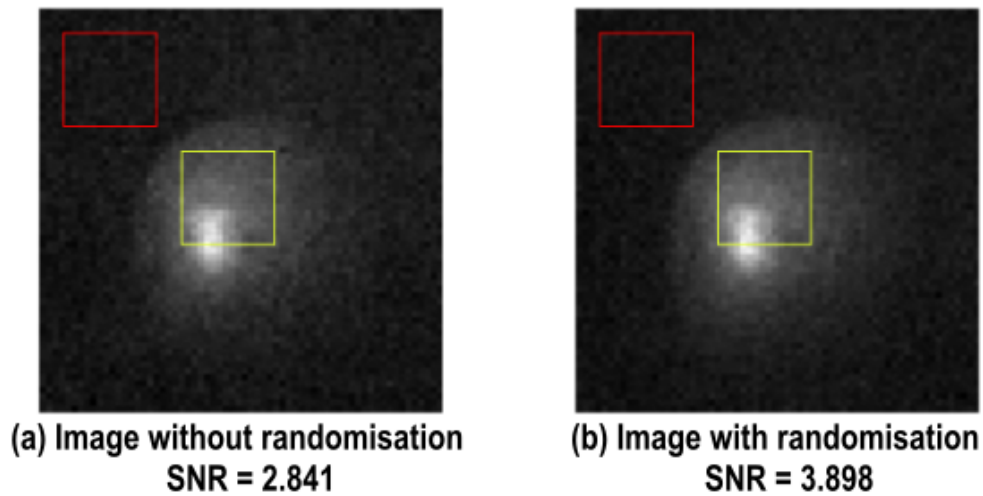


FIGURE 5.5: NIR imaging results with and without randomisation. Both NIR single-pixel images (a) and (b) were obtained through a 1400 ± 35 nm bandpass filter at a resolution of 128x128 using micro-scanning and 20 averaging.

To reduce the impact of this noise, we have randomised the order of illumination patterns in the Hadamard set every time so each pattern was tended to have an equal opportunity under the light of all fluctuating intensities. Figure 5.5 shows a comparison of the NIR imaging results generated with and without randomising the pattern sequence during signal averaging. For each image, we calculated the SNR using [135, 136]:

$$SNR = \frac{P_f - P_b}{[(\delta_f + \delta_b)/2]} \quad (5.2)$$

where P_f and P_b are, respectively, the average intensities of the image feature and the image background, δ_f and δ_b are, correspondingly, the standard deviations of the intensities of the image feature and the image background. Here we chose a region highlighted by a yellow square to represent the image feature and a same-size region highlighted by a red square to represent the image background for calculation. By comparison, the image obtained with randomisation of pattern sequence showed a better SNR than the one without using it.

5.4.3 Regularisation

Typical images of objects tend to exhibit strong correlations between adjacent pixels. Besides signal averaging and randomisation of pattern sequence, we have further tested to improve the image SNR with a regularisation approach based on minimisation of the image spatial curvature. In this approach, the intensity value of each individual pixel was kept changing with randomised noise level until the program finds a derived image with the lowest value of the cost function, C , given by [137–139]

$$C = \frac{\chi^2}{N} + \lambda R, \quad (5.3)$$

and C is subject to $\chi^2/N \leq 1$. In this equation, the first term (χ^2/N) corresponds to degree of discrepancy between the measured signals (S_m) and the predicted signals (S_p) based on the estimated noise level in the measurements. χ^2 is given by

$$\chi^2 = \sum_{i=1}^N \left[\frac{S_{m_i} - S_{p_i}}{\delta} \right]^2 \quad (5.4)$$

where N is the total number of measurements and δ is the standard deviation of estimated noise in the measurements. The second term (λR) corresponds to the level of total curvature squared in the estimated reconstruction and R is given by

$$R = \left[\sum_{j=1}^n \left(\left| \frac{d^2 I_j}{dx^2} \right| + \left| \frac{d^2 I_j}{dy^2} \right| \right) \right]^2 \quad (5.5)$$

where n is the total number of pixels in the image and I_j is the intensity value of pixel j . λ is an empirical factor to balance the importance of both terms. It was manually adjusted through experience to prevent the program over-smoothing the image.

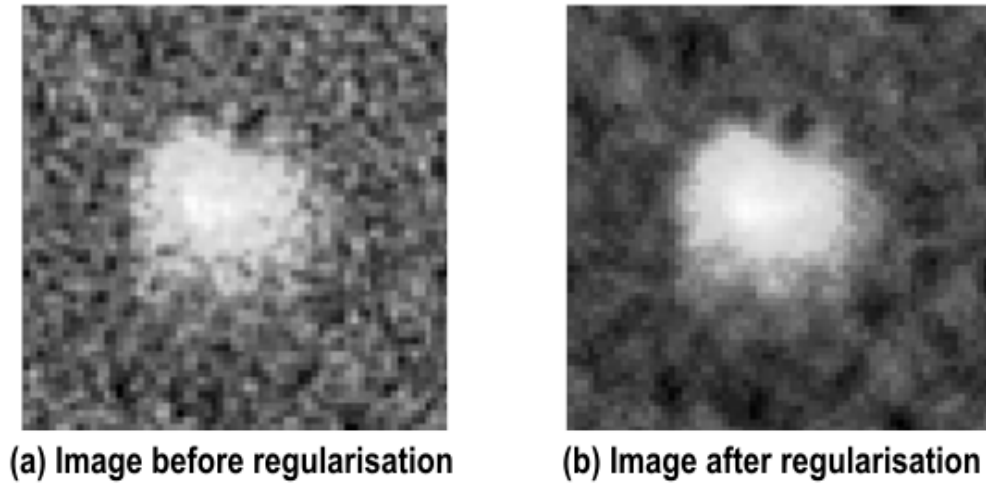


FIGURE 5.6: NIR imaging results before and after regularisation. Both images (a) and (b) were obtained through a 1000 ± 5 nm bandpass filter at a resolution of 64×64 using micro-scanning with no averaging.

In Figure 5.6, it shows a comparison of the NIR imaging results of an apple before and after employing the regularisation optimisation. By increasing the number of iteration steps, gradual image noise reduction was observed. However, with this approach, λ needs to be readjusted artificially each time based on the prior knowledge of the imaging object to avoid distortion caused by over-smoothing regularisation.

5.5 Conclusion

In conclusion, we have shown a modified single-pixel system with VIS and NIR dual-band detection by making use of the DMD more productively. The imaging results of various fruits obtained from the NIR detector at different IR bands were compared with the images from the VIS single-pixel detector and the ones from a reference colour camera to conclude a better understanding of the overripe or rotten regions hidden under the surface of those fruits.

In this chapter, we have also demonstrated three optimisation approaches that had been applied or tested in our systems for improving the NIR image SNR. In those approaches, signal averaging helped to increase the strength of measured signals over the hardware noise with a trade-off of longer settling time. By randomising the pattern sequence of each time, we were able to reduce the noise caused by the fixed illumination discrepancies between individual patterns. A regularisation approach based on minimising the image spatial curvature was tested to reduce the image noise. It's been observed that increasing the number of iterations of the best solution for the cost function lowered the noise level in the resulting image. Further work would be performed on optimising the cost function with a self-adjusting λ value.

Chapter 6

Conclusion and Future Work

Computational imaging is a highly active interdisciplinary research field which aims to enhance and extend the capabilities of traditional imaging devices and solutions. Numerous techniques have been proposed to extract different types of information from an imaged scene. In this thesis, the work has been focused on investigating practical applications of imaging systems utilising two of those computational imaging techniques, photometric stereo and single-pixel imaging.

Photometric stereo is a 3D imaging technique that captures the shape of a 3D solid object based on several images of a scene including the object taken from a fixed viewpoint under different illumination conditions. Typically, it recovers the object's surface gradients by estimating the surface normal vectors from the shading information in those images, measured as pixel intensities by a digital camera, and then integrates the gradient field to determine the 3D shape. Based on this approach, we have developed a fast 3D camera system with a low-cost accessory consisting of four white LEDs controlled by an Arduino board. Besides the typical calculation procedure, I have proposed a fast surface gradient calculation model to reduce the computing complexity. By comparing the 3D reconstructions on a selection of objects with the known reference data, our camera system has shown a reasonable level of accuracy in estimating the depth information. With further

tests at several public exhibitions, our inexpensive 3D camera system has demonstrated its capability of providing realistic 3D images of objects and faces on a 3D-enabled TV with a wide range of viewing angles, despite both the camera and the subject staying static during the image acquisition. We have also observed an increased estimation error at regions where the surface normals were pointing towards a direction perpendicular to the camera perspective. Further improvements can be made by optimising the depth estimation at those regions and recovering the information at obstructed surface areas.

In general, 3D imaging system using photometric stereo requires the scene to remain still whilst acquiring images under different lighting conditions in order to prevent significant distortion in 3D reconstructed images caused by pixel-matching errors, which limits its applicability to real-time applications. To tackle this limitation, we then introduced a revised 3D imaging system by combining photometric stereo with single-pixel imaging, allowing continuous operation for 3D reconstruction of dynamic scenes. Single-pixel imaging is an emerging technique that enables images to be generated by using a simple detector which have only one pixel to sense light rather than millions of pixels by conventional digital cameras. By making use of a fast DMD to provide a set of micro-structured light patterns onto the scene and four single-pixel detectors in different spatial locations to record the signals associated to each pattern simultaneously, our system can obtain four images of a scene containing different shading profiles to perform photometric stereo without pixel-matching error.

Comparing to projecting totally random patterns or using raster scanning, we chose a set of patterns derived from orthogonal Hadamard matrices to perform the structured illumination in our system, which improved the image SNR along with the sampling efficiency. With Hadamard patterns, each element is either '+1' or '-1' rather than '0' or '1' which are respectively represented by the black or white illumination blocks on the scene. To solve this, a differential signal acquisition approach was employed, in which the differential intensity value from

each black and white illumination pattern and its inverse was used to correspond to the Hadamard pattern. Under the restriction of a finite modulation rate of the DMD, an evolutionary compressed sensing strategy was employed in our 3D single-pixel system to further improve the imaging speed by reducing the number of sampling for recovering an image. Instead of a complete pattern set, a subset of Hadamard patterns were chosen based on the ranking of the corresponding signal intensities to represent the images. By comparing the 3D reconstruction results at the same image resolution with different compression ratios, we have shown an increase of the frame rate when using a smaller subset of the Hadamard patterns at the cost of only a modest reduction in image quality.

As a remarkable alternative to conventional imaging techniques, single-pixel technique is promising for applications to explore unusual spectral bands for imaging with a highly sensitive light sensor. With an InGaAs detector, we developed a NIR single-pixel imaging system trying to expose the hidden information underneath surfaces of plant leaves and fruits. To improve the SNR when increasing the image resolution with single-pixel imaging, a micro-scanning approach was applied in our system, which produces a high resolution image with a deduced noise level by combining a set of lower resolution images based on modified Hadamard patterns. For a given resolution, images reconstructed with and without using micro-scanning require the same total amount of Hadamard patterns. However, the image SNR can be improved significantly by using this approach at the expense of a modest reduction in the contrast of high spatial frequencies. By imaging different objects through a set of NIR bandpass filters with the sunlight at sea level, we have noticed that light near the 1400 nm wavelength was fully absorbed by the water vapour in atmosphere, based on which we then chose an IR lamp as the light source and a bandpass filter around this wavelength to detect the water information underneath surfaces of plant leaves and fruits. We have also had a rough comparison of the imaging results between our system and an expensive InGaAs camera, in which both showed the capability of water detection with leaves and fruits.

As both micro-scanning technique and evolutionary compressed sensing approach require no extra hardware in single-pixel imaging system, further research can be carried on by combining those two methods in our system to optimise the imaging quality together with a better efficiency.

Instead of using only one detector, we have modified our single-pixel imaging system by taking advantage of the light from both reflection arms of the DMD with two detectors, one for visible light and the other for light in the NIR spectrum, enabling dual-band imaging. Besides averaging detected signal and randomising pattern sequence, we have also tried to improve the image SNR with a regularization technique based on minimisation of the image spatial curvature. Further investigation can be undertaken to improve this method and combine it with the evolutionary compressed sensing technique in single-pixel imaging. Our dual-band single-pixel system can also be combined with a telescope to invest the influence of atmospheric dispersion and atmospheric turbulence on long-distance imaging.

Appendix A

Extra NIR Single-Pixel Imaging Data of Different Objects

Effects of atmospheric absorption of radiation on sea-level solar NIR imaging at a resolution of 128 x 128 with micro-scanning and 50-frame averaging:

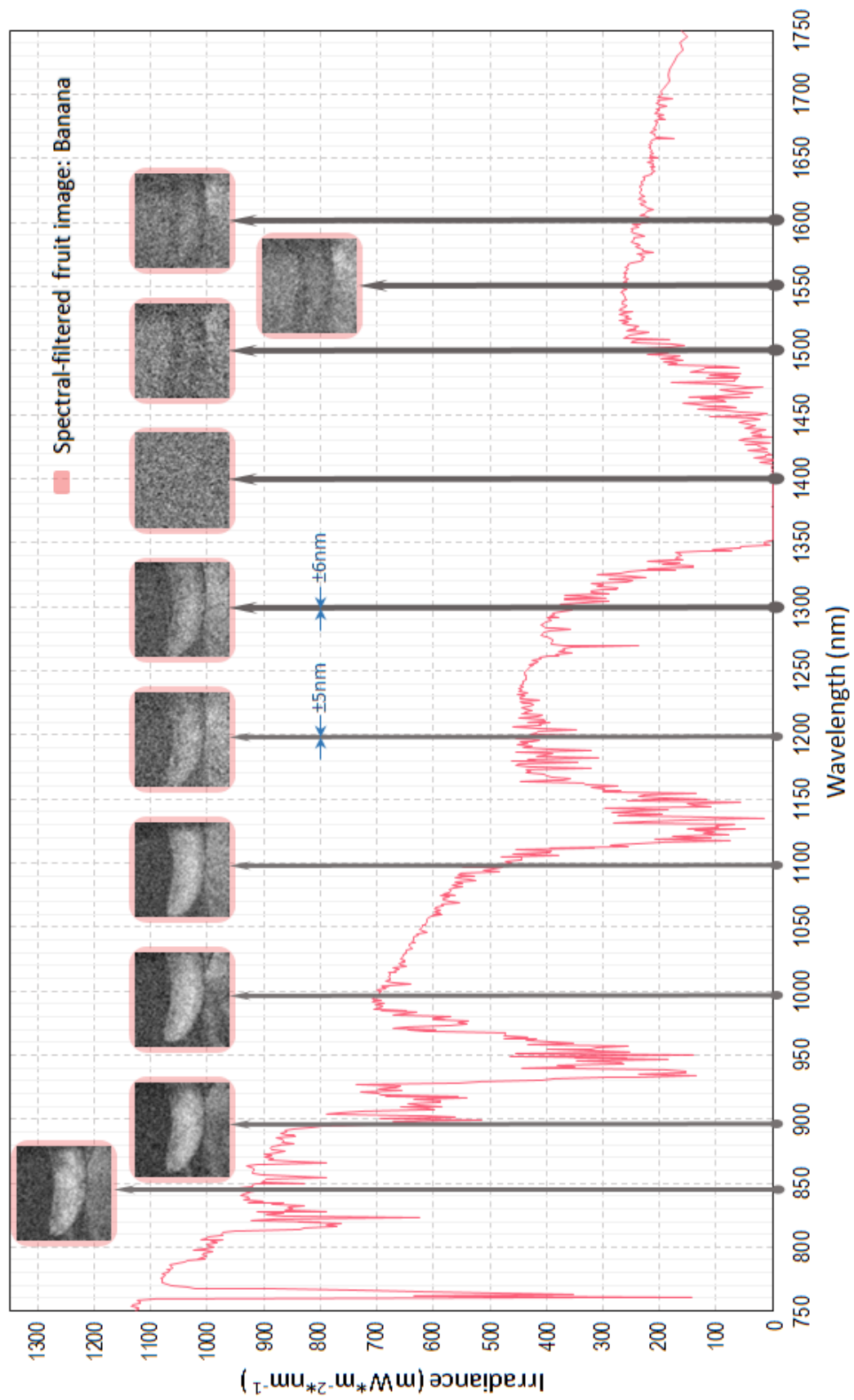


FIGURE A.1: Sea-level solar NIR imaging outcomes of a banana.

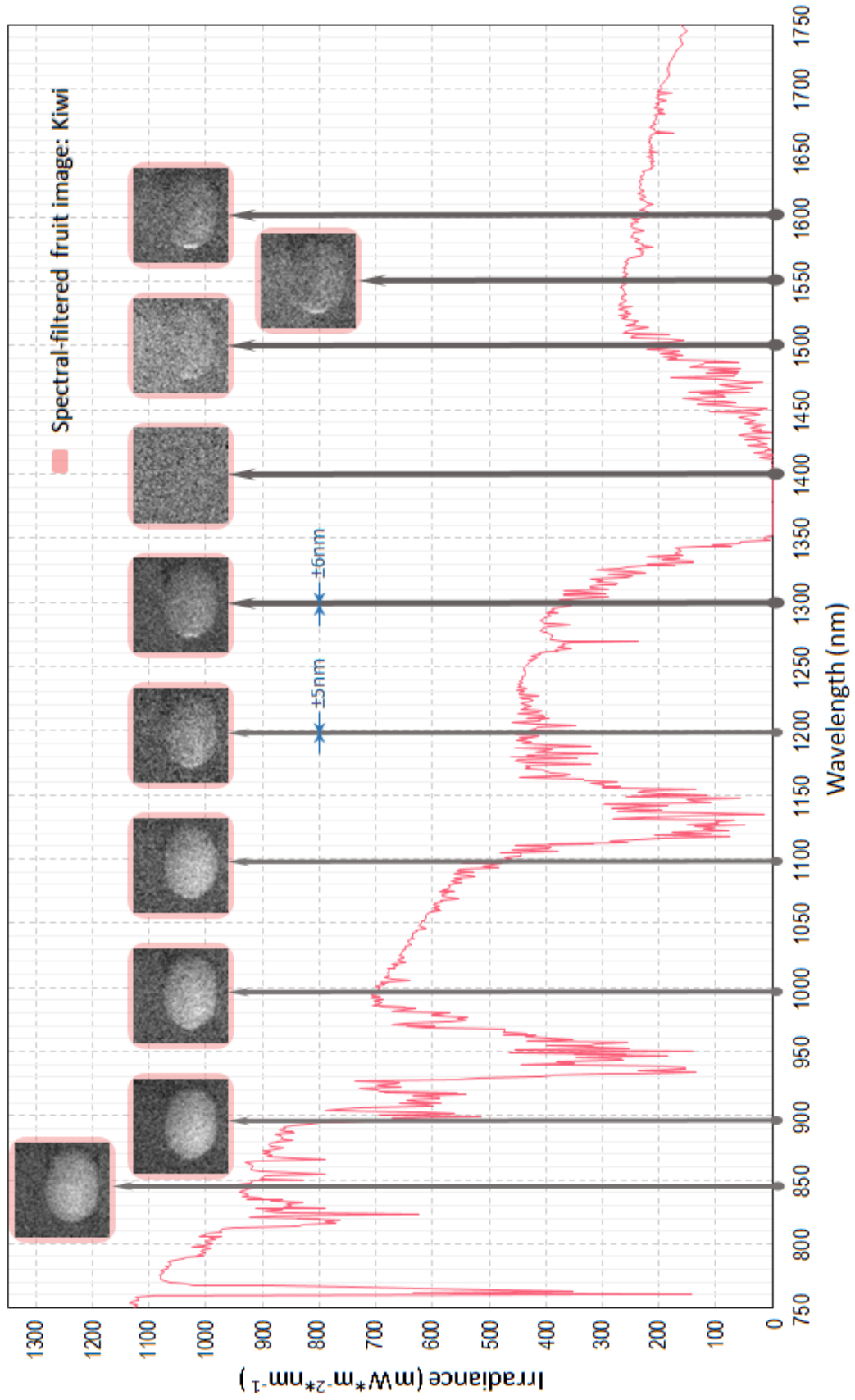


FIGURE A.2: Sea-level solar NIR imaging outcomes of a kiwi.

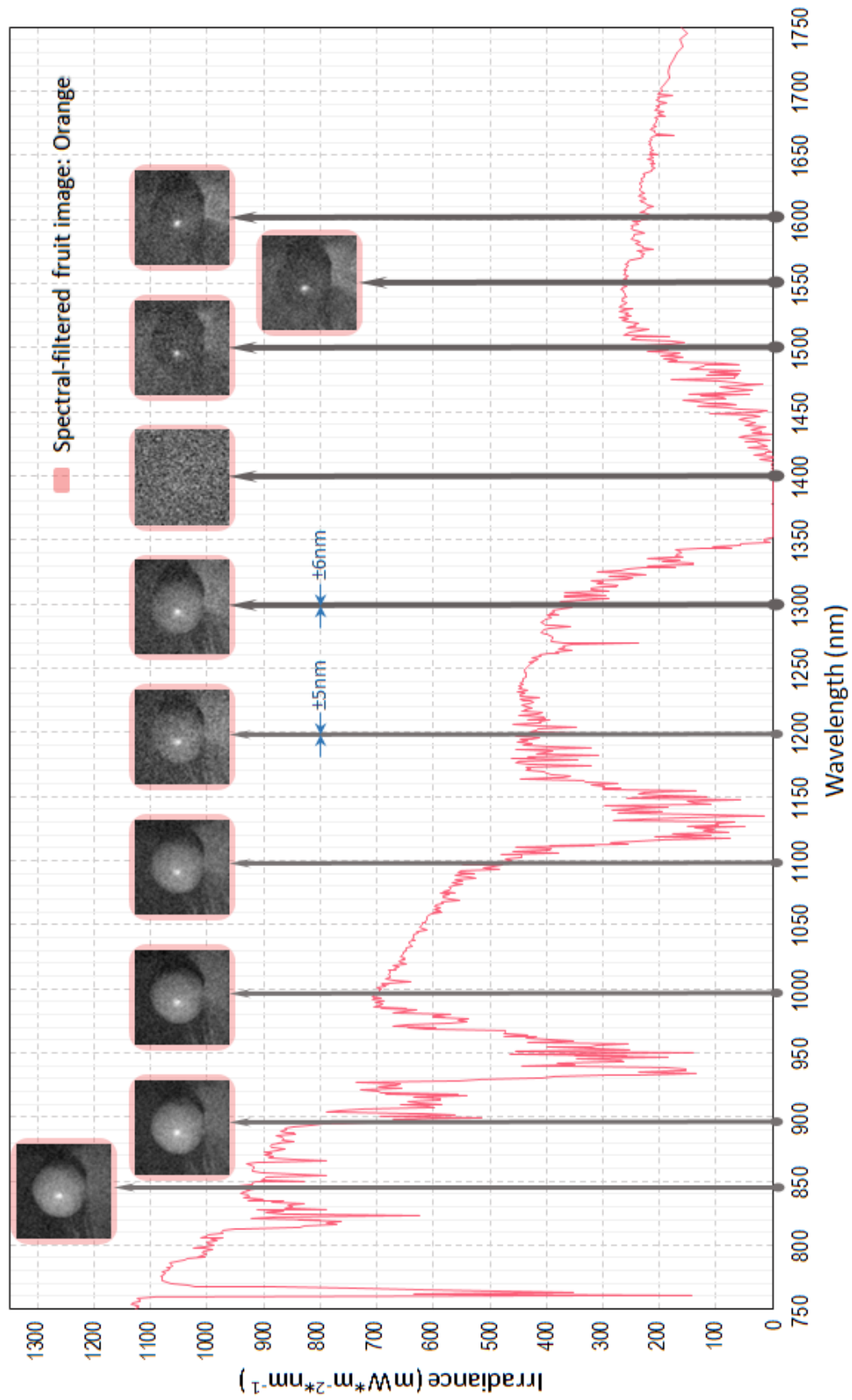


FIGURE A.3: Sea-level solar NIR imaging outcomes of an orange.

Appendix B

Extra Data of Single-Pixel Dual-Band Detection

Images from the industrial colour camera (CC) are adjusted as the reference to compare with the VIS & IR single-pixel images at a resolution of 128x128 with micro-scanning. By increasing the number of averaging with the NIR signals, the noise level in the imaging results has been deduced.


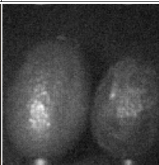
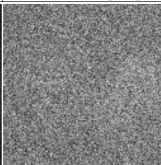
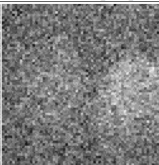
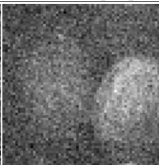
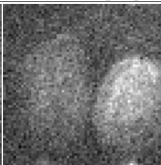
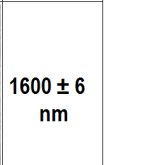

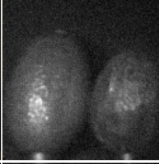
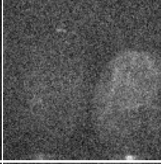

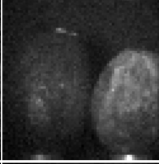
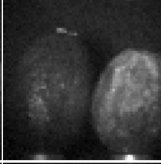
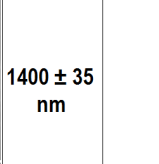


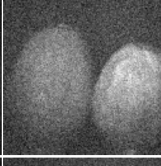

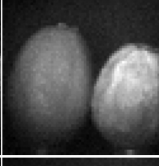
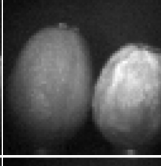
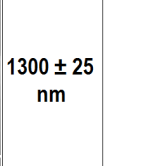

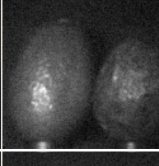
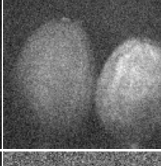
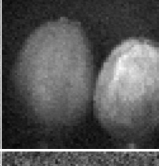
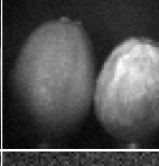
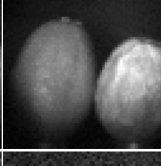
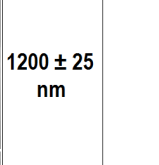

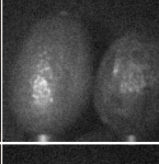
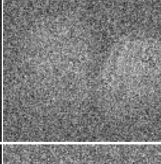

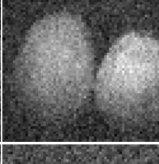
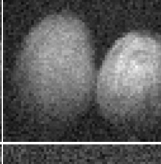
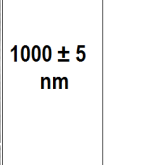

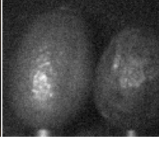
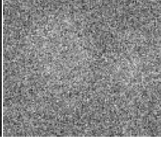


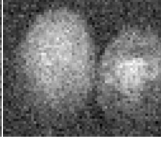
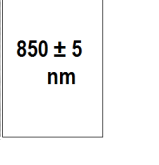
CC	VIS	No averaging	10 averaging	IR	50 averaging	100 averaging	Bandpass Filter
							1600 ± 6 nm
							1400 ± 35 nm
							1300 ± 25 nm
							1200 ± 25 nm
							1000 ± 5 nm
							850 ± 5 nm

FIGURE B.1: NIR imaging results of avocados using different averaging values.


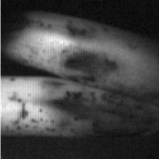
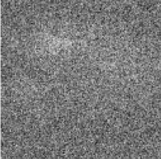




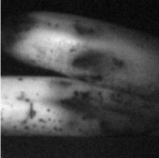



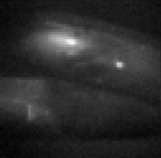

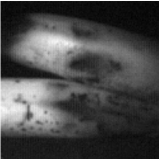
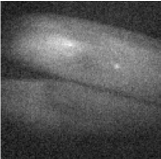
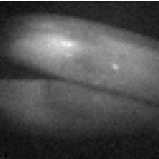
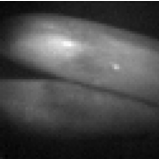
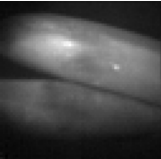

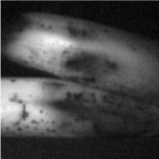
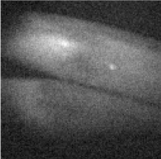
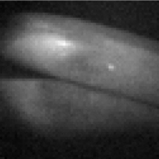
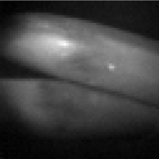
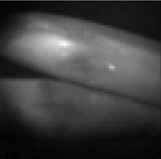

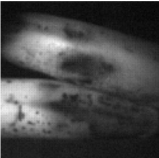
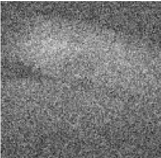


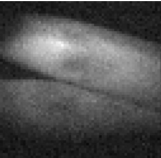

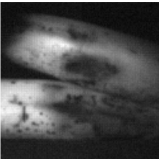
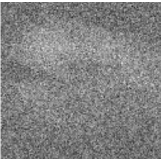


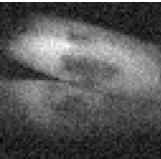
CC	VIS	IR				Bandpass Filter
		No averaging	10 averaging	50 averaging	100 averaging	
						1600 ± 6 nm
						1400 ± 35 nm
						1300 ± 25 nm
						1200 ± 25 nm
						1000 ± 5 nm
						850 ± 5 nm

FIGURE B.2: NIR imaging results of bananas using different averaging values.


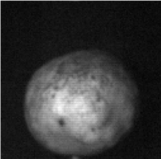
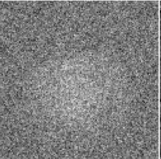
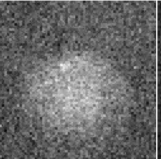
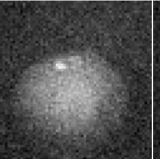
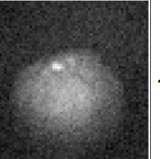


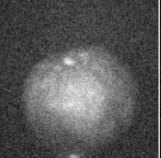


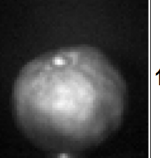

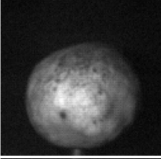
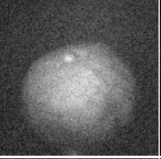

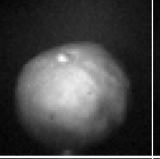
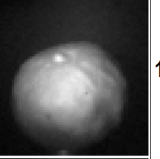
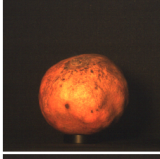
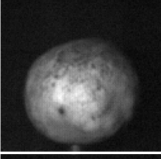
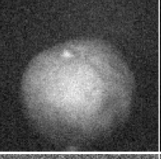
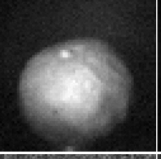
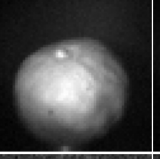
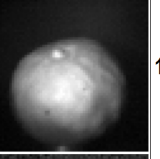
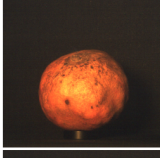
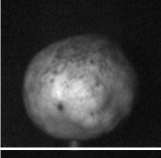
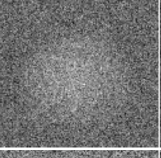
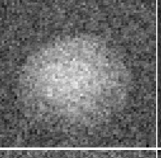
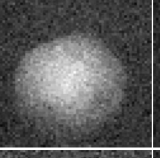
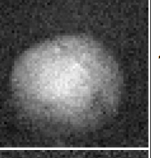

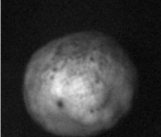
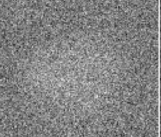


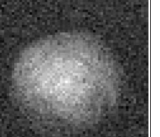
CC	VIS	No averaging	10 averaging	IR	50 averaging	100 averaging	Bandpass Filter
						1600 ± 6 nm	
						1400 ± 35 nm	
						1300 ± 25 nm	
						1200 ± 25 nm	
						1000 ± 5 nm	
						850 ± 5 nm	

FIGURE B.3: NIR imaging results of a pomegranate using different averaging values.

Bibliography

- [1] F. S. Pavone and P. J. Campagnola. *Second harmonic generation imaging*. Taylor & Francis, 2013.
- [2] H. Gernsheim and A. Gernsheim. *The History of Photography from the Earliest Use of the Camera Obscura in the Eleventh Century up to 1914*. Oxford University Press, 1955.
- [3] H. Gernsheim. The 150th anniversary of photography. *History of photography*, 1(1):3–8, 1977.
- [4] B. Newhall. *Latent image: the discovery of photography*, volume 54. Doubleday, 1967.
- [5] B. Coe. *George Eastman and the early photographers*. Hodder Wayland, 1973.
- [6] M. D. McFarlane. Digital pictures fifty years ago. *Proceedings of the IEEE*, 60(7):768–770, 1972.
- [7] M. Tripsas and G. Gavetti. Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic management journal*, pages 1147–1161, 2000.
- [8] T. Makimoto. The age of the digital nomad: Impact of cmos innovation. *IEEE Solid-State Circuits Magazine*, 5(1):40–47, 2013.
- [9] R. Hain, C. J. Kahler, and C. Tropea. Comparison of ccd, cmos and intensified cameras. *Experiments in fluids*, 42(3):403–411, 2007.

- [10] L. Quan. *Image-Based Modeling*. Springer, 2010.
- [11] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1):191139–191139, 1980.
- [12] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. E. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2):83, 2008.
- [13] S. Rout and S. Sonkusale. *Active Metamaterials: Terahertz Modulators and Detectors*. Springer, 2017.
- [14] N. Marin and J. M. Buszka. *Alternate light source imaging: forensic photography techniques*. Routledge, 2013.
- [15] A. A. Michelson. Recent advances in spectroscopy. *Nobel Lecture*, 12:1–13, 1907.
- [16] M. Hill, C. OMahony, H. Berney, P. J. Hughes, E. Hynes, and W. A. Lane. Verification of 2-d mems model using optical profiling techniques. *Optics and lasers in engineering*, 36(2):169–183, 2001.
- [17] W. H. Steel. *Interferometry*, volume 1. CUP Archive, 1983.
- [18] P. Hariharan. *Basics of Interferometry*. Academic Press, 2010.
- [19] P. Hariharan. *Optical interferometry*. Academic press, 2003.
- [20] Jason Geng. Structured-light 3d surface imaging: a tutorial. *Advances in Optics and Photonics*, 3(2):128–160, 2011.
- [21] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 1, pages I–195. IEEE, 2003.

- [22] F. Pipitone and R. Hartley. A structured light range imaging system using a moving correlation code. In *3D Data Processing, Visualization, and Transmission, Third International Symposium on*, pages 931–937. IEEE, 2006.
- [23] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan. Structured light 3d scanning in the presence of global illumination. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 713–720. IEEE, 2011.
- [24] L. Iocchi and K. Konolige. A multiresolution stereo vision system for mobile robots. In *New Trends in Robotics Research, AIIA '98 Workshop on*, pages 317–321. AIIA, 1998.
- [25] E. Keppel. Approximating complex surfaces by triangulation of contour lines. *IBM Journal of Research and Development*, 19(1):2–11, 1975.
- [26] E. Izquierdo. On the epipolar geometry in stereo vision. In *Proceeding Third International Conference on Approximation and Optimization in the Caribbean, Puebla, Mexico*, 1995.
- [27] R. A. Hamzah and H. Ibrahim. Literature survey on stereo vision disparity map algorithms. *Journal of Sensors*, 2016, 2015.
- [28] A. David and P. Jean. Computer vision: a modern approach. *Prentice Hall*, pages 654–659, 2002.
- [29] L. E. Kruger, C. Wohler, A. Wurz-Wessel, and F. Stein. In-factory calibration of multiocular camera systems. In *Photonics Europe*, pages 126–137. International Society for Optics and Photonics, 2004.
- [30] D. Chrysostomou and A. Gasteratos. Three-dimensional scene reconstruction: A review of approaches. *Depth Map and 3D Imaging Applications: Algorithms and Technologies: Algorithms and Technologies*, page 142, 2011.

- [31] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Computer vision and pattern recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 519–528. IEEE, 2006.
- [32] H. Hirschmuller and D. Scharstein. Evaluation of cost functions for stereo matching. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.
- [33] R. A. Hamzah and H. Ibrahim. Literature survey on stereo vision disparity map algorithms. *Journal of Sensors*, 2016, 2015.
- [34] B. K. P. Horn and M. J. Brooks. *Shape from Shading*. MIT Press, Cambridge, MA, USA, 1989.
- [35] D. Nandy and J. Ben-Arie. Shape from recognition: a novel approach for 3-d face shape recovery. *Image Processing, IEEE Transactions on*, 10(2): 206–217, 2001.
- [36] S. Romdhani and T. Vetter. Estimating 3d shape and texture using pixel intensity, edges, specular highlights, texture constraints and a prior. volume 2, pages 986–993. IEEE, 2005.
- [37] M. Song, D. Tao, X. Huang, C. Chen, and J. Bu. Three-dimensional face reconstruction from a single image by a coupled rbf network. *Image Processing, IEEE Transactions on*, 21(5):2887–2897, 2012.
- [38] K. Lee and C. Kuo. Shape reconstruction from photometric stereo. In *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR'92., 1992 IEEE Computer Society Conference on*, pages 479–484. IEEE, 1992.
- [39] T. Malzbender, B. Wilburn, D. Gelb, and B. Ambricol. Surface enhancement using real-time photometric stereo and reflectance transformation. *Rendering techniques*, 2006:17th, 2006.

- [40] T. Okatani and K. Deguchi. On uniqueness of solutions of the three-light-source photometric stereo: Conditions on illumination configuration and surface reflectance. *Computer Vision and Image Understanding*, 81(2):211–226, 2001.
- [41] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. volume 72, pages 239–257. Springer, 2007.
- [42] P. Tan, S. Lin, and L. Quan. Subpixel photometric stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(8):1460–1471, 2008.
- [43] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan. Self-calibrating photometric stereo. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1118–1125. IEEE, 2010.
- [44] M. F. Hansen, G. A. Atkinson, L. N. Smith, and M. L. Smith. 3d face reconstructions from photometric stereo using near infrared and visible light. *Computer Vision and Image Understanding*, 114(8):942–951, 2010.
- [45] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Computer Vision–ACCV 2010*, pages 703–717. Springer, 2011.
- [46] Y. Sun, J. Dong, M. Jian, and L. Qi. Fast 3d face reconstruction based on uncalibrated photometric stereo. *Multimedia Tools and Applications*, pages 1–16, 2013. doi: 10.1007/s11042-013-1791-3.
- [47] M. J. E. Golay. Multi-slit spectrometry. *JOSA*, 39(6):437–444, 1949.
- [48] J. H. Shapiro. Computational ghost imaging. *Physical Review A*, 78(6):061802, 2008.
- [49] Y. Bromberg, O. Katz, and Y. Silberberg. Ghost imaging with a single detector. *Physical Review A*, 79(5):053840, 2009.

- [50] M. Harwit and N. J. A. Sloane. *Hadamard Transform Optics*. Academic Press, 1979.
- [51] M. I. Kolobov and C. Fabre. Quantum limits on optical resolution. *Physical review letters*, 85(18):3789, 2000.
- [52] B. J. Lawrie and R. C. Pooser. Toward real-time quantum imaging with a single pixel camera. *Optics express*, 21(6):7549–7559, 2013.
- [53] V. Durán, P. Clemente, M. Fernández-Alonso, E. Tajahuerce, and J. Lancis. Single-pixel polarimetric imaging. *Optics letters*, 37(5):824–826, 2012.
- [54] S. S. Welsh, M. P. Edgar, R. Bowman, P. Jonathan, B. Sun, and M. J. Padgett. Fast full-color computational imaging with single-pixel detectors. *Optics express*, 21(20):23068–23074, 2013.
- [55] Ahmed A. Kirmani, A. Colaço, F. N. C. Wong, and V. K. Goyal. Exploiting sparsity in time-of-flight range acquisition using a single time-resolved sensor. *Optics Express*, 19(22):21485–21507, 2011.
- [56] B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, and M. J. Padgett. 3d computational imaging with single-pixel detectors. *Science*, 340(6134):844–847, 2013.
- [57] P. Clemente, V. Durán, E. Tajahuerce, P. Andrés, V. Climent, and J. Lancis. Compressive holography with a single-pixel detector. *Optics letters*, 38(14):2524–2527, 2013.
- [58] W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman. A single-pixel terahertz imaging system based on compressed sensing. *Applied Physics Letters*, 93(12):121105, 2008.
- [59] F. J. Taylor. *Principles of signals and systems*. McGraw-Hill Singapore, 1994.

- [60] B. I. Erkmen and J. H. Shapiro. Signal-to-noise ratio of gaussian-state ghost imaging. *Physical Review A*, 79(2):023833, 2009.
- [61] Jonathan J. Dunayevsky and D. M. Marom. Mems spatial light modulator for spectral phase and amplitude modulation. In *Optical MEMS and Nanophotonics, 2011 International Conference on*, pages 133–134. IEEE, 2011.
- [62] D. Dudley, W. M. Duncan, and J. Slaughter. Emerging digital micromirror device (dmd) applications. In *Micromachining and Microfabrication*, pages 14–25. International Society for Optics and Photonics, 2003.
- [63] W. P. Bleha and L. A. Lei. Advances in liquid crystal on silicon (lcos) spatial light modulator technology. In *SPIE Defense, Security, and Sensing*, pages 87360A–87360A. International Society for Optics and Photonics, 2013.
- [64] D. L. Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.
- [65] E. J. Candès and M. B. Wakin. An introduction to compressive sampling. *IEEE signal processing magazine*, 25(2):21–30, 2008.
- [66] J. Ma. Single-pixel remote sensing. *IEEE Geoscience and Remote Sensing Letters*, 6(2):199–203, 2009.
- [67] A. Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(8):1254–1264, 2005.
- [68] J. H. Lambert. *Photometria: sive de mensura et gradibus luminis, colorum et umbrae*. sumptibus viduae E. Klett, typis CP Detleffsen, 1760.
- [69] M. Oren and S. K. Nayar. Generalization of lambert’s reflectance model. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 239–246. ACM, 1994.

- [70] B. K. P. Horn. Shape from shading: A method for obtaining the shape of a smooth opaque object from one view. 1970.
- [71] B. K. P. Horn. The psychology of computer vision chapter“obtaining shape from shading information”, 1975.
- [72] B. K. P. Horn. Understanding image intensities. *Artificial intelligence*, 8(2): 201–231, 1977.
- [73] V. S. Ramachandran. Perception of shape from shading. *Nature*, 1988.
- [74] R. Zhang, P. Tsai, J. E. Cryer, and M. Shah. Shape-from-shading: a survey. *IEEE transactions on pattern analysis and machine intelligence*, 21(8):690–706, 1999.
- [75] K. Ikeuchi and B. K. P. Horn. Numerical shape from shading and occluding boundaries. *Artificial intelligence*, 17(1-3):141–184, 1981.
- [76] Y. Zhang, G. M. Gibson, R. Hay, R. W. Bowman, M. J. Padgett, and M. P. Edgar. A fast 3d reconstruction system with a low-cost camera accessory. *Scientific reports*, 5, 2015.
- [77] M. J. Chantler and J. Wu. Rotation invariant classification of 3d surface textures using photometric stereo and surface magnitude spectra. In *British Machine Vision Conference (BMVC)*, pages 1–10. British Machine Vision Association, 2000.
- [78] R. J. Stone. Improved statistical procedure for the evaluation of solar radiation estimation models. *Solar Energy*, 51(4):289–291, 1993.
- [79] K. Tong and M. H. Granat. A practical gait analysis system using gyroscopes. *Medical engineering & physics*, 21(2):87–94, 1999.
- [80] F. Gao, K. Shi, and S. Li. *Computational Methods for Molecular Imaging*. Lecture Notes in Computational Vision and Biomechanics. Springer International Publishing, 2015. ISBN 9783319184319.

- [81] R. J. Stone. Improved statistical procedure for the evaluation of solar radiation estimation models. *Solar energy*, 51(4):289–291, 1993.
- [82] I. Horovitz and N. Kiryati. Depth from gradient fields and control points: Bias correction in photometric stereo. *Image and Vision Computing*, 22(9):681–694, 2004.
- [83] C. H. Esteban, G. Vogiatzis, and R. Cipolla. Multiview photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3):548–554, 2008.
- [84] T. Kuparinen and V. Kyrki. Optimal reconstruction of approximate planar surfaces using photometric stereo. *IEEE transactions on pattern analysis and machine intelligence*, 31(12):2282–2289, 2009.
- [85] M. Chandraker, J. Bai, and R. Ramamoorthi. On differential photometric reconstruction for unknown, isotropic brdfs. *IEEE transactions on pattern analysis and machine intelligence*, 35(12):2941–2955, 2013.
- [86] G. Fyffe and X. Yu and P. Debevec. Single-shot photometric stereo by spectral multiplexing. In *Computational Photography (ICCP), 2011 IEEE International Conference on*, pages 1–6. IEEE, 2011.
- [87] J. H. Shapiro. Computational ghost imaging. *Physical Review A*, 78(6):061802, 2008.
- [88] Y. Bromberg, O. Katz, and Y. Silberberg. Ghost imaging with a single detector. *Physical Review A*, 79(5):053840, 2009.
- [89] T. S. Francis and X. Yang. *Introduction to optical engineering*. Cambridge University Press, 1997.
- [90] T. S. Francis and Suganda S. Jutamulia. *Optical pattern recognition*. Cambridge University Press, 1998.

- [91] A. R. Jha. *MEMS and nanotechnology-based sensors and devices for communications, medical and aerospace applications*. CRC Press, 2008.
- [92] L. J. Hornbeck. From cathode rays to digital micromirrors: A history of electronic projection display technology. *Texas Instruments Technical Journal*, 15(3):7–46, 1998.
- [93] L. A. Yoder, W. M. Duncan, E. M. Koontz, J. So, T. A. Bartlett, B. L. Lee, B. D. Sawyers, D. Powell, and P. Rancuret. Dlp technology: applications in optical networking. In *International Symposium on Optical Science and Technology*, pages 54–61. International Society for Optics and Photonics, 2001.
- [94] Y. Lu and S. Chen. Direct write of microlens array using digital projection photopolymerization. *Applied Physics Letters*, 92(4):041109, 2008.
- [95] Y. Ren, M. Li, K. Huang, J. Wu, H. Gao, Z. Wang, and Y. Li. Experimental generation of laguerre-gaussian beam using digital micromirror device. *Applied optics*, 49(10):1838–1844, 2010.
- [96] T. Ota, S. Kawata, T. Sugiura, M. J. Booth, M. A. A. Neil, R. Juskaitis, and T. Wilson. Dynamic axial-position control of a laser-trapped particle by wave-front modification. *Optics letters*, 28(6):465–467, 2003.
- [97] P. Mertz and F. Gray. A theory of scanning and its relation to the characteristics of the transmitted signal in telephotography and television. *Bell Labs Technical Journal*, 13(3):464–515, 1934.
- [98] J. Porter, H. Queener, J. Lin, K. Thorn, and A. A. S. Awwal. *Adaptive optics for vision science: principles, practices, design and applications*, volume 171. John Wiley and Sons, 2006.
- [99] T. Sun and K. Kelly. Compressive sensing hyperspectral imager. In *Computational Optical Sensing and Imaging*, page CTuA5. Optical Society of America, 2009.

- [100] J. Hadamard. Resolution dune question relative aux determinants. *Bull. sci. math*, 17(1):240–246, 1893.
- [101] K. J. Horadam. *Hadamard matrices and their applications*. Princeton university press, 2007.
- [102] A. Hedayat and W. D. Wallis. Hadamard matrices and their applications. *The Annals of Statistics*, 6(6):1184–1238, 1978.
- [103] J. Williamson. Hadamards determinant theorem and the sum of four squares. *Duke Math. J*, 11(1):65–81, 1944.
- [104] J. J. Sylvester. Lx. thoughts on inverse orthogonal matrices, simultaneous signsuccessions, and tessellated pavements in two or more colours, with applications to newton’s rule, ornamental tile-work, and the theory of numbers. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 34(232):461–475, 1867.
- [105] K. G. Beauchamp. *Walsh functions and their applications*, volume 3. Academic press, 1975.
- [106] K. W. Busch and M. A. Busch. *Multielement detection systems for spectrochemical analysis*. John Wiley & Sons, 1990.
- [107] T. Beth, D. Jungnickel, and H. Lenz. *Design theory*, volume 69. Cambridge University Press, 1999.
- [108] M. Ye, H. Ye, and G. Yan. Hadamard transform sample matrix used in compressed sensing super-resolution imaging. In *International Conference on Intelligent Robotics and Applications*, pages 796–807. Springer, 2017.
- [109] S. A. Burns, A. E. Elsner, M. B. Mellem-Kairala, and R. B. Simmons. Improved contrast of subretinal structures using polarization analysis. *Investigative ophthalmology & visual science*, 44(9):4061–4068, 2003.

- [110] A. Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1254–1264, 2005.
- [111] P. Milanfar. *Super-resolution imaging*. CRC press, 2010.
- [112] Philips. Philips Infrared Heat Lamps. URL http://images.100y.com.tw/pdf_file/21-PHILIPS-InfraredHeatLamp.pdf. Accessed: 2015-11-20.
- [113] Thorlabs. PDA20-(CS) InGaAs Switchable Gain Detector User Guide. <https://www.thorlabs.com/drawings/b03118b3426174f7-74EA272D\ -E7D7-6548-EB30876F13602720/PDA20CS-Manual.pdf>. Accessed: 2016-01-10.
- [114] P. A. Jansson. *Deconvolution of images and spectra*. Courier Corporation, 2014.
- [115] L. Reimer. *Scanning electron microscopy: physics of image formation and microanalysis*. 2000.
- [116] A. L. Baert, K. J. Johnson, and E. Bache. *Imaging in pediatric skeletal trauma: techniques and applications*. Springer Science and Business Media, 2007.
- [117] R. C. Hardie, K. J. Barnard, and R. Ordonez. Fast super-resolution with affine motion using an adaptive wiener filter and its application to airborne imaging. *Optics express*, 19(27):26208–26231, 2011.
- [118] R. C. Hardie and K. J. Barnard. Fast super-resolution using an adaptive wiener filter with robustness to local motion. *Optics express*, 20(19):21053–21073, 2012.
- [119] M. Sun and K. Yu. A sur-pixel scan method for super-resolution reconstruction. *Optik-International Journal for Light and Electron Optics*, 124(24):6905–6909, 2013.

- [120] G. Carles, J. Downing, and A. R. Harvey. Super-resolution imaging using a camera array. *Optics letters*, 39(7):1889–1892, 2014.
- [121] M. Sun, M. P. Edgar, D. B. Phillips, G. M. Gibson, and M. J. Padgett. Improving the signal-to-noise ratio of single-pixel imaging using digital microscanning. *Optics express*, 24(10):10476–10485, 2016.
- [122] P. Kner, B. B. Chhun, E. R. Griffis, L. Winoto, and M. G. L. Gustafsson. Super-resolution video microscopy of live cells by structured illumination. *Nature methods*, 6(5):339–342, 2009.
- [123] D. J. Schroeder. *Astronomical optics*. Academic press, 1999.
- [124] J. T. Bushberg and J. M. Boone. *The essential physics of medical imaging*. Lippincott Williams and Wilkins, 2011.
- [125] A. Rose. *Vision: human and electronic*. Springer Science and Business Media, 2013.
- [126] E. V. Appleton. Departure of long-wave solar radiation from black-body intensity. *Nature*, 156(3966):534–535, 1945.
- [127] M. Iqbal. *An introduction to solar radiation*. Elsevier, 2012.
- [128] ASTM. Standard. G173-03 (reapproved 2012), standard tables for reference solar spectral irradiances: Direct normal and hemispherical on 37 tilted surface. *West Conshohocken, PA: ASTM International*, 2012.
- [129] R. A. Rohde. Solar radiation spectrum. *Global Warming Art*, 9, 2007.
- [130] M. Planck. *The theory of heat radiation*. Courier Corporation, 2013.
- [131] M. Massoud. *Engineering thermofluids*. Springer, 2005.
- [132] T. P. Ackerman and G. M. Stokes. The atmospheric radiation measurement program. *Physics Today*, 56(1):38–44, 2003.

-
- [133] G. W. Petty. *A first course in atmospheric radiation*. Sundog Pub, 2006.
- [134] Allied Vision. Goldeye Datasheet G-008. <https://www.alliedvision.com/en/products/cameras/detail/Goldeye/G-008/action/pdf.html>. Accessed: 2016-02-16.
- [135] R. N. Bryan. *Introduction to the science of Medical Imaging*. Cambridge University Press, 2009.
- [136] B. Redding, M. A. Choma, and H. Cao. Speckle-free laser imaging using random laser illumination. *Nature photonics*, 6(6):355–359, 2012.
- [137] I. B. Vapnyarskii. Lagrange multipliers. *Hazewinkel, Michiel, Encyclopedia of Mathematics, Springer, ISBN*, pages 978–1, 2001.
- [138] M. P. Edgar, G. M. Gibson, R. W. Bowman, B. Sun, N. Radwell, K. J. Mitchell, S. S. Welsh, and M. J. Padgett. Simultaneous real-time visible and infrared video with single-pixel detectors. *Scientific reports*, 5:10669, 2015.
- [139] L. Mertens, M. Sonnleitner, J. Leach, M. Agnew, and M. J. Padgett. Image reconstruction from photon sparse data. *Scientific reports*, 7, 2017.