
The Geometric Documentation of Painted Surfaces

M.A.Sc. Thesis

Abhijit Dhanda

THE GEOMETRIC DOCUMENTATION OF PAINTED
SURFACES

ABHIJIT DHANDA

A thesis submitted to
the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of
the requirements for the degree of
Master of Applied Science
in
Civil Engineering
Department of Civil & Environmental Engineering
Carleton University
Ottawa, Ontario, Canada

ABSTRACT

Painted surfaces are complex heritage assets. Digital documentation for the conservation of 2.5D painted surfaces requires high-resolution data. Non-geometric approaches such as Reflectance Transformation Imaging (RTI) are used by conservators to analyze the brush strokes and condition of a painted surface through real-time relighting; it allows them to separate the colour and surface information to reveal details that might otherwise be difficult to see. Geometric detail information can provide the same insight as RTI, but with the added asset of metric information which can be used for monitoring purposes, to create facsimiles, and for valorization purposes. The issue with geometric detail documentation lies in its democratization; many methods are much more expensive than qualitative methods like RTI. This thesis explores the use of data fusion as a way to achieve high-resolution detail documentation (less than $100 \mu m$ lateral spatial resolution) at a low-cost. The data fusion combines two image-based methods: photogrammetry and photometric stereo. Photogrammetric data is effective at capturing the low-frequency information of surfaces (surface shape), but it is subject to high-frequency noise. Photometric stereo data, on the other hand, can resolve high-frequency information (surface details) well but has trouble with lower-frequencies. An optimal result can be achieved by using a Discrete Fourier Transform (DFT) and fusing the relevant frequencies from both methods. A case study comparing photogrammetry, photometric stereo, data fusion, and data from a triangulation-based scanner showed that the data fusion improves the results of the photogrammetry and photometric stereo, but is inferior in some aspects when

compared to the triangulation-based scanner. Nonetheless, data fusion is a good low-cost alternative to more expensive commercial methods.

Stopping by Woods on a Snowy Evening

Whose woods these are I think I know.

His house is in the village though;

He will not see me stopping here

To watch his woods fill up with snow.

My little horse must think it queer

To stop without a farmhouse near

Between the woods and frozen lake

The darkest evening of the year.

He gives his harness bells a shake

To ask if there is some mistake.

The only other sound's the sweep

Of easy wind and downy flake.

The woods are lovely, dark and deep,

But I have promises to keep,

And miles to go before I sleep,

And miles to go before I sleep.

-Robert Frost

ACKNOWLEDGMENTS

This thesis was supported by funding from the NSERC Create Heritage Engineering and New Paradigm New Tools programs at Carleton University. That funding resulted in research placements at the 3D Optical Metrology Unit (3DOM) of the Bruno Kessler Foundation (FBK) in Trento, Italy; and the Factum Foundation for Digital Technology in Conservation in Madrid, Spain. The Factum Foundation funded the construction of the relighting dome prototype.

The last two years and three months have been a blur of incredible experiences, and I would like to do my best to give my sincere gratitude to the people who helped me along the way.

To Fabio Remondino and the rest of my friends at the 3DOM unit of FBK – Fabio Menna, Erica Nocerino, Isabella Toschi, Danielle Morabito, Eleonora Grilli, Domenica Dininno, Filippo Carraro, Matthew Vincent, and Michele Welponer: thank you for starting my journey into research and teaching me that anyone can code (not without some frustration of course). I will never forget the mountains of Trento.

To my tutors at the CIPA Summer School – Andreas Georgopolous, Stratos Stylianidis, Francesco Fassi, and the others from 3DOM: you helped lay the foundations of my knowledge, and for that I am grateful.

To the people at the DOA in Bagan: you showed me the beautiful wall paintings in your temples, which sparked my love affair with paintings and the stories that we tell. Thank you.

To my friends and colleagues at the Factum Foundation – Adam Lowe, Otto Lowe, Pedro Miro, Carlos Bayod, Gabriel Scarpa, Enrique Esteban, Jorge Cano, Quinner Baird and many others: Thank you for your engagement, guidance, and collaboration in my research. You brought technical knowledge and a sense of creativity that put new life into the work.

To my dear friends Adam Weigert, Miquel Reina Ortiz, and Lara Chow: thank you for talking through ideas and sharing in research endeavours along the way. I would not be where I am today without your friendship.

To my supervisors Mario Santana Quintero and Stephen Fai (and honorary supervisor Laurie Smith): thank you for your curation, guidance, and for pushing me to do more.

Lastly, I would like to thank my family – from the Dhandas to the Waldocks and everyone in between: thank you for your love and support. You make every day better.

*This thesis is dedicated to my grandmother, Indu Dhanda,
whose belief in me never faltered.*

We miss you.

CONTENTS

Abstract [iii](#)

Acknowledgements [vi](#)

List of Figures [xii](#)

Acronyms [xix](#)

I INTRODUCTION

1	INTRODUCTION	2
1.1	Motivation and Scope	2
1.2	Thesis Outline	3

II LITERATURE REVIEW

2	LITERATURE REVIEW: FUNDAMENTALS	5
2.1	Scanning Systems	6
2.1.1	Laser Scanning	6
2.1.2	Structured Light Scanning	9
2.2	Photogrammetry	10
2.2.1	Data Capture	10
2.2.2	Image Pre-Processing	12
2.2.3	Feature Point Extraction and Image Matching	13
2.2.4	Bundle Adjustment	13
2.2.5	Dense Point Cloud Generation	18
2.2.6	DEM Generation	19
2.2.7	Mesh Generation	19
2.3	Photometric Stereo	20

2.3.1	BRDF	21
2.3.2	Basic Photometric Stereo	22
2.3.3	Other Photometric Stereo Approaches	24
2.3.4	Surface to Shape Integration	26
2.4	Reflectance Transformation Imaging	27
3	LITERATURE REVIEW: STATE OF THE ART	31
3.1	Photogrammetry	31
3.2	Scanning Techniques	34
3.3	Photometric Stereo	42
3.4	Data Fusion	44
III RESEARCH		
4	APPROACH	53
4.1	Photometric Stereo	55
4.2	Data Fusion	61
5	TESTING THE RELIGHTING DOME	63
5.1	The One-Arm Test	63
5.2	Testing Methodology	65
5.3	Test One	66
5.4	Test Two	68
6	CASE STUDY	77
6.1	Data Capture	77
6.1.1	Photogrammetric Acquisition	77
6.1.2	Photometric Stereo Acquisition	78
6.2	Processing	79
6.2.1	Pre-processing	79
6.2.2	Photogrammetry	80

6.2.3	Photometric Stereo	83
6.3	Data Fusion	85
6.4	Workflow and Discussion	98
IV CONCLUSION		
7	CONCLUSION	102
7.1	Conclusion	102
7.2	Future Work	104
V APPENDIX		
A	RELIGHTING DOME: DESIGN AND CONSTRUCTION	106
A.1	Introduction	106
A.2	Dome Designs	107
A.3	Prototype Design	110
A.4	Prototype Construction	115
	BIBLIOGRAPHY	125

LIST OF FIGURES

Figure 2.1	The number of articles published related to documentation technologies from 1975-2009. (Eppich and Hadzic, 2013)	5
Figure 2.2	Heritage documentation techniques organized by scale and complexity. (Boehler, Heinz and Marbs, 2001; Moitinho, 2018)	6
Figure 2.3	A point cloud with pseudocolour intensity values (left), grayscale intensity values (middle), and colour values (right). (Grussenmeyer and Landes, 2016)	7
Figure 2.4	The geometry of a triangulation scanning system. (Grussenmeyer and Landes, 2016)	8
Figure 2.5	The setup of a Digital Fringe Projection (DFP) scanner. (Zhang, 2010)	9
Figure 2.6	The photogrammetric process. (Historic England, 2017)	11
Figure 2.7	The interior orientation of a camera. (Luhmann et al., 2011)	14
Figure 2.8	The collinearity condition. (Han, Guo and Chou, 2011) - modified	17
Figure 2.9	The coplanarity condition. (Luhmann et al., 2011)	18
Figure 2.10	The projection of a DEM and orthophoto. (Historic England, 2017)	20
Figure 2.11	The Digital Elevation Model (DEM) (left), orthophoto (middle), and textured mesh (right) of a landscape. (Historic England, 2017)	20

Figure 2.12	An overview of photometric stereo. (Commons, 2015)	21
Figure 2.13	The components of a Bidirectional Reflectance Distribution Function (BRDF). (Herbort and Wöhler, 2011)	22
Figure 2.14	Light transport from the point light source, reflected by the object surface, and sensed by the camera. (Trucco and Verri, 1998)	24
Figure 2.15	An example of Reflectance Transformation Imaging (RTI) of a painting. The painting lit virtually from the front (left), and from the right (right). (Padfield, Saunders and Malzbender, 2005)	27
Figure 2.16	An example of lighting capture for RTI. (Duffy et al., 2018)	28
Figure 2.17	An overlaid comparison (using horizontal stripes) of Polynomial Texture Mapping (PTM) vs. Hemispherical Harmonics (HSH), showing that the hemispherical harmonic representation better preserves contrast from the original images. (Mudge et al., 2008)	30
Figure 3.1	Analysis of the Saint John the Baptist painting. (Abate et al., 2014)	33
Figure 3.2	A 3D image of Dame en prière scanned at a depth resolution of 10 μm . (Lahanier et al., 2005)	38
Figure 3.3	Simplified model of the paint layer including a translucent layer of finely cracked old varnish and/or transparent glaze. (Blais et al., 2008)	39
Figure 3.4	The documentation of The Triumph of the Eucharist over Idolatry by Paul Rubens. (Factum Foundation, 2015)	40
Figure 3.5	The documentation of a self-portrait by Rembrandt using a DFP scanner. (Zaman, 2013)	41

Figure 3.6	Image rendering results for a glossy oil painting (varnished). (Hasegawa et al., 2011)	43
Figure 3.7	The comparison of the far-light and near-light photometric stereo models on a woodblock created by Paul Gauguin. (Huang et al., 2016)	44
Figure 3.8	Quality assessment from photometric stereo, a triangulation scanner, and data fusion. (c-d) The combination of the data significantly reduces the error from both methods. (Nehab et al., 2005)	49
Figure 3.9	The results of the fusion algorithms on a coin: (a) Initial scanner depth; (b) gradient-based; (c) Method of Nehab; (d) Generalized Nehab; (e) Total Generalized Variation (TGV). (Antensteiner, Štolc and Pock, 2018)	50
Figure 3.10	A comparison of Poisson surface reconstruction (left), Nehab et al., 2005 (middle), and Berkiten, Fan and Rusinkiewicz, 2014 (right) on Egyptian scarabs. (Berkiten, Fan and Rusinkiewicz, 2014)	50
Figure 3.11	A Comparison of reconstruction methods. (Yeh et al., 2016)	51
Figure 4.1	Comparing different documentation methods.	53
Figure 4.2	The scaled depth and calibration target from the photometric stereo of Paul Gauguin's Nativity using the method of Huang et al., 2016. It can be seen in the calibration target that there is still a low-frequency warping in the data – the base should be flat. (Cossairt et al., 2015)	55
Figure 4.3	The anonymous 17th century painting of the Virgin Mary.	57
Figure 4.4	A comparison of the Huang et al., 2016 near-light algorithm and the Wu et al., 2011 far-light algorithm.	59

Figure 4.5	Results from Huang et al., 2016 near-light algorithm.	60
Figure 4.6	Qualitative comparison of the high-frequency data of the Virgin Mary using the Huang et al., 2016 photometric stereo algorithm (left) and the Lucida (right). © Oak Taylor Smith for the Factum Foundation for Digital Technology in Conservation	61
Figure 5.1	Documenting the Virgin Mary painting using one arm of the relighting dome. © Oak Taylor Smith for the Factum Foundation for Digital Technology in Conservation	63
Figure 5.2	Comparing the results of free-hand and arm-based photometric stereo approaches using a near-light model.	64
Figure 5.3	The Standard Board for testing scanning techniques. The board contains challenging areas (dark and glossy) that are common in paintings.	65
Figure 5.4	Comparing the calibration target from Cossairt et al., 2015 (left) and the first relighting dome test (right).	66
Figure 5.5	The setup for test one. © Otto Lowe for the Factum Foundation for Digital Technology in Conservation	67
Figure 5.6	The results of test one on the Standard Board.	67
Figure 5.7	The second iteration of height calibration targets.	68
Figure 5.8	One image from the light calibration set (left) and the detected light directions on the black sphere (right).	69
Figure 5.9	Comparing the uncalibrated (left) and calibrated light positions (right).	70
Figure 5.10	The depth from test two (top) and the calibrated normals (bottom).	71

Figure 5.11	The tested lighting configurations (red is active, and black is inactive).	72
Figure 5.12	The angular differences between tested lighting configurations with 16 lights and the ground truth (all 32 light positions).	73
Figure 5.13	The angular differences between tested lighting configurations with 24 lights and the ground truth (all 32 light positions).	74
Figure 5.14	The integrated depth from test two on the Standard Board.	76
Figure 6.1	The photometric stereo documentation of the painting of the Virgin Mary © Osama Dawod for the Factum Foundation for Digital Technology in Conservation.	79
Figure 6.2	The image poses in relation to the Virgin Mary painting. The dark blue images are the photometric stereo tiles.	80
Figure 6.3	The coloured point cloud of the Virgin Mary painting captured using photogrammetry.	81
Figure 6.4	The DEM of the Virgin Mary painting captured using photogrammetry.	82
Figure 6.5	The calculated normals for two adjacent photometric stereo tiles of the Virgin Mary painting.	84
Figure 6.6	The orthophoto normal blending of the Virgin Mary painting.	86
Figure 6.7	An enlarged view of the angular differences between the two normal blending methods.	88
Figure 6.8	A shaded render of the depth maps from pixel-based scaling (top) and target-based scaling (bottom).	89

- Figure 6.9 A comparison of the different depth scaling methods for the photometric stereo data. 90
- Figure 6.10 Depth profiles comparing the photometric stereo depth and the photogrammetric DEM. The red line is the photometric stereo depth and the green line is the DEM. 91
- Figure 6.11 The log power distribution of the spatial frequencies of the DEM (left) and photometric stereo depth map (right). 92
- Figure 6.12 The cumulative sum of the log power of spatial frequencies of the DEM (left) and photometric stereo depth map (right) measured in radial pixels from the centre. 92
- Figure 6.13 The difference between data fusion of the DEM and photometric stereo scaled using the calibration target and pixel methods. 93
- Figure 6.14 The depth map of the Virgin Mary painting captured using the Lucida. 94
- Figure 6.15 A depth profile comparing the data fusion (top of the profile) to the Lucida data (bottom of the profile) at an area with steep edges. 95
- Figure 6.16 A depth profile comparing the photometric stereo depth (red) to the Lucida data (yellow). 96
- Figure 6.17 A depth profile comparing the data fusion (top profile) to the Lucida data (middle profile) and the photometric stereo depth (bottom profile). 96
- Figure 6.18 A depth profile comparing the data fusion with a radius of 40 pixels (top profile) to one with a radius of 100 pixels (bottom profile). 97

Figure 6.19	Comparing the details of the Lucida (top) and the data fusion (bottom) in a shaded render of the depth maps.	98
Figure 6.20	The overall workflow developed from the case study.	100
Figure A.1	An example of a photometric stereo dome. (Watteeuw et al., 2016)	107
Figure A.2	The relighting dome from Palma et al., 2014.	108
Figure A.3	A 3D printed relighting dome designed for macro RTI. (Kinsman, 2016)	109
Figure A.4	The RTI dome designed by Pawlowicz, 2016.	109
Figure A.5	The RTI dome from Mudge et al., 2006.	112
Figure A.6	The initial design of the relighting dome.	113
Figure A.7	The first design of the relighting dome prototype.	114
Figure A.8	The mitre boxes used to cut the aluminum extrusions.	115
Figure A.9	One arm of the prototype relighting dome.	116
Figure A.10	The LEDs tested.	118
Figure A.11	The capture setup of the first LED test.	119
Figure A.12	The results of the second LED test.	119
Figure A.13	The two arms connected to the test circuit.	120
Figure A.14	The arms connected to the circular plate.	120
Figure A.15	The circuit board design and construction.	122
Figure A.16	The dome tested in its downward position. A large tripod was used to cantilever over the dome and through the centre. © Otto Lowe for the Factum Foundation for Digital Technology in Conservation	123
Figure A.17	The L-Bracket.	123
Figure A.18	The relighting dome in vertical configurations.	124
Figure A.19	The cost of the relighting dome prototype.	124

ACRONYMS

BRDF	Bidirectional Reflectance Distribution Function
CCD	Charge-Coupled Device
CHI	Cultural Heritage Imaging
DEM	Digital Elevation Model
DFP	Digital Fringe Projection
DFT	Discrete Fourier Transform
DIM	Dense Image Matching
DSLR	Digital Single Lens Reflex
FFT	Fast Fourier Transform
FOV	Field of View
GCP	Ground Control Point
GSD	Ground Sample Distance
HSH	Hemispherical Harmonics
ICP	Iterative Closest Point
IR	Infrared
MOSFET	Metal-Oxide-Semiconductor Field-Effect Transistor

MVS	Multi-View Stereo
NIR	Near-infrared
NRC	National Research Council
PTM	Polynomial Texture Mapping
RANSAC	Random Sample Consensus
RTI	Reflectance Transformation Imaging
SfM	Structure-from-Motion
SfS	Shape from Shading
SIFT	Scale-Invariant Feature Transform
SURF	Speeded Up Robust Features
TGV	Total Generalized Variation
ToF	Time-of-Flight
TLS	Terrestrial Laser Scanning
UNESCO	United Nations Educational, Scientific, and Cultural Organization
UV	Ultraviolet
VIS	Visible Spectrum

Part I

INTRODUCTION

“Art is a lie that makes us realize truth, at least the truth that is given us to understand.”

– Pablo Picasso

INTRODUCTION

1.1 MOTIVATION AND SCOPE

This research aimed to test and develop a workflow for the low-cost digital documentation of planar painted surfaces – 2.5D surfaces. To be useful for conservation purposes, the capture of a painted surface requires a spatial resolution and precision of 0.01 mm to 1 mm (Remondino et al., 2011). Smaller municipalities and museums do not have money for expensive commercial scanning solutions that can produce those resolutions. The method proposed by this research involves data fusion between photometric stereo and photogrammetry – both low-cost documentation techniques that require one camera and several lights. It should be noted that low-cost refers to the hardware required. A data fusion approach would require a more experienced user. This thesis will discuss the design and testing of an inexpensive relighting dome as well as the data fusion methodology.

1.2 THESIS OUTLINE

Part 2 presents a literature review that is divided into the fundamental concepts ([chapter 2](#)) and state of the art methods for painting documentation ([chapter 3](#)). Part 3 discusses the research and case studies in detail and is divided into the research approach ([chapter 4](#)), the testing of the relighting dome ([chapter 5](#)), and a case study and data fusion workflow. ([chapter 6](#)). Part 4 presents the conclusions and future work for this thesis ([chapter 7](#)). Lastly, details on the design and construction of the dome prototype can be found in the Appendix ([Appendix A](#)).

Part II

LITERATURE REVIEW

“The introduction of numbers as coordinates is an act of violence”

– Hermann Weyl

LITERATURE REVIEW: FUNDAMENTALS

The easiest way to divide 3D heritage documentation is by the scales of the heritage place or asset and the density of the capture method (Figure 2.2).

The focus of this thesis is with the documentation of painted surfaces typically at a scale of less than ten metres. The following sections will introduce methods of documentation within that scales range.

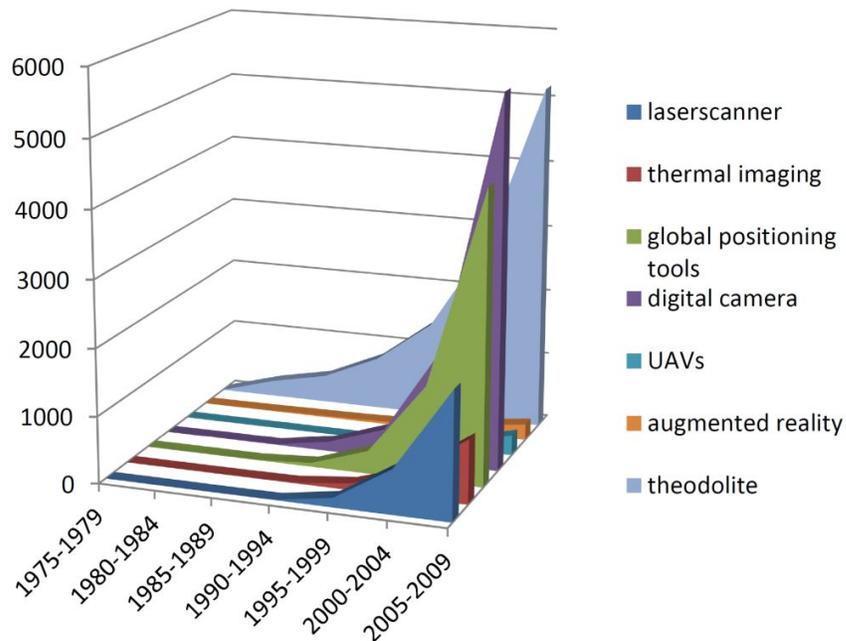


Figure 2.1: The number of articles published related to documentation technologies from 1975-2009. (Eppich and Hadzic, 2013)

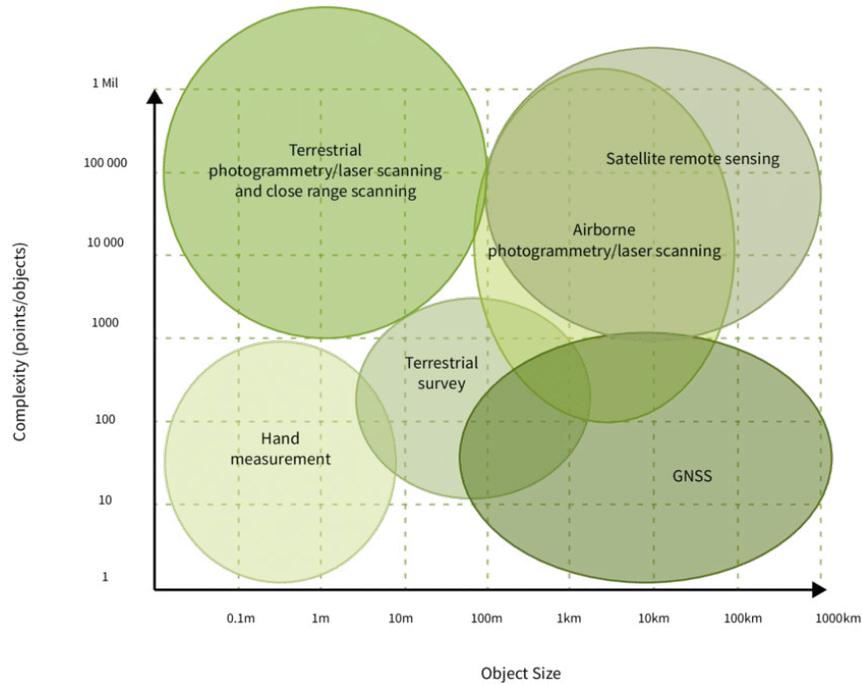


Figure 2.2: Heritage documentation techniques organized by scale and complexity. (Boehler, Heinz and Marbs, 2001; Moitinho, 2018)

2.1 SCANNING SYSTEMS

Scanning systems use active measuring techniques to make measurements between the scanner and an object or surface. The scanners make discrete point measurements by calculating angles and distances. Each point has a 3D coordinate with the possible addition of an intensity and/or colour value. The total set of points from a scanner produce a point cloud. (Figure 2.3).

2.1.1 Laser Scanning

Grussenmeyer and Landes, 2016 defines laser scanning as:



Figure 2.3: A point cloud with pseudocolour intensity values (left), grayscale intensity values (middle), and colour values (right). (Grussenmeyer and Landes, 2016)

“... an active, fast and automatic acquisition technique using laser light for measuring, without any contact and in a dense regular pattern, 3D coordinates of points on surfaces.”

There are three main types of laser scanners: pulse-based scanners, phase-based scanners, and triangulation-based scanners (Grussenmeyer and Landes, 2016). Pulse-based laser scanners emit laser pulses that reflect from a surface and return to the scanner. When the signal returns, the scanner calculates the point distance based on the time between the pulse emission and return (Equation 2.1) (Grussenmeyer and Landes, 2016). Phase-based scanners emit a continuous wave and calculate the distance based on the shift in phase of the wave (Grussenmeyer and Landes, 2016). Because both of their methods are related to time, pulse-based and phase-based laser scanners are commonly categorized labelled as Time-of-Flight (ToF) scanners. ToF scanners are suited for larger scale measurements because their maximum resolution is usually limited to the range of several millimetres (Grussenmeyer and Landes, 2016).

$$D = c \frac{t}{2} \quad (2.1)$$

where

D distance

c speed of light in medium

t time

Triangulation-based laser scanners use the relationship between a fixed camera and laser to make measurements (Figure 2.4). The distance between the scanner and object can be calculated using Equation 2.2 (Grussenmeyer and Landes, 2016). Triangulation scanners are suited for close up measurements because the range accuracy of the scanner decreases by the square of the distance between the scanner and the object (Lerma Garcia et al., 2008).

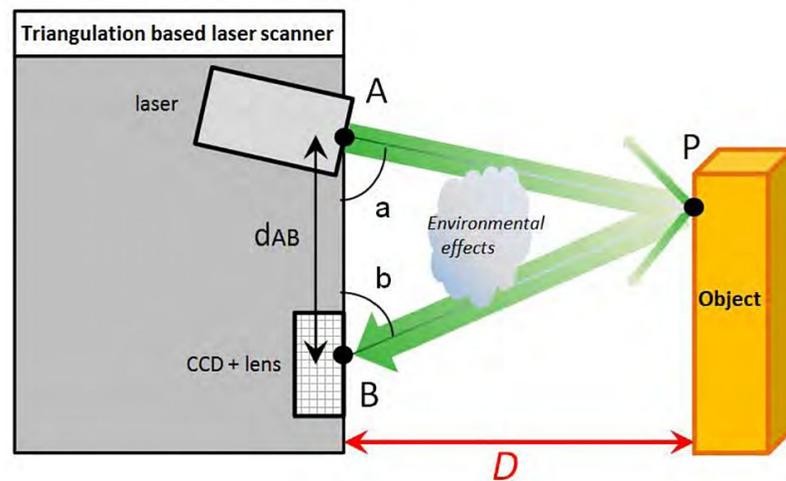


Figure 2.4: The geometry of a triangulation scanning system. (Grussenmeyer and Landes, 2016)

$$D = d_{AB} \frac{1}{\cot \alpha + \cot \beta} \quad (2.2)$$

where

D distance

d_{AB} base length between receiver and transmitter

α angle between emitted beam direction and the AB axis

β angle between received beam direction and the AB axis

2.1.2 Structured Light Scanning

Structured light scanning is similar to triangulation-based scanning except instead of the laser it uses a projector to project a pattern onto the object (Grussenmeyer and Landes, 2016). The warping of the patterns on the object allow the scanner to calculate the geometry. One of the popular structured light methods, *DFP*, projects a sinusoidal pattern onto the object (Figure 2.5) (Webster et al., 2016).

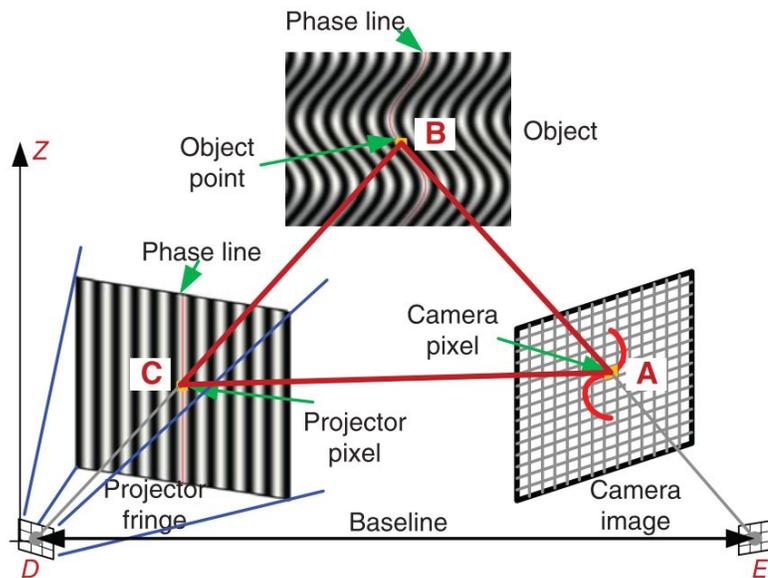


Figure 2.5: The setup of a *DFP* scanner. (Zhang, 2010)

2.2 PHOTOGRAMMETRY

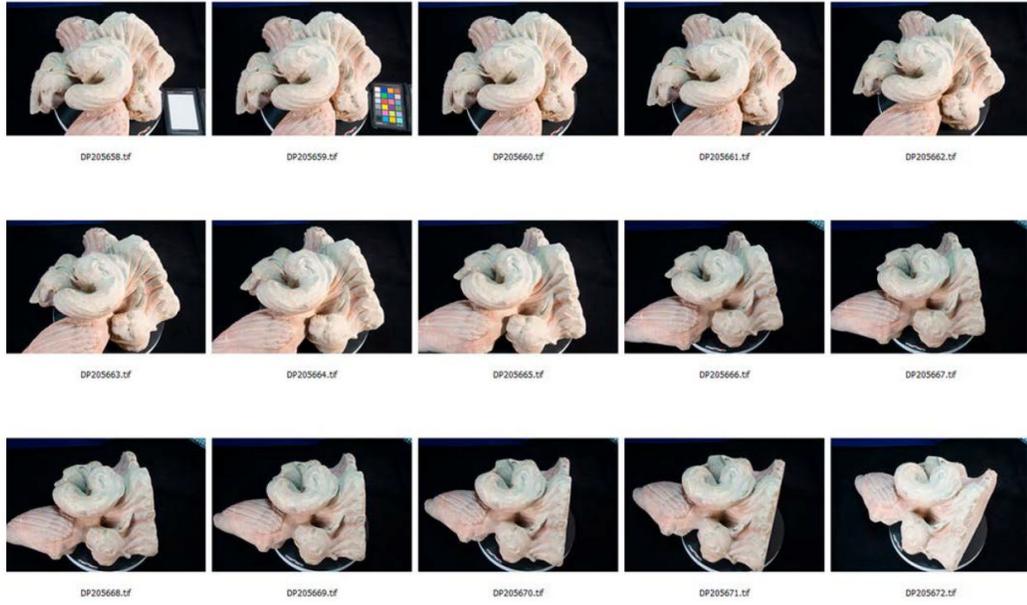
Photogrammetry is a passive, optical documentation technique that uses photographs to make measurements. Photogrammetry started in the middle of 19th century with a French military officer, Aimé Laussedat, and a German architect, Albrecht Meydenbauer (Luhmann et al., 2011). As digital cameras have developed and computational power has increased, photogrammetry has moved from a manual process to one that is semi-automatic. The semi-automated photogrammetric process is referred to as Structure-from-Motion (SfM) and Multi-View Stereo (MVS).

The SfM MVS photogrammetric pipeline involves several main steps (Figure 2.6) (Historic England, 2017):

1. Data capture
2. Image pre-processing
3. Feature point extraction and image matching
4. Bundle adjustment
5. Dense point cloud generation
6. DEM generation (optional)
7. Mesh generation (optional)

2.2.1 Data Capture

Photogrammetric data is acquired by capturing a series of overlapping images – with at least 60% overlap – of the target object (Waldhausl and Ogleby, 1994). The



(a) Overlapping images of the subject.



(b) Tie points from image matching.



(c) Camera locations from the bundle adjustment.



(d) The dense point cloud.

Figure 2.6: The photogrammetric process. (Historic England, 2017)

distance between the camera and object during the acquisition is dependent on the Ground Sample Distance (GSD). The GSD is the size of an image pixel in real world measurements and is dependent on the requirements for a particular project Equation 2.3. In practice, due to imaging artifacts and triangulation errors, the GSD that is captured is two to three times smaller than the project requirements (Stylianidis, Georgopoulos and Remondino, 2016).

$$\frac{GSD}{ps} = \frac{D}{f} \quad (2.3)$$

where

GSD Ground Sample Distance

ps pixel size of the camera sensor

D distance between the camera and the object

f focal length of the camera

A set of images is enough to reconstruct a *relative* version of the object; it has no scale. A set of measurements is needed to make the measurements *absolute*: a set of two-point, distance measurements to scale the model; or at least three 3D point measurements to place the model in a coordinate system (Luhmann et al., 2011).

2.2.2 Image Pre-Processing

This step is typically done by the software without the user's knowledge. It involves converting the image from colour to greyscale to improve the performance and efficiency of the feature detection algorithms (Verhoeven et al., 2015). For example, Speeded Up Robust Features (SURF) and Scale-Invariant Feature Transform (SIFT), two of the most popular feature detection algorithms were designed to work on greyscale images (Bay et al., 2008; Lowe, 1999).

2.2.3 *Feature Point Extraction and Image Matching*

Feature points, also called interesting points and key points, are found algorithmically in each image. These points are usually corners – areas of high contrast (Stylianidis, Georgopoulos and Remondino, 2016). Each feature point is given a unique descriptor that describes the point using the pixels surrounding it. Image matching algorithms use the descriptors to look for the same points in the other images of the data set. The matched points are called tie points (Stylianidis, Georgopoulos and Remondino, 2016). The matches will inevitably contain false positives, which are filtered out using algorithms like the Random Sample Consensus ([RANSAC](#)) algorithm (Hartmann, Havlena and Schindler, 2016).

2.2.4 *Bundle Adjustment*

The bundle adjustment solves for the exterior orientation parameters of the cameras. The exterior orientation is the position in 3D space of each image relative to the documented object (Luhmann et al., 2011). This process, however, requires the interior orientation parameters of the camera to be known (Luhmann et al., 2011). Most consumer cameras are not metric; the interior orientation – the relationship between the sensor and the lens – is variable and unknown. Digital Single Lens Reflex ([DSLR](#)) cameras, for example, have interchangeable lenses, as well as lenses that can change their focal length – zoom lenses. So, if the camera has not been pre-calibrated, as is often the case, a self-calibrating bundle adjustment needs to be performed. This type of bundle adjustment simultaneously solves for the interior and exterior orientation parameters of the cameras (Luhmann et al., 2011).

2.2.4.1 Parameters of Interior Orientation

The interior orientation of a camera is defined by the location of the perspective centre (O'), the principal distance (c), and the location of the principal point (H') (Luhmann et al., 2011). A central perspective is used as an ideal case when determining the interior parameters of camera (Stylianidis, Georgopoulos and Remondino, 2016). The distortions are measured as deviations from this model and are also considered as parameters of interior orientation (radial distortion, tangential distortion, etc.) (Luhmann et al., 2011). Figure 2.7 shows the interior orientation of a camera as a ray as it passes through the lens.

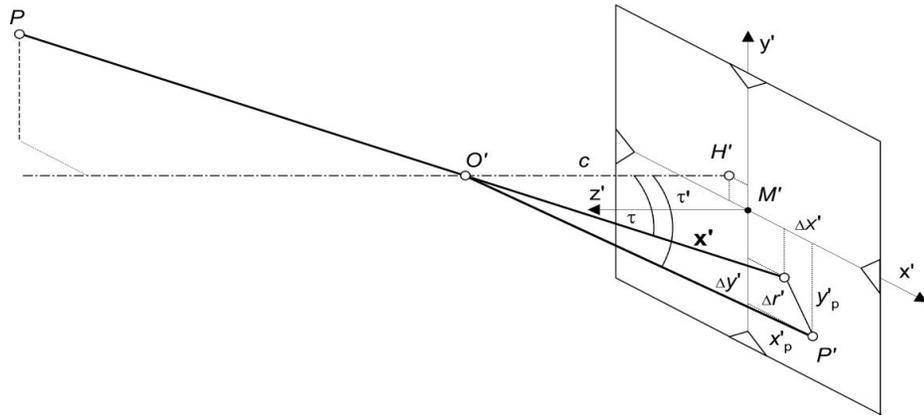


Figure 2.7: The interior orientation of a camera. (Luhmann et al., 2011)

RADIAL DISTORTION Radial symmetric distortion, the largest of the lens distortion effects, occurs because of different refractive properties of elements within the lens (Luhmann et al., 2011). The properties of radial distortion change with each lens, but also change as the focus of an individual lens is changed (Luhmann et al., 2011). Radial distortion increases as the image radius (r') increases (as you get further from the principal point and get closer to the edge of the image) (Luhmann et al., 2011). The radial distortion is modelled to a polynomial that is repres-

ented by the distortion parameters K_n (Equation 2.4 and Equation 2.5) (Luhmann et al., 2011).

$$\Delta r' = K_0 r' + K_1 r'^3 + K_2 r'^5 + K_3 r'^7 \quad (2.4)$$

$$\begin{aligned} \Delta x'_{rad} &= x' \frac{\Delta r'_{rad}}{r'} \\ \Delta y'_{rad} &= y' \frac{\Delta r'_{rad}}{r'} \end{aligned} \quad (2.5)$$

TANGENTIAL DISTORTION Tangential distortion, which is also known as radial-asymmetric distortion, occurs when lens elements are misaligned (Stylianiadis, Georgopoulos and Remondino, 2016). The effects of tangential distortion can be described by Equation 2.6, and are minimal when compared to radial distortion (Luhmann et al., 2011).

$$\begin{aligned} \Delta x'_{tan} &= B_1(r'^2 + 2x'^2) + 2B_2x'y' \\ \Delta y'_{tan} &= B_2(r'^2 + 2y'^2) + 2B_1x'y' \end{aligned} \quad (2.6)$$

AFFINITY AND SHEAR Affinity and shear occur in DSLR cameras when sensor cells are rectangular instead of square (Luhmann et al., 2011). The effects can be calculated from Equation 2.7.

$$\begin{aligned} \Delta x'_{aff} &= C_1 x' + C_2 y' \\ \Delta y'_{aff} &= 0 \end{aligned} \quad (2.7)$$

The total correction of distortion effects can be described by [Equation 2.8](#) (Luhmann et al., 2011).

$$\begin{aligned}\Delta x' &= \Delta x'_{rad} + \Delta x'_{tan} + \Delta x'_{aff} \\ \Delta y' &= \Delta y'_{rad} + \Delta y'_{tan} + \Delta y'_{aff}\end{aligned}\tag{2.8}$$

2.2.4.2 Parameters of Exterior Orientation

The exterior orientation of an image consists of six unknowns: its position (X, Y, Z) and rotation $(\omega, \varphi, \kappa)$ relative to the object (Stylianidis, Georgopoulos and Remondino, 2016). The poses of the images are determined using two main factors: collinearity and coplanarity.

Most photogrammetric calculations use central perspective projection as the fundamental model used to connect the image coordinates (x, y) and the object coordinates (X, Y, Z) . According to central perspective, each object point seen from an image is collinear – on a line with – a point on the image and the perspective centre (O) ([Figure 2.8](#), [Equation 2.9](#)) (Stylianidis, Georgopoulos and Remondino, 2016).

$$\begin{aligned}\frac{x_a}{-c} &= \frac{R_{11}(X_A - X_0) + R_{12}(Y_A - Y_0) + R_{13}(Z_A - Z_0)}{R_{31}(X_A - X_0) + R_{32}(Y_A - Y_0) + R_{33}(Z_A - Z_0)} \\ \frac{y_a}{-c} &= \frac{R_{21}(X_A - X_0) + R_{22}(Y_A - Y_0) + R_{23}(Z_A - Z_0)}{R_{31}(X_A - X_0) + R_{32}(Y_A - Y_0) + R_{33}(Z_A - Z_0)}\end{aligned}\tag{2.9}$$

where

X_A, Y_A, Z_A the coordinates of point A in object space

X_0, Y_0, Z_0 the coordinates of the perspective centre in world coordinates

x_a, y_a the coordinates of point A in image space

c principal distance

R rotation matrix

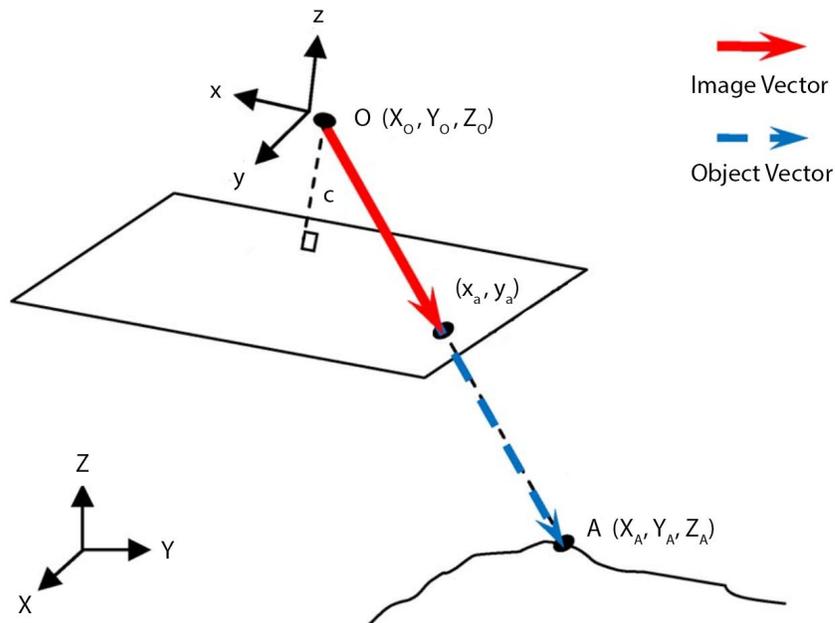


Figure 2.8: The collinearity condition. (Han, Guo and Chou, 2011) - modified

The relative exterior orientation of two images can be explained using the coplanarity condition. The coplanarity condition states that a point (P) that can be seen from a stereo pair will create a plane (known as the epipolar plane) between the point, and the two perspective centres of the cameras. On this plane, there are two epipolar lines (k' and k'') (Luhmann et al., 2011). The epipolar lines describe the parallax (difference in position of a point) between the two images. A homologous point (or tie point) is a point that can be found in both images. Using the epipolar lines, a homologous point (Q) can be found quickly by searching along the epipolar line (Figure 2.9) (Luhmann et al., 2011).

The self-calibrating bundle adjustment can then use the image matching results to apply these principles to the block of images to determine the interior and exterior orientation parameters of the cameras relative to the object. First, two images are oriented relatively, like a stereo pair, and then other images from the image set are oriented to that stereo pair and adjusted in an iterative process (Luhmann et

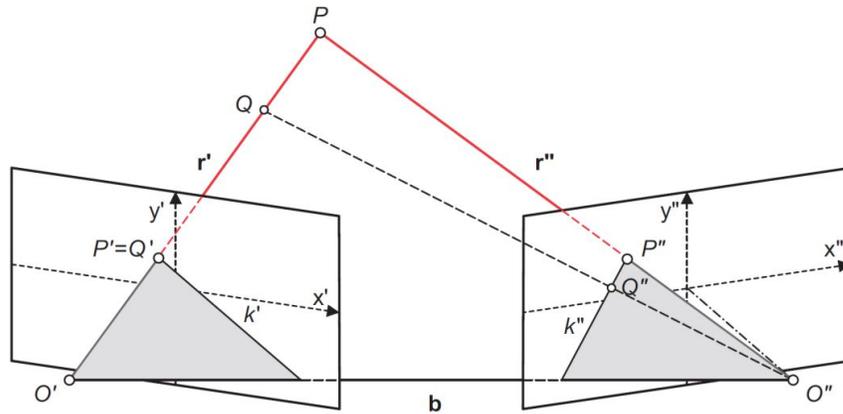


Figure 2.9: The coplanarity condition. (Luhmann et al., 2011)

al., 2011). Once the parameters have been found, the externally measured points (each called a Ground Control Point (GCP)) can be used to move the cameras and object model from a model coordinate system to an object or global coordinate system (Stylianidis, Georgopoulos and Remondino, 2016). At least three GCPs are needed to solve for the seven unknowns – three translation, three rotation, and one scaling – and move to the defined coordinate system (Luhmann et al., 2011).

2.2.5 Dense Point Cloud Generation

When position and interior orientation are known, the remaining problem becomes that of 3D reconstruction. In this step, called Dense Image Matching (DIM), the initial image matching is expanded on by increasing the matching density between corresponding images to include most of the pixels of the images (Stylianidis, Georgopoulos and Remondino, 2016). These matches are turned into individual depth maps for each image and projected into 3D space to create discrete points (Stylianidis, Georgopoulos and Remondino, 2016). The discrete points have

position and colour values and are collectively called a dense point cloud. This point cloud is similar in nature to the point cloud produced from ToF scanners.

2.2.6 DEM Generation

A DEM is a 2D, orthographic representation of a 3D surface. A DEM uses pixel locations to represent X and Y coordinates, and pixel values to represent depth (Stylianidis, Georgopoulos and Remondino, 2016). DEMs – also sometimes called depth images, depth maps, or height maps – are primarily used for terrain mapping. A DEM can be created by sampling the dense point cloud to a grid based on a projection direction. Orthophotos are created in similar way (Figure 2.10) (Stylianidis, Georgopoulos and Remondino, 2016).

2.2.7 Mesh Generation

A triangulated mesh is generated from a dense point cloud by sampling and interpolating between the points of the object (Historic England, 2017). Using the sampled points and their normal direction, a series of triangles are created to follow the shape and structure of the capture object.

Figure 2.11 shows the relationship between the DEM, orthophoto, and mesh of a surface.

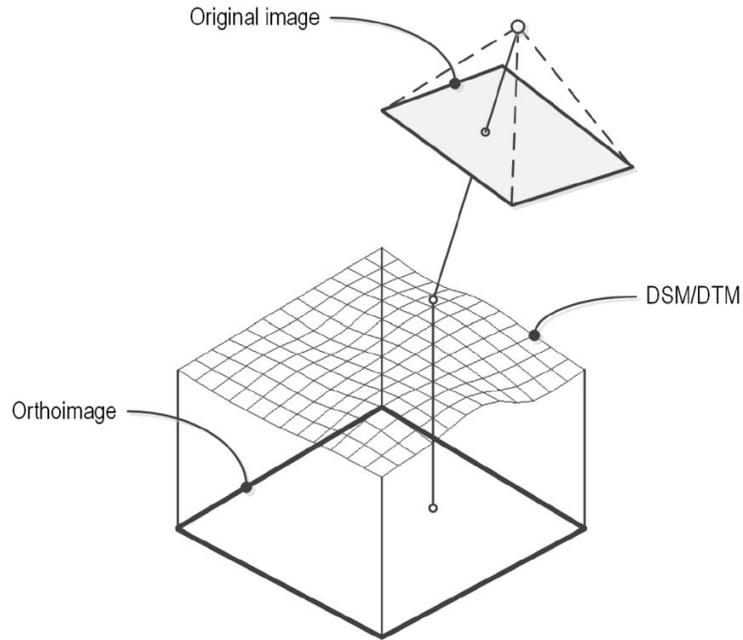


Figure 2.10: The projection of a DEM and orthophoto. (Historic England, 2017)



Figure 2.11: The DEM (left), orthophoto (middle), and textured mesh (right) of a landscape. (Historic England, 2017)

2.3 PHOTOMETRIC STEREO

Photometric Stereo is a Shape from Shading (SfS) technique in which multiple images are captured from a single camera position with varying light positions – typically in a hemispherical pattern – in order to estimate the normals of a surface at each pixel location (Figure 2.12). The surface normal can then be converted to surface gradients, and the 3D shape of the object can be calculated. The goal of most SfS algorithms is to reconstruct a shape from a single image with a single light. Single image reconstruction can be tricky; it requires many assumptions to

be made. Photometric stereo overcomes some of those assumptions by instead using multiple images to create an overdetermined system of equations with N light positions for N photographs.

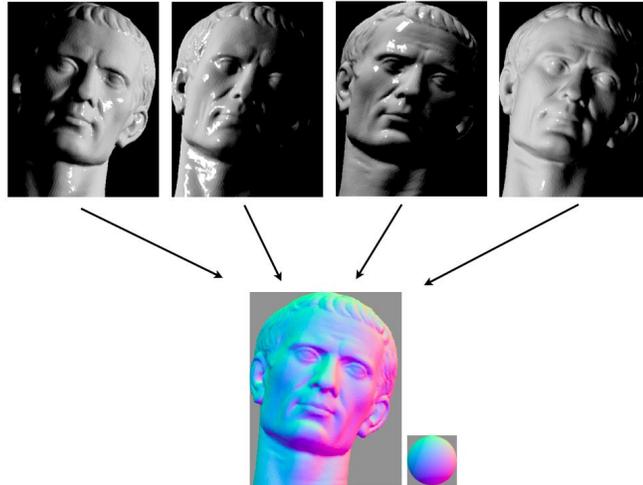


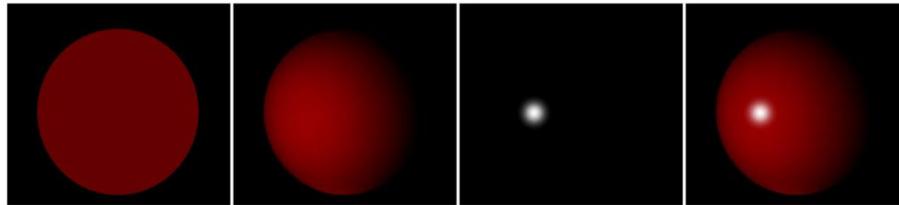
Figure 2.12: An overview of photometric stereo. (Commons, 2015)

2.3.1 BRDF

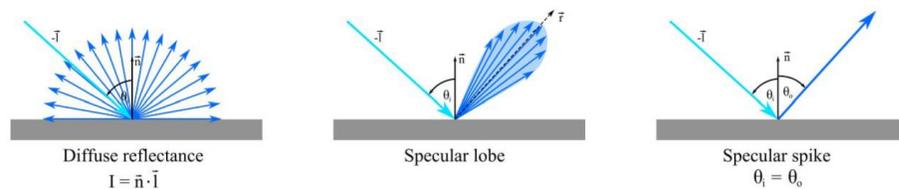
To truly understand photometric stereo, one must first understand the basics of surface properties. The **BRDF** is a common method for describing the reflectance of a surface. In its fundamental form, the **BRDF** has four variables: the amount of light coming to the surface (irradiance E_i), the amount of light leaving the surface (radiance L_0), the incident angles (azimuth ϕ_i , elevation θ_i), and the emission angles (ϕ_r, θ_r) (Equation 2.10) (Herbort and Wöhler, 2011).

$$BRDF = \frac{L_0}{E_i} = f_r(\theta_i, \phi_i, \theta_r, \phi_r) \quad (2.10)$$

BRDFs can be divided into two types of models: empirically based models (Lambertian model, and Phong model), or physically based models (Ward model, Torrance-Sparrow Model, and Cook-Torrance model) (Herbort and Wöhler, 2011). The BRDF models describe the surface reflections using a combination of a diffuse reflection, and/or a specular lobe, and a specular spike (Figure 2.13).



(a) The appearance achieved by the Phong model (right) which consists of an ambient component (left), a diffuse component (middle-left), and a specular component (middle-right).



(b) Reflectance phenomena in many photometric BRDF models.

Figure 2.13: The components of a BRDF. (Herbort and Wöhler, 2011)

2.3.2 Basic Photometric Stereo

The basic photometric stereo approach assumes that the surface being photographed has Lambertian reflectance; all of the radiance is evenly diffuse. This assumption places constraints on the surface and light behaviour during acquisition: the light is infinitely far away, or at least several times further (also known as the far-light model); the lights are the same distance from the surface; the lighting directions are known; the surface is not subject to inter-reflections or external shadows; and that the camera is orthographic (Herbort and Wöhler, 2011). Equation 2.11 describes the relationship between the photographed pixels, the light

positions, the surface albedo, and the surface normal at each pixel (Figure 2.14). The normal and surface albedo are both unknowns in the equation. The normal and albedo are solved for simultaneously in photometric stereo, which is why the three normal components (x , y , and z) constrain the unknowns in the equation; at least three images are needed to solve a photometric stereo problem (Macdonald, 2015).

$$I = \rho L n \tag{2.11}$$

where

I pixel intensities

ρ surface albedo

L lighting direction vectors

n surface normal

Equation 2.11 can be rearranged to simultaneously solve for the normals and albedo – taking advantage of the fact that the normals are unit vectors (Equation 2.12).

$$\rho n = L^{-1} I \tag{2.12}$$

Some of the base photometric stereo assumptions are easily met or overcome. In order to have known light positions, photometric lighting usually comes from a rig made of fixed lights – also known as a relighting dome. The light directions of the rig can be calibrated using reflections on a glossy sphere, light fall off in images, or other calibration methods. The orthographic assumption is easily overcome by camera calibration. In reality, however, surfaces – especially painted surfaces – are rarely Lambertian, and often contain inter-reflections.

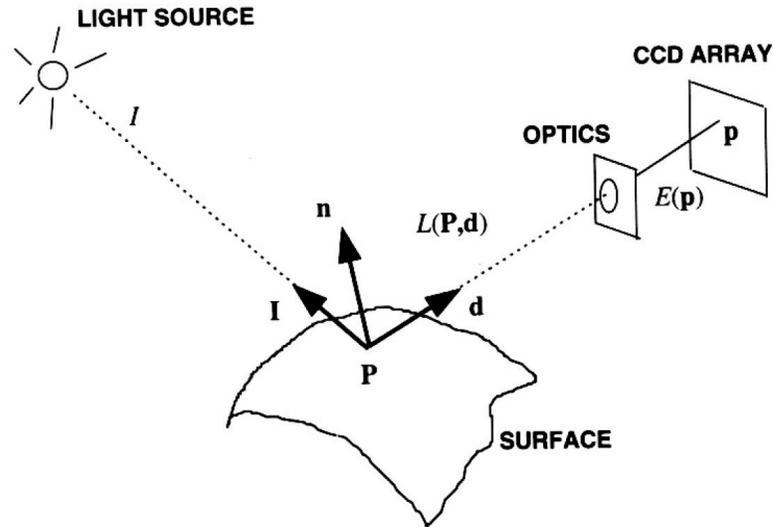


Figure 2.14: Light transport from the point light source, reflected by the object surface, and sensed by the camera. (Trucco and Verri, 1998)

2.3.3 Other Photometric Stereo Approaches

Many variations on the base photometric stereo model have introduced since the base photometric stereo model was introduced by Woodham, 1980. The contributions tend to vary from the base model in three main ways: introducing reflectance models that account for specularity and/or inter-reflections, a near-light model that allows lights to be close and act as point lights rather than uniform lights, and uncalibrated photometric stereo that allows for quicker capture using unfixed lights.

Several authors have introduced photometric stereo models that utilize a near-light rather than a far-light model (Huang et al., 2016; Logothetis et al., 2016; Quéau et al., 2018; Xie, Dai and Wang, 2015). If the energy of the light is considered in a far-light model – to account for light fall off – the intensity at each pixel can be described by Equation 2.13. The near light model in Huang et al.,

2016 uses a more complicated approach that accounts for the larger fall off that occurs from closely positioned lights (Equation 2.14). The near-light model can overcome some of the warping of the shape that is caused by the far-light assumption (Huang et al., 2016).

$$I = e\rho Ln \quad (2.13)$$

$$I = n \frac{(l - \rho)}{\|l - \rho\|^3} e\rho' \quad (2.14)$$

where

- I pixel intensities
- ρ surface albedo
- ρ' surface albedo and vignetting
- L lighting direction vectors
- l light positions
- n surface normal

Tan, Lin and Long Quan, 2008, Goldman et al., 2010, Hasegawa et al., 2011, Zhou, Wu and Tan, 2013, Quéau et al., 2017 use approaches that expand on the Lambertian reflectance model to introduce specular components and more accurately represent real surfaces. Approaches like that of Hasegawa et al., 2011 extract the albedo and the specular components, while other approaches – like that of Quéau et al., 2017 – model specularities or inter-reflections as outliers. The latter approach can lead to a robust extraction of surface and normal, but it cannot render the surface as accurately as the non-Lambertian approaches because it does not extract a specular component.

Photometric stereo requires known light positions. Though those positions are usually pre-calibrated, an uncalibrated photometric stereo approach allows for data capture using a single light instead of N lights. In an uncalibrated approach, the light positions are estimated from the images themselves. Shi et al., 2015, Huang et al., 2016, and Yeh et al., 2017 all introduce methods for uncalibrated photometric stereo. The implementation in Huang et al., 2016 combines an uncalibrated approach with a near-light model. The approach uses the light falloff in the images and the assumption that the lights are all approximately the same distance from the surface in order to calculate the light directions and positions (Huang et al., 2016).

2.3.4 *Surface to Shape Integration*

Though normals are enough for some photometric stereo applications, others require a 3D surface. Frankot and Chellappa, 1988 introduced a method to force the integrability of surfaces obtained from SfS algorithms. The x and y surface gradients are integrable if they meet a condition (Equation 2.15). The condition forces the height at a point to be independent of the direction of integration along the surface. Frankot and Chellappa, 1988 approached this problem by projecting the calculated surface gradients (z_x, z_y) – that might not be integrable – onto integrable surface gradients $(z - x, z - y)$ that meet the condition described by Equation 2.15 while minimizing the distance between the two (Equation 2.16). The implementation utilized a Discrete Fourier Transform (DFT) via a Fast Four-

ier Transform (FFT) algorithm (Frankot and Chellappa, 1988). Other approaches to normal integration can be seen in Quéau, Durou and Aujol, 2018.

$$z_{xy}(x, y) = z_{yx}(x, y) \quad (2.15)$$

$$d\{(\hat{z}_x, \hat{z}_y), (\tilde{z}_x, \tilde{z}_y)\} = \iint |\tilde{z}_x - \hat{z}_x|^2 + |\tilde{z}_y - \hat{z}_y|^2 dx dy \quad (2.16)$$

2.4 REFLECTANCE TRANSFORMATION IMAGING

Light shapes understanding of the things we see. The details of an object look different when light is at a grazing angle as opposed to perpendicular. RTI, known initially as PTM, is a visualization technique used for relighting surfaces in real-time. The technique is used by conservators, historians, and other heritage professionals to understand painted surfaces through an analysis of smaller details (Figure 2.15). In RTI, much like photometric stereo, a series of images are taken from



Figure 2.15: An example of RTI of a painting. The painting lit virtually from the front (left), and from the right (right). (Padfield, Saunders and Malzbender, 2005)

a fixed point with varying lighting directions that form a hemispherical shape around the object (Figure 2.16). The difference between the two techniques lies in their purpose: the goal of photometric stereo is accurate normals and metric

surface reconstruction, while the goal of RTI is to extract normals to facilitate re-lighting. PTM, the first RTI method, introduced by Malzbender et al., 2000, uses a simplified BRDF that reduces the variable dependence only to the incoming light position and angle (irradiance E_i and incident angles azimuth ϕ_i , elevation θ_i). With this assumption, Malzbender et al., 2000 uses the image set to calculate a per-pixel reflectance function that is fitted using a biquadratic polynomial (Equation 2.17). In addition, the surface normals are estimated using Equation 2.18 and Equation 2.19. The fitted reflectance function allows for the changing of properties – specularity, for example – in order to enhance details and features to the viewer. It also allows for the interpolation of new lighting positions enabling the viewer to move a virtual light to any angle (Malzbender et al., 2000).

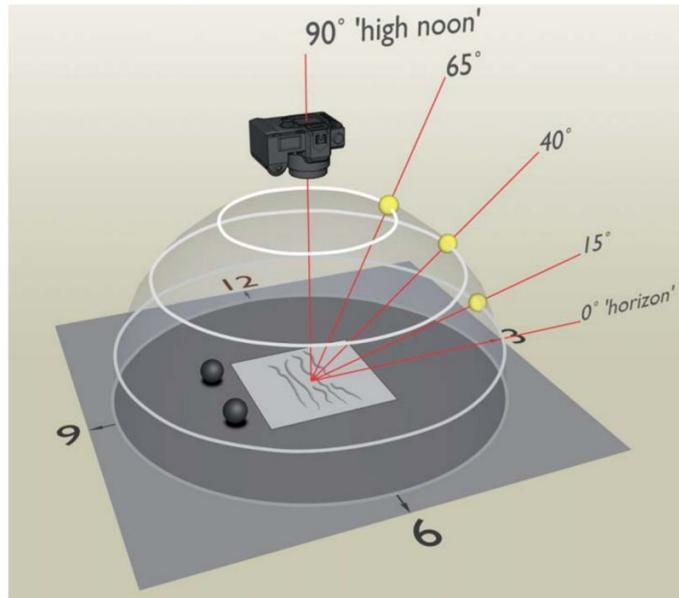


Figure 2.16: An example of lighting capture for RTI. (Duffy et al., 2018)

$$L(x, y; l_x, l_y) = a_0(x, y) l_x^2 + a_1(x, y) l_y^2 + a_2(x, y) l_x l_y + a_3(x, y) l_x + a_4(x, y) l_y + a_5(x, y) \quad (2.17)$$

where

L surface luminance

x, y the image pixel

l_x, l_y light direction

a_0, a_5 polynomial coefficients

$$l_{x0} = \frac{a_2 a_4 - 2a_1 a_3}{4a_0 a_1 - a_2^2} \quad (2.18)$$

$$l_{y0} = \frac{a_2 a_3 - 2a_0 a_4}{4a_0 a_1 - a_2^2}$$

$$n = \left(l_{x0}, l_{y0}, \sqrt{1 - l_{x0}^2 - l_{y0}^2} \right) \quad (2.19)$$

[HSH](#), an alternative method to [PTM](#), maps the [RTI](#) lights to a hemisphere rather than using the biquadratic polynomials. The result is faster to obtain and more accurate ([Figure 2.17](#)) ([Macdonald, 2015](#)).

[Mudge et al., 2006](#) introduced a new method of [RTI](#) capture called [Highlight-RTI](#). The method uses a black sphere to estimate light positions and allows for manual [RTI](#) capture with a single moveable light rather than a fixed, calibrated dome – much like uncalibrated photometric stereo.

The simplified [BRDF](#) and normal calculation of [RTI](#) results in data that is qualitatively

ively useful, but not as robust quantitatively as those obtained from photometric stereo calculations (MacDonald, 2011).



Figure 2.17: An overlaid comparison (using horizontal stripes) of PTM vs. HSH, showing that the hemispherical harmonic representation better preserves contrast from the original images. (Mudge et al., 2008)

LITERATURE REVIEW: STATE OF THE ART

The following sections describe several state-of-the-art methods related to the 2.5D documentation of painted surfaces: photogrammetry, scanning-based methods, photometric stereo, and data fusion methods. The approaches discussed focus on whole surface documentation and do not include techniques used to document local areas like conoscopic holography. The methods discussed can achieve a spatial resolution and precision within the range required for conservation work of paintings – 0.01 *mm* to 1 *mm* (Remondino et al., 2011).

3.1 PHOTOGRAMMETRY

Photogrammetry has been used as a low-cost way to document painted surfaces for over a decade. It has been used to monitor changes, as a way to merge multispectral data sets and to produce orthophotos for 2D measurement and analysis. Remondino et al., 2011 states that the high-frequency details from photogrammetry of painted surfaces are often too noisy due to reflections, point matching errors, or dark areas. It is for this reason that photogrammetry is typically used as a rough 3D measurement or an inexpensive, and robust scaling tool for quick 2D measurements (orthophotos). Some examples are described in the following paragraphs.

Robson et al., 2004 used photogrammetry to monitor the changes to *the Westminster Retable*, a 13th-century wood-based art piece, during its restoration. The goal of the project was to understand the physical changes of the Retable under several factors to help guide the restoration and final display of the piece. For example: how does the removal of dirt layers change the Retable; how do changes in humidity and temperature affect the supports; and, will placing it on its side damage the piece during restoration? The precision of the features of the final model was to be between 0.2 mm and 1 mm, and the precision of the targets was to be less than 100 μm (Robson et al., 2004). The documentation allowed Robson et al., 2004 to understand the global warping and shifting of the Retable.

Remondino et al., 2011 discuss how photogrammetry was used to monitor the deformation of a canvas – due to its weak structure – during its restoration. The artwork, *Pala Trivulzio* by Andrea Mantegna, was documented during four different stages with targets placed on the holding frame to allow comparison between the epochs (Remondino et al., 2011). The bundle adjustment resulted in a coordinate accuracy of 0.7 mm laterally and 1.2 mm axially (Remondino et al., 2011).

Barbetti et al., 2013 used photogrammetry to monitor the 3D changes in a 13th-century wall painting before and after cleaning treatments were applied and to ensure that the cleaning did not change the surface roughness. The authors report a precision of 80 μm laterally, and 60 μm axially (Barbetti et al., 2013). In this case, the target is a matte surface with a lot of texture, not representing the problems that many painted surfaces pose for high-resolution surface reconstruction. Abate et al., 2014 used a similar approach to capture the surface and colour of the 16th-century painting *Saint John the Baptist* by Giacomo Francia (Figure 3.1). The painting was in the centre of the room during documentation to allow for a

robust connection of the front and back parts of the painting; there is often a weak connection between these two areas because of thin canvases or supports (Abate et al., 2014).

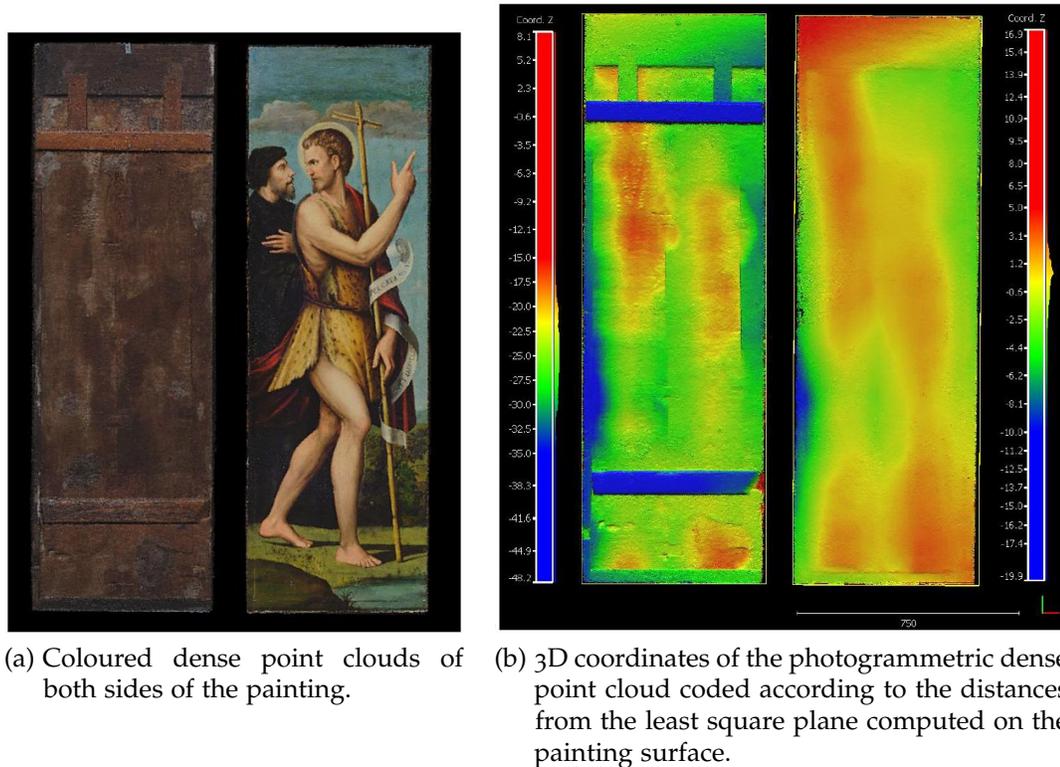


Figure 3.1: Analysis of the Saint John the Baptist painting. (Abate et al., 2014)

Bennett, 2015 used photogrammetry and Near-infrared (NIR) images to analyze a 19th-century British landscape painting. The photogrammetry was used to make an orthophoto for 2D planar measurements and to align the different spectrums of data, rather than as a 3D documentation (Bennett, 2015). The NIR orthophoto led to the discovery of a stamp on the back of the painting that helped to determine its origins (Bennett, 2015). Pamart et al., 2017 applied a similar approach with multispectral photogrammetry in the documentation of wall paintings. They used the photogrammetry more to align the data sets and make orthophotos for 2D analysis rather than for 3D measurements. Alternatively, Nocerino et al., 2018

used Visible Spectrum (VIS) and Ultraviolet (UV) photogrammetry to document a painted vase. The output was a 3D model within a web viewer with the VIS and UV results as changeable textures (Nocerino et al., 2018).

3.2 SCANNING TECHNIQUES

Structured light and triangulation-based scanning are the most common methods used to document painting surfaces. Both methods can achieve high precision, and though commercial scanners are still costly, they are becoming less expensive. ToF methods, however, produce too much noise to be used for anything other than deformation analysis. The following paragraphs describe both commercial, and custom scanning systems used to document paintings.

Lahanier et al., 2005 from the National Research Council (NRC) Canada, introduced a method to capture multispectral 2D information as well as 3D information of easel paintings. The 13 bands of multispectral data were captured using a specialized Charge-Coupled Device (CCD) camera called the CRISATEL Camera (Lahanier et al., 2005). The 3D data was captured using two scanners developed by the NRC; a large scanner used for the quick capture of a painting surface for deformation analysis, and smaller laser scanner which captures the colour and surface of the painting in higher detail (Lahanier et al., 2005). According to Lahanier et al., 2005, the large scanner has a lateral spatial resolution (x and y) of 200 μm and a depth uncertainty (z) of 100 μm . The smaller, colour laser scanner can capture areas of 20 cm x 4 cm per scan. It has a lateral spatial resolution of 50 μm and a depth precision of 10 μm . Figure 3.2 shows results from the NRC's smaller scanner in the documentation of *Dame en prière*, a 16th-century oil painting at the Louvre Museum. As a comparison, Lahanier et al., 2005 also scanned *Dame en*

prière using a Breuckmann GmbH triTOS-HE-100 structured light scanner that uses DFP. The scan was completed in 30 overlapping sections that covered a lateral size of 8 cm x 6 cm. Each scan had a lateral spatial resolution of 60 μm and a depth precision of 2 μm Lahanier et al., 2005. The scans were stitched together using the overlapping areas between them.

Blais et al., 2008 used the smaller laser scanner to document Leonardo Da Vinci's *Mona Lisa*. The painting, from the early 16th-century, was done in oil on a wooden panel. The scanning endeavour required 72 overlapping scans for the front of the painting, which were subsequently stitched together (Blais et al., 2008). Blais et al., 2008 discusses systematic errors that could occur due to varnish, step discontinuities (edges), and changes in colour. Varnish is a transparent or translucent layer placed on top of the paint to protect it. When light hits the varnish, it may not fully penetrate to the paint layer, causing a measurement of the varnish layer rather than the paint (Figure 3.3). Blais et al., 2008 measured the index of refraction of the varnish based on reflections on the painting observed under a microscope. These observations allowed them to calculate the approximate height of the varnish.

Structured light scanners are a popular choice for the documentation of 2.5D painted surfaces. Guidi, Atzeni and Lazzari, 2003 used a commercial DFP scanner to document Leonardo Da Vinci's *Adoration of the Magi*. The 15th-century, oil on wood painting was documented in over 800 scans with a maximum overlap of 35% (Guidi, Atzeni and Lazzari, 2003). To align successive scans, Guidi, Atzeni and Lazzari, 2003 used an Iterative Closest Point (ICP) algorithm. The algorithm takes an initial manual alignment, performed by the user, and iteratively minimizes the distance between the matching points on the two surfaces. The alignment

can be error prone when scanning surfaces that have very few textural details, as was the case with front portions of the painting (Guidi, Atzeni and Lazzari, 2003). The manual nature of the alignment procedure also increased processing times. A similar study by Pelagotti, Del Mastio and Razionale, 2007 used multispectral photography and a commercial structured light scanner to capture the surface of *Saint Catherine's Polyptyc*. The scanner captured individual depth maps that were then converted to a high-resolution mesh. The data was captured in 67 scans at a lateral spatial resolution of 400 μm and a spatial depth resolution of 100-200 μm (Pelagotti, Del Mastio and Razionale, 2007). The large, high-resolution mesh was converted to a depth map to make the data easier to work with. Other structured light examples for painting documentation can be found in Remondino et al., 2011 and Abate et al., 2014. Remondino et al., 2011 also shows an example in which a ToF scanner was used for the deformation analysis of painting surfaces at a resolution of 3 mm.

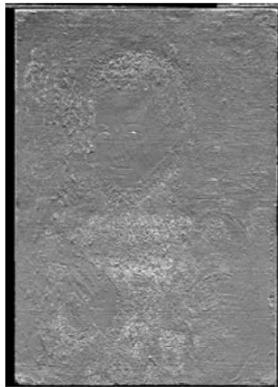
In 2011, Factum Foundation, 2015 introduced the Lucida, a laser-based triangulation system for scanning 2.5D surfaces. The scanner uses two cameras pointed inwards at a laser. The cameras and laser track in the x and y-direction along the painting surface as the Lucida scans. The calibrated cameras measure the deviations of the laser from its original shape, a line. The software calculates the deviations and reconstructs the surface as a depth map with a resolution of 100 μm (Factum Foundation, 2015). The scanner can capture an area of close to 48 cm^2 ; the various sections are stitched together using methods similar to those previously listed (Factum Foundation, 2015). The Lucida has been used to scan over a hundred paintings from *The Triumph of the Eucharist over Idolatry* by Rubens, to *The Holy family with Lamb* by Raphael (Figure 3.4). Though the Lucida works well for its purpose, the scanner is slow; it can take up to an hour for one section

(Factum Foundation, 2015).

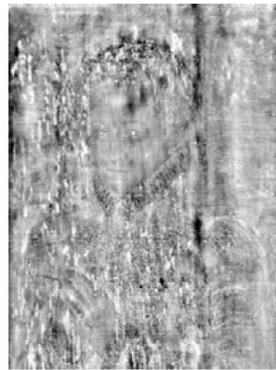
Zaman, 2013 introduced a low-cost, DFP scanner. The scanner consists of two DSLR cameras pointed inwards at a 45-degree angle with a projector separating them. The scanning system has a lateral spatial resolution of $50 \mu m$ and a depth precision of $9 \mu m$ (Zaman, 2013). Cross polarization was used to overcome the varnish problem mentioned earlier; polarizing filters were placed in the same direction on both cameras and in the opposite direction on the projector. The system can capture $136 cm^2$ in 150 seconds (Zaman, 2013). The scanning system, not considering the cameras and projector, cost aaround \$1200 CAD. The advantage of using a DSLR based DFP method is that it results in precise depth maps as well as precise colour; both of which are needed by conservators (Zaman, 2013). Zaman, 2013 used the scanner to scan a self-portrait by Rembrandt and *Flowers in a Blue Vase* by Van Gogh (Figure 3.5).



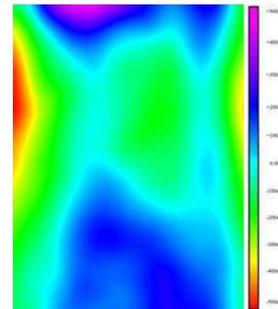
(a) The painting in colour.



(b) Shaded monochrome image illustrating surface relief.



(c) Grey coded depth image of the panel.



(d) Colour coded image illustrating the overall shape of the slightly warped panel.



(e) Colour detail of the face showing craquelure.



(f) Grey coded depth image of the face.



(g) Shaded monochrome image of the face.

Figure 3.2: A 3D image of Dame en prière scanned at a depth resolution of $10 \mu\text{m}$. (Lahaniier et al., 2005)

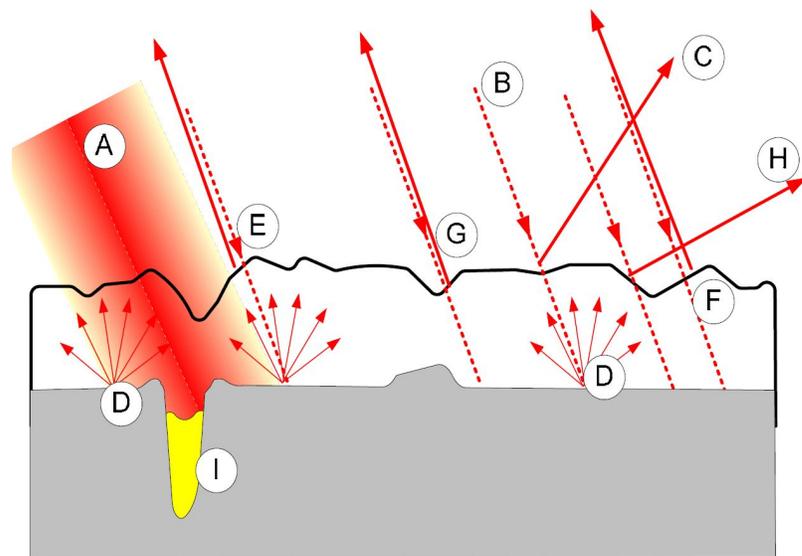
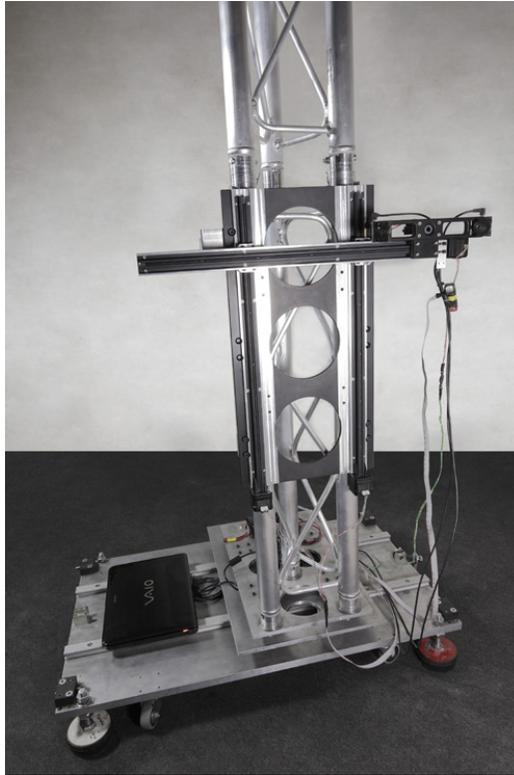


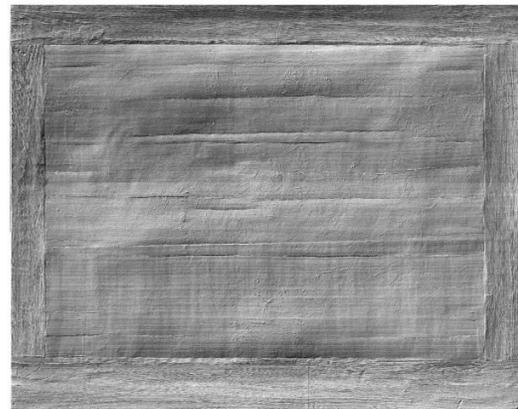
Figure 3.3: Simplified model of the paint layer including a translucent layer of finely cracked old varnish and/or transparent glaze. (Blais et al., 2008)



(a) The Lucida scanner.



(b) Colour photography.



(c) 3D render of the depth map.

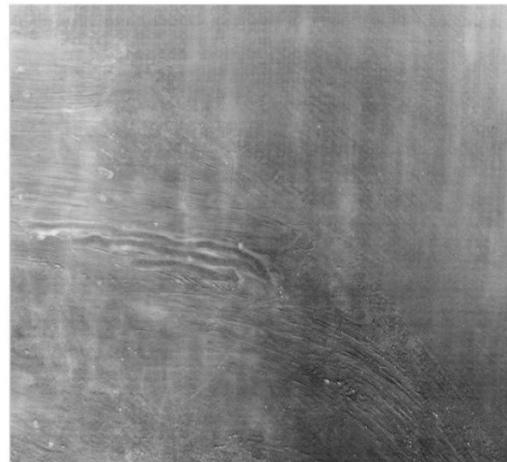
Figure 3.4: The documentation of *The Triumph of the Eucharist over Idolatry* by Paul Rubens. (Factum Foundation, 2015)



(a) The DSLR based [DFP](#) scanner.



(b) A detail of the colour.



(c) 3D render of the depth map.

Figure 3.5: The documentation of a self-portrait by Rembrandt using a [DFP](#) scanner. (Zaman, 2013)

3.3 PHOTOMETRIC STEREO

There are some, but not many, examples of photometric stereo being used to document 2.5D painted surfaces. The technique comes from the computer science community, which focuses more on research and synthetic data for testing. RTI has made it over to heritage applications and has been used widely on painted surfaces, but it is only a relighting technique. Photometric stereo, on the other hand – if it can overcome challenges such as large surface areas, low-frequency bias, specularities, and edge discontinuities – is an effective, low-cost technique for documenting 2.5D painted surfaces.

Hasegawa et al., 2011 used photometric stereo to document the surface of paintings for digital visualization purposes. Hasegawa et al., 2011 modified the traditional photometric stereo assumption – that the surface was Lambertian – and added a specular reflection component, based on the Torrance-Sparrow model, to the BRDF to more accurately reconstruct oil paintings. They captured several paintings using a 16-light photometric stereo setup and compared renderings to photographs of the painting (Hasegawa et al., 2011). The result was a separated normal, diffuse albedo, and specular component for the painted surface (Figure 3.6).

Huang et al., 2016 showed that the traditional RTI or uncalibrated photometric stereo approach uses a far-light model, which is difficult to reproduce; the light would have to be at the same angle and intensity for the whole Field of View (FOV). The far-light assumption often results in a potato chip shape during normal integration and a spot-light effect during RTI relighting. Huang et al., 2016 created a near-light photometric stereo that overcomes the potato chip effect

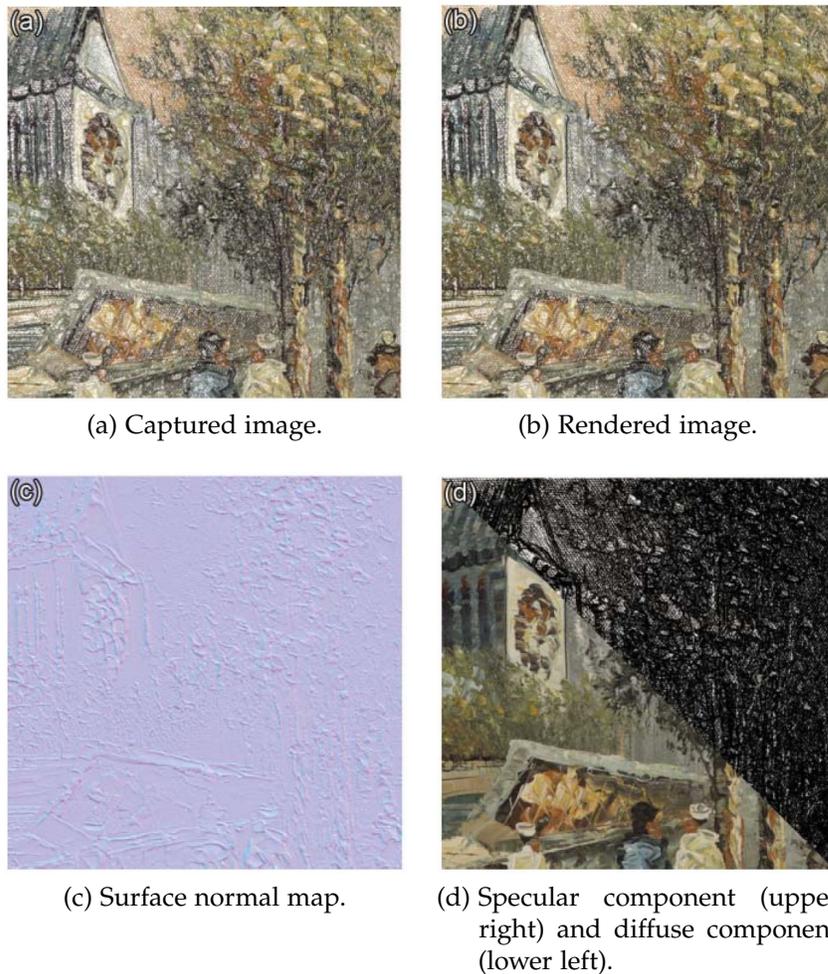


Figure 3.6: Image rendering results for a glossy oil painting (varnished). (Hasegawa et al., 2011)

and includes correction of the spot-light effect to create more uniformly lit images for relighting (Figure 3.7). Cossairt et al., 2015 used the same approach as Huang et al., 2016 to document transfer drawings by Paul Gauguin. The ability to separate the colour and the surface information helped the authors understand the artist's methods for making transfer drawings (Cossairt et al., 2015). Salvant et al., 2019 used the same photometric stereo approach but applied it to UV images of paintings by Georgia O'Keeffe. The UV method makes it easier to identify and monitor the protrusions in the paintings (Salvant et al., 2019).

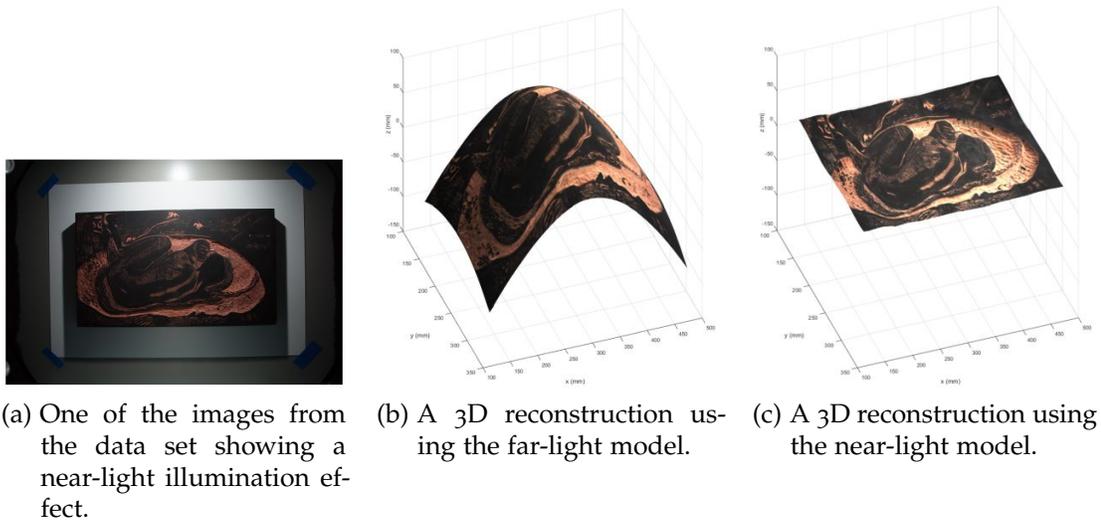


Figure 3.7: The comparison of the far-light and near-light photometric stereo models on a woodblock created by Paul Gauguin. (Huang et al., 2016)

Watteuw et al., 2016 introduced the design of a portable, multispectral light dome. The dome contains five different types of LED lights: UV, red, green, blue, and Infrared (IR). The goal of the research is to integrate multispectral data into a virtual relighting interface so that conservators can view RTI type relighting data in multiple spectrums along with the 3D surface shape (Watteuw et al., 2016).

3.4 DATA FUSION

Data fusion combines two or more different methods to overcome the shortcomings of each method. A popular approach is to combine range data or photogrammetry with photometric stereo. Photogrammetry and low-cost scanners capture low-frequency data (shape) well but struggle with high-frequency details. Photometric stereo approaches are the opposite: good high-frequency details, but bias in the overall shape.

Nehab et al., 2005 introduced a method to combine low-frequency data from a range scanner with the high-frequency data from photometric stereo. Nehab et al., 2005 built their own range scanner, which was integrated with their photometric stereo rig. The data is perfectly aligned. The approach uses the range data to first improve the low-frequency bias in the measured photometric stereo normals. Nehab et al., 2005 then used a linear least squares approach to improve the high-frequency details by minimizing the normal and position errors (Figure 3.8). The approach works on meshes and is meant for a 2.5D implementation but can be applied to an arbitrary 3D mesh as well. The 2.5D approach had errors less than 0.5 mm, while the 3D approach had more significant errors (Nehab et al., 2005).

MacDonald, 2015 used a height measurement gauge to make 16 point measurements on a terracotta bas relief of Frederic Chopin. The point measurements were used to create the low-frequency shape of the bas relief (MacDonald, 2015). The photometric stereo normals of the relief – captured using a custom designed dome – were integrated into a surface using the Frankot and Chellappa, 1988 formula. Both shapes were converted into the frequency space using a DFT through a FFT algorithm (MacDonald, 2015). The transformation divides a signal, or an image, into its sine and cosine frequencies. MacDonald, 2015 utilized the frequency space to extract the low and high frequencies and combine them. The resulting model was better than either of the individual models, with a maximum error of 2 mm (MacDonald, 2015). MacDonald et al., 2016 improved on the process by using a DEM from photogrammetry as the low frequency when documenting a calibrated test target. The disadvantage of this approach is that it is only applicable to 2.5D surfaces.

Antensteiner, Štolc and Pock, 2018 compares several different approaches for depth and normal fusion: a traditional photometric stereo to depth integration method (also known as a gradient-based method), the method of Nehab et al., 2005, the method of Heber, 2015, an introduced method Generalized Nehab, and another new method based on TGV. According to Antensteiner, Štolc and Pock, 2018, the gradient-based method has issues with steep angles. The Nehab et al., 2005 approach improves on that by adding a weighting factor for the Z depth. Their Generalized Nehab approach adds an exponent on the Z depth weighting factor that can improve the results with steep slopes over the Nehab method (Antensteiner, Štolc and Pock, 2018). The TGV method is an improvement on total variation – a regularization method which is used to prevent overfitting and thus overly smoothed edges (Antensteiner, Štolc and Pock, 2018). The results from the compared methods can be seen in Figure 3.9.

Berkiten, Fan and Rusinkiewicz, 2014 improved on the Nehab et al., 2005 approach for data fusion on 3D meshes. The approach aligns multiple normal maps, captured using photometric stereo, to a 3D mesh using a modified SIFT algorithm (Berkiten, Fan and Rusinkiewicz, 2014). The method is more versatile because the data can be captured using two different methods, or at different times. Berkiten, Fan and Rusinkiewicz, 2014 improves the combination of the various normal maps by implementing a minimization that takes the normal with their corrected low frequency (as in Nehab et al., 2005) and optimizes them using two variables: a weighted average at each vertex and the differences between neighbouring normals. For the improvement of the positions on a 3D mesh, Nehab et al., 2005 assumed that the surface of the mesh is locally planar in order to keep the problem linear, and thus the calculations faster – this is not true for many heritage objects. Berkiten, Fan and Rusinkiewicz, 2014 optimizes the po-

sitions for the normal and position error using a non-linear approach, resulting in a more detailed mesh (Figure 3.10). Other versions of a combined SfM/MVS and photometric stereo approach can be found in Sabzevari, Del Bue and Murino, 2011, Grochulla and Thormählen, 2015, and Park et al., 2017.

Tominaga, Ujike and Horiuchi, 2010 introduced a method to reconstruct the surfaces of paintings that combines photometric stereo with point measurements. The authors used nine, point measurements from a laser displacement meter to reconstruct an overall shape of a painting (low frequency) (Tominaga, Ujike and Horiuchi, 2010). The painting was then captured using photometric stereo at a lateral spatial resolution of around $40 \mu m$. The integrated photometric stereo is combined with the low-frequency information using a FFT algorithm. The result is an error of just over $100 \mu m$ (Tominaga, Ujike and Horiuchi, 2010).

Yeh et al., 2016 introduced a method to capture perfectly aligned photogrammetric and photometric stereo data simultaneously. The acquisition involves two cameras: one stationary photometric camera, and a second photogrammetric camera with a ring flash on it. The two camera shutters are synced, and as the user moves the photogrammetric camera around different light positions are obtained for the photometric stereo (Yeh et al., 2016). The data from the two cameras are aligned in a photogrammetric bundle adjustment, and the camera positions from the photogrammetric camera are used as the light positions for the photometric stereo calculation (Yeh et al., 2016). They combine the depth obtained from photogrammetry (as a DEM) with the normals from photometric stereo using a weighted least-squares approach (Figure 3.11).

Aure, O'Dowd and Padfield, 2017 used a Lucida scanner and a relighting dome to document paintings in the National Gallery in London. The goal of the developed workflow was to align colour data with depth data for visualization purposes (Aure, O'Dowd and Padfield, 2017). The Lucida was used to capture the surface of the painting, and an RTI algorithm was used to extract the albedo and specular maps of the surface. The data was then aligned and layered for renderings (Aure, O'Dowd and Padfield, 2017).

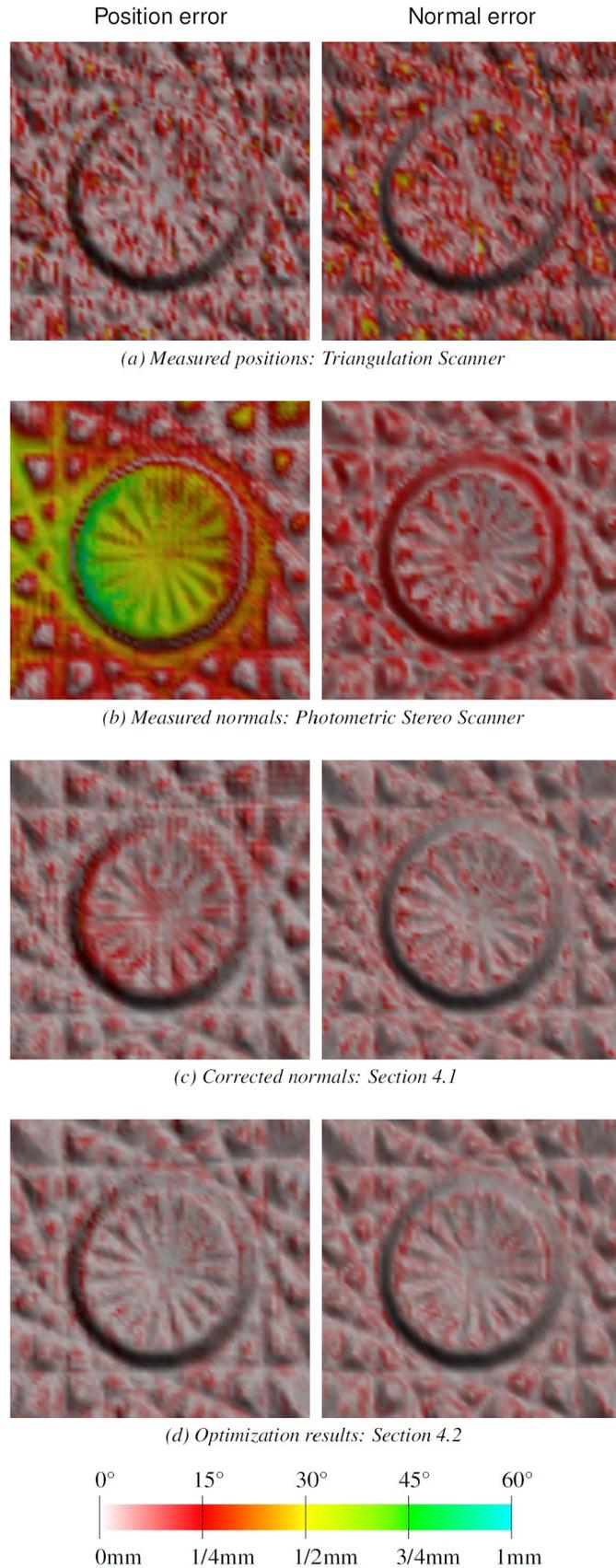


Figure 3.8: Quality assessment from photometric stereo, a triangulation scanner, and data fusion. (c-d) The combination of the data significantly reduces the error from both methods. (Nehab et al., 2005)

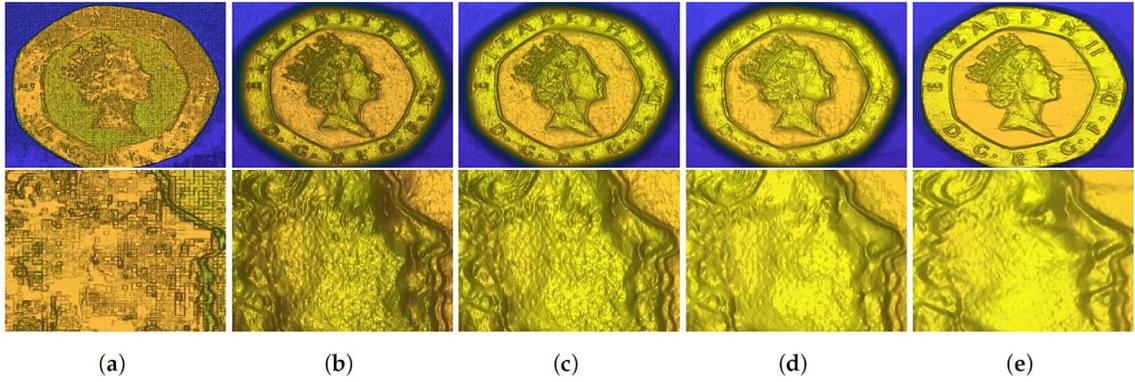


Figure 3.9: The results of the fusion algorithms on a coin: (a) Initial scanner depth; (b) gradient-based; (c) Method of Nehab; (d) Generalized Nehab; (e) TGV. (Antensteiner, Štolc and Pock, 2018)

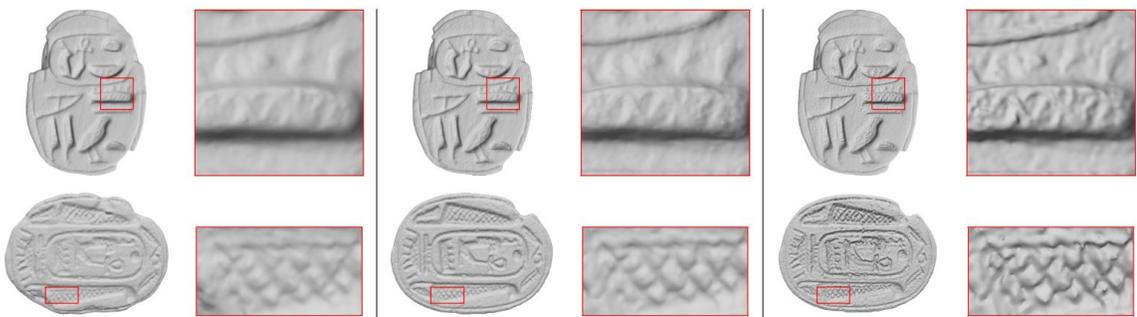
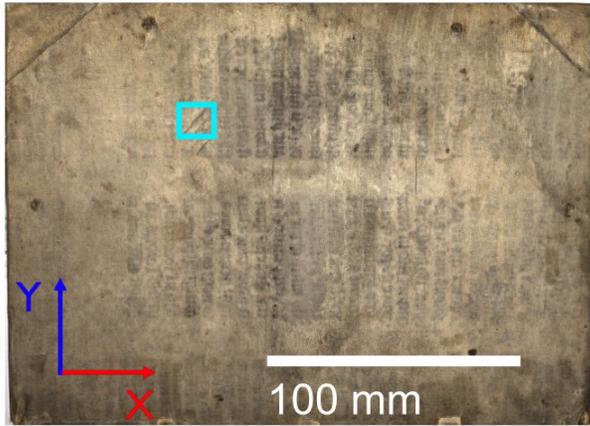
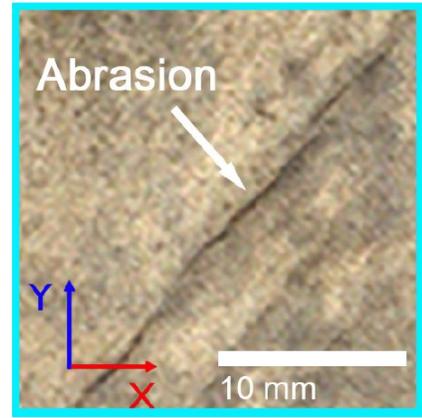


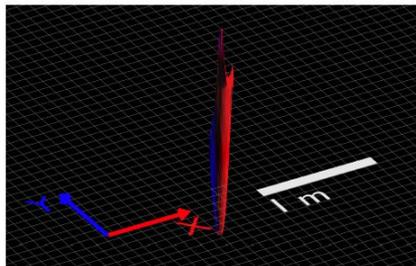
Figure 3.10: A comparison of Poisson surface reconstruction (left), Nehab et al., 2005 (middle), and Berkiten, Fan and Rusinkiewicz, 2014 (right) on Egyptian scarabs. (Berkiten, Fan and Rusinkiewicz, 2014)



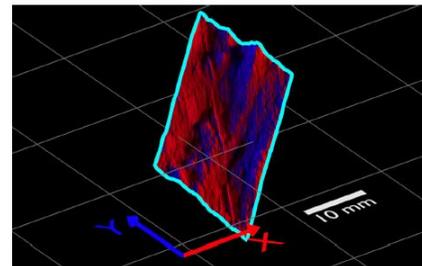
(a) Reference Image



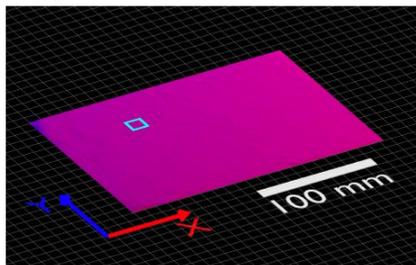
(b) Reference Image Inset



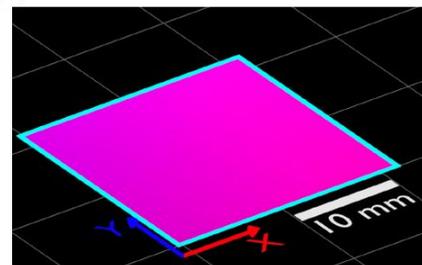
(c) PS Near-Light Surface



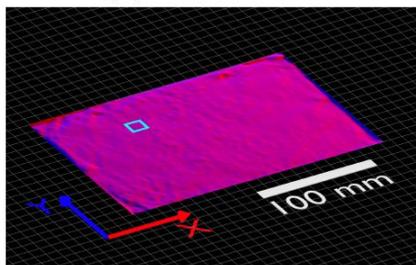
(d) PS Near-Light Surface Inset



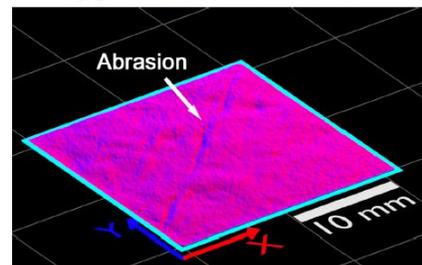
(e) PG Surface



(f) PG Surface Inset



(g) Fusion Surface



(h) Fusion Surface Inset

Figure 3.11: A Comparison of reconstruction methods. (Yeh et al., 2016)

Part III

RESEARCH

"The substance of painting is light."

– André Derain

APPROACH

When choosing a scanning system for 2.5D surfaces it is difficult to balance cost with other metrics. [Figure 4.1](#) compares different scanning methods according to their cost, and ability to resolve low and high frequency details. The structured light scanner built by Zaman, [2013](#) is the only one that is cost effective, and can resolve high and low frequency information well. The downsides of this approach, however, are that it requires two of the same [DSLR](#) and lens setup which many people do not have, and that it is limited to documenting only a few centimetres in depth. The goal of this thesis is to explore a low-cost and portable method for documenting the geometry of 2.5D painted surfaces that matches the resolution of other approaches – less than $100 \mu m$. It should be noted that low-cost refers to the hardware required, not the skills involved to operate. A data fusion approach would acquire a more skilled user when compared to scanning methods.

Documentation Method	Cost	Ability to Resolve Low Frequency Information	Ability to Resolve High Frequency Information
Photogrammetry	● ○ ○ ○ ○	● ● ● ● ● ● ● ●	● ● ● ● ● ● ● ●
Terrestrial laser scanning	● ● ● ● ● ● ● ●	● ● ● ● ● ● ● ●	● ● ● ● ● ● ● ●
Triangulation-based scanning - Commercial	● ● ● ● ● ○ ○	● ● ● ● ● ● ● ●	● ● ● ● ● ● ● ●
Structured light scanning - Commercial	● ● ● ● ● ○ ○	● ● ● ● ● ● ● ●	● ● ● ● ● ● ● ●
Uncalibrated Photometric Stereo	● ○ ○ ○ ○ ○ ○	● ● ● ● ● ● ● ●	● ● ● ● ● ● ● ●
Calibrated Photometric Stereo - Relighting Dome	● ● ● ● ● ○ ○ ○	● ● ● ● ● ● ● ●	● ● ● ● ● ● ● ●
Triangulation-based scanner - <i>Lucida</i> - Factum Foundation, 2015	● ● ● ● ● ● ● ●	● ● ● ● ● ● ● ●	● ● ● ● ● ● ● ●
Structured light scanning - Zaman, 2013	● ● ● ● ● ○ ○ ○	● ● ● ● ● ● ● ●	● ● ● ● ● ● ● ●

Figure 4.1: Comparing different documentation methods.

Based on the methods discussed in (chapter 3), a data fusion approach seems to be the most robust, cost-effective, and widely applicable approach. A data fusion approach could be adapted in the future to include a focus on colour and 3D applications. It is clear from the methods presented in previous chapters that a combination of low and high-frequency data can lead to improved results (Figure 4.2) (Berkiten, Fan and Rusinkiewicz, 2014; MacDonald, 2015; Nehab et al., 2005; Tominaga, Ujike and Horiuchi, 2010; Yeh et al., 2016). Tominaga, Ujike and Horiuchi, 2010 showed how photometric stereo, combined with a low frequency from something as simple as point measurements is enough for the documentation of some painted surfaces. With photometric stereo as a solution for high-frequency data, photogrammetry is a logical choice as a low-frequency method – both methods require only a single camera, so there would be little extra equipment required. The photographic approach of both methods makes it possible to capture a 2.5D surface using multiple tiles of photometric stereo at the desired resolution. Those tiles can be aligned to the photogrammetric data during the bundle adjustment.

Two questions remain to be answered – what method will be adopted for photometric stereo and what is the most appropriate data fusion approach?

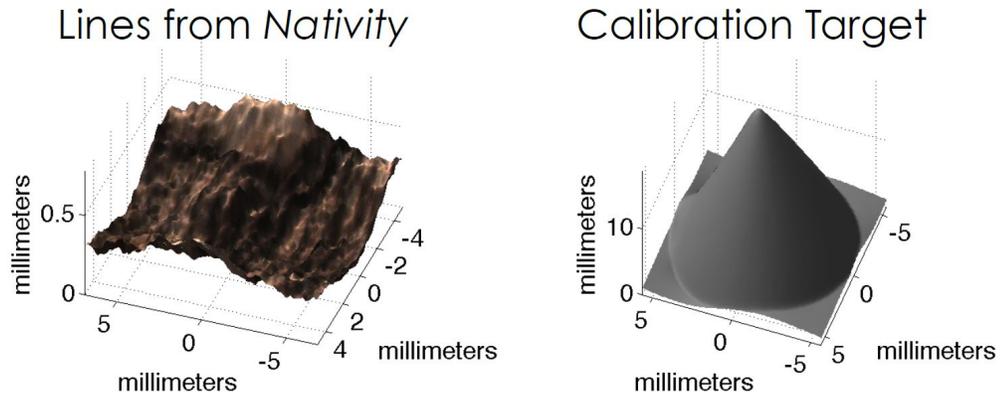


Figure 4.2: The scaled depth and calibration target from the photometric stereo of Paul Gauguin’s Nativity using the method of Huang et al., 2016. It can be seen in the calibration target that there is still a low-frequency warping in the data – the base should be flat. (Cossairt et al., 2015)

4.1 PHOTOMETRIC STEREO

The goal of the photometric stereo approach is not to create a new approach but to implement an existing one into the overall painting documentation workflow. The technique of photometric stereo stems more from the computer vision field than from a heritage point of view. Because of this distinction, many photometric stereo algorithms introduced in literature (conference proceedings, journals, etc.) have only been tested on synthetic data; the algorithms are not optimized for the larger size of real images or data sets. So, while some approaches may produce satisfactory results, they would not be viable in practice. For this reason, after testing the capacity of several available algorithms, initial tests were performed using two different algorithms: the traditional far-light photometric stereo approach of Wu et al., 2011, and the uncalibrated, near-light approach of Huang et al., 2016.

Photometric stereo can be captured in an uncalibrated, semi-calibrated, or calibrated manner. In uncalibrated photometric stereo, the light positions are un-

known and are estimated from the images. Though the lighting can come from a fixed lighting rig, uncalibrated photometric stereo is typically done manually with a single, portable light source and a string – used to keep the light at a constant distance. A fixed lighting rig allows for a semi-calibrated or calibrated lighting setup – known light directions or known directions and positions respectively. Calibrated approaches are more robust; surface normals can be calculated more accurately. For practical reasons, an uncalibrated, manual approach – if the results are acceptable – is the ideal choice; it is inexpensive and easily reproducible. The initial tests were performed on an unnamed, anonymous painting of the Virgin Mary from the 17th century ([Figure 4.3](#)).

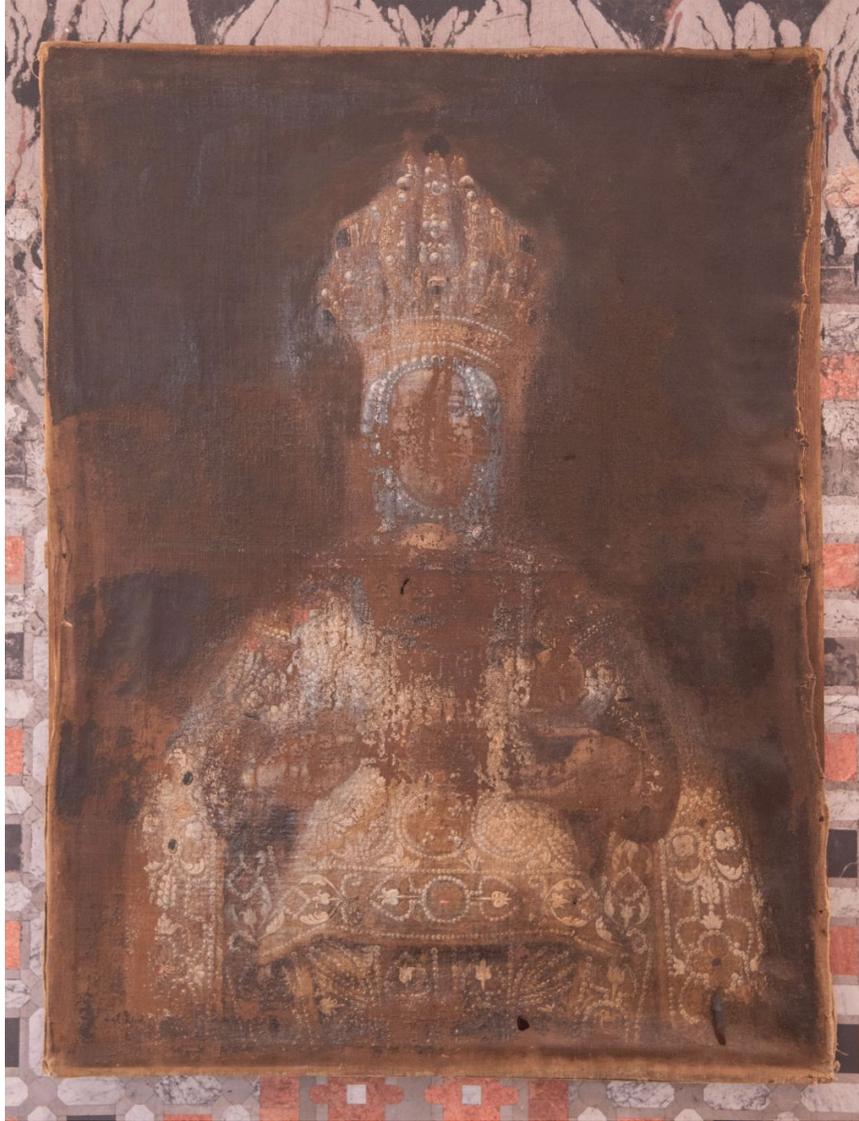


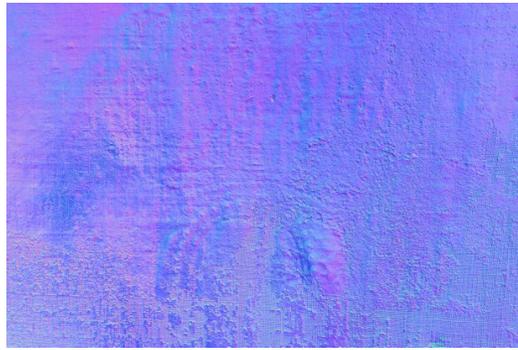
Figure 4.3: The anonymous 17th century painting of the Virgin Mary.

The photometric stereo of the painting was captured with a Canon 5DS R with a 50 *mm* lens at a distance of 65 *cm* from the surface of the painting – at a lateral spatial resolution of 50 μm . The light source was a handheld lamp that was used to create 24 lighting positions, every 45° radially at angles of approximately of 15°, 40°, and 65° from the surface (eight columns and three rows). First, the data was processed using the algorithm from Huang et al., 2016. The code estimated the light positions using the falloff of light in each image from the spotlight effect of near lights. The calculated light positions were then used as the light positions

for the Wu et al., 2011 calculation. The Huang et al., 2016 code took more than 20 minutes to process and required 64 GBs of RAM to process fully – due to the size of the 50-megapixel images. The Wu et al., 2011 code, however, took less than one minute to solve the photometric stereo problem (though it was given light positions calculated from the other method). Figure 4.4 shows the normal maps resulting from the test on the Virgin Mary painting.

The low-frequency bias of the far-light model becomes immediately clear when looking at the colour shifts in Figure 4.4. There are also differences between the two models in the high-frequency domain. There were inaccuracies in the estimation of the light positions caused by inconsistencies in the manual capture – lights that were taken from small, raking angles were estimated to be coming from larger angles (Figure 4.5). The inaccuracies cause the near-light model to overcompensate and create a bumpier high-frequency result; the affect on the far-light model was less pronounced.

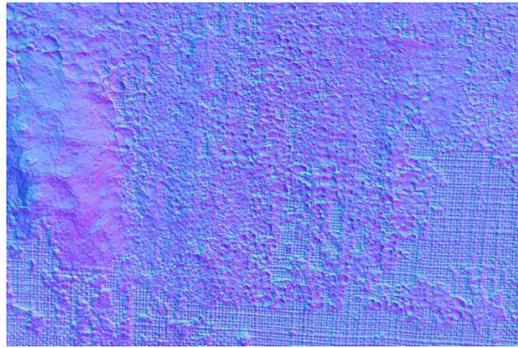
The near-light normals were integrated into a depth map using the Frankot and Chellappa, 1988 formula and scaled using the approximate distance from the camera to the surface (GSD). The Lucida was then also used to scan the same section of the Virgin Mary painting. The goal was to compare only the high-frequency details between the photometric stereo and the Lucida because the photometric stereo still had low-frequency bias (Figure 4.5). The high-frequency details of the two methods were printed by Océ using elevated printing. It is clear, from the printed area that the photometric stereo data has distorted high-frequency details (Figure 4.6). As can be seen, the photometric stereo details have sharper edges; the height of the details are incorrectly exaggerated. The errors are due to inaccurate light position estimations as mentioned above.



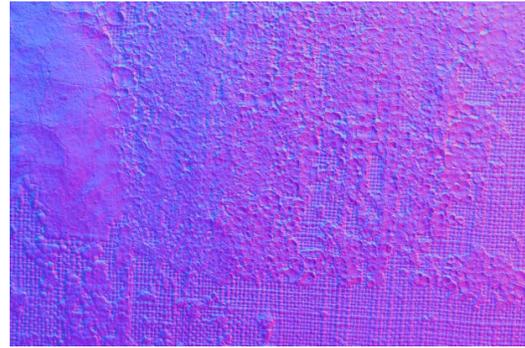
(a) The full near-light normal map.



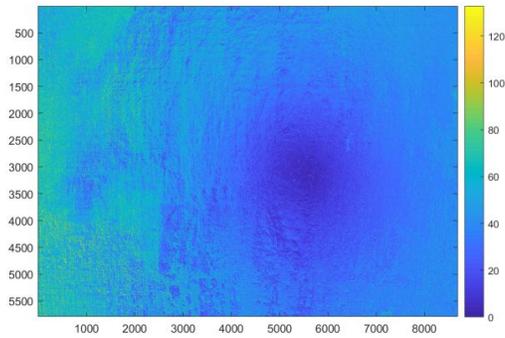
(b) The full far-light normal map.



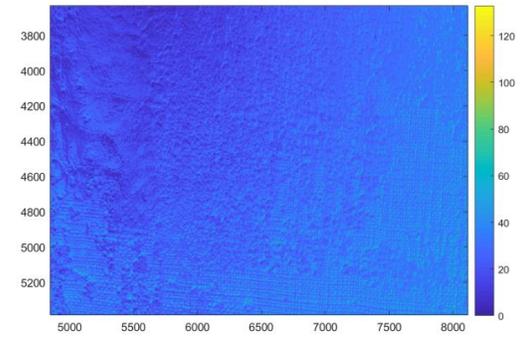
(c) An enlarged view of the near-light normal map.



(d) An enlarged view of the far-light normal map.

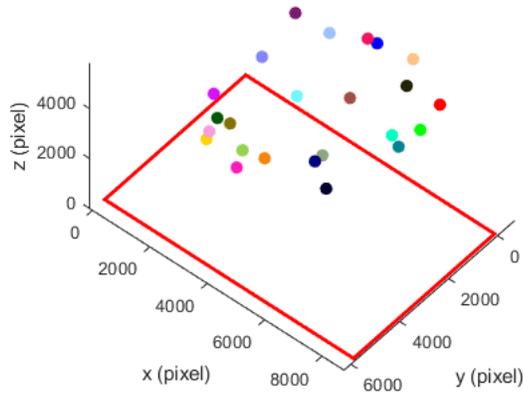


(e) The angular difference (degrees) between the near-light and far-light normals.

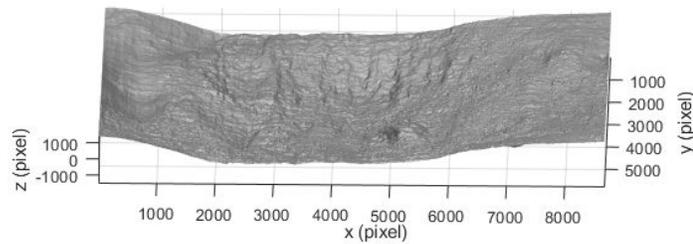


(f) An enlarged view of the angular difference (degrees) between the near-light and far-light normals.

Figure 4.4: A comparison of the Huang et al., 2016 near-light algorithm and the Wu et al., 2011 far-light algorithm.



(a) The estimated light positions.



(b) The integrated depth of the near-light data using the Frankot and Chellappa, 1988 algorithm.

Figure 4.5: Results from Huang et al., 2016 near-light algorithm.

Is a relighting dome necessary for robust photometric stereo results? A fixed lighting rig – a relighting dome – should help to overcome the inaccuracies of the high-frequency data from the manual approach by using consistent lighting positions that could be calibrated. A relighting dome should also speed up the data capture. [chapter 5](#) explores this hypothesis.



(a) The prints without colour.



(b) The prints with colour.

Figure 4.6: Qualitative comparison of the high-frequency data of the Virgin Mary using the Huang et al., 2016 photometric stereo algorithm (left) and the Lucida (right). © Oak Taylor Smith for the Factum Foundation for Digital Technology in Conservation

4.2 DATA FUSION

As seen in Nehab et al., 2005, Tominaga, Ujike and Horiuchi, 2010, MacDonald, 2015, and other data fusion approaches photometric stereo data is not sufficient by itself. However, it is not clear if they used photometric stereo algorithms with

a far-light model or a near-light model. A data fusion approach was chosen so that the different methods could be compared. Data fusion methods from Nehab et al., 2005, Berkiten, Fan and Rusinkiewicz, 2014, and Antensteiner, Štolc and Pock, 2018 come from the field of computer vision; while they improve the geometric results, the high-frequency data is used in its normal form when the data is combined. The high-frequency data is not given a scale; instead, its contribution is decided based on somewhat arbitrary weighting factors (from a metric point of view). For this reason, it was decided to use the method described in Tominaga, Ujike and Horiuchi, 2010 and MacDonald, 2015. This approach uses a FFT algorithm to combine the photometric stereo after it has been integrated with low-frequency depth data. This allows for the photometric stereo data to be scaled using a calibration target before the data fusion, and thus, metrically accurate.

The following chapters discuss the testing of the relighting dome (chapter 5), and the data fusion workflow introduced through a case study (chapter 6). The case study compares results from photogrammetry, photometric stereo, data fusion, and the Lucida. More details on the design and construction of the relighting dome prototype is discussed in detail in Appendix A.

TESTING THE RELIGHTING DOME

5.1 THE ONE-ARM TEST

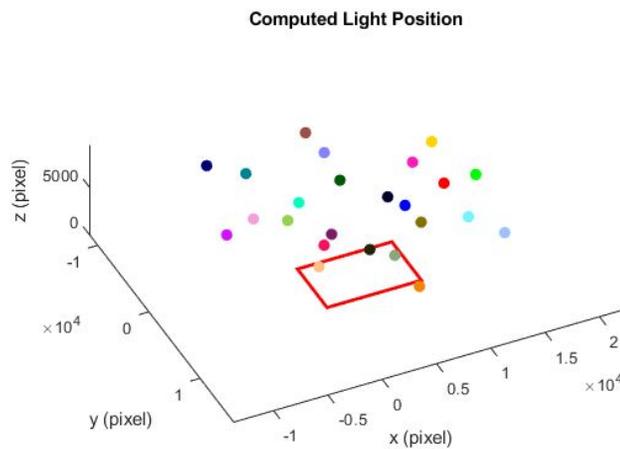
The following chapter outlines the tests performed on the relighting dome. These tests were done before the testing of the data fusion methodology.

During the construction of the dome, a single-arm was used to test the rigidity of the dome in comparison to the free-hand method shown in [Figure 4.4](#). With the camera, distance, and target area kept similar to the previous test, the single-arm was rotated radially to the approximate angles of the final dome ([Figure 5.1](#)).

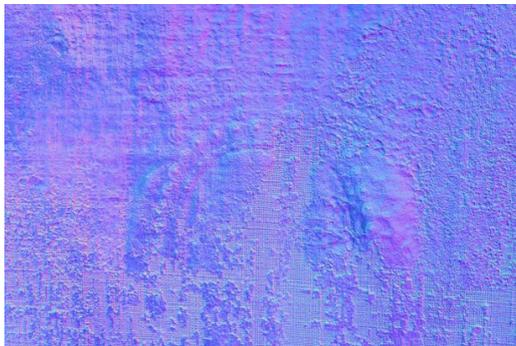


Figure 5.1: Documenting the Virgin Mary painting using one arm of the relighting dome.
© Oak Taylor Smith for the Factum Foundation for Digital Technology in Conservation

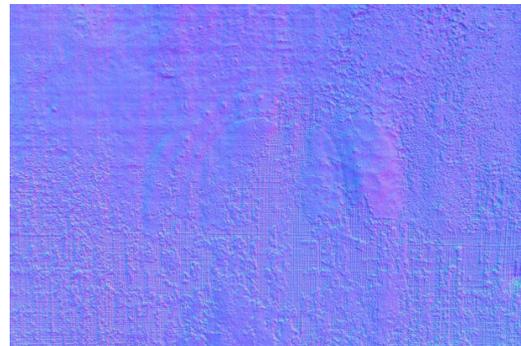
The data from the single-arm test was processed using the uncalibrated method from Huang et al., 2016. Figure 5.2 shows the normals from the free-hand test and the one arm test, as well as the estimated light positions. The light positions from the one-arm test represent the shape of the dome much better than the those from the free-hand test (Figure 4.5), and as a result, the normals appear flatter – particularly in the details on the right side of the image.



(a) Estimated light positions.



(b) The normals from the free-hand test.



(c) The normals from the one arm test.

Figure 5.2: Comparing the results of free-hand and arm-based photometric stereo approaches using a near-light model.

5.2 TESTING METHODOLOGY

The remaining tests on the dome used the Standard Board (Figure 5.3). The board, created by Factum Foundation for Digital Technology in Conservation, contains many of the types of surfaces that some scanning systems have difficulty with: black surfaces, glossy surfaces, and areas of high contrast. The purpose of these tests was to determine the relighting dome's ability to capture these difficult surfaces, as well as to test different lighting configurations, light calibration methods, and the height calibration target.



Figure 5.3: The Standard Board for testing scanning techniques. The board contains challenging areas (dark and glossy) that are common in paintings.

In this stage of developing the prototype, the dome was placed in a downward configuration for the tests, and a power supply and laptop were used to control the camera and the lights of the dome.

5.3 TEST ONE

A calibration target was created for the first test. The target was used to calibrate the height range of the depth map after the normals are integrated. The calibration target was designed based on the one used by Huang et al., 2016 and was 3D printed in resin using a stereolithography-based printer at a precision of $50 \mu\text{m}$ (Figure 5.4).

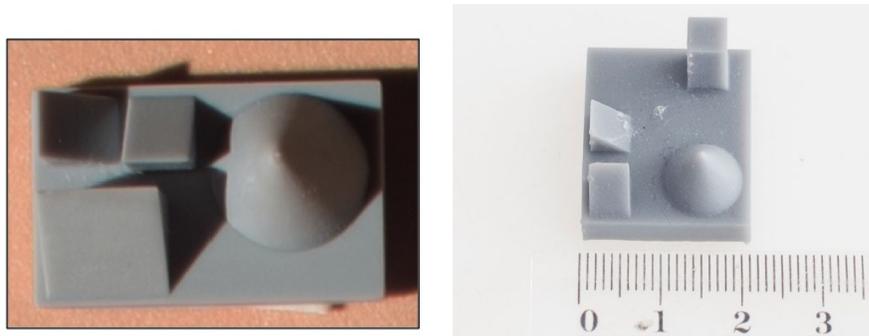
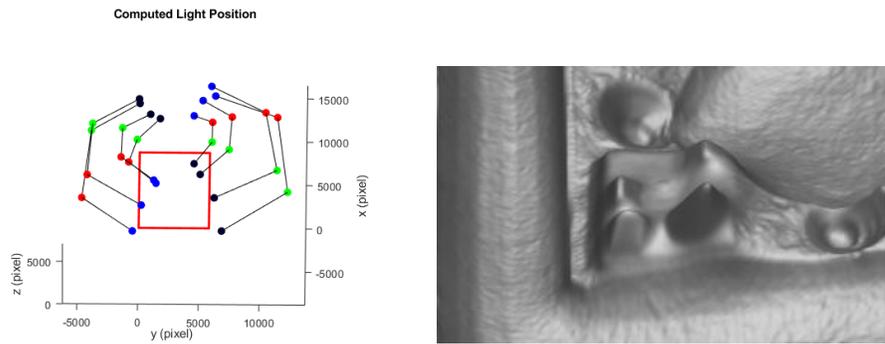


Figure 5.4: Comparing the calibration target from Cossairt et al., 2015 (left) and the first relighting dome test (right).

The setup for the first test can be seen in Figure 5.5. A Canon 5DS R was used with a 50 mm lens at a distance of 55 cm to the board surface. This resulted in a lateral spatial resolution of $45 \mu\text{m}$. Figure 5.6 shows the results from the first test on the Standard Board. The normals show that the photometric stereo was able to resolve the details in the dark and glossy areas. The light positions however, are incorrectly estimated. Furthermore, the light position estimates were not consistent between the three tiles captured on the Standard Board. Another issue was the calibration target – the shape of the base is warped. Our research suggests that this is a result of the way that normal integration works. Successful normal to depth conversion requires a continuous gradient. Many integration algorithms produce errors when the normal field drops – at an edge.



Figure 5.5: The setup for test one. © Otto Lowe for the Factum Foundation for Digital Technology in Conservation



(a) The estimated light positions.

(b) An enlarged image of the integrated depth of the calibration target.



(c) The calculated normals

Figure 5.6: The results of test one on the Standard Board.

5.4 TEST TWO

The second test was designed to improve the two shortcomings of the first test: the light positions and the calibration target. Bevels were added to each edge of the calibration target to create a continuous surface during the depth integration. In the first test, reconstruction of the cone proved the most successful. The new target was designed in two iterations, each with two main features: One with two hemispheres and one with two cones [Figure 5.7](#).

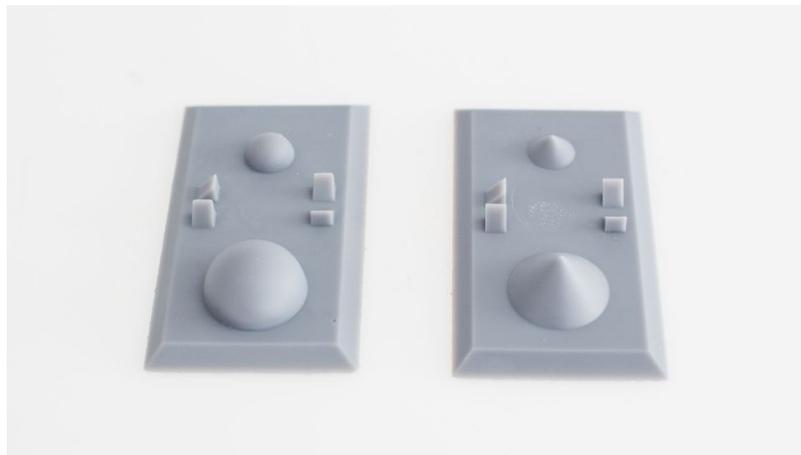


Figure 5.7: The second iteration of height calibration targets.

The L-bracket was used to fix the camera to the dome for test two. Because the L-bracket connects to a specific place on the circular plate, the arms, and therefore the lights were constant relative to the camera. The light calibration was done in a two-step process. First, a photometric stereo tile was captured using a black reflective sphere. The light directions were calculated using the reflections from the sphere and the [RTI software \(Figure 5.8\)](#) ((Cultural Heritage Imaging, 2015)). Second, the absolute light positions – in pixels – were calculated using the known

geometry of the dome and distance of the capture setup (Equation 5.1). Figure 5.9 shows the difference between the estimated and calibrated light positions.

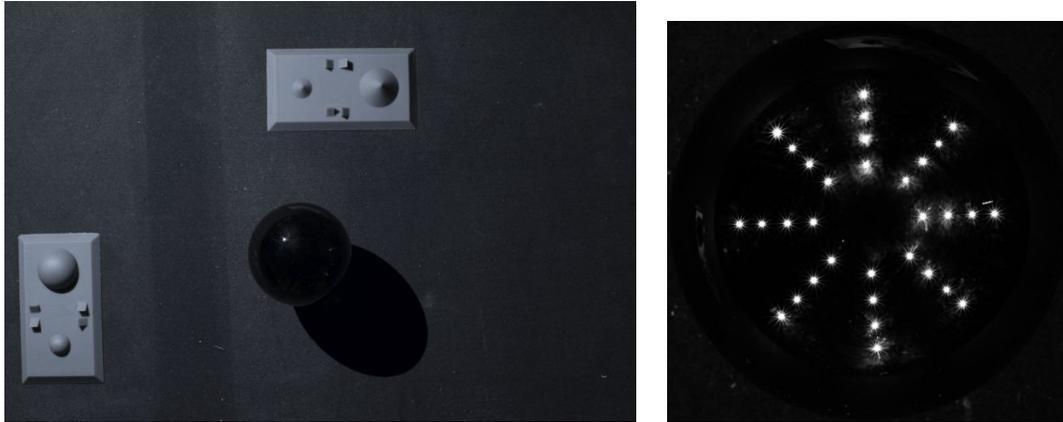


Figure 5.8: One image from the light calibration set (left) and the detected light directions on the black sphere (right).

$$Lp = \frac{Ld \times D_L}{GSD} = \frac{Ld \times D_L}{\left(\frac{D_c \times ps}{f}\right)} = \frac{Ld \times 500 \text{ mm}}{\left(\frac{550 \text{ mm} \times 0.004 \text{ mm}}{50 \text{ mm}}\right)} \quad (5.1)$$

where

- Lp light position
- Ld light direction vector
- D_L light distance
- GSD ground sample distance
- D_c camera to surface distance
- ps pixel size
- f focal length

The improved light positions also improve the calculated normals. As can be seen in Figure 5.10 the normals from the near-light calibrated light positions are flatter and produce sharper edges – notice the calibration targets.

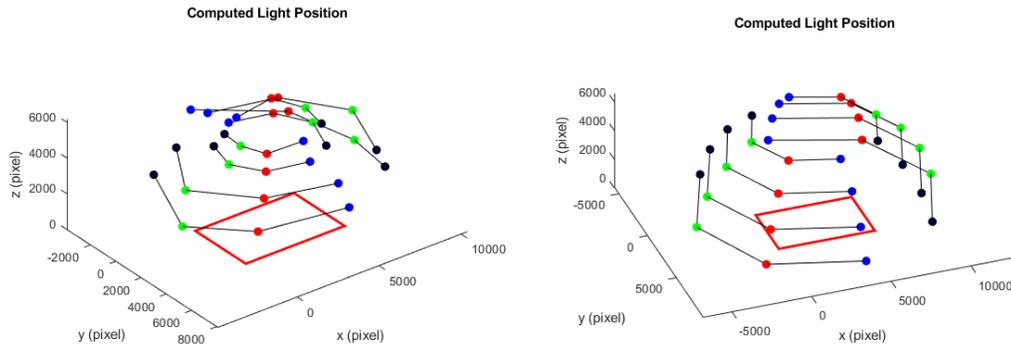


Figure 5.9: Comparing the uncalibrated (left) and calibrated light positions (right).

With the light positions calibrated, four different lighting configurations were tested to determine if all of the 32 LEDs are required to produce accurate normals (Figure 5.11). The normals from these configurations were compared against the full set of 32 LEDs, treating that as a ground truth. The angular difference between each configuration and the ground truth is shown in Figure 5.12 and Figure 5.13.

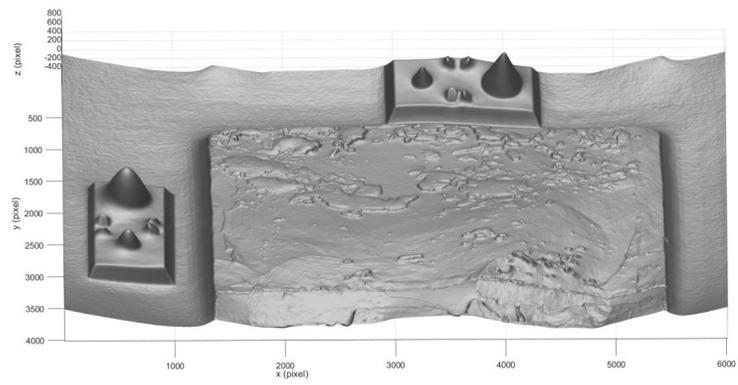
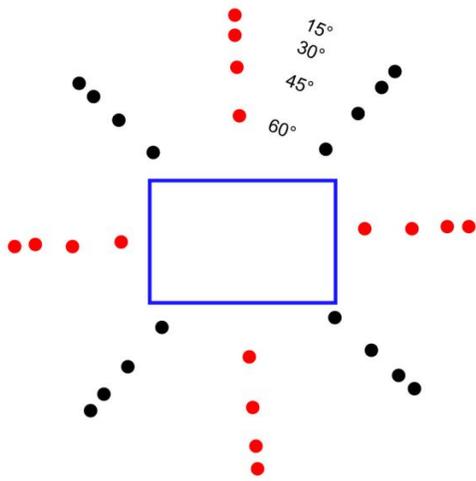
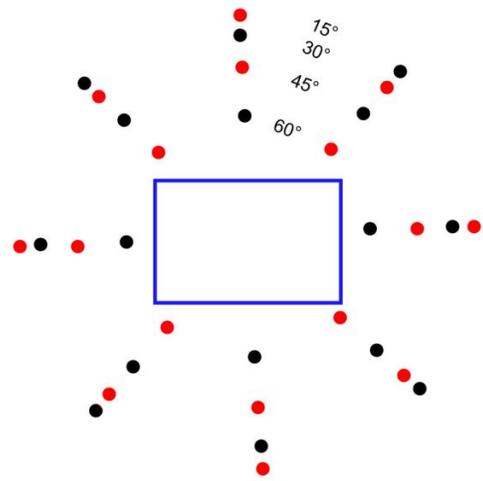


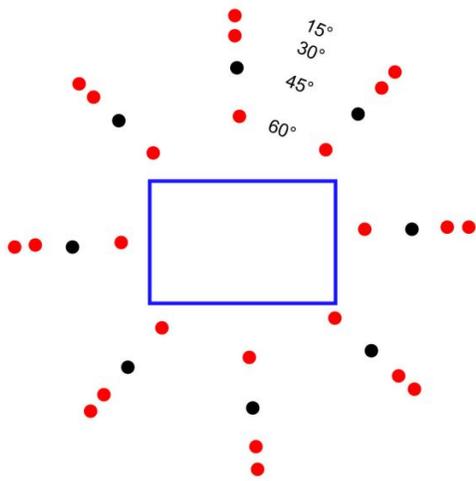
Figure 5.10: The depth from test two (top) and the calibrated normals (bottom).



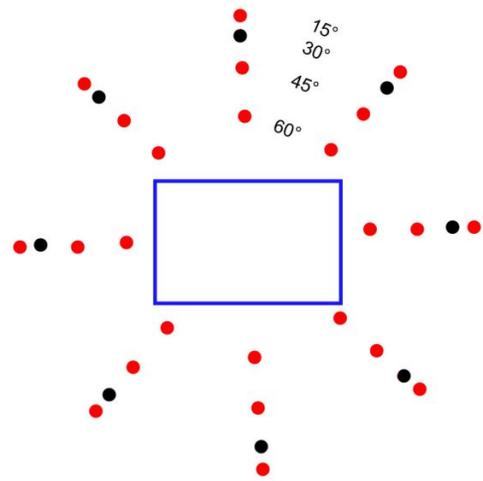
(a) 16 - every second column.



(b) 16 - alternating columns

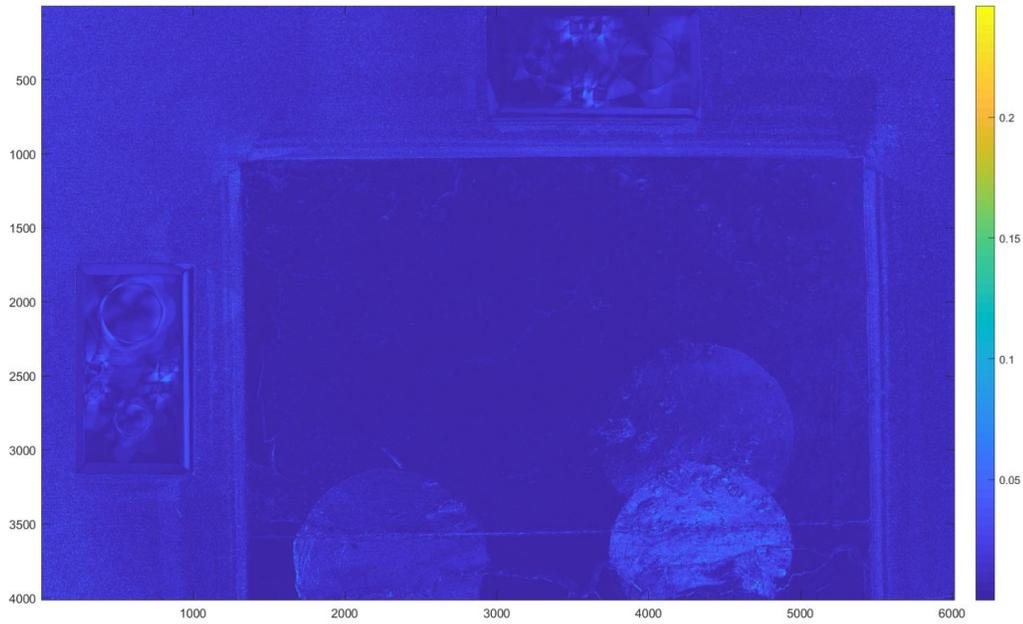


(c) 24 - rows one, two, and four.

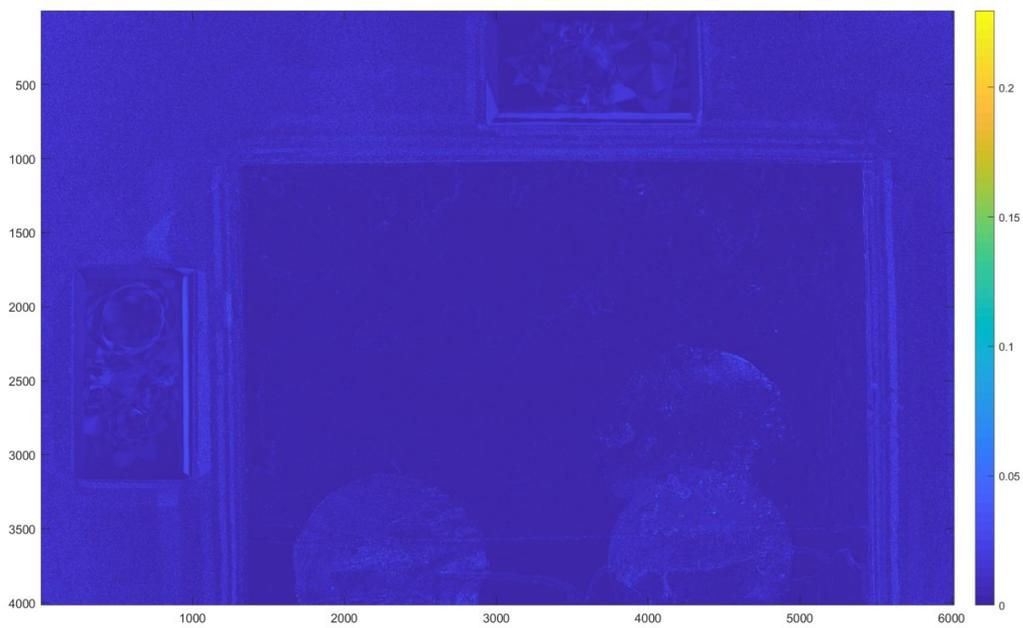


(d) 24 - rows one, three, and four

Figure 5.11: The tested lighting configurations (red is active, and black is inactive).

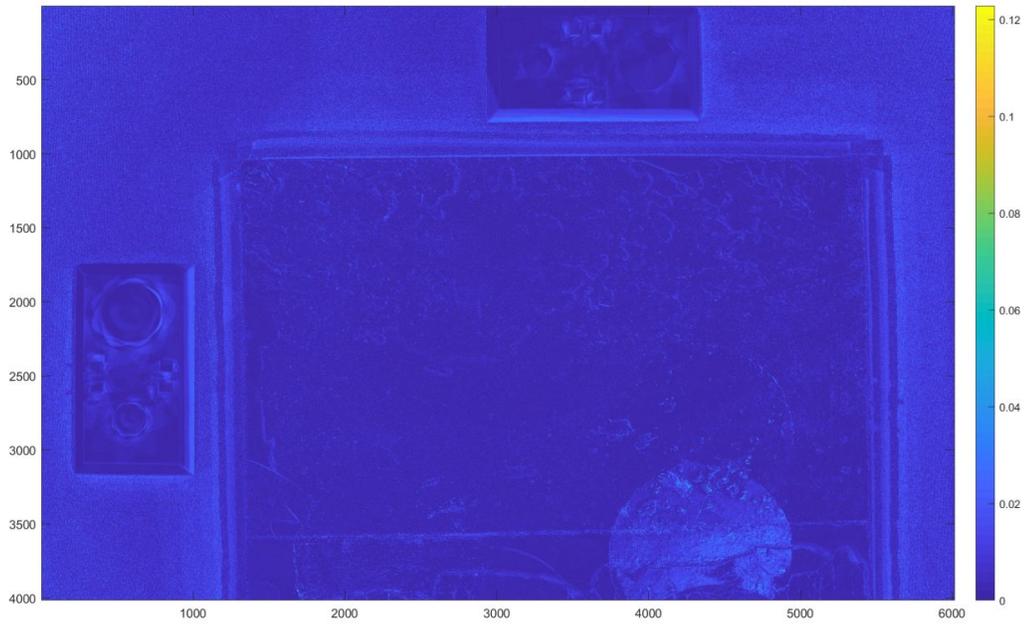


(a) 16 - every second column.

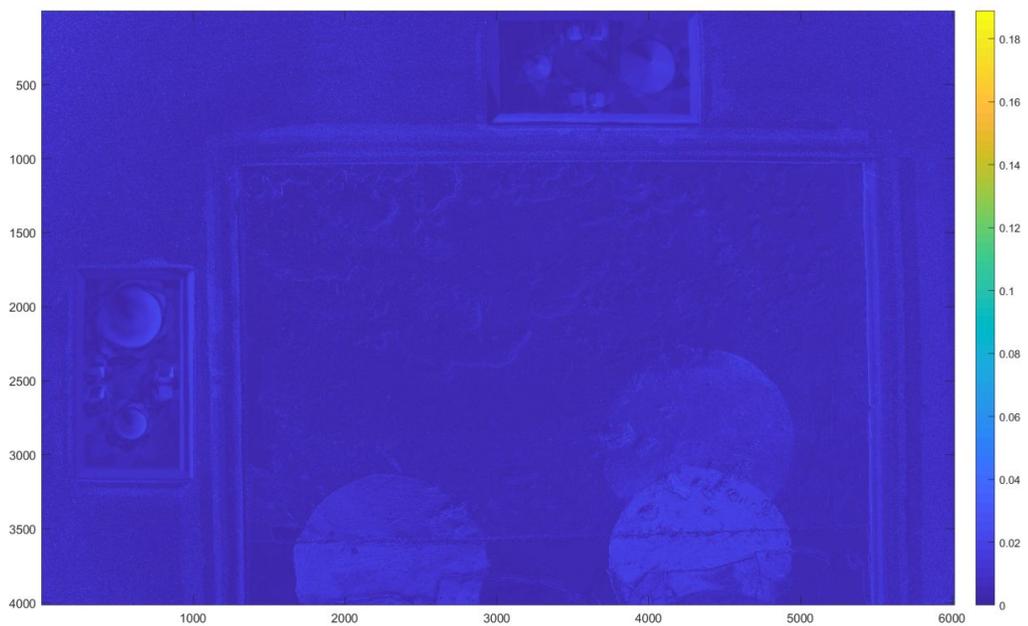


(b) 16 - alternating columns

Figure 5.12: The angular differences between tested lighting configurations with 16 lights and the ground truth (all 32 light positions).



(a) 24 - rows one, two, and four.



(b) 24 - rows one, three, and four

Figure 5.13: The angular differences between tested lighting configurations with 24 lights and the ground truth (all 32 light positions).

The angular differences were all low – just over 0.05° for the 16-light configurations and just below 0.05° for the 24-light configurations. In the 16-light configurations, the radial variation of alternating columns – as opposed to every second column – produces fewer errors in the dark and glossy areas, as well as some areas of high relief (on the calibration targets). In the 24-light configurations, the omission of the third row (45°) results in more angular variation in the darkest patch on the Standard Board as well as around the edges of objects. Conversely, the omission of the second row (30°) leads to a higher angular variation on the faces of the calibration targets as well as on the grey and specular circular patches. Though the full configuration produces the best results, if speed, file-size, or computing power are constraints on a project, fewer lights could be used to process the data resulting only in small errors. The angular errors in the glossy areas are likely caused by specular effects; the photometric stereo model assumes a Lambertian surface and when the light positions change between the configurations, so does the perceived specularity. This could be improved by either implementing an algorithm to discount bright points, or implementing a reflectance model that also calculated the specular components of the surface.

Figure 5.14 shows the depth acquired by integrating the normals from the full set of calibrated lights. The integration used the Frankot and Chellappa, 1988 approach. The sphere calibration target (left) does not reconstruct well; it became more conical in shape. On the cone target (top), though the base is still somewhat warped, the reconstruction of the conical geometry was successful. The difference between the two cone heights can be used for the height calibration.

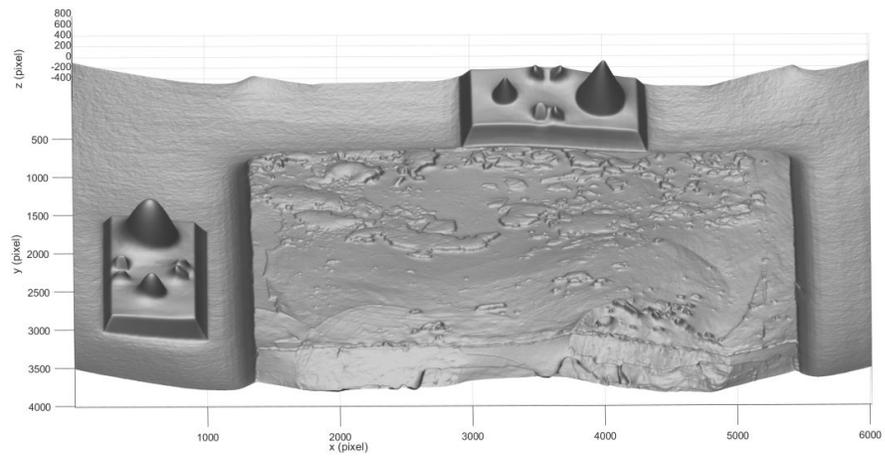


Figure 5.14: The integrated depth from test two on the Standard Board.

CASE STUDY

The following sections discuss the case study used to develop the workflow for documenting painted surfaces developed as part of this research. The 17th-century painting of the Virgin Mary (Figure 4.3) was used as the case study to test and compare the data fusion results with photogrammetry, photometric stereo, and the Lucida.

6.1 DATA CAPTURE

The image-based data for the Virgin Mary painting was captured in two steps: first, the photogrammetric acquisition, and then the photometric stereo acquisition. Both data types were captured with the same camera and lens – A Canon 5DS R with a 50 mm lens. The painting was hung on a wall to simplify the placement of targets and provide continuity for the normal integration. Photogrammetric scale bars and photometric stereo calibration targets were placed close to the painting. The top left quarter of the painting was also scanned using the Lucida scanner over two hours.

6.1.1 *Photogrammetric Acquisition*

Cross polarized lighting was used to reduce the reflections during the photogrammetric capture: a polarizing filter on the camera lens and polarizing paper on a

top-mounted flash. The photogrammetric images were taken at grazing angles to the surface from a distance of around 55 cm – the same distance as the photometric stereo. This improves the camera calibration between images from the two data types and keeps the camera at good proximity for the flash. 174 images were captured during the 30-minute photogrammetric acquisition at a GSD of approximately $50\ \mu\text{m}$.

6.1.2 *Photometric Stereo Acquisition*

The photometric stereo data was captured using the relighting dome. First, one set of photometric stereo was captured with a reflective sphere to calibrate the light directions. Because the painted surface was larger than the capture area (36 cm wide by 24 cm high), the photometric stereo was captured in multiple tiles. The dome and camera were connected to the aluminum rails used by the Lucida (Figure 6.1). The rails allowed for easy horizontal movement and guaranteed that there was no rotation between the overlapping tiles. The camera was placed 55 cm from the surface as perpendicularly as possible, resulting in a lateral spatial resolution of $46\ \mu\text{m}$. A total of 24 tiles were captured with an overlap of 30%. The overlap amount was enough to ensure redundancy in case there were issues aligning the tiles together. With a 0.5 second exposure, each tile took just over one minute to capture the set of 32 images. However, the total acquisition took one and a half hours due to difficulty with the vertical movement of the dome on the rails.

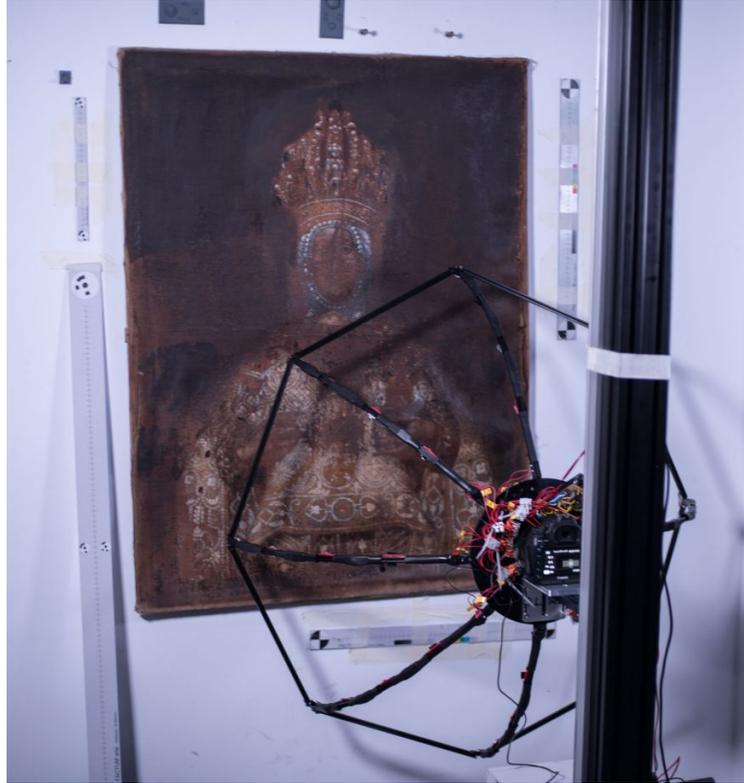


Figure 6.1: The photometric stereo documentation of the painting of the Virgin Mary © Osama Dawod for the Factum Foundation for Digital Technology in Conservation.

6.2 PROCESSING

6.2.1 *Pre-processing*

First, the exposure on the images was corrected. Each image in a photometric stereo tile was adjustment by the same amount. The RAW photogrammetric images were converted to JPG files. The photometric stereo files were converted to DNG files, with the addition of one JPG image from each tile.

6.2.2 Photogrammetry

The JPG images (from photogrammetry and one from each photometric stereo tile) were processed using the photogrammetric pipeline discussed in [chapter 2](#) in Agisoft Photoscan (1.4.1). A self-calibrating bundle adjustment was used to calibrate the camera and calculate the exterior orientation of each camera position. [Figure 6.2](#) shows the camera poses of the photogrammetric and photometric stereo data. The data was scaled using the scale bars around the painting to an average precision of $120 \mu\text{m}$. The point cloud from [DIM](#) can be seen in [Figure 6.3](#). The projection plane for the [DEM](#) was created by manually picking four points on the painting surface to define the horizontal and vertical directions. This was sufficient for the case study, but a plane fitting approach should be implemented in the future. The DEM was created at a resolution of $174 \mu\text{m}$ based on the resolution of the point cloud ([Figure 6.4](#))

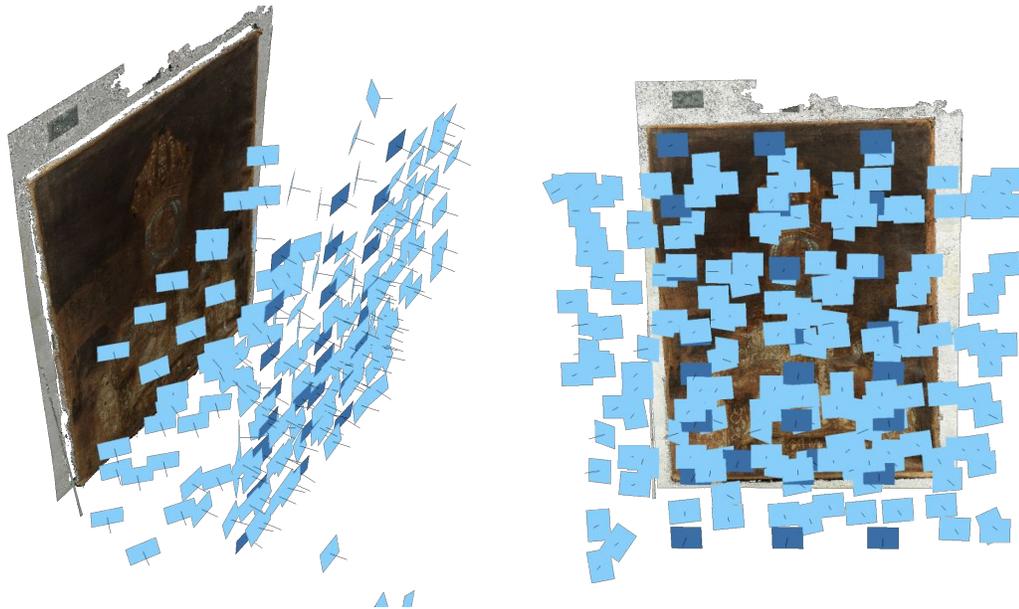


Figure 6.2: The image poses in relation to the Virgin Mary painting. The dark blue images are the photometric stereo tiles.

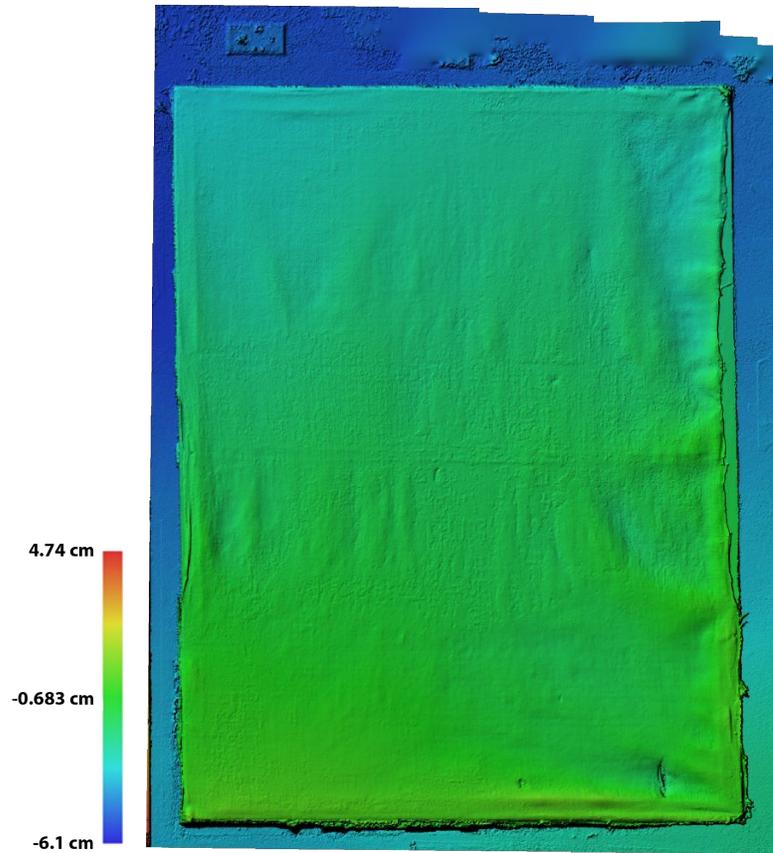


(a) The point cloud of the whole painting.

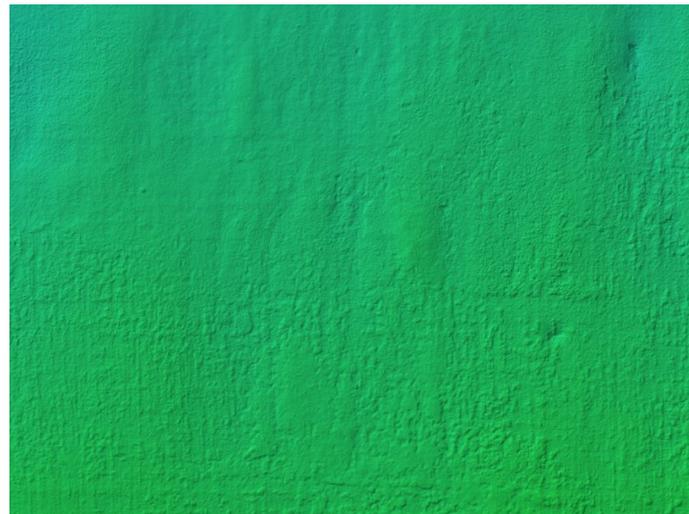


(b) An enlarged view of the point cloud.

Figure 6.3: The coloured point cloud of the Virgin Mary painting captured using photogrammetry.



(a) The DEM of the whole painting.



(b) An enlarged view of the DEM.

Figure 6.4: The DEM of the Virgin Mary painting captured using photogrammetry.

6.2.3 Photometric Stereo

The lighting directions of the relighting dome were calibrated using the method described in [chapter 5](#) – using a reflective sphere and the RTIBuilder software Cultural Heritage Imaging, 2015. The near-light photometric stereo algorithm from Huang et al., 2016 was modified so that it accepts .lp files, the pixel size of the camera, and the camera to surface distance; it used that information to calculate the absolute light positions (in pixels) from the surface. Once the light positions were calibrated, the modified algorithm was used to run a batch process on all 24 photometric stereo tiles. At around twelve minutes per tile, the processing took five hours. The normal calculation required a powerful computer with a lot of memory – 64 GB of RAM because of the size of the 50 megapixel images. Though the relighting dome itself is low cost, it is clear that a powerful computer is required to process the data. This reduces the universal applicability of the overall approach.

[Figure 6.5](#) shows the calculated normals for two adjacent tiles. There is a difference in the low-frequency normals of the overlapping area – as can be seen on the right side of the top image and the left side of the bottom image. This is caused by a combination of two factors. The areas at the edge of the capture area are farther from the centre of the dome and receive less even lighting from the dome. The uneven lighting effect causes a different low frequency when the light comes from predominantly one side on one tile and another side on the adjacent tile. The photometric stereo algorithm also assumes that the surface is Lambertian. However, the Virgin Mary painting is not Lambertian. The specularities in the images affect the normals at the edges of the capture area. In the future, pixel filtering or a different surface reflection model – one that includes specularities –

should be added to the near-light algorithm. A test should be done to determine an ideal overlap percentage for photometric stereo capture with the dome.

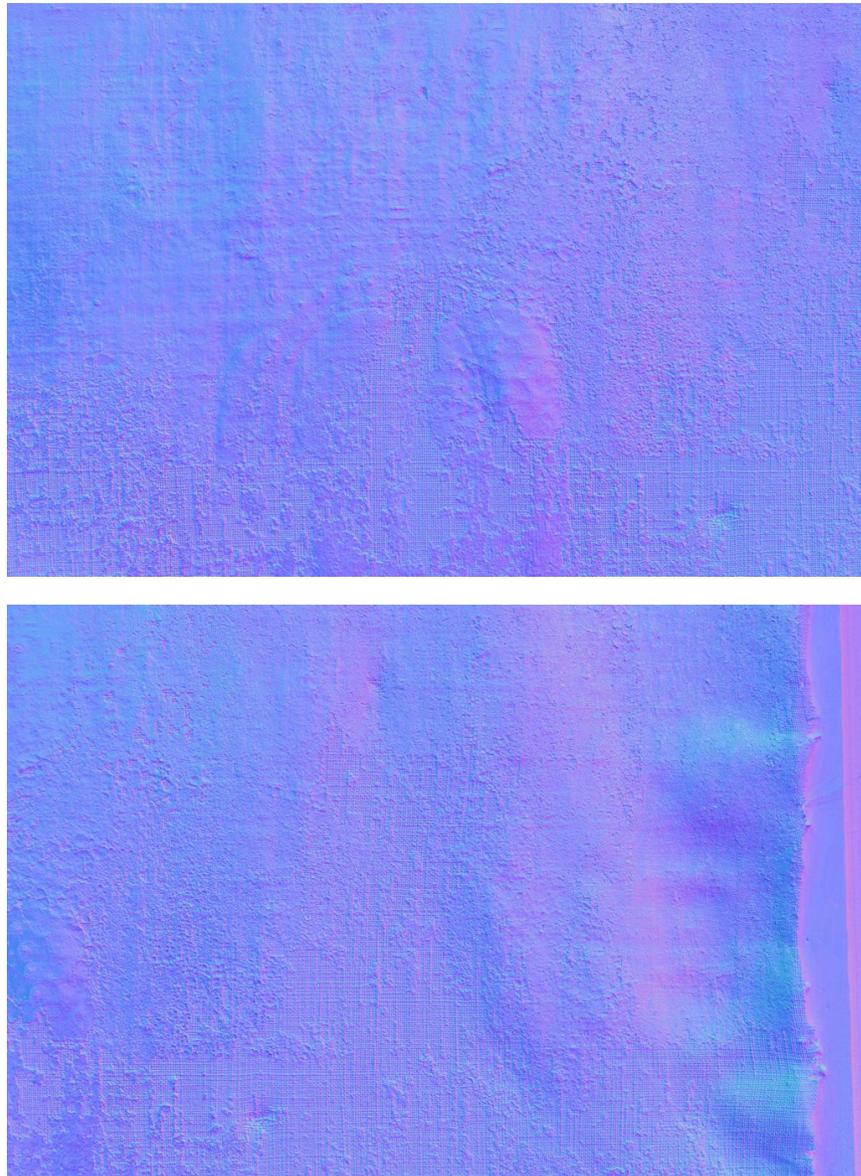


Figure 6.5: The calculated normals for two adjacent photometric stereo tiles of the Virgin Mary painting.

6.3 DATA FUSION

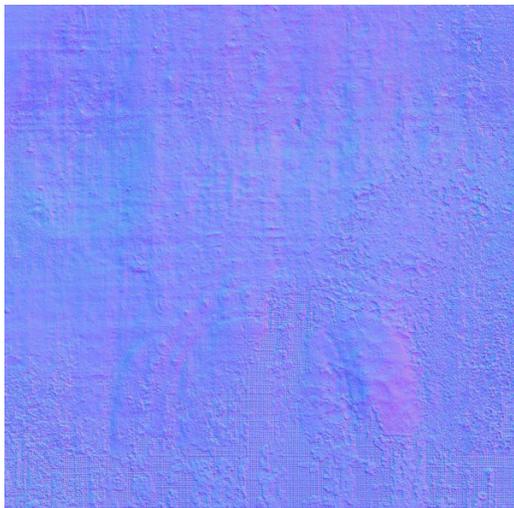
A Python script was used in Photoscan to calculate the rotation between each camera and the projection plane. Because the normals are relative to the camera position, these rotations were applied to the normals of each tile to account for any differences in angle between the projection plane and the camera. The normal maps were then placed in Photoscan, replacing the JPG images from each tile that were used in the bundle adjustment. An orthophoto was made by projecting each tile on the same projection plane as the [DEM](#). This orthographic projection uses the interior orientation parameters from the camera calibration to remove the distortion from the images. The overlapping regions were blended using two different methods: mosaic blending and average blending ([Figure 6.6](#)). Average blending averages the pixel values in the overlapping regions between the images. The issue with average blending is that it causes visible seamlines. These seamlines create ridges when the normals are integrated into depth values. The ridges contain both high and low-frequency information, creating unwanted artifacts after data fusion. Mosaic blending creates a seamline between two images and then transitions the values on either side of the seam to create an 'invisible' blend. The issue, however, is that when the overlapping normals are contradictory in the low-frequencies – which they were at the edges of some of the tiles – it creates hill or valley effect.



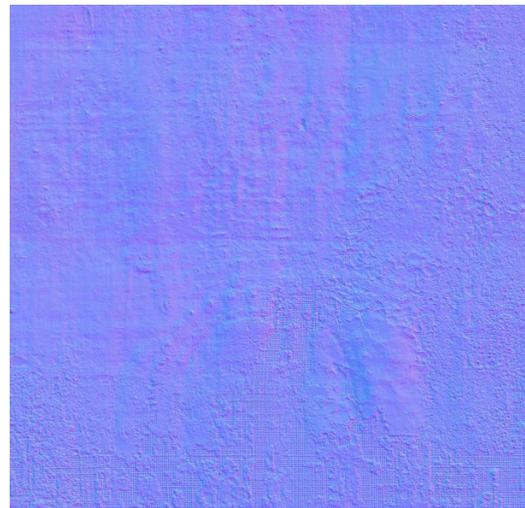
(a) The whole mosaic normal map.



(b) The whole average normal map.



(c) An enlarged view of the mosaic normal map.



(d) An enlarged view of the average normal map.

Figure 6.6: The orthophoto normal blending of the Virgin Mary painting.

Figure 6.7 shows the angular differences between the two blending methods. Mosaic blending was used for this case study because most of the adverse effects are limited to the low-frequencies. However, a more sophisticated blending method (like the one introduced by Berkiten, Fan and Rusinkiewicz, 2014) should be implemented in the future.

The mosaic blended normals were integrated into a depth map using the algorithm from Frankot and Chellappa, 1988. The resulting depth values from the integration were in pixel units; they needed to be scaled. Two scaling methods were tested: scaling using the pixel size, and scaling using the calibration target. The first method multiplies the Z-depth values by the pixel size, which was $46 \mu\text{m}$ for the case study. For the second method, the depth at the peak of each cone of the calibration target was used to calculate a scaling factor. This scaling factor was based on the knowledge that the height between the peaks was 6 mm . An enlarged render of the two depth maps can be seen in Figure 6.8.

The two depth maps were compared in the mapping software, Global Mapper. Figure 6.9 shows the difference between the pixel scaling method and the depth scaling method for the photometric stereo depth. The taller cone should have been 12 mm tall, and the shorter cone should have been 6 mm tall. The pixel scaling method under scaled the depth (by a factor of about 1.5).

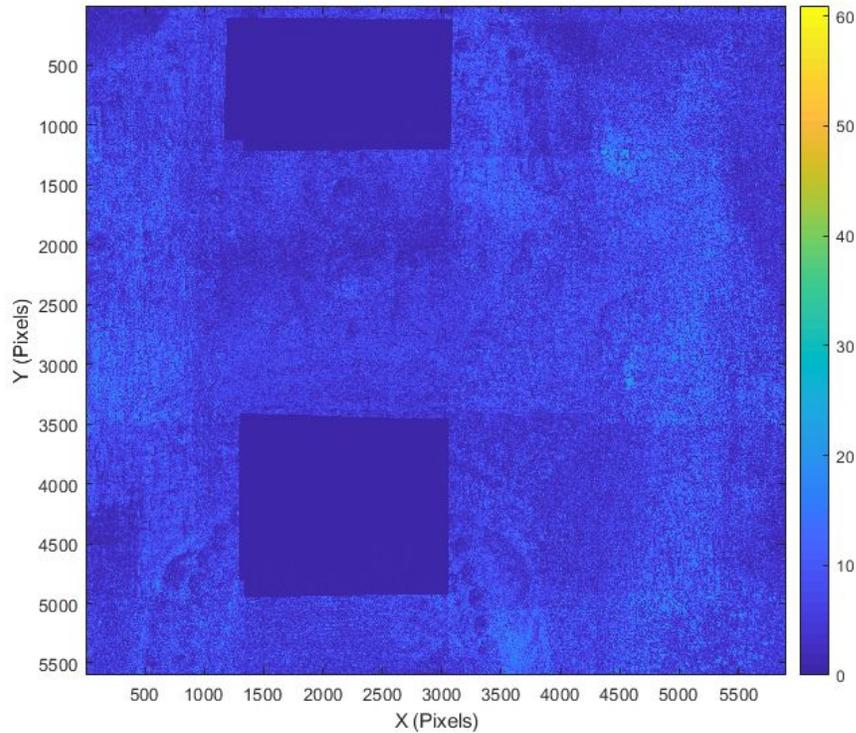


Figure 6.7: An enlarged view of the angular differences between the two normal blending methods.

The following depth map comparisons are comparisons of shape and magnitude, but not the exact 3D height. It is difficult to align 2D depth maps to one another in the depth plane (Z), and converting the high-resolution data creates meshes that are too large to work with. So, for analysis, the height values were brought close together, and profiles of the surface were compared to one another. For comparison, the data from the Lucida was considered as a ground truth. However, the Lucida is not a *true* ground truth because the data fusion captures data at a higher resolution than it.

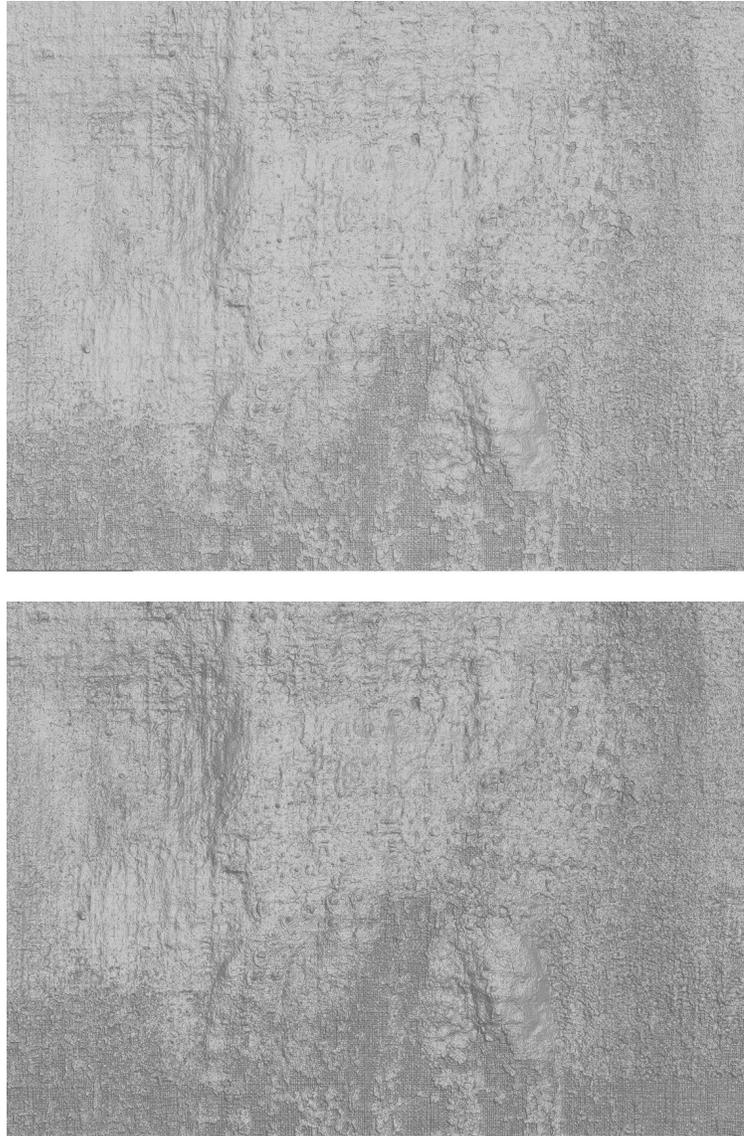
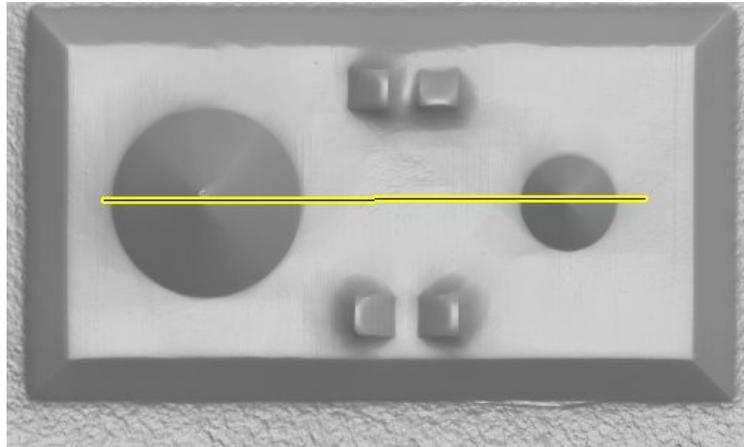
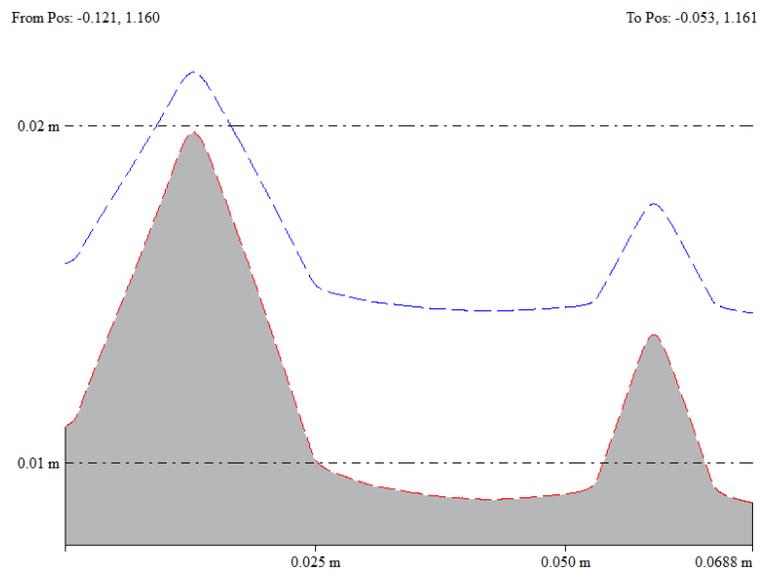


Figure 6.8: A shaded render of the depth maps from pixel-based scaling (top) and target-based scaling (bottom).

Figure 6.10 shows three profile comparisons between the DEM and photometric stereo depth scaled using the calibration target. The figure shows that the photometric stereo data contained more high-frequency details than the DEM and was able to reconstruct the calibration target well. The low frequency of the photometric stereo was warped by the mosaic blending in the overlapping areas between the tiles.



(a) The profile drawn through the calibration target.



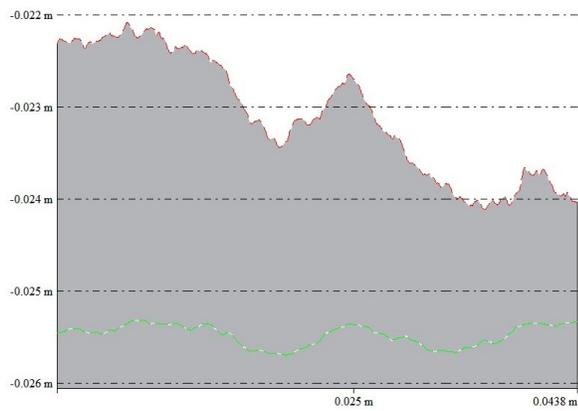
(b) The depth at the profile drawn through the calibration target.

The red line is the depth scaled with the calibration target, and the blue line is the depth scaled with the pixel values.

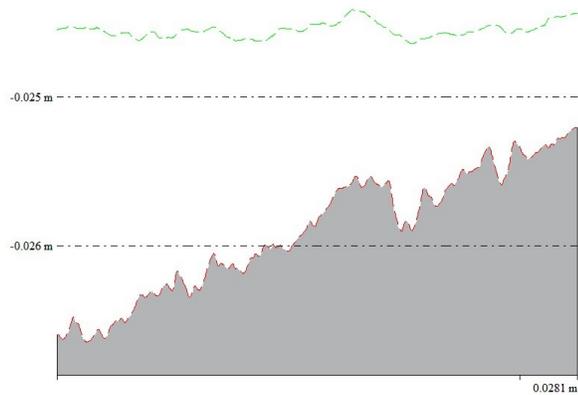
Figure 6.9: A comparison of the different depth scaling methods for the photometric stereo data.



(a) Depth comparison at the calibration target.



(b) Depth comparison in the middle of a photometric stereo tile.



(c) Depth comparison at the overlapping region between photometric stereo tiles.

Figure 6.10: Depth profiles compare the photometric stereo depth and the photogrammetric DEM. The red line is the photometric stereo depth and the green line is the DEM.

The DEM and the photometric stereo depth map were converted to the frequency domain using a FFT. The log power distribution of the spatial frequencies can be seen in Figure 6.11. The DEM log power distribution had mostly higher power values (low-frequencies), while the photometric stereo depth map had mostly lower power values (high-frequencies). Figure 6.12 shows the cumulative sum of the log power distributions of the DEM and photometric stereo depth map. Again, the sharper curve of the photometric stereo data shows that it contained more high-frequency data than the DEM.

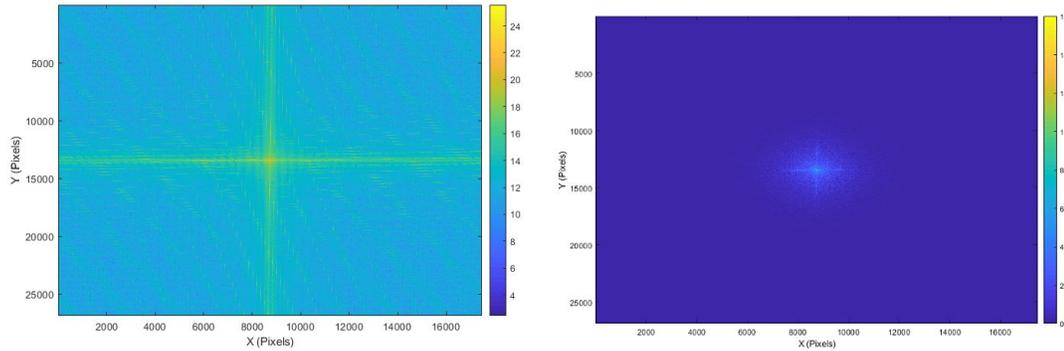


Figure 6.11: The log power distribution of the spatial frequencies of the DEM (left) and photometric stereo depth map (right).

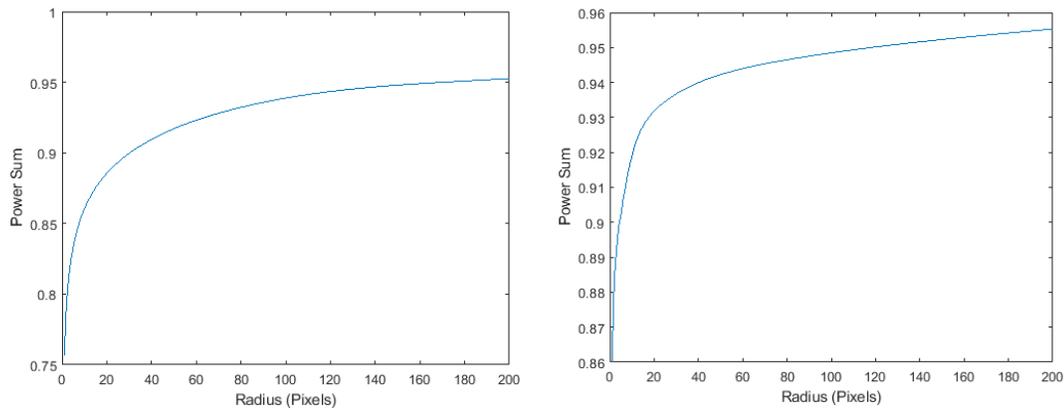


Figure 6.12: The cumulative sum of the log power of spatial frequencies of the DEM (left) and photometric stereo depth map (right) measured in radial pixels from the centre.

The log power distribution makes it easy to extract frequencies from each depth map. The low frequencies start in the middle and extend radially outwards towards the edge of the image. The low-frequency and high-frequency were combined using a linear interpolation at two different pixel radii: 40 pixels and 100 pixels. After the FFT, the combined depth was scaled using the pixel size. Figure 6.13 shows the difference between data fusion of photometric stereo depth – scaled using both methods – and the DEM. Because the FFT combines relative frequencies, the difference between the photometric stereo scaling methods was smaller than the required threshold – close to $10 \mu m$. As a result, if data fusion is being used, then a calibration target is not needed for the photometric stereo.

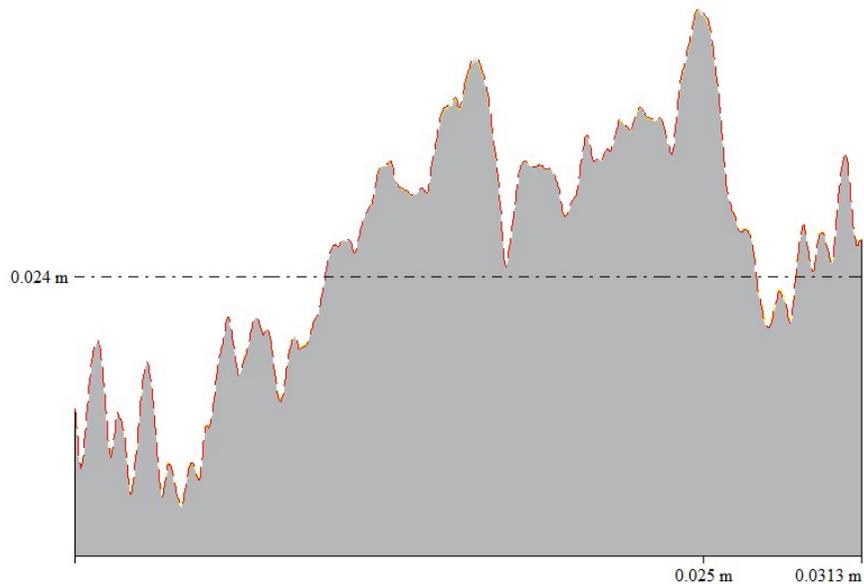


Figure 6.13: The difference between data fusion of the DEM and photometric stereo scaled using the calibration target and pixel methods.

The two fused depth maps were compared to the data from the Lucida (Figure 6.14). Even without measuring, it can be seen that the Lucida data was sharper at small edges. This is because photometric stereo normal integration

requires a continuous surface, so it creates ridges in areas with sharp drop-offs. [Figure 6.15](#) shows the depth profiles in the area to the right of the Virgin Mary's face. The shallower hills and valleys of the data fusion differed from the Lucida data in the range of 100 to 500 μm . However, the photometric stereo depth – before the data fusion – reproduced those details within a difference of around 200 μm ([Figure 6.16](#)). The data fusion flattened the details slightly.



Figure 6.14: The depth map of the Virgin Mary painting captured using the Lucida.

Though the photometric stereo depth produced details closer to the Lucida, it was also subject to areas of low frequency bias. [Figure 6.17](#) shows a hill effect caused by the mosaic blending of the normals. The low frequency bias could be improved with more robust normal blending and accounting for specularities in the photometric stereo model. It is possible that fixing those issues could remove the need for data fusion with the [DEM](#).

When comparing the data fusion, though the photometric stereo depth had sharper edges by itself (that were closer to the Lucida data), the larger radius (100 pixels) produced sharper details when compared to the smaller radius (40 pixels) (Figure 6.18). The larger contribution of the low-frequency extracted sharper details from the high-frequency data.

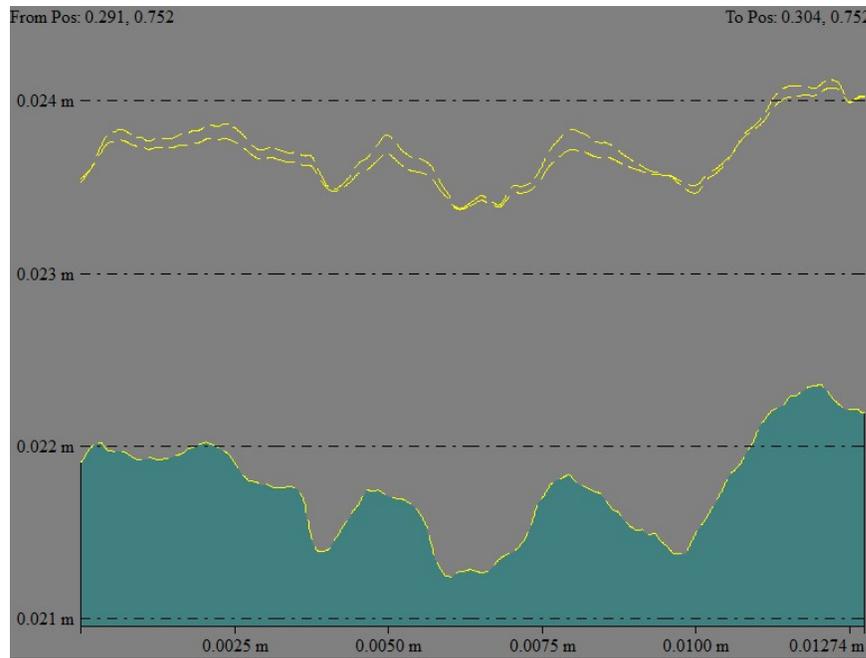


Figure 6.15: A depth profile comparing the data fusion (top of the profile) to the Lucida data (bottom of the profile) at an area with steep edges.

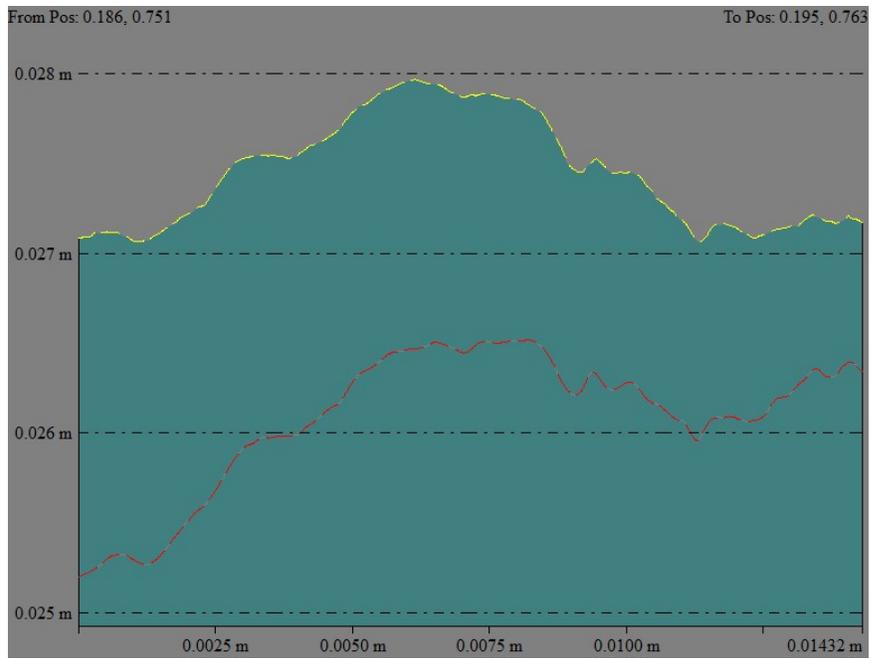


Figure 6.16: A depth profile comparing the photometric stereo depth (red) to the Lucida data (yellow).

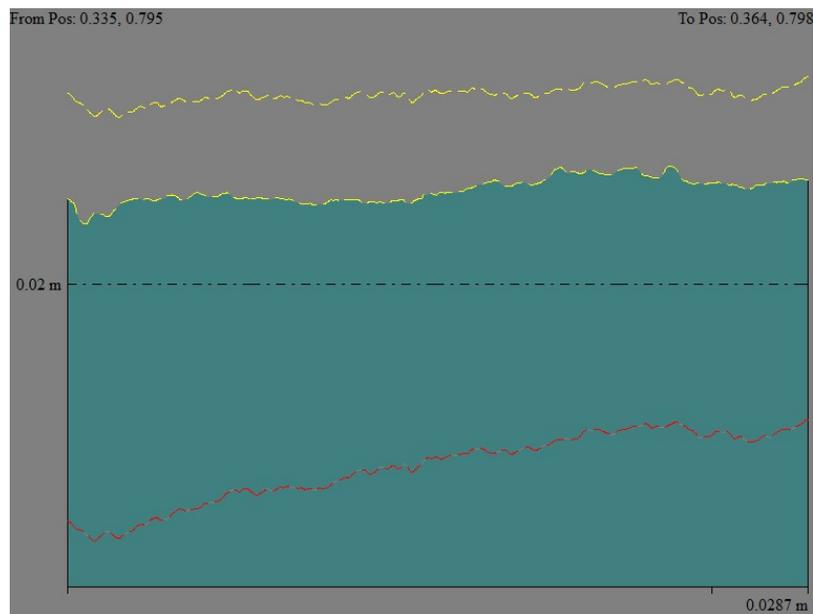


Figure 6.17: A depth profile comparing the data fusion (top profile) to the Lucida data (middle profile) and the photometric stereo depth (bottom profile).

Though there were some differences in depth, the photometric stereo and data fusion approaches captured more detail than the Lucida, which is limited to a

lateral spatial resolution of $100\ \mu\text{m}$ (Figure 6.19). According to the case study, the low-cost approaches are accurate on continuous surfaces – they are able to accurately reconstruct surface details like brush strokes. The largest errors occurred in areas with sharp edges, making the current method less suited for painted surfaces with significant deterioration.

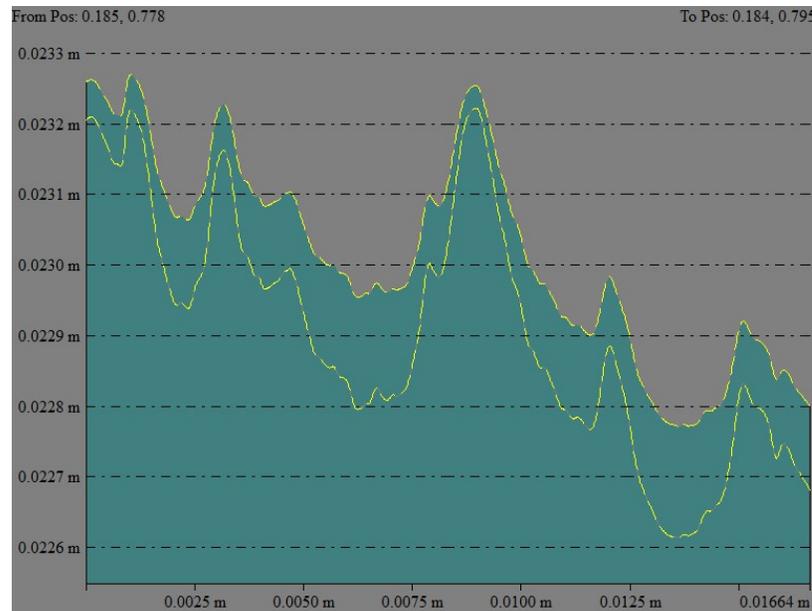


Figure 6.18: A depth profile comparing the data fusion with a radius of 40 pixels (top profile) to one with a radius of 100 pixels (bottom profile).

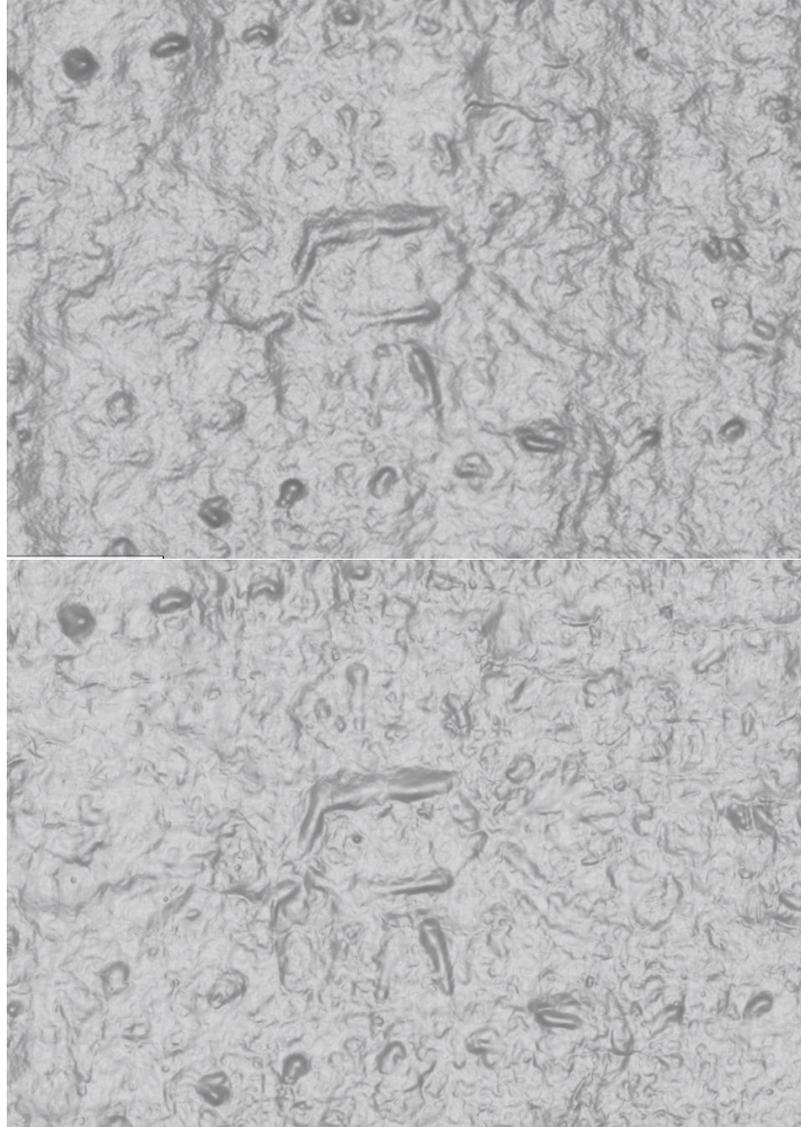


Figure 6.19: Comparing the details of the Lucida (top) and the data fusion (bottom) in a shaded render of the depth maps.

6.4 WORKFLOW AND DISCUSSION

A workflow – outlined in [Figure 6.20](#) – was developed based on the Virgin Mary case study. Though the high-level workflow is well defined, some smaller steps need refinement before it can be widely applicable. As mentioned in the previous section, the overlapping areas of the tiles are one of the largest sources of error due to light distances, specularities, and simplistic normal blending. Adding

a specular component to the photometric stereo model would help to solve these problems. A robust approach would also include a more sophisticated normal blending method like the one in Berkiten, Fan and Rusinkiewicz, 2014 which takes into account the average normal as well as its continuity with neighbouring normals. Better normal blending would also allow for photometric stereo capture using a tripod as opposed to a rigid rail system – it could overcome slight rotations between adjacent tiles. If these aspects were improved, it is possible that the DEM data fusion would not be needed at all: the low frequency of the photometric stereo could be robust enough. In this way, the photogrammetry could be captured at a lower GSD and just be used for scaling and image matching with the camera poses of the photometric stereo tiles.

During the case study, the pixel radius used for the data fusion involved some iteration and comparison to the Lucida data to find the optimal value. This would not be possible in practice if it were being used as the sole documentation method. A thresholded approach could help to overcome this by iterating the pixel radius until it varied by a certain threshold from the DEM.

Compared to other methods, this workflow reduces time on site by half, while increasing processing time. However, most of the processing is semi-automatic and can be run overnight. The hardware for the data fusion approach was around 15% of the cost of the Lucida.

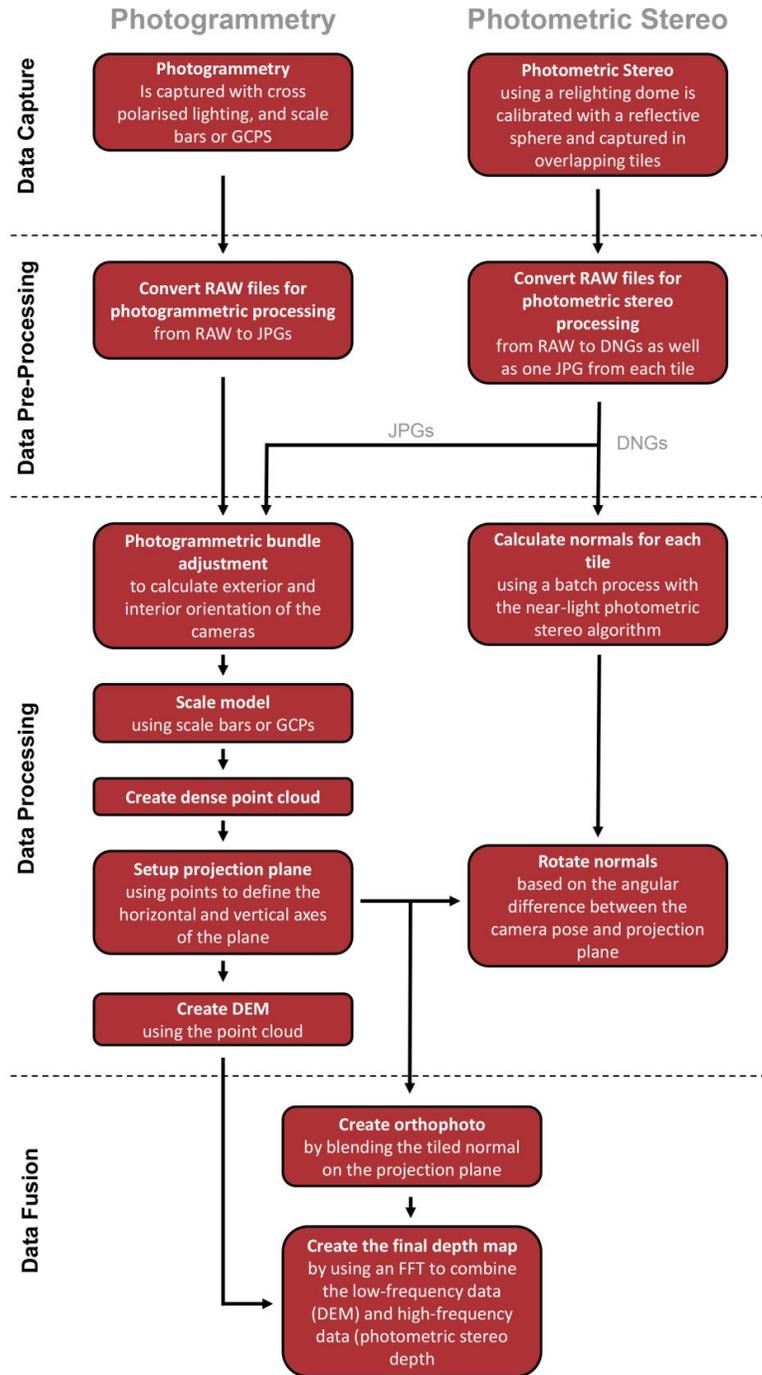


Figure 6.20: The overall workflow developed from the case study.

Part IV

CONCLUSION

“I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it...”

– William Thomson

CONCLUSION

7.1 CONCLUSION

The subtlety of the layers of a painted surface requires high-resolution documentation for conservators. Current commercial methods for documenting planar painted surfaces are either too expensive, too slow or are limited in their ability to capture both high and low-frequency information in sufficient detail. This thesis discussed a low-cost method for documenting planar (2.5D) painted surfaces by fusing photogrammetry (low-frequency information) and photometric stereo (high-frequency information). Both methods require nothing more than lights and a digital camera. [chapter 5](#) discussed the testing of photometric stereo approaches. An uncalibrated approach did not provide accurate estimates of the light positions, resulting in inaccurate normals. A relighting dome prototype was designed and built to allow for the calibration of the light positions. The relighting dome was designed to be operated in a vertical orientation, which allows the dome to capture any painted surface regardless of its size – it can be captured in a series of overlapping tiles. The light positions, when calibrated – using a reflective sphere along with the known geometry of the dome – improved the calculated normals. [chapter 6](#) compared the results of the data fusion approach with that of the Lucida – a triangulation-based scanner. The data fusion approach proved to be over two times as fast in the data capture; a number that could be increased with some optimizations to the system. The hardware required for the data fusion

approach was around 15% of the cost of the Lucida. The data processing, however, took longer with the data fusion approach, though most of it was semi-automatic. In the case study, the data fusion method produced good details but had errors where photometric stereo tiles overlapped, particularly in non-Lambertian areas. The data fusion produces more detail than the Lucida, but the edges of details are not as sharp. Softer edges occurred because the normal integration of the photometric stereo normals forced the gradient to conform a steep drop into a continuous surface. The integration constraint results in the largest errors of data between the two methods – on the order of 0.5 *mm* and less. The results of this research show that photogrammetry and photometric stereo data fusion is a good low-cost alternative to more expensive commercial approaches, particularly for surfaces that have few sharp edges.

7.2 FUTURE WORK

The research presented in this thesis is not complete. Some suggestions for future work are:

- As shown, the normal blending is flawed due to both specular effects and low-frequency shifts in the overlapping edges. Further tests should be done to determine an ideal overlap between adjacent photometric stereo tiles. A specular component should be introduced into the photometric stereo model to account for the non-Lambertian nature of many painted surfaces. With these adjustments, an improved normal blending algorithm like the ones introduced in Nehab et al., 2005 or Berkiten, Fan and Rusinkiewicz, 2014 should be implemented.
- The Frankot and Chellappa, 1988 algorithm has difficulties when integrating normals with a sharp fall off. Different normal integration algorithms should be tested to see if they can reconstruct ridged surfaces more accurately.
- The methodology was only tested on two surfaces. A wider variety of painted surfaces should be tested to determine the varietal limits of the data fusion approach. Multiple tests should also be performed on the same surface to determine the robustness and repeatability of the approach.
- Plane fitting methods should be introduced to allow direct metric comparisons between data fusion and other approaches.
- To add more value to conservators, the workflow should also be expanded to produce one large RTI of the whole painting – made up of the individual tiles.

Part V

APPENDIX

RELIGHTING DOME: DESIGN AND CONSTRUCTION

The relighting dome was designed and built, in collaboration with Jorge Cano, Enrique Esteban, and Quinner Baird of the Factum Foundation for Digital Technology in Conservation.

A.1 INTRODUCTION

The fundamental photometric stereo model assumes that the lights are infinitely far away. Though physically impossible, the far-light model can be approximated if all the lights are the same distance from the centre of capture – which is typically at a ratio of 3:1 between light distance and object size (Mudge et al., 2006). A similar distance constraint is assumed in near-light photometric stereo models; the problem is much simpler to solve without distance variation. This constraint led to the birth of the *relighting dome* – whether that be for photometric stereo or RTI. The dome is hemispherical with fixed lights on the inside and a camera pointing through the dome towards the open side (Figure A.1). With fixed – and potentially calibrated – light positions, a relighting dome can produce consistent results.

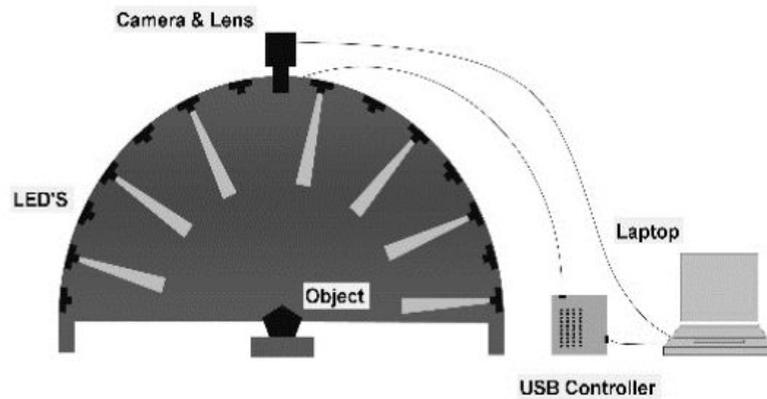


Figure A.1: An example of a photometric stereo dome. (Watteeuw et al., 2016)

A.2 DOME DESIGNS

The following are relighting domes that influenced the design of the prototype.

Palma et al., 2014 created a relighting dome with a one metre diameter for RTI constructed of an aluminum shell divided into four pieces (Figure A.2). The dome contains 116 6W LEDs and uses an Arduino to simultaneously turn on a light and release the camera shutter (Palma et al., 2014). Like many relighting domes, this dome is built for a downward configuration. There is no connection between the camera and dome that would allow for vertical capture.

Kinsman, 2016 introduced the design for a low cost, 3D printable RTI dome. The design uses 50 10 mm LED lights, an LED matrix to help control the lights, and an Arduino Uno microprocessor to turn the lights on and off and release the camera shutter. Kinsman has made the wiring design, 3D printable dome file, and Arduino code publicly available (Figure A.3). The dome, however, is designed to be downward facing and is meant for small, macro objects – so it fits within the

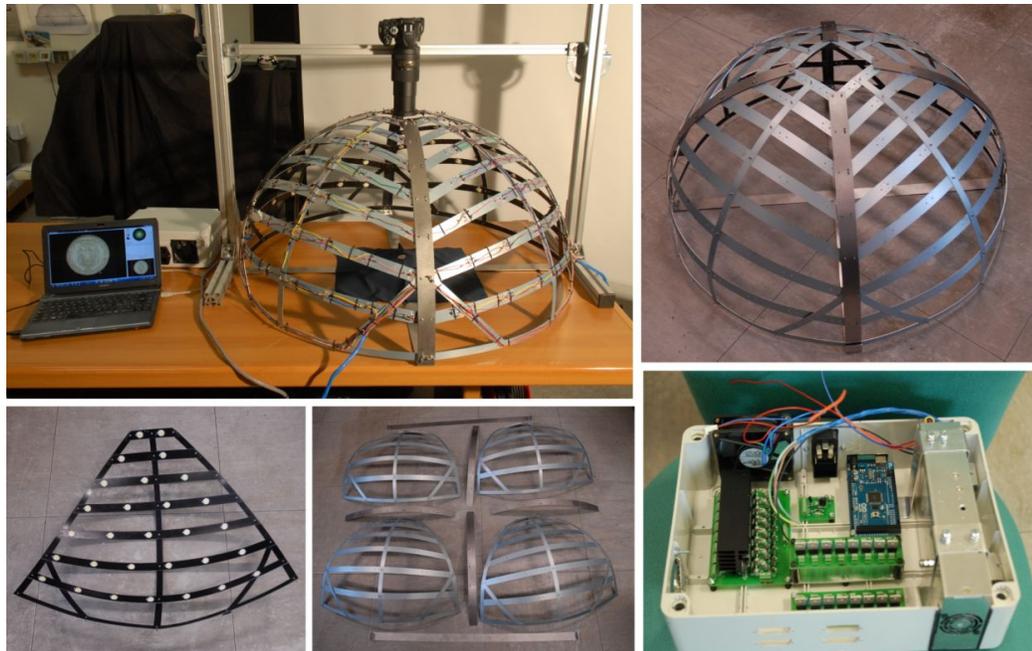


Figure A.2: The relighting dome from Palma et al., 2014.

bounds of a 3D printer.

Pawlowicz, 2016 introduced a portable, low-cost, open-design dome for RTI. The dome is larger than the Kinsman, 2016 dome but smaller than the Palma et al., 2014 one with a diameter of 30 cm. The dome is made using a pre-made hemisphere that is cut to accommodate the lights. The design uses 48 3W LEDs powered by an Arduino Mega, a Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) for each column, and an LED driver for each row in the form of a matrix. Pawlowicz, 2016 designed the dome so it could be used in both a downward and vertical facing positions (Figure A.4). The price of the dome was under \$800 USD, and Pawlowicz, 2016 estimated that a 100 cm dome would cost just under \$1200 USD.



(a) The 3D printed dome.



(b) The inside of the dome with the LEDs.

Figure A.3: A 3D printed relighting dome designed for macro RTI. (Kinsman, 2016)



(a) The 3W LEDs inside the dome.



(b) The dome in a vertical position to document a painting.



(c) The specularity from the RTI of a painting.

Figure A.4: The RTI dome designed by Pawlowicz, 2016.

A.3 PROTOTYPE DESIGN

As seen above, most relighting domes are designed to be operated in a downward-facing configuration. This approach limits the capture area to an area smaller than the size of the dome. A vertical facing dome, however, is more versatile; it can capture a large surface in multiple, overlapping sections. The dome limits only the size of one tile, rather than the size of the object.

The premise of a vertical relighting dome requires the dome and the camera to be attached in some way; this constrains the design to balance resolution and portability between the two. Even with a near-light model, the capture area for a relighting dome should not be larger than half the dome to avoid warping at the edges – though smaller would be better. A smaller relighting dome would be more portable, but it would result in a small capture area and thus longer capture times for large surfaces. If the camera used a lens with a long focal length – greater than 50 *mm* – it would have to be further away and require a large dome, or it would capture too small of an area per tile. Conversely, a wider focal length – less than 50 *mm* – would have a larger capture area, but would need a larger dome and is also subject to larger amounts of barrel distortion. A 50 *mm* lens balances the capture area, dome size, and resolution. With a 50 *mm* lens, the minimal focal distance between the surface and the camera sensor is usually between 40 and 50 *cm*. The [GSD formula \(Equation 2.3\)](#) can be used to calculate the lateral spatial resolution of the setup. At 55 *cm* from the surface – to allow for some minor axial movement – a 22-megapixel full-frame camera with a 50 *mm* lens could achieve a lateral spatial resolution of 69 μm : within the requirements. Many, newer [DSLRs](#) have higher megapixel counts and could achieve even higher resolutions.

A 55 *cm* sensor to surface distance would require a dome with a 50 *cm* radius. Some smaller relighting domes remain portable with solid construction. However, a dome of this size – like the dome built by Palma et al., 2014 – would be impractical to transport as a solid piece. To address this, the dome being developed for this prototype is designed to be broken down into smaller pieces. Because the goal is to document 2.5D surfaces, the dome does not need as many lights as some relighting domes – particularly those focused on RTI (Aure, O’Dowd and Padfield, 2017; Huang et al., 2016; Palma et al., 2014; Watteeuw et al., 2016). Photometric stereo problems are typically solved with fewer lights (Hasegawa et al., 2011; Logothetis, Mecca and Cipolla, 2017; Xie, Dai and Wang, 2015). Instead of quarter hemispheres, the dome prototype is divided into individual, detachable arms (columns at specific radial angles) with LEDs at rows (different angles from the surface along each arm). Like the relighting domes of Mudge et al., 2006, Artal-Isbrand, Klausmeyer and Murray, 2011, and Mytum and Peterson, 2018 the prototype dome has eight arms – one at every 45°. The arms are attached to a base plate that connects the dome to the camera and electronics and has bracing at the end to keep them stable – much like the dome in Figure A.5. The structure can be easily dismantled and transported.

With the radial angles designed, the row angles were next. Duffy et al., 2018 suggests lighting the surface for RTI at angles of 15°, 40°, and 65° (Figure 2.16). Macdonald, 2015 designed a lighting dome with lights at 20°, 40°, 60° and 80°. The dome prototype takes cues from both designs mentioned – the lower start angle of the Duffy et al., 2018 dome, and the consistent angles of the Macdonald, 2015 dome. This dome starts at a 15° angle for the first light and proceeds at 15° increments to favour the lights more towards the raking angles in order to pick up the low relief details of the painting: 15°, 30°, 45°, and 60° (Figure A.6). Each

dome arm ends at the 15° light to ensure that it does not touch the painted surface – it leaves more than 10 *cm* of clearance between the end of the dome and the painting. The drawing of the first design of the dome can be seen in [Figure A.7](#).



Figure A.5: The RTI dome from Mudge et al., 2006.

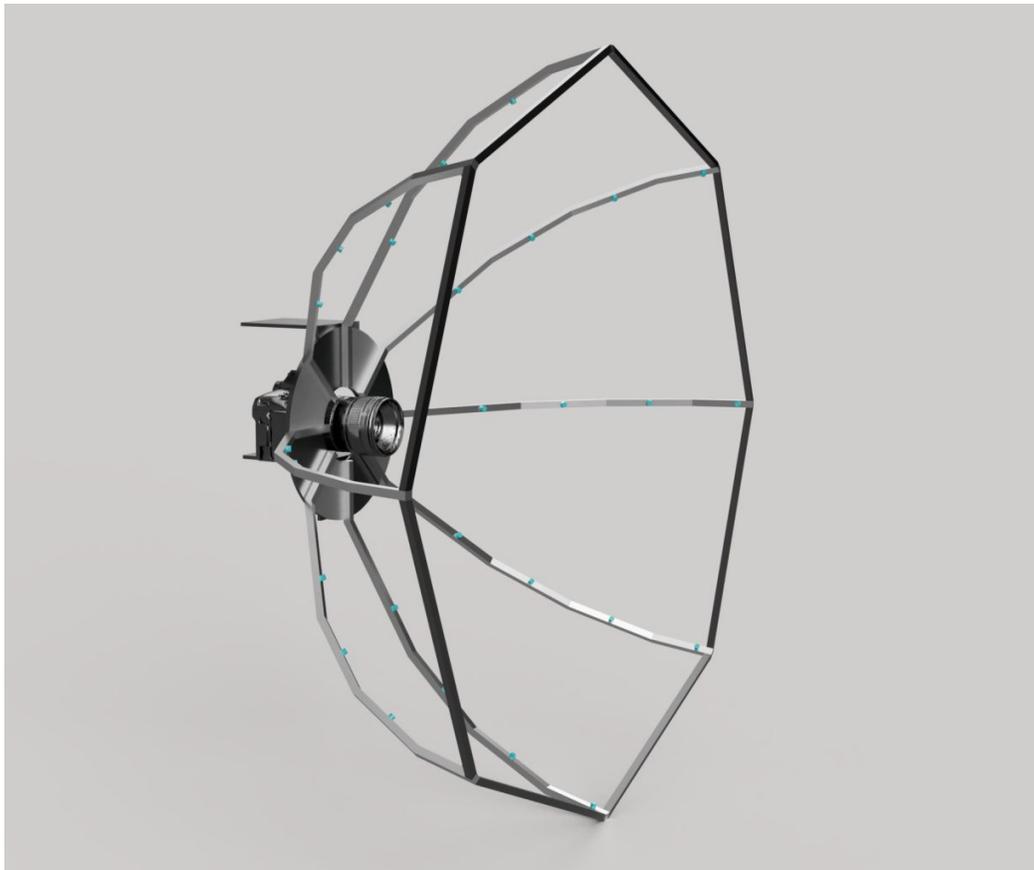


Figure A.6: The initial design of the relighting dome.

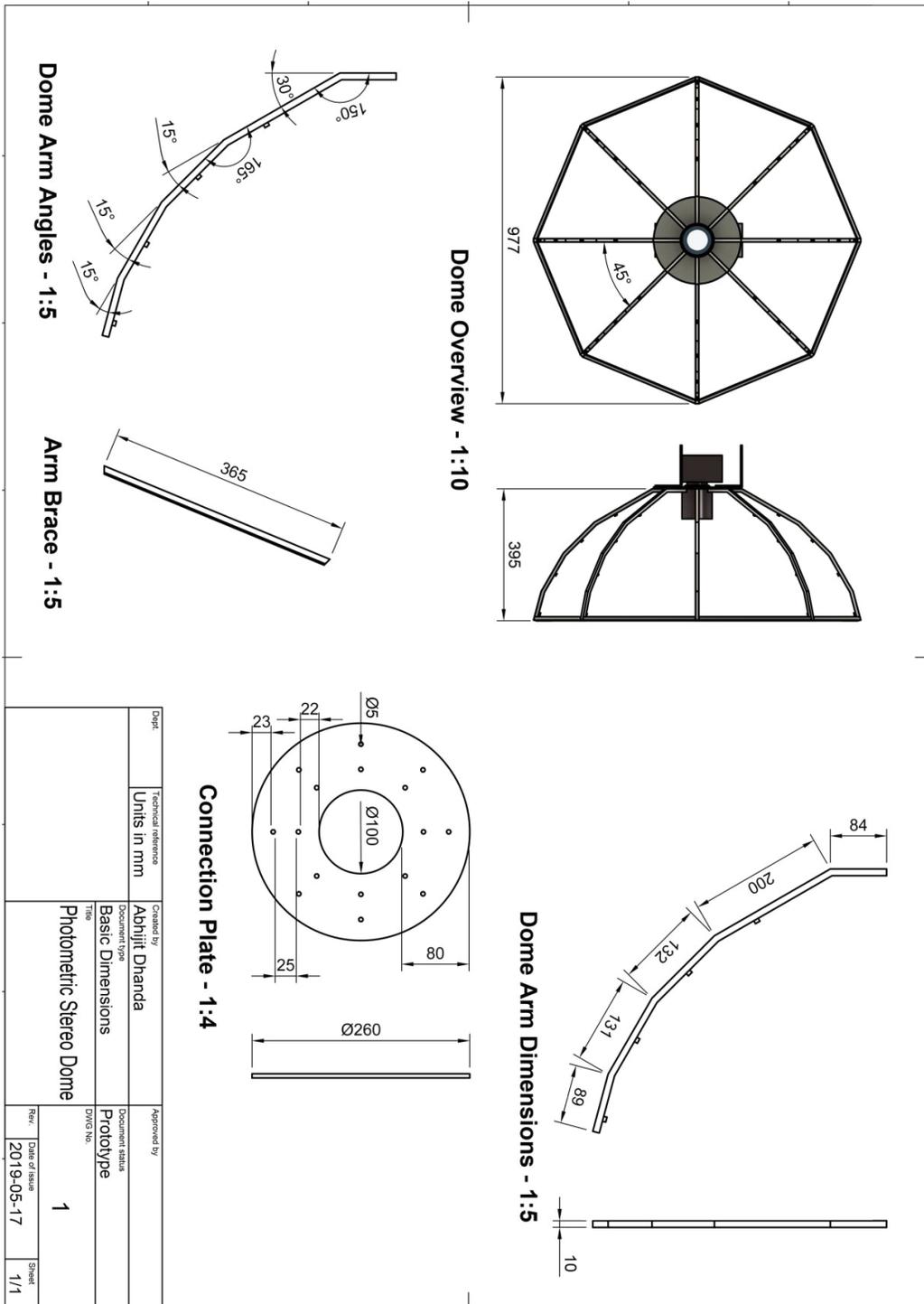


Figure A.7: The first design of the relighting dome prototype.

A.4 PROTOTYPE CONSTRUCTION

The relighting dome prototype was constructed using 10 *mm* aluminum extrusions. The extrusions are easy to cut and well suited for prototyping. Two custom mitre boxes was designed and 3D printed so that the sections of each arm could be cut to exact lengths (Figure A.8). Brackets for the aluminum extrusions were 3D printed because they could not be found at the required angles: 15° and 30°. Each connection used two brackets with four M3 screws per bracket (Figure A.9).



Figure A.8: The mitre boxes used to cut the aluminum extrusions.

Six different LED configurations were tested to determine which configuration would be suitable for the dome: a 10 *mm* LED, a 1W LED, a 3W LED, a 4-LED array, a single segment of an LED strip (with three LEDs on it), and a double segment of an LED strip (with six LEDs on it) (Figure A.10). Four of the LEDs were connected to a breadboard and wired to relays that were connected to an Arduino Uno. The LED strip segments were controlled by alligator clips connected to a 12V power supply. The arm and a camera were placed at the distances dictated by design: 55 *cm* from a white piece of paper with a small reflective sphere in the middle. This setup was used to test the brightness and spread of light for each LED at every tow angle along the arm.



(a) The 3D printed bracket.

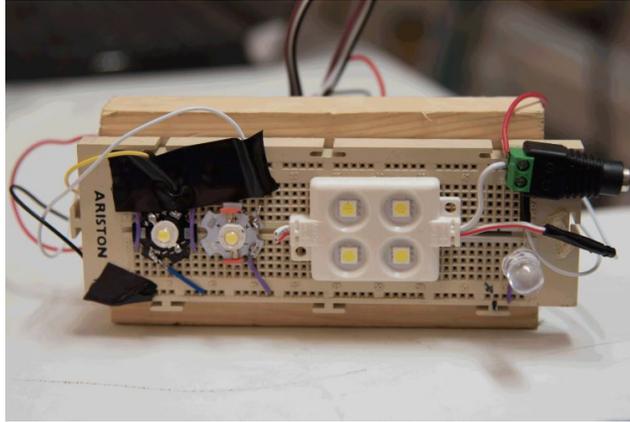


(b) The first constructed arm of the prototype.

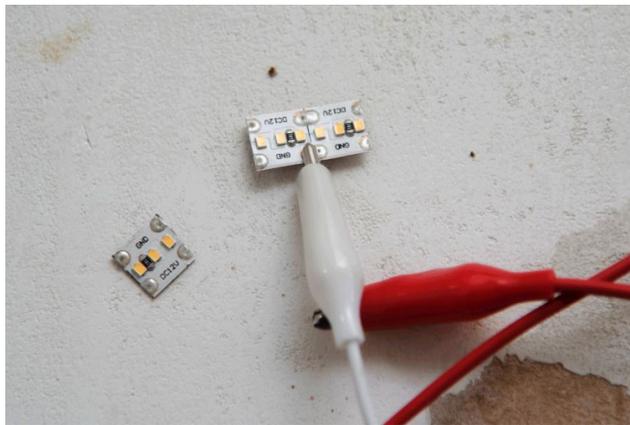
Figure A.9: One arm of the prototype relighting dome.

The first LED test was performed during the day, with ambient sunlight, but no direct light, coming through the window (Figure A.11). The camera lens was set to an aperture of $f8.0$ with an ISO of 100 and a shutter speed of $1/2$ second. It became apparent from this test that the LEDs did not produce enough light to have sufficient exposure and a fast enough shutter speed so that an image with no LED would be black – eliminating the effects of the sunlight. To get accurate normal calculations, the light from the dome should be the only light in the images. The first test made several things clear, the 10 mm LED did not have a wide enough spread of light for relighting purposes, the LED array was too large physically to use on the arms, and that the LED strip segments were flimsy – they could not be relied on to consistently light the surface.

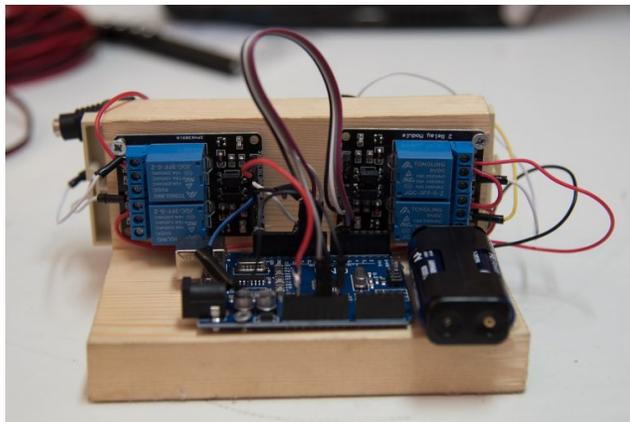
A second LED test was done at night to mimic the effect of a cover over the dome used to block out the ambient light. It was clear that this type of cover would be needed for capture in uncontrolled environments. [Figure A.12](#) shows the comparison between the 1W and 3W LEDs. It is clear that the 3W LED is stronger, but only just. A 3W LED with a higher power (greater than double the lumen output at around 500 *lm*) was selected for the dome prototype.



(a) The 3W LED (left), 1W LED (middle-left), 4-LED array (middle-right), and 10 mm LED (left).



(b) The tested segments of the LED strip.



(c) The relays and Arduino used to control the four LEDs on the breadboard.

Figure A.10: The LEDs tested.

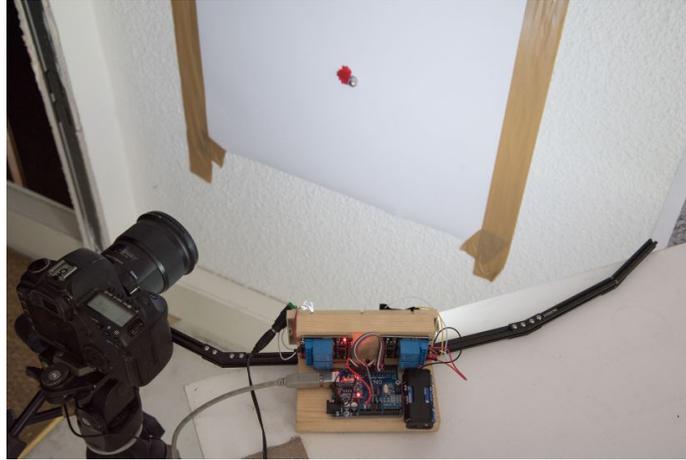


Figure A.11: The capture setup of the first LED test.



(a) The 1W LED tested at the row angles of 15° , 30° , 45° , and 60° .



(b) The 3W LED tested at the row angles of 15° , 30° , 45° , and 60° .

Figure A.12: The results of the second LED test.

Once the lights were specified, a second arm was built to test the circuit with a two column by four row matrix. The negative side of the four LEDs on each arm were wired in series to make the column, and the positive sides were wired in parallel to make the rows – a total of five wires from each arm. The test circuit, made on a breadboard, can be seen in [Figure A.13](#).

The remaining six arms were built following the circuit test. Each arm segment was wrapped in heat shrink tubing to protect the wires and keep them close to

the arm. Each arm was then connected to the circular plate – laser cut in 5 *mm* acrylic – using two M3 screws (Figure A.14).

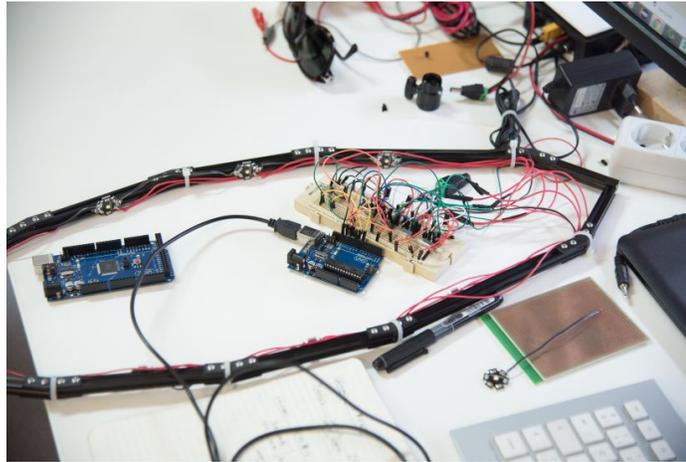


Figure A.13: The two arms connected to the test circuit.

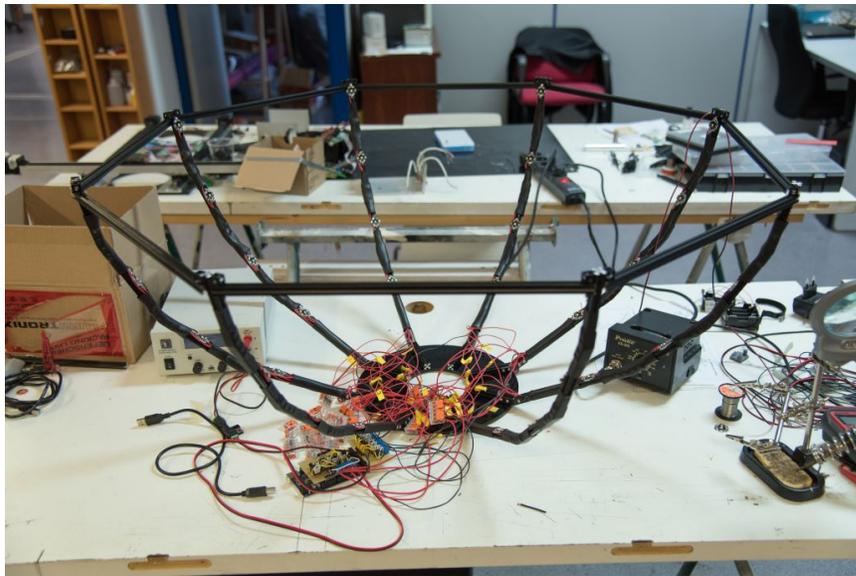
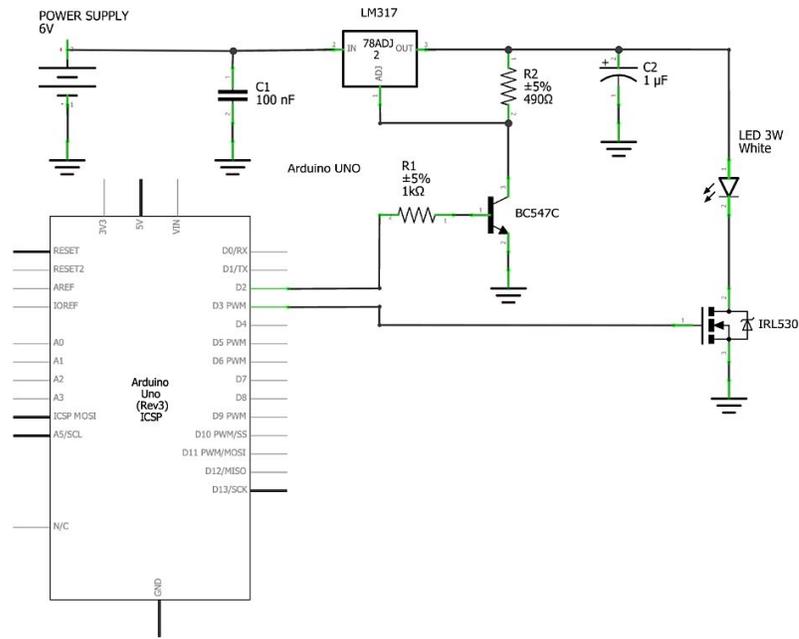


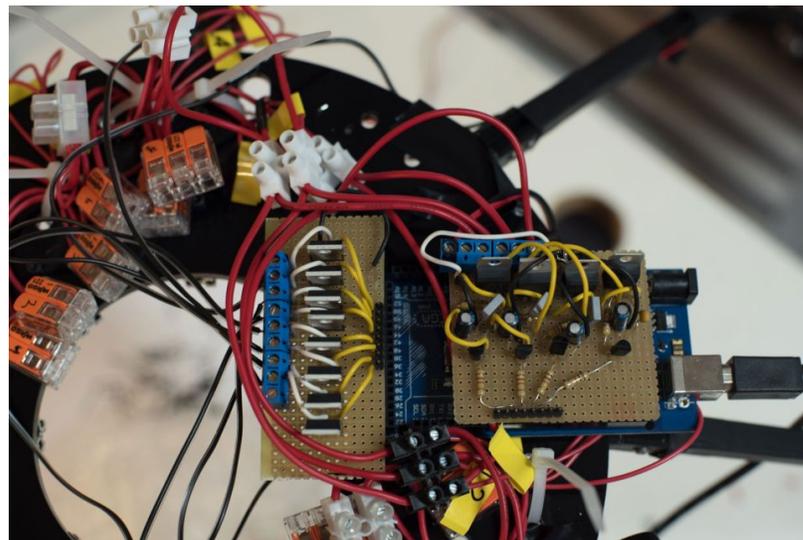
Figure A.14: The arms connected to the circular plate.

Two Arduino shields were hand soldered based on the circuit design (Figure A.15). The boards contained all of the components to control the matrix but were still missing an optocoupler (to control the camera shutter), a button (to operate it

without a laptop), and a battery set as a power supply. The dome prototype was also still missing the L-bracket (to connect it to the camera and tripod), and an electronics box to protect the Arduino and the shields. However, there were enough components to test the dome in a downward-facing position by using a power supply for the LEDs and a laptop to control the Arduino and camera shutter ([Figure A.16](#)).



(a) The schematic of the circuit using a 1x1 LED matrix. The circuit was designed by Jorge Cano and Enrique Esteban.



(b) The hand soldered circuit board.

Figure A.15: The circuit board design and construction.

The L-bracket (Figure A.17) allows the dome to be oriented in a vertical position. This was tested in two different configurations: one on a tripod, and one on the aluminum extrusion rails used by the Lucida (Figure A.18). The cost of the

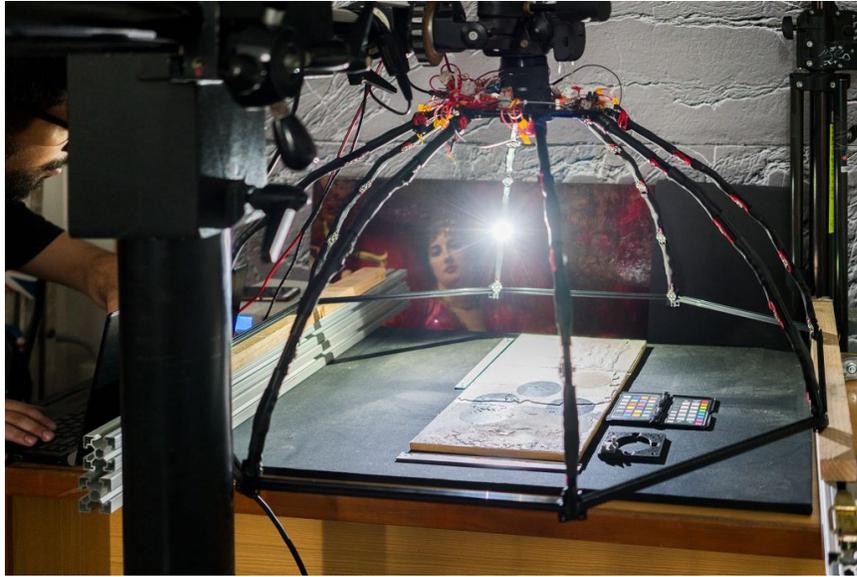


Figure A.16: The dome tested in its downward position. A large tripod was used to cantilever over the dome and through the centre. © Otto Lowe for the Factum Foundation for Digital Technology in Conservation



Figure A.17: The L-Bracket.

relighting dome prototype is shown in [Figure A.19](#). The price is equivalent to around \$1200 CAD.



(a) The relighting dome on a tripod.

(b) The relighting dome on aluminum rails.

Figure A.18: The relighting dome in vertical configurations.

Cost of Relighting Dome Prototype		
10x10mm Aluminum Extrusions	€	100.00
T-slots and screws	€	250.00
LEDs	€	50.00
LED Control Circuit (Arduino, MOSFETs, etc.)	€	50.00
Circular Plate	€	40.00
3D Printed Brackets	€	-
L Bracket	€	170.00
Black Sphere	€	8.00
Wiring, Heatshrink Tubing, Connectors	€	50.00
Total	€	718.00
Total + Tax + Shipping	€	883.14

Figure A.19: The cost of the relighting dome prototype.

BIBLIOGRAPHY

- Abate, D., Menna, F., Remondino, F. and Gattari, M.G., 2014. 3D Painting documentation: Evaluation of conservation conditions with 3D imaging and ranging techniques. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 40(5), pp.1–8. Available from: <https://doi.org/10.5194/isprsarchives-XL-5-1-2014>.
- Antensteiner, D., Štolc, S. and Pock, T., 2018. A Review of depth and normal fusion algorithms. *Sensors (Switzerland)*, 18(2), pp.1–24. Available from: <https://doi.org/10.3390/s18020431>.
- Artal-Isbrand, P., Klausmeyer, P. and Murray, W., 2011. An Evaluation of Decorative Techniques on a Red-Figure Attic Vase from the Worcester Art Museum using Reflectance Transformation Imaging (RTI) and Confocal Microscopy with a Special Focus on the “Relief Line”. *MRS Proceedings [Online]*, 1319, p.31. Available from: <https://doi.org/10.1557/opl.2011.793>.
- Aure, X., O’Dowd, P. and Padfield, J., 2017. Generating 3D Models of Paintings through the Combination of 2D, 3D and RTI Data. *EVA London 2017*, pp.25–32. Available from: <https://doi.org/10.14236/ewic/eva2017.4>.
- Barbetti, I., Felici, A., Magrini, D., Manganelli Del Fa, R. and Riminesi, C., 2013. Ultra Close-Range Photogrammetry to Assess the Roughness of the Wall Painting Surfaces After Cleaning Treatments. *International Journal of Conservation Science*, 4, pp.525–534.

- Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L., 2008. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding* [Online], 110(3), pp.346–359. Available from: <https://doi.org/10.1016/j.cviu.2007.09.014>.
- Bennett, T., 2015. Photogrammetry and Transmitted Infrared Imaging to Document the Support of a 19th C. British Landscape Painting. *COSCH e-Bulletin*, No. 2.
- Berkiten, S, Fan, X and Rusinkiewicz, S, 2014. Merge2-3D: Combining Multiple Normal Maps with 3D Surfaces. *2014 2nd International Conference on 3D Vision*, 1, pp.440–447. Available from: <https://doi.org/10.1109/3DV.2014.22>.
- Blais, F., Cournoyer, L., Beraldin, J.-A. and Picard, M., 2008. 3D imaging from theory to practice: the Mona Lisa story. *Current Developments in Lens Design and Optical Engineering IX*, 7060(May 2014), pp.1–10. Available from: <https://doi.org/10.1117/12.795456>.
- Boehler, W., Heinz, G. and Marbs, A., 2001. The potential of non- contact close range laser scanner for cultural heritage recording. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, XXXIV-5/C7.
- Commons, W., 2015. *File:Photometric stereo.png* [Online]. Available from: https://commons.wikimedia.org/wiki/File:Photometric_stereo.png.
- Cossairt, O., Huang, X., Matsuda, N., Stratis, H., Broadway, M., Tumblin, J., Berman, G., Doehne, E., Katsaggelos, A. and Walton, M., 2015. Surface shape studies of the art of Paul Gauguin. *2015 Digital Heritage International Congress, Digital Heritage 2015*, pp.13–20. Available from: <https://doi.org/10.1109/DigitalHeritage.2015.7419447>.

- Cultural Heritage Imaging, 2015. *RTIBuilder* [Online]. Available from: http://culturalheritageimaging.org/What_We_Offer/.
- Duffy, S., Bryan, P., Earl, G., Beale, G., Pagi, H. and Kotoula, E., 2018. *Multi-light Imaging for Cultural Heritage*. Swindon: Historic England, p.66.
- Eppich, R. and Hadzic, L., 2013. Heritage Conservation–Aligning Technologies. *5th International Conference on Hazards and Modern Heritage* [Online], pp.1–20. Available from: https://www.academia.edu/3445619/Heritage_Conservation_-_Aligning_Technologies_CICOP_Sarajevo_2013.
- Factum Foundation, 2015. *Lucida 3D Scanner: Selected Projects* [Online], p.87. Available from: http://www.factum-arte.com/resources/files/ff/manuals/05_printer_projects_i_lucida_booklet_2018_factum_foundation_2_pdf.
- Frankot, R.T. and Chellappa, R., 1988. A Method for Enforcing Integrability in Shape from Shading Algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4), pp.439–451. Available from: <https://doi.org/10.1109/34.3909>.
- Goldman, D.B., Curless, B., Hertzmann, A. and Seitz, S.M., 2010. Shape and spatially-varying BRDFs from photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(6), pp.1060–1071. Available from: <https://doi.org/10.1109/TPAMI.2009.102>.
- Grochulla, M. and Thormählen, T., 2015. Combining Photometric Normals and Multi-view Stereo for 3D Reconstruction. *Proceedings of the 12th European Conference on Visual Media Production*, pp.1–8. Available from: <https://doi.org/10.1145/2824840.2824846>.

- Grussenmeyer, P. and Landes, T., 2016. Basics of Range-Based Modelling Techniques in Cultural Heritage 3D Recording. In: Stylianidis, E. and Remondino, F. eds. *3d recording, documentation and management of cultural heritage*. Dunbeath: Whittles Publishing. Chap. 6, pp.305–368.
- Guidi, G., Atzeni, C. and Lazzari, S., 2003. 3D optical scanning diagnostics for Leonardo da Vinci's "adorazione dei Magi" conservation. *Proceedings of International Conference on 3-D Digital Imaging and Modeling, 3DIM*, 2003-Janua(January 2015), pp.110–115. Available from: <https://doi.org/10.1109/IM.2003.1240239>.
- Han, J.-Y., Guo, J. and Chou, J.-Y., 2011. A Direct Determination of the Orientation Parameters in the Collinearity Equations. *IEEE Geoscience and Remote Sensing Letters* [Online], 8(2), pp.313–316. Available from: <https://doi.org/10.1109/LGRS.2010.2066955>.
- Hartmann, W., Havlena, M. and Schindler, K., 2016. Recent developments in large-scale tie-point matching. *ISPRS Journal of Photogrammetry and Remote Sensing* [Online], 115, pp.47–62. Available from: <https://doi.org/10.1016/j.isprsjprs.2015.09.005>.
- Hasegawa, T., Tsumura, N., Nakaguchi, T. and Iino, K., 2011. Photometric approach to surface reconstruction of artist paintings. *Journal of Electronic Imaging*, 20(1), p.11. Available from: <https://doi.org/10.1117/1.3533329>.
- Heber, S., 2015. *Variational Shape from Unconventional Imaging Devices*. PhD thesis. Graz University of Technology.
- Herbort, S. and Wöhler, C., 2011. An introduction to image-based 3D surface reconstruction and a survey of photometric stereo methods. *3D Research*, 2(3), pp.1–17. Available from: [https://doi.org/10.1007/3DRes.03\(2011\)4](https://doi.org/10.1007/3DRes.03(2011)4).

- Historic England, 2017. *Photogrammetric Applications for Cultural Heritage*. Swindon: Historic England, pp.1–124. Available from: <https://doi.org/10.1212/01.WNL.0000132885.83350.45>.
- Huang, X., Walton, M., Bearman, G. and Cossairt, O., 2016. Near light correction for image relighting and 3D shape recovery. *Digital Heritage*, pp.215–222. Available from: <https://doi.org/10.1109/digitalheritage.2015.7413874>.
- Kinsman, T., 2016. An Easy to Build Reflectance Transformation Imaging (RTI) System. *Journal of Biocommunication*, 40(1), pp.10–14. Available from: <https://doi.org/10.5210/jbc.v40i1.6625>.
- Lahanier, C. et al., 2005. Two-dimensional multi-spectral digitization and three-dimensional modelling of easel paintings. *14th triennial meeting the hague, rome: international council of museums (icom), 2005*, May 2014. Rome: International Council of Museums (ICOM), p.19.
- Lerma Garcia, J., Van Genechten, B., Heine, E. and Quintero, M.S., 2008. *3D RiskMapping - Theory and Practice on Terrestrial Laser Scanning*. Universidad Politecnica De Valencia, (technical report June), pp.1–241. Available from: http://jllerma.webs.upv.es/pdfs/Leonardo_Tutorial_Final_vers5_ENGLISH.pdf.
- Logothetis, F, Mecca, R, Quéau, Y and Cipolla, R, 2016. Near-Field Photometric Stereo in Ambient Light. *British Machine Vision Conference (BMVC)* [Online], pp.1–12. Available from: <https://doi.org/10.17863/CAM.4580>.
- Logothetis, F, Mecca, R. and Cipolla, R., 2017. Semi-Calibrated Near Field Photometric Stereo. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [Online]. IEEE, pp.4521–4530. Available from: <https://doi.org/10.1109/CVPR.2017.481>.

- Lowe, D., 1999. Object recognition from local scale-invariant features. *Proceedings of the seventh IEEE international conference on computer vision* [Online]. IEEE, pp.1150–1157. Available from: <https://doi.org/10.1109/ICCV.1999.790410>.
- Luhmann, T., Robson, S., Kyle, S. and Harley, I., 2011. *Close range photogrammetry sensors*. Dunbeath: Whittles Publishing.
- MacDonald, L., 2011. Visualising an Egyptian artefact in 3D: Comparing RTI with laser scanning. *EVA London 2011. Electronic Visualisation and the Arts. London, UK, 6-8 July 11* [Online], pp.155–162. Available from: http://ewic.bcs.org/upload/pdf/ewic_ev11_s8paper4.pdf.
- MacDonald, L.W., 2015. Surface reconstruction from photometric normals with reference height measurements. *Optics for Arts, Architecture, and Archaeology V*, 9527(June 2015), pp.1–16. Available from: <https://doi.org/10.1117/12.2184980>.
- MacDonald, L., Toschi, I., Nocerino, E., Hess, M., Remondino, F. and Robson, S., 2016. Accuracy Of 3D reconstruction in an illumination dome. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 41(July), pp.69–76. Available from: <https://doi.org/10.5194/isprsarchives-XLI-B5-69-2016>.
- Macdonald, L.W., 2015. *Realistic Visualisation of Cultural Heritage Objects*. PhD thesis. University College London, p.316.
- Malzbender, T., Malzbender, T., Gelb, D., Wolters, H. and Zuckerman, B., 2000. Enhancement of Shape Perception by Surface Reflectance Transformation [Online]. [Online]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.3.1834>.

- Moitinho, V., 2018. Case Study 4: Cantabria – rock art in El Mirón Cave. *3d laser scanning for heritage: advice and guidance on the use of laser scanning in archaeology and architecture*. 3rd. Swindon: Historic England. Chap. 4, pp.61–63.
- Mudge, M., Malzbender, T., Schroer, C. and Lum, M., 2006. New Reflection Transformation Imaging Methods for Rock Art and Multiple-Viewpoint Display. *The 7th International Symposium on Virtual Reality Archaeology and Cultural Heritage VAST2006*, 6, (January), pp.195–202. Available from: <https://doi.org/10.2312/VAST/VAST06/195-202>.
- Mudge, M. et al., 2008. Image-Based Empirical Information Acquisition, Scientific Reliability , and Long-Term Digital Preservation for the Natural Sciences and Cultural Heritage. *Eurographics* [Online], (January). Available from: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MUDGE/EG-mudge-tutorial-notes-final.pdf.
- Mytum, H. and Peterson, J.R., 2018. The Application of Reflectance Transformation Imaging (RTI) in Historical Archaeology. *Historical Archaeology*, 52(2), pp.489–503. Available from: <https://doi.org/10.1007/s41636-018-0107-x>.
- Nehab, D., Rusinkiewicz, S., Davis, J. and Ramamoorthi, R., 2005. Efficiently combining positions and normals for precise 3D geometry. *ACM SIGGRAPH 2005 Papers on - SIGGRAPH '05* [Online], p.536. Available from: <https://doi.org/10.1145/1186822.1073226>.
- Nocerino, E., Rieke-Zapp, D.H., Trinkl, E., Rosenbauer, R., Farella, E.M., Morabito, D. and Remondino, F., 2018. Mapping VIS and UVL imagery on 3D geometry for non-invasive, non-contact analysis of a vase. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS*

Archives, 42(2), pp.773–780. Available from: <https://doi.org/10.5194/isprs-archives-XLII-2-773-2018>.

Padfield, J., Saunders, D. and Malzbender, T., 2005. Polynomial texture mapping: a new tool for examining the surface of paintings. *14th Triennial Meeting The Hague Preprints* [Online], 1, pp.504–510. Available from: http://shiftleft.com/mirrors/www.hpl.hp.com/personal/Tom_Malzbender/papers/icom14_079.pdf.

Palma, G., National, I., Scopigno, R. and National, I., 2014. Storytelling of a Coin Collection by Means of RTI Images : the Case of the Simoneschi Collection in Palazzo Blu. In: Proctor, N. and Cherry, R. eds. *Museums and the web 2013* [Online]. Available from: <https://mwf2014.museumsandtheweb.com/paper/storytelling-of-a-coin-collection-by-means-of-rti-images-the-case-of-the-simoneschi-collection-in-palazzo-blu/>.

Pamart, A., Guillon, O., Faraci, S., Gattet, E., Genevois, M., Vallet, J.M. and De Luca, L., 2017. Multispectral photogrammetric data acquisition and processing for wall paintings studies. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(2W3), pp.559–566. Available from: <https://doi.org/10.5194/isprs-archives-XLII-2-W3-559-2017>.

Park, J., Sinha, S.N., Matsushita, Y., Tai, Y.-W. and Kweon, I.S., 2017. Robust Multiview Photometric Stereo Using Planar Mesh Parameterization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* [Online], 39(8), pp.1591–1604. Available from: <https://doi.org/10.1109/TPAMI.2016.2608944>.

Pawlowicz, L., 2016. *Affordable Reflectance Transformation Imaging Dome* [Online]. Available from: <https://artid.readthedocs.io/en/latest/index.html>.

- Pelagotti, A., Del Mastio, A. and Razonale, A.V., 2007. Active and passive sensors for art works analysis and investigations. *Videometrics IX*, 6491(January 2007), 64910R. Available from: <https://doi.org/10.1117/12.703811>.
- Quéau, Y., Durou, J.D. and Aujol, J.F., 2018. Normal Integration: A Survey. *Journal of Mathematical Imaging and Vision*, 60(4), pp.576–593. Available from: <https://doi.org/10.1007/s10851-017-0773-x>.
- Quéau, Y., Wu, T., Lauze, F., Durou, J.D. and Cremers, D., 2017. A non-convex variational approach to photometric stereo under inaccurate lighting. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-Janua, pp.350–359. Available from: <https://doi.org/10.1109/CVPR.2017.45>.
- Quéau, Y., Durix, B., Wu, T., Cremers, D., Lauze, F. and Durou, J.D., 2018. LED-Based Photometric Stereo: Modeling, Calibration and Numerical Solution. *Journal of Mathematical Imaging and Vision*, 60(3), pp.313–340. Available from: <https://doi.org/10.1007/s10851-017-0761-1>.
- Remondino, F., Rizzi, A., Barazzetti, L., Scaioni, M., Fassi, F., Brumana, R. and Pelagotti, A., 2011. Review of geometric and radiometric analyses of paintings. *Photogrammetric Record*, 26(136), pp.439–461. Available from: <https://doi.org/10.1111/j.1477-9730.2011.00664.x>.
- Robson, S., Bucklow, S., Woodhouse, N. and Papadakia, H., 2004. Periodic photogrammetric monitoring and surface reconstruction of a historical wood panel painting for restoration purposes. *International archives of photogrammetry and remote sensing* [Online], 35(B5), pp.395–401. Available from: <http://www.cartesia.org/geodoc/isprs2004/comm5/papers/585.pdf>.

- Sabzevari, R., Del Bue, A. and Murino, V., 2011. Structure from motion and photometric stereo for dense 3D shape recovery. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6978 LNCS(PART 1), pp.660–669. Available from: https://doi.org/10.1007/978-3-642-24085-0_67.
- Salvant, J., Walton, M., Kronkright, D., Yeh, C.-K., Li, F., Cossairt, O. and Katsaggelos, A.K., 2019. Photometric Stereo by UV-Induced Fluorescence to Detect Protrusions on Georgia O’Keeffe’s Paintings. In: Casadio, F., Keune, K., Noble, P., Van Loon, A., Hendriks, E., Centeno, S. and Osmond, G. eds. *Metal soaps in art* [Online]. Springer, Cham, pp.375–391. Available from: https://doi.org/10.1007/978-3-319-90617-1_22.
- Shi, B., Inose, K., Matsushita, Y., Tan, P., Yeung, S.K. and Ikeuchi, K., 2015. Photometric stereo using internet images. *Proceedings - 2014 International Conference on 3D Vision, 3DV 2014*, pp.361–368. Available from: <https://doi.org/10.1109/3DV.2014.9>.
- Stylianidis, E., Georgopoulos, A. and Remondino, F., 2016. Basics of Image-Based Modelling Techniques in Cultural Heritage 3D Recording. In: Stylianidis, E. and Remondino, F. eds. *3d recording, documentation and management of cultural heritage*. Dunbeath: Whittles Publishing. Chap. 5, pp.253–304.
- Tan, P., Lin, S. and Long Quan, 2008. Subpixel Photometric Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence* [Online], 30(8), pp.1460–1471. Available from: <https://doi.org/10.1109/TPAMI.2007.70789>.
- Tominaga, S., Ujike, H. and Horiuchi, T., 2010. Surface reconstruction of oil paintings for digital archiving. *Proceedings of the IEEE Southwest Symposium on Im-*

- age Analysis and Interpretation*, (January), pp.173–176. Available from: <https://doi.org/10.1109/SSIAI.2010.5483891>.
- Trucco, E. and Verri, A., 1998. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, p.343.
- Verhoeven, G., Karel, W., Štuhec, S., Doneus, M., Trinks, I. and Pfeifer, N., 2015. Mind Your Grey Tones – Examining The Influence of Decolourization Methods on Interest Point Extraction and Matching for Architectural Image-Based Modelling. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* [Online], XL-5/W4, pp.307–314. Available from: <https://doi.org/10.5194/isprsarchives-XL-5-W4-307-2015>.
- Waldhausl, P. and Ogleby, C., 1994. 3X3 Rules for Simple Photogrammetric Documentation of Architecture. *International Archives of Photogrammetry and Remote Sensing*, XXX(Part 5), pp.426–429.
- Watteeuw, L., Hameeuw, H., Vandermeulen, B., Van der Perre, A., Boschloos, V., Delvaux, L., Proesmans, M., Van Bos, M. and Van Gool, L., 2016. Light, shadows and surface characteristics: the multispectral Portable Light Dome. *Applied Physics A: Materials Science and Processing* Watteeuw, L., Hameeuw, H., Vandermeulen, B., Van der Perre, A., Boschloos, V., Delvaux, L., Proesmans, M., Van Bos, M., Van Gool, L., 2016. *Light, shadows and surface characteristics: the multispectral Po*, 122(11), pp.1–7. Available from: <https://doi.org/10.1007/s00339-016-0499-4>.
- Webster, J.G., Bell, T., Li, B. and Zhang, S., 2016. Structured Light Techniques and Applications. *Wiley Encyclopedia of Electrical and Electronics Engineering*, (8), pp.1–24. Available from: <https://doi.org/10.1002/047134608x.w8298>.

- Woodham, R.J., 1980. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1), pp.139–144.
- Wu, L., Ganesh, A., Shi, B., Matsushita, Y., Wang, Y. and Ma, Y., 2011. Robust Photometric Stereo via Low-Rank Matrix Completion and Recovery. In: Kimmel R., Klette R., S.A. ed. *Computer vision – accv 2010* [Online]. Lecture No. Springer, pp.703–717. Available from: https://doi.org/10.1007/978-3-642-19318-7_55.
- Xie, W., Dai, C. and Wang, C.C., 2015. Photometric stereo with near point lighting: A solution by mesh deformation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 07-12-June, pp.4585–4593. Available from: <https://doi.org/10.1109/CVPR.2015.7299089>.
- Yeh, C.-K., Matsuda, N., Huang, X. and Li, F., 2016. A Streamlined Photometric Stereo Framework. *European Conference on Computer Vision*, 1, pp.738–752. Available from: <https://doi.org/10.1007/978-3-319-46604-0>.
- Yeh, C.K., Li, F., Pastorelli, G., Walton, M., Katsaggelos, A.K. and Cossairt, O., 2017. Shape-from-shifting: Uncalibrated photometric stereo with a mobile device. *Proceedings - 13th IEEE International Conference on eScience, eScience 2017*, pp.551–558. Available from: <https://doi.org/10.1109/eScience.2017.89>.
- Zaman, T., 2013. *Development of a Topographic Imaging Device: for the Near-Planar Surfaces of Paintings* [Online]. PhD thesis. Delft University of Technology, p.122. Available from: <https://repository.tudelft.nl/islandora/object/uuid:bd71a192-eea8-4f90-8778-b18f86cac79c?collection=research>.
- Zhang, S., 2010. Recent progresses on real-time 3D shape measurement using digital fringe projection techniques. *Optics and Lasers in Engineering* [Online],

48(2), pp.149–158. Available from: <https://doi.org/10.1016/j.optlaseng.2009.03.008>.

Zhou, Z., Wu, Z. and Tan, P., 2013. Multi-view photometric stereo with spatially varying isotropic materials. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.1482–1489. Available from: <https://doi.org/10.1109/CVPR.2013.195>.