

Realistic Visualisation of Cultural Heritage Objects

Lindsay William MacDonald

3DIMPact Research Group

Department of Civil, Environmental and Geomatic Engineering

University College London

Supervisor:

Professor Stuart Robson

Thesis for the Degree of Doctor of Philosophy (Ph.D.)

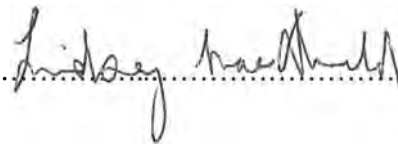
October 2014

Revised April 2015



Declaration of Authorship

I, Lindsay William MacDonald, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed.....

DateOctober 2014.....

Acknowledgements

This has been a long journey, after eight years of study (part-time). The research was informed by my two careers, the first for 20 years in industrial R&D and the second for 15 years in academia. I acknowledge the benign and liberal spirit of Jeremy Bentham, who permeates the whole institution of UCL and enlivens inter-disciplinary collaboration. I am deeply indebted to my supervisor, Professor Stuart Robson, first for taking me on as a mature student and then for his infinite patience, encouragement and friendly guidance throughout. I have also benefitted greatly from discussions, interactions and practical assistance from other students in my cohort, notably Mona Hess, John Hindmarch and Ali Hosseiniaveh Ahmadabadian. Most important of all has been the unwavering support from my wife Sandra, without whose positive attitude this extensive project would probably never have been completed.

Set in 11-point Gill Sans MT

Abstract

Digital imaging finds its perfect application in the domain of cultural heritage, enabling objects to be captured, rendered and displayed in many ways. It facilitates public access to objects that would otherwise be too rare, valuable or fragile to be widely seen, and it enables conservators and palæographers literally to see the objects in a new light.

A key issue for reproductions, both digital and physical, of cultural heritage objects is how realistic they should be. There are broadly two modes of reproduction – accurate and pleasing – and there are deep implications of this dichotomy for the digitising and rendering of objects in museum collections. For documentation and recording purposes, metric accuracy is important, and ideally the archival digital dataset should contain all the information of every aspect of the actual object. For presenting the object through digital media, such as displays in computers and mobile devices, however, the user experience depends on ease of access, interactivity, quality of annotation, as well as an attractive appearance. What matters for realistic visualisation is to convey the illusion of looking at a real 3D object, rather than the accuracy of the underlying representation.

This research programme was motivated by a desire to understand the limits of quality that could be achieved in the digitisation of 3D objects and their visual rendering on displays. Attributes of interest included colour, tonal range, surface finish, spatial resolution and dimensionality. Much of the investigation stemmed from photography in a hemispherical dome, enabling a set of 64 photographic images of an object to be captured in perfect pixel register, with each image illuminated from a different direction. This representation turns out to be much richer than a single 2D image, because it contains information at each point about both the 3D shape of the surface (gradient and local curvature) and the directionality of reflectance (gloss and specularities). So it enables not only interactive visualisation through viewer software, giving the illusion of 3D, but also the reconstruction of an actual 3D surface and highly realistic rendering of a wide range of materials.

The following seven outcomes of the research are claimed as novel and therefore as representing contributions to knowledge in the field:

- A method for determining the geometry of an illumination dome;
- An adaptive method for finding surface normals by bounded regression;
- Generating 3D surfaces from photometric stereo;
- Relationship between surface normals and specular angles;
- Modelling surface specularities by a modified Lorentzian function;
- Determining the optimal wavelengths of colour laser scanners;
- Characterising colour devices by synthetic reflectance spectra.

Twenty-one papers relevant to the topic have been written and presented at international conferences (all published in the conference proceedings) since 2009, during the course of the research leading to this thesis.

Ithaca

When you start on your journey to Ithaca,
then pray that the road is long,
full of adventure, full of knowledge.
Do not fear the Lestrygonians
and the Cyclopes and the angry Poseidon.
You will never meet such as these on your path,
if your thoughts remain lofty, if a fine
emotion touches your body and your spirit.
You will never meet the Lestrygonians,
the Cyclopes and the fierce Poseidon,
if you do not carry them within your soul,
if your soul does not raise them up before you.

Then pray that the road is long.
That the summer mornings are many,
that you will enter ports seen for the first time
with such pleasure, with such joy!
Stop at Phoenician markets,
and purchase fine merchandise,
mother-of-pearl and corals, amber and ebony,
and pleasurable perfumes of all kinds,
buy as many pleasurable perfumes as you can;
visit hosts of Egyptian cities,
to learn and learn from those who have knowledge.

Always keep Ithaca fixed in your mind.
To arrive there is your ultimate goal.
But do not hurry the voyage at all.
It is better to let it last for long years;
and even to anchor at the isle when you are old,
rich with all that you have gained on the way,
not expecting that Ithaca will offer you riches.
Ithaca has given you the beautiful voyage.
Without her you would never have taken the road.
But she has nothing more to give you.
And if you find her poor, Ithaca has not defrauded you.
With the great wisdom you have gained, with so much experience,
you must surely have understood by then what Ithacas mean.

– C.P. Cavafy (1911)
Collected Poems, tr. Rae Dalven

Contents

Abstract

Contents

Chapter 1 – Introduction	1
1.1 Objectives of Research	2
1.2 Context of Research Programme	2
1.3 Contributions to Knowledge	3
1.4 Organisation of this Thesis	4
1.5 Related Publications	4
Chapter 2 – Literature Review	7
2.1 Photogrammetry	7
2.1.1 Principles	7
2.1.2 Software Utilities	9
2.2 Reflectance Transform Imaging	11
2.2.1 Polynomial Texture Mapping (PTM)	11
2.2.2 Implementation of PTM	12
2.2.3 Hemispherical Harmonics	15
2.2.4 Highlight RTI	17
2.2.5 Related Studies	20
2.3 Photometric Stereo	22
2.3.1 Surface Normals	22
2.3.2 Shape from Shading	23
2.3.3 Surface Height Reconstruction	26
2.4 Specular Reflectance	31
2.4.1 Reflectance Models	31
2.4.2 Gloss	35
2.4.3 Bidirectional Reflectance Distribution Function (BRDF)	37
2.4.4 Specular Distribution Models	38
2.5 Colour Specification	40
2.5.1 The CIE System	40
2.5.2 Colour Measurement	45
2.5.3 Measurement of Angular Reflectance	47
2.5.4 Imaging Techniques for Measurement of Reflectance	48
2.5.5 Metrology	50

2.5.6 Spectral Sampling and Principal Components	52
2.5.7 Spectral Sensitivity	54
2.5.8 Colorimetric Characterisation	57

Chapter 3 – Dome Metrology and Calibration

3.1 The Illumination Dome	61
3.2 Shadow Casting	63
3.3 Photogrammetry of Retroreflective Targets	65
3.4 Photogrammetry using Flash Lamps as Targets	69
3.5 Comparison of Lamp Coordinate Methods	71
3.5.1 Pin Shadows vs VMS Targets	71
3.5.2 VMS Targets vs Flashes	72
3.5.3 Discussion	73
3.6 Targets on Baseboard	75
3.7 Reflections off Glass Sheet	80
3.8 Reflections off Billiard Ball	83
3.8.1 Outline Detection	83
3.8.2 Highlight Detection	84
3.8.3 Horizon Effect	84
3.8.3 Locating Ball Centre	86
3.8.5 Determining Lamp Coordinates	88
3.9 Analysis of Errors	90
3.9.1 Glass Sheet	90
3.9.2 Billiard Ball	91
3.9.3 Summary	93
3.9.4 Finding Lamp Coordinates in RTI	95
3.10 Conclusion	96

Chapter 4 – Reflectance Transform Imaging

4.1 Image Acquisition	98
4.1.1 Image Processing Workflow	98
4.1.2 Pixel Intensity Vectors	99
4.1.3 Vector Projection	100
4.1.4 Hemispherical Projection	100
4.1.5 Polar Projection	100
4.1.6 Making an Outline Mask	101
4.2 Polynomial Texture Mapping	103

4.3	Extensions of PTM	104
4.3.1	Separating and Combining PTM Files	104
4.3.2	Extracting Normals from PTM	105
4.3.3	Image Detail Enhancement	107
4.4	Hemispherical Harmonics (HSH)	108
4.4.1	Basis Functions	108
4.4.2	Fitted HSH Components	109
4.4.3	Quality of Fit	110
4.4.4	Angular Resolution	111
4.4.5	Deriving Normals from HSH	112
4.4.6	Comparing Normals with Arius 3D Geometry	113
4.5	Conclusion	114

Chapter 5 – 3D Surface Reconstruction

5.1	Lambertian Surfaces	117
5.2	Effect of Specularity	119
5.3	Verifying Normals with Digital Test Objects	124
5.3.1	Image Noise	124
5.3.2	Square Pyramid	125
5.3.3	Hemi-ellipsoid	129
5.4	Height Reconstruction by Integration of Gradients	131
5.5	Height Reconstruction with Reference Points	135
5.6	Height Reconstruction by Fourier Transform	138
5.7	Height Reconstruction with Surface from a 3D Scanner	144
5.8	Process Flow for 3D Reconstruction	156

Chapter 6 – Specularity

6.1	Contrast of Specular Surfaces	157
6.1.1	Specular Highlight for one Light Source	158
6.1.2	Intensity Distribution in Multi-light Image Sets	159
6.2	Determining Specular Angle	162
6.3	Modelling Angular Distribution	164
6.3.1	Lorentzian Function	166
6.3.2	Specular Quotient Distributions	167
6.3.3	Cluster Lorentzian Model	170
6.3.4	Power Lorentzian Model	174
6.3.5	Comparison of Lorentzian Models	178

6.4 Image Rendering	182
6.4.1 Lorentzian Reflectance Distribution	182
6.4.2 Image Computation	184
6.4.3 Image Viewer	185
Chapter 7 – Colour	193
7.1 Spectral Characterisation of Camera	194
7.1.1 Monochromator	194
7.1.2 LED Spectral Target	196
7.1.3 Multispectral Filters on Illumination	198
7.1.4 Multispectral Filters on Camera	200
7.1.5 Camera White Balance	202
7.1.6 Correction of Illumination Profile	204
7.1.7 Prediction of Camera RGB	207
7.1.8 Image Noise	209
7.2 Optimum vs Actual Wavelengths of Colour Laser Scanner	210
7.2.1 Method	210
7.2.2 Arius 3D Scanner	212
7.2.3 Influence of White Balance	217
7.2.4 Effect on Colour Images	220
7.2.5 Revision of Optimal Red Laser Wavelength	223
7.2.6 Relationship to Prime Wavelengths of Vision	226
7.3 Principal Components Analysis of Reflectance Spectra	230
7.3.1 Introduction	230
7.3.2 GretagMacbeth DC Chart	230
7.3.3 Synthetic Reflectance Spectra	232
7.4 Constructing Colour Transformation by Synthetic Spectra	233
7.4.1 Device Characterisation	237
7.4.2 Using Synthetic Spectra	239
7.4.3 Colour Transformation by Lookup Table	241
7.4.4 Linearisation of Laser Signals	243
7.4.5 LUT Transformation	246
7.4.6 Evaluation	248
Chapter 8 – Case Studies and Future Work	253
8.1 Studies of Objects	254
8.1.1 Hunters Palette	254

Contents

8.1.2	Egyptian Scarab	258
8.1.3	Courtauld Bag	262
8.1.4	Painted Ceiling	264
8.1.5	Degraded Parchment	267
8.2	Contributions to Knowledge	270
8.2.1	Geometry of Dome	270
8.2.2	Photometric Normals	270
8.2.3	Specular Angle	271
8.2.4	Modelling Specularity	272
8.2.5	3D Surface Reconstruction	273
8.2.6	Colour Laser Scanner Wavelengths	274
8.2.7	Synthetic Spectra and Colour Transformation	274
8.3	Directions for Further Research	275
8.3.1	Dome Design	275
8.3.2	Dome Size and Automation	275
8.3.3	Goniometric Adjustment	276
8.3.4	Anisotropic Surfaces	277
8.3.5	RTI Algorithms	277
8.3.6	High Dynamic Range (HDR) Imaging	277
8.3.7	3D Multiview within Dome	278
8.3.8	Hybrid 3D Reconstruction	279
8.3.9	Multispectral RTI	279
8.4	Final Word	279

References	281
-------------------	-----

Appendices

Glossary of Abbreviations	293
Appendix 1 – Procedure for image capture from dome	294
Appendix 2 – Pin shadow coordinates with dome uplift	296
Appendix 3 – Principal component analysis of spectra	302

Chapter 1 – Introduction

“... and what is the use of a book”, thought Alice,
“without pictures or conversations?”

– Lewis Carroll (1865)
Alice's Adventures in Wonderland

Digital imaging finds its perfect application in the domain of cultural heritage, enabling objects to be captured, rendered and displayed in many ways. It facilitates public access to objects that would otherwise be too rare, valuable or fragile to be widely seen. A key issue for reproductions, both digital and physical, of cultural heritage objects of all types is how realistic they should be. Broadly there are two modes of reproduction:

1. Something that closely matches the original. If the veridical reproduction were placed alongside the original, ideally one could not tell them apart. A perfect reproduction would in every way be a facsimile of the original, and would faithfully reproduce all its defects and flaws. Moreover every attribute could be measured and metrics could be applied against quantitative criteria, such as dimensions, shape, colour, weight, gloss, etc.



Figure 1.1: The rare situation of assessing two printed reproductions side-by-side with the original painting. 'Adoration of the Kings', by the Master of Liesborn c.1475, in the National Gallery London. MARC project, 1993.

2. Something that looks good. It may not be a perfect match with the original in a metric sense, because defects may have been removed and attributes enhanced. When assessed in isolation by a panel of observers on multiple visual quality scales, such a reproduction will often be preferred and ranked more highly than a veridical reproduction. Examples are to be found in every museum shop, where the customer can only see the facsimile object in isolation and cannot make a side-by-side comparison with the original, so the makers are inclined to render the reproduction in such a way to make it more attractive.

There are deep implications of this dichotomy for digitisation and rendering of objects in museum collections. Clearly for documentation and recording purposes, metric accuracy is important. Ideally the archival digital dataset should contain all the information of every aspect of the actual object. For presenting the object through digital media, such as displays in computers and mobile devices, however, the user experience depends on ease of access, interactivity, quality of annotation, and relevance to context, as well as an attractive appearance. Different levels of accuracy may be required for different types of user. What matters most for exhibitions and public access to collections is to convey the illusion of looking at a real 3D object, rather than the accuracy of the underlying representation.

1.1 Objectives of Research

This research programme was initially motivated by a desire to understand the limits of quality that could be achieved in the digitisation of 3D objects and their visual rendering on displays. Attributes of interest included colour, tonal range, surface finish, spatial resolution and dimensionality. The stimulus for the research was the availability of a hemispherical dome with 64 flash lights (Fig. 1.2), which the author had specified and fabricated in 2003, inspired by the development of the Polynomial Texture Mapping (PTM) technique by Malzbender in 2001. The apparatus enables a set of 64 photographic images of an object to be captured in perfect pixel register, with each image illuminated from a different direction. This representation turns out to be much richer than a single 2D image, because it contains information at every point about both the 3D shape of the surface (local curvature) and the directionality of reflectance (gloss and specularity). So it enables not only interactive visualisation of an object through the PTM viewer software, giving the illusion of 3D, but also the extraction of an actual 3D surface and highly realistic rendering of a wide range of materials.



Figure 1.2: The author measuring the uplift of an Egyptian scarab above the baseboard, prior to photography in the illumination dome at UCL.

The over-arching research questions from the outset were:

1. How accurately can spatial dimension and colour be quantified from imagery?
2. Can a realistic rendering of real objects be achieved with illumination from any angle?
3. Can better 3D representations be obtained from dome imaging than from a laser scanner?

1.2 Context of Research Programme

The research leading to this thesis has taken the form of a series of technical studies, each seeking to understand the capabilities of a different aspect of the system. Central always have been the image sets obtained from photography in the dome, and the image processing pipeline. The benchmark for assessment has for the most part been the original PTM viewer, processing image sets with the software developed by Tom Malzbender at HP.

In general my approach has always been to work from first principles rather than using existing tools, because it enhances understanding and leads to more significant developments. Thus I prefer to derive mathematical relationships and transformations from thinking about a problem rather than using someone else's formulation. Yet of course I have relied on the powerful infrastructure that all researchers now have available, in the form of computers, digital cameras, email, file transfer via internet, web services such as Wikipedia and Google Scholar, and desktop software from Microsoft and Adobe. Newton's observation about "standing on the shoulders of giants" is apt at every level.

The greater depth of the research during the past four years, since transfer from the MPhil stage, has been enabled primarily by adopting the Matlab programming environment to develop algorithms and process images. Particularly the ability to generate images, graphs and diagrams has given me confidence to work in a much more visual way. Initially only small image sizes could be processed, but recently the availability of computers with 64-bit Windows, very fast multi-core processors and huge amounts of RAM and disk memory has enabled images at full camera resolution to be processed efficiently.

Also the availability at UCL of laboratory facilities for experimentation has been important, with a space to use the dome and to equip a photographic studio with cameras, a copystand, lighting, and various measuring instruments. These have enabled the research to be conducted in a more exploratory way, by setting up an imaging scenario, measuring parameters, capturing and processing images to extract data, and visualise the results. The colour characterisation studies described in Section 7.1 are good examples of this way of working.

Most important of all has been the existence of a community of scholars with whom I have interacted in many ways. Within UCL the other students and staff within Geomatics provide a lively



Figure 1.3: Discussion around the dome in the lab at UCL in June 2010. (left to right) Ivor Pridden of the Petrie Museum, the author, Mona Hess of Geomatics, and Cristina Re of the University of Padova.

forum for discussing relevant issues. There have been many opportunities to engage with researchers in other UCL departments, notably in Computer Science, Medical Physics, Digital Humanities and the Institute of Archaeology. Also through conferences, symposia and online forums I have been able to meet and communicate regularly with outstanding researchers and experts in related disciplines. Particularly important for this project have been colour scientists in the International Colour Association (AIC), RTI practitioners through the CHI mailing list, interactive media colleagues through the EVA Conference, classicists at the Centre for the Study of Ancient Documents (CSAD) in Oxford, and conservators in the European network Colour and Space in Cultural Heritage (COSCH).

1.3 Contributions to Knowledge

The following outcomes of this research programme are claimed as novel and therefore as representing contributions to knowledge in the field of computational imaging:

- A method for determining the geometry of an illumination dome;
- An adaptive method for finding surface normals by bounded regression;
- Generating 3D surfaces from photometric stereo;
- Relationship between specular and normal angles;

- Modelling surface specularities by a modified Lorentzian function;
- Determining the optimal wavelengths of colour laser scanners;
- Characterising colour devices by synthetic reflectance spectra.

These are summarised in Section 8.2.

1.4 Organisation of this Thesis

After the literature review (Chapter 2) there is a detailed examination of issues related to establishing the geometry of the dome and its lighting (Chapter 3), followed by three studies in Chapters 4 to 6 dealing with aspects of how to extract useful information from image sets captured in the dome:

- Visualisation of objects under movable lighting (PTM and RTI techniques);
- Reconstruction of 3D surfaces;
- Modelling the specularities of shiny surfaces.

These are followed in Chapter 7 by an investigation of the colorimetry of a colour laser scanner and what the optimum wavelengths would be to minimise colorimetric errors. A new method is described for characterising colour devices by large sets of synthetic reflectance spectra. Finally in Chapter 8 the results of the project are summarised with descriptions of five related projects conducted during the period of study, and an outline of further research directions.

1.5 Related Publications

The predecessor to this work was the book *Digital Heritage: Applying Digital Imaging to Cultural Heritage*, published by Elsevier, Oxford, in 2006. This was an edited collection of 20 chapters, many of which were related to the European FP5-IST project 'Veridical Imaging of Reflective and Transmissive Artefacts (VITRA)', which I led from 2002 to 2005.

The MPhil–PhD Transfer Report produced as the first part of this research degree programme was *3D Imaging of Cultural Objects*, October 2010, 182 pp. It covered much of the underlying physics and techniques in 3D computer graphics, colorimetry and image processing, with an extensive review of the literature in relevant areas. It explored the limits of fine detail that can be rendered onto the surface of an artefact by human hand with a tool. It also described at some length the characterisation of the Nikon camera and Arius 3D laser scanner as imaging devices.

The following papers relevant to the topic have been written and presented at conferences and published in conference proceedings during the course of research leading to this thesis.



2014

MacDonald, L.W., Ahmadabadian, A.H., Robson, S. and Gibb, I. (2014) High Art Revisited – A Photogrammetric Approach, *Proc. BCS Conf. on Electronic Visualisation in the Arts (EVA)*, London, July 2014.

MacDonald, L.W., Guerra, M.F., Pillay, R., Hess, M., Quirke, S., Robson, S. and Ahmadabadian, A.H. (2014) Practice-based comparison of imaging methods for visualization of toolmarks on an Egyptian Scarab, *Proc. Intl. Conf. on Image and Signal Processing (ICISP)*, Cherbourg, July 2014.

MacDonald, L.W., Hindmarch, J. and Robson, S. (2014), Modelling the Appearance of Heritage Metallic Surfaces, *Proc. ISPRS Tech. Comm. V Symp. 'Close-range imaging, ranging and applications'*, Riva del Garda, June 2014.

Robson, S. and MacDonald, L.W. (2014), Multispectral Characterisation of Lenses for Photogrammetry, *Proc. ISPRS Tech. Comm. V Symp. 'Close-range imaging, ranging and applications'*, Riva del Garda, June 2014.

MacDonald, L.W., Hosseinaveh Ahmadabadian, A. and Robson, S. (2014), Photogrammetric Analysis of a Heritage Ceiling, *Proc. ISPRS Tech. Comm. V Symp. 'Close-range imaging, ranging and applications'*, Riva del Garda, June 2014.

MacDonald, L.W. (2014) Colour and Directionality in Surface Reflectance, *Proc. Conf. on Artificial Intelligence and the Simulation of Behaviour (AISB)*, Goldsmiths College, April 2014.

2013

MacDonald, L.W. and Roque, T. (2013) Chromatic Adaptation in an Immersive Viewing Environment, *Proc. 12th Congress of the International Colour Association (AIC)*, Newcastle, July 2013, pp.623-626.

MacDonald, L.W., Giacometti, A., Weyrich, T., Terras, M. and Gibson, A. (2013) Colour Analysis of Degraded Parchment. *Proc. 12th Congress of the International Colour Association (AIC)*, Newcastle, July 2013, pp.1809-1812.

MacDonald, L.W., Giacometti, A., Campagnolo, A., Robson, S., Terras, M. and Gibson, A. (2013) Multispectral Imaging of Degraded Parchment. *Proc. Computational Color Imaging Workshop (CCIW)*, Chiba, Japan, March 2013. In *Computational Color Imaging*, ed. S. Tominaga et al, Berlin: Springer, pp.143-157.

2012

MacDonald, L.W., Gibb, I. and Robson, S. (2012) Conservation Monitoring of a Heritage Ceiling by Photometric Stereo. *Proc. XXII Congress of the Intl. Soc. for Photogrammetry and Remote Sensing (ISPRS)*, Melbourne, August 2012, pp.109-114.

Giacometti, A., Campagnolo, A., MacDonald, L.W., Mahony, S., Terras, M., Robson, S., Weyrich, T., Gibson, A. (2012) Cultural Heritage Destruction: Documenting Parchment Degradation via Multi-spectral Imaging. *Proc. BCS Conf. on Electronic Imaging and the Visual Arts (EVA)*, London, July 2012, pp.301-308.

MacDonald, L.W., Gibb, I., Robson, S. and Vlachou-Mogire, C. (2012) High Art: Visualising Damage on a Heritage Ceiling. *Proc. BCS Conf. on Electronic Visualisation and the Arts (EVA)*, London, July 2012, pp.309-316.

MacDonald, L.W. (2012) Colour Laser Scanner Characterisation by Enhanced LUT. *Proc. 6th Eur. Conf. on Colour in Graphics, Imaging & Vision (CGIV)*, Amsterdam, May 2012, pp.137-142.

2011

MacDonald, L.W. (2011) Choosing Optimal Wavelengths for Colour Laser Scanners, *Proc. 19th IS&T/SID Color Imaging Conf.*, San Jose, November 2011, pp.357-362.

MacDonald, L.W. (2011) Optimal Wavelengths of Colour Laser Scanners, *Proc. AIC Mid-term meeting on 'Interaction of Colour and Light in the Arts and Sciences'*, Zurich, Switzerland, June 2011, pp.82-85.

MacDonald, L.W. (2011) Visualising an Egyptian Artefact in 3D: comparing RTI with laser scanning. *Proc. BCS Conf. on Electronic Visualisation and the Arts (EVA)*, London, July 2011, pp. 155-162.

Robson, S., Beraldin, J-A., Brownhill, A. and MacDonald, L.W. (2011) Artefacts for optical surface measurement. *Proc. SPIE Conf. 8085 on 'Videometrics, Range Imaging and Applications XI'*, May 2011, p. 80850C.

2010

MacDonald, L.W. and Robson, S. (2010) Polynomial texture mapping and 3D representation, *Proc. ISPRS Commission V Symp. 'Close Range Image Measurement Techniques'*, Newcastle, June 2010.

MacDonald, L.W. (2010) The Limits of Resolution, *Proc. BCS Conf. on Electronic Visualisation and the Arts (EVA)*, London, July 2010, pp.149-156.

Pending journal publications

Piquette, K.E. and MacDonald, L.W., 3D Representations from Reflectance Transformation Image Datasets: A case study from early Egypt. *J. Computing and Cultural Heritage*.

MacDonald, L.W., Hosseininaveh-Ahmadabadian, A. and Robson, S., Determining the Geometry of Lamp Coordinates in an Illumination Dome, *Image Science Journal*.

MacDonald, L.W., Re C. and Robson S., Constructing Digital Surface Models of an Egyptian Artefact, *Photogrammetric Record*.

MacDonald, L.W. and Mylonas, D., Augmenting Basic Colour Terms in English, accepted for publication subject to minor revision by *Color Research & Application*.

Chapter 2 – Literature Review

Reading furnishes the mind only with materials of knowledge;
it is thinking that makes what we read ours.

- John Locke (1706)
Of the Conduct of the Understanding

This chapter focuses on the literature most relevant to the experimental investigation, and is organised into five sections corresponding to the topics of Chapters 3 to 7.

2.1 Photogrammetry

2.1.1 Principles

Photogrammetry is a non-contact optical method for measurement of complex objects. Its main purpose is the 3D reconstruction of an object in digital or graphical form, with accuracy as a key objective. Under ideal conditions target coordinates can be determined to a precision of 1 part in 10^5 , for example 50 microns at a range of 5 metres (Pegg *et al*, 2009). All photogrammetric methods proceed in three stages: image acquisition, image measurement and object reconstruction, and for each stage many techniques exist (Luhmann *et al*, 2008). Photogrammetric methods may be conveniently classified into single-image, double-image (stereo) and multi-image types.

Single-image photogrammetry is a method of capturing geometrical information (e.g. length, area and volume) about an object from a single image. This method uses linear perspective, i.e. straight lines in the object space, to obtain image rotations (Euler angles) and interior orientation parameters of the camera (focal length and principal points). From these parameters, distances can be calculated (Tommaselli and Reiss, 2005). Single-image photogrammetry is restricted because it cannot be used for 3D modelling of an object. In order to determine ground points and obtain a 3D model, both the exterior orientation and the scale factor of every light ray should be known (the scale factor is not necessarily stable for different regions of an image).

In stereo photogrammetry, two images are taken from different positions in order to capture the same scene (at least in part). To achieve 3D point coordinates, corresponding points in the two images are identified and the rays from object points to image pixels are modelled by the collinearity equations:

$$\begin{aligned} x_a &= x_0 - f \left[\frac{m_{11}(X_A - X_L) + m_{12}(Y_A - Y_L) + m_{13}(Z_A - Z_L)}{m_{31}(X_A - X_L) + m_{32}(Y_A - Y_L) + m_{33}(Z_A - Z_L)} \right] \\ y_a &= y_0 - f \left[\frac{m_{21}(X_A - X_L) + m_{22}(Y_A - Y_L) + m_{23}(Z_A - Z_L)}{m_{31}(X_A - X_L) + m_{32}(Y_A - Y_L) + m_{33}(Z_A - Z_L)} \right] \end{aligned} \quad (2-1)$$

where: (x_a, y_a) are the photo coordinates of image point a ; (X_A, Y_A, Z_A) are the object space coordinates of object point A ; (X_L, Y_L, Z_L) are the object space coordinates of the camera; f is the camera focal length; (x_0, y_0) are the coordinates of the principal point of the lens; and m_{ij}

are the coefficients of the rotation matrix M , derived from the three camera rotation angles $(\omega, \varphi, \kappa)$. These are illustrated in Fig. 2.1 (left). Because the Collinearity equations are non-linear, they are usually solved by first linearising by a Taylor expansion, ignoring higher-order terms, and iterating until the corrections reduce below a specified threshold. By using at least five control points, the relative orientation parameters and finally 3D point coordinates can be calculated (Schenk, 2005).

A second mathematical condition underpinning stereo photogrammetry is the coplanarity condition, illustrated in Fig. 2.1 (right). Any object point and its corresponding image points in the two photos must lie on the same plane. Applying this condition enables the exterior orientations of the two cameras to be determined. More commonly, however, the external orientation parameters are found by the process of resection, using a minimum of three known points in object space to be imaged by both cameras (Wolf and Dewitt, 2000).

Stereo-photogrammetry has been developed over many decades as a fundamental technique in aerial photogrammetry. By using a rigid space bar for mounting the cameras at a specific distance from each other, moreover, the method can also be used in terrestrial and close-range photogrammetry (Cyganek *et al*, 2009). This technique has widely been used in the manufacturing industry (Luhmann, 2010) and for determination of changes in position and attitude for navigation purposes (Byczkowski and Lang, 2009).

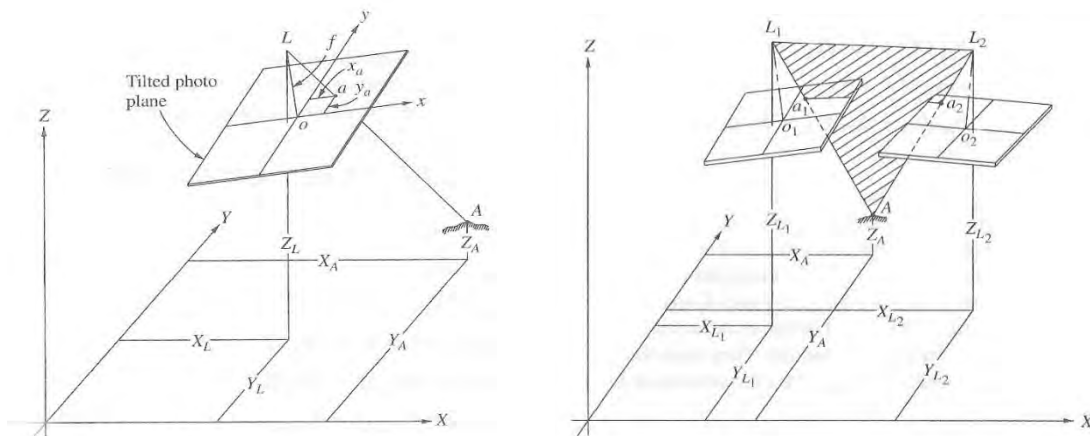


Figure 2.1: (left) Collinearity and (right) coplanarity conditions in photogrammetry (from Wolf and Dewitt, 2004, pp.235-236).

More than two images taken from different positions can be used in the collinearity equations, which leads to a redundancy in observations. Least Square Estimation (LSE) methods can then be exploited to achieve more accurate 3D point coordinates of the object with greater reliability. This theory is the foundation of iterative multi-parameter regression, known as bundle adjustment, in multi-view photogrammetry.

Photogrammetric methods do not have any limitations in range and so can be used for object acquisition from distances from very close (less than one metre) to very far (tens of kilometres from aerial and space). Moreover, the data of the object is captured very quickly, which is essential for obtaining information from unstable objects such as the human body (Fua *et al*, 1998) or from dynamic structures in test rigs. Once the images have been captured, the 3D-coordinates of object points can be estimated with a high level of accuracy.

2.1.2 Software Utilities

Multi-view photogrammetry is widely used in industrial applications (Luhmann, 2010) and many software packages have been produced for computing the surface topography of specific locations on an object or for 3D reconstruction of the whole object, examples being iWitness, Photomodeler, Canoma, ShapeCapture. For photogrammetric analysis, targets are affixed to suitable positions on the object, and the coordinate data from corresponding points in each image are registered to produce a complete 3D network (Faugeras and Hebert, 1986). The multi-view photogrammetric method is effective because of its ability to locate homologous 'features' to sub-pixel accuracy and to assess the dimensional stability of the digital camera and object during imaging. For remote non-contact surface measurement of protective metallic tiles in an experimental fusion device, Brownhill *et al* (2009) used photogrammetry for the measurement of gap and step features of deposition and erosion and achieved a plane-fitting accuracy of the order of 20 μm rms.

The Vision Measurement System (VMS) is a photogrammetric software package developed over the past 20 years, which has been used effectively in industrial applications where very accurate measurements are required with a discrimination range between 1:10000 and 1:100000 (Harvey *et al*, 2003; Shortis *et al*, 2001). The software analyses multiple images captured from an object marked with targets, usually retro-reflective dots or coded markers. The targets are very conspicuous in the image when the illumination direction is close to the optical axis, for example when using a flash or ring light around the lens (Fig. 2.2).

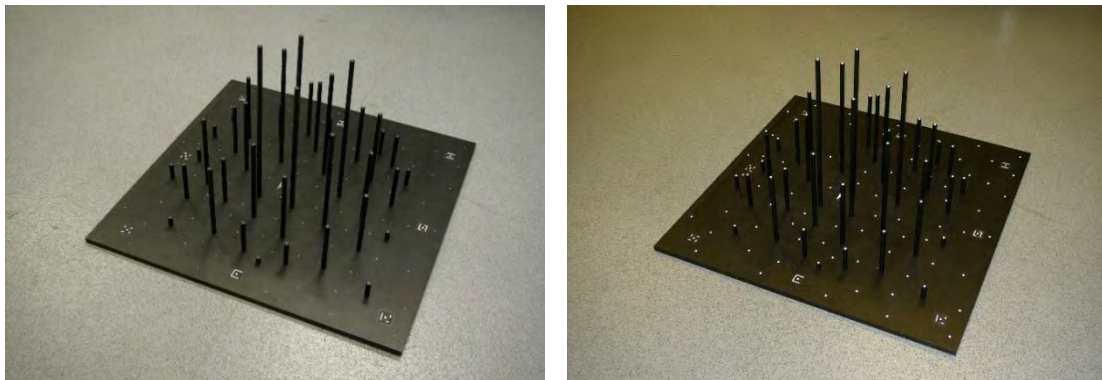


Figure 2.2: Two views of a large 'Manhattan' object at UCL with targets on vertical rods on an aluminium baseplate: (left) under overhead room lights; (right) illuminated by the inbuilt camera flash, close to the lens.

Starting with an initial estimation of camera locations and orientation, resection and intersection, VMS performs iterative bundle adjustment to achieve precise results simultaneously for camera calibration parameters, camera locations and orientations, and target coordinates (Granshaw, 1980). The more targets there are, and the more views of the object from independent camera positions, the more accurate the results can be. Shortis *et al* (1995) analysed the best performance that could be expected from a range of algorithms including ellipse-fitting, binary centroiding, and Gaussian shape fitting. They concluded that errors of the order of 0.1 pixels were common and that the results would be largely unaffected by quantisation or threshold levels, and that a marginal increase in target location precision would be expected with larger targets. Any method using grey-scale images was likely to perform an order of magnitude better than with binary (black/white) images.

Table 2.1: VMS camera calibration parameters, specified in units of millimetres (Shortis, 2001).

Number	Parameter	Description
1	PP _x	X image coordinate of the principal point
2	PP _y	Y image coordinate of the principal point
3	PD	Principal distance of the camera
4	k_3	3rd power term of radial lens distortion
5	k_5	5th power term of radial lens distortion
6	k_7	7th power term of radial lens distortion
7	p_1	first term of decentring (tangential) lens distortion
8	p_2	second term of decentring (tangential) lens distortion
9	a_1	orthogonality of the image coordinate system
10	a_2	affinity of the image coordinate system

The ten calibration parameters generated by VMS for the camera-lens combination are described in Table 2.1. All values are of the parameters, not the corrections. For example, radial distortion is specified as the actual distortion from the ideal location to the distorted location. The principal distance (PD) is the separation between the lens perspective centre and the focal plane. Radial lens distortion is computed using the following formula:

$$dR = k_3 r^3 + k_5 r^5 + k_7 r^7 \quad (2-2)$$

The components of the decentring (tangential) distortion are computed as:

$$dD_x = p_1(r^2 + 2x) + 2p_2xy \quad (2-3)$$

$$dD_y = 2p_1xy + p_2(r^2 + 2y)$$

The actual image location relative to the ideal location is given by:

$$x = x + \left(\frac{x}{r}\right) dR + dD_x + a_1y + a_2x \quad (2-4)$$

$$y = y + \left(\frac{y}{r}\right) dR + dD_y$$

where:

r = radial distance from the principal point = $\sqrt{(x^2 + y^2)}$

x, y = coordinates of the image relative to the principal point.

VMS has been used successfully on a wide variety of 3D measurement applications, including the assessment of shape accuracies of solar concentrators and their components (Pottler *et al*, 2004), underwater for fish and marine fauna (Shortis and Harvey, 1998) and deformation of rotor blades in large wind turbines (Robson and Setan, 1996).

VMS has also been applied in cultural heritage applications, for example in determining the surface geometry and measuring the dimensional stability of the oak support of the Westminster Retable during an extended conservation process (Robson *et al*, 2004). Discrete points on the surface of the Retable were monitored and a 3D surface model was generated enabling the art conservator to make measurements and shape comparisons between different elements of the structure. The approach was first to establish the network of points corresponding to the centroids of retro-reflective targets, for deformation monitoring.

Photogrammetric bundle adjustment in VMS provided the information necessary to ensure calibrated image geometry and agreement to a specified datum. The relative precision for the network was 1:202,000, with a mean target precision of 20 micron and worst case of 90 micron. These points also formed the basis of the initial sparse Delaunay triangulation of the structure. The process of densification was then initialised by applying an interest operator (Forstner, 1987) to create a seed point cloud (Fig. 2.3). Multiple image point correspondence was reliably established for a number of these seed points by defining criteria from various radiometric and geometric properties. Precise correspondence was achieved with a least-squares image patch matching routine (Gruen, 1988), from comparison of image and object triangle shapes.

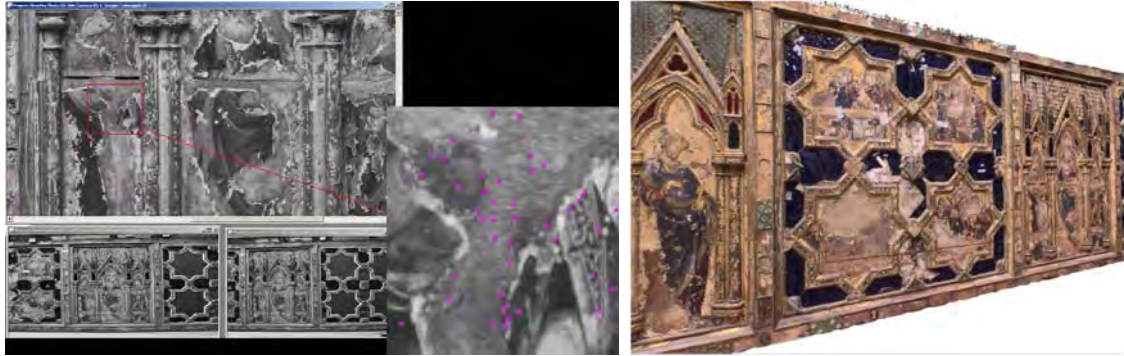


Figure 2.3: (left) Interest points computed during second iteration; (right) Final rendered surface model.

2.2 Reflectance Transform Imaging

2.2.1 Polynomial Texture Mapping (PTM)

Malzbender et al (2001) invented the polynomial texture map (PTM), a novel image-based relighting technique for visualising a surface under a spatially variable source of illumination. It takes a set of images captured from a fixed camera position, with each image lit by a point source at a known coordinate position. The algorithm fits a parametric polynomial function to the set of intensities at each pixel location. The interactive viewer software uses the cursor position, representing the coordinates of a ‘virtual light source’, to generate the intensity of each pixel as if it had been illuminated from that direction. The effect is of a ‘virtual torch’ moving over a static 3D object surface, although there is no inherent 3D representation of the surface. Malzbender et al were motivated by bidirectional texture function (BTF) models, but simplified the formulation by holding the exitant direction constant, i.e. with the reflected angle always toward the fixed camera position. The pixel intensity is a function of the angular coordinates (θ_i, ϕ_i) of the incident light source and two spatial variables (u, v) :

$$I = I_{r,g,b}(\theta_i, \phi_i, u, v) \quad (2-5)$$

By not including dependence on the exitant direction, they sacrificed the ability to capture view-dependent effects such as specularity, but retained the ability to represent arbitrary geometric shadowing and diffuse shading effects. On the assumption of a Lambertian surface, only the luminance of each pixel varies with light source direction and the chromaticity is taken to be constant. This enables separability of the reconstruction function, with a constant colour per pixel modulated by an angle-dependent luminance factor $L(u, v)$:

$$I = L(\theta_i, \phi_i, u, v)R(u, v) \quad (2-6)$$

and similarly for $G(u, v)$ and $B(u, v)$. The dependence of luminance on light direction is modelled by a biquadratic function:

$$L(u, v; I_u, I_v) = a_0(u, v)I_u^2 + a_1(u, v)I_v^2 + a_2(u, v)I_u I_v + a_3(u, v)I_u + a_4(u, v)I_v + a_5(u, v) \quad (2-7)$$

where (I_u, I_v) are projections of the normalised light vector into the local texture coordinate system (u, v) and L is the resultant surface luminance. Given $n+1$ images, the best fit at each pixel is computed using singular value decomposition (SVD) to solve the system of equations for a_0 - a_5 :

$$\begin{bmatrix} I_{u0}^2 & I_{v0}^2 & I_{u0}I_{v0} & I_{u0} & I_{v0} & 1 \\ I_{u1}^2 & I_{v1}^2 & I_{u1}I_{v1} & I_{u1} & I_{v1} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ I_{un}^2 & I_{vn}^2 & I_{un}I_{vn} & I_{un} & I_{vn} & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_5 \end{bmatrix} = \begin{bmatrix} L_0 \\ L_1 \\ \vdots \\ L_n \end{bmatrix} \quad (2-8)$$

Thus a separate set of coefficients $(a_0$ - $a_5)$ is fitted to the image data for each pixel and stored as a spatial map referred to as a Polynomial Texture Map (PTM). The PTM has the same spatial resolution as each of the original images, but has a low resolution in the angular space of incident illumination, because the n directions of the image set are approximated by only 6 coefficients at each pixel. An estimate of the surface normal at each pixel can be extracted from the PTM by setting $\frac{\partial L}{\partial u} = \frac{\partial L}{\partial v} = 0$ to solve for the maximum of the biquadratic in Eq. (3.34). This represents the angle of illumination at which the reflected luminance is maximum, bisected by the surface normal \mathbf{N} :

$$\mathbf{N} = (I_u, I_v, \sqrt{1 - I_u^2 - I_v^2}) \quad (2-9)$$

where:

$$I_u = \frac{a_2 a_4 - 2a_1 a_3}{4a_0 a_1 - a_2^2} \quad I_v = \frac{a_2 a_3 - 2a_0 a_4}{4a_0 a_1 - a_2^2} \quad (2-10)$$

The dependence of reflectance on the angle of incident illumination dependence can also be used to represent anisotropic materials, with different BRDFs at different locations, and Fresnel effects for surfaces such as metals and glass where reflectance increases greatly for grazing angles of incidence. By making a separate PTM for each channel (R, G, B), changes of colour with angle (i.e. wavelength dependence) can be represented.

Malzbender later designated the function in Eq. 2-6 as a Unidirectional Texture Function (UTF). By sacrificing the generality of viewing angle of a BTF, the UTF has the advantages of a compact texture representation, well-matched to the rendering process, which can also be directly employed for synthesis. PTMs derived in this way can later be used in place of conventional texture maps and applied to 3D objects, providing interactive and realistic control of lighting effects, such as shading, self-shadowing, inter-reflections and surface scattering (Hel-Or, 2003).

2.2.2 Implementation of PTM

Malzbender designed and built the original dome at HP Labs from an acrylic hemisphere of diameter 18 inches (45 cm) with 24 flash lights connected by a wiring loom to an external controller (Fig. 2.4 left). He demonstrated the power of the PTM technique for the visual representation of ancient artefacts, such as inscriptions on clay tablets (Fig. 2.4 right).



Figure 2.4: (left) First dome developed at HP Labs; (right) PTM reconstruction of a neo-Sumerian clay tablet (c.2100 BC), illuminated from the top left by a virtual light source (from Malzbender, 2001).

The interactive control of lighting direction in the viewer facilitates perception of the surface structure compared to static photographs, thereby enhancing the legibility of surface relief and inscriptions. In a study on the palaeontological illustration of fossils, Hammer *et al* (2002) found that PTM gave better results than laser scanning for specimens with very low surface relief. They noted that spatial resolution was compromised by the computation of geometric surface normals from laser point cloud data, because of convolution with a kernel having a spatial extent, whereas for PTM the normal estimation for each pixel is performed independently.

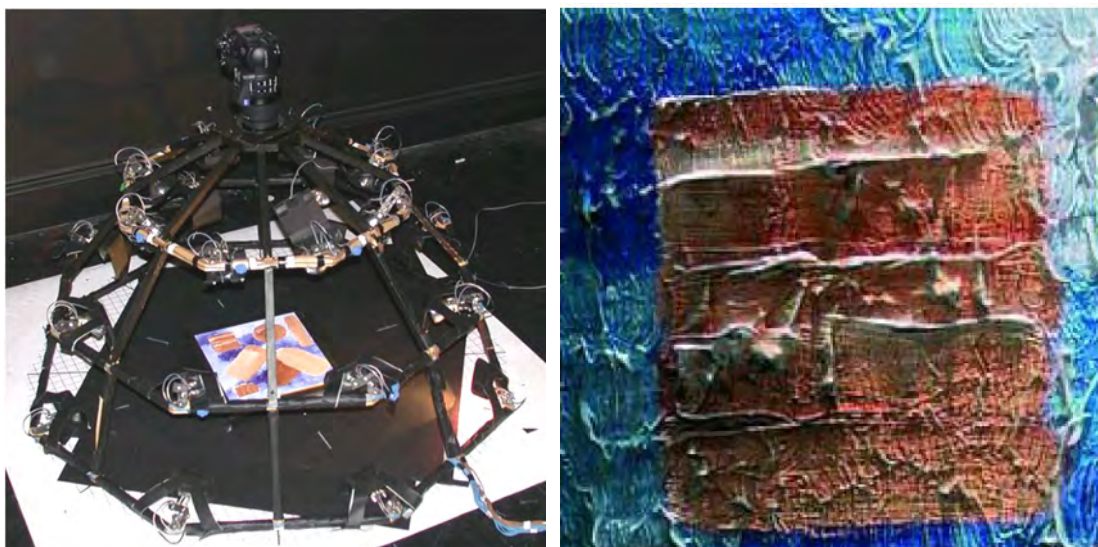


Figure 2.5: (left) Experimental apparatus for PTM at the National Gallery (London) with 24 tungsten spotlights mounted on a rigid framework; (right) surface relief of a painted test panel, lit from the top right.

PTM was applied at the National Gallery in London to investigate the surface structure of paintings. Twenty-four tungsten lamps were mounted onto an open framework in three tiers of eight lamps each (Fig. 2.5 left). The camera was mounted at the top of the framework, pointing down to the painting on the floor. The lamps were turned on and off manually for each image in the sequence to be captured. With the variable ‘virtual light’ in the PTM viewer more features were visible than could be seen by raking light from one direction alone, enabling the study of features in the painting resulting from ageing, such as craquelure and distortion of the support. Comparing PTM renderings made before and after physical changes to the painting facilitated examination of alterations in its texture and shape (Padfield *et al*, 2005).

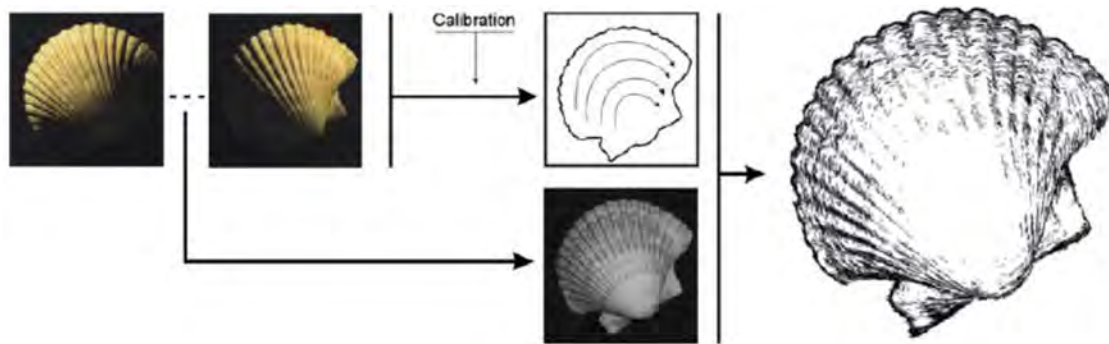


Figure 2.6: Generation of non-photorealistic rendering from multiple images, extracting principal directions of objects in the scene (from Bartesaghi et al, 2005).

Bartesaghi et al (2005) demonstrated a method for automatic 3D shape and geometry rendering from multiple images, using non-photorealistic rendering (NPR) techniques in the style of pen-and-ink illustrations. From a set of images in the multiple-light photometric stereo domain they extracted the tonal information to be mimicked by the NPR synthesis algorithm, and from the object domain they extracted the geometry in the form of principal directions, obtained from the image set without explicitly using 3D models, to convey shape to the drawings (Fig. 2.6).



Figure 2.7: Tom Malzbender using his illumination dome in an upright orientation to image a fragment of the Antikythera Mechanism; (right) PTM reconstruction of fragment, lit from the upper left.

In a famous investigation, PTM was used to acquire images of the fragments of the Antikythera Mechanism at the labs of the National Archaeological Museum in Athens. This unique geared bronze device was constructed around the end of the third century BC, to calculate and display celestial information, particularly cycles such as the phases of the moon and a luni-solar calendar. Named after its place of discovery in 1901 in a Roman shipwreck, the Antikythera Mechanism is technically more complex than any known device for at least a millennium afterwards and indications are that Archimedes was involved in its design. Because the fragile fragments could not be taken out of the museum, the PTM dome was taken there and used in a vertical orientation in front of the camera on a tripod (Fig. 2.7 left). Samples were carefully positioned on holders to enable the imaging to be done without any physical contact. The resulting 82 image sets (Fig. 2.7 right) were used for analysis of the inscriptions, and enabled a better understanding of the structure and function of the mechanism (Freeth et al, 2006).

2.2.3 Hemispherical Harmonics

Spherical harmonics (SH) are based on the associated Legendre polynomials and were originally developed for geodesic and planetary modelling. They consist of a set of orthonormal basis functions defined in two dimensions over the surface of a sphere. A weighted sum of these basis functions can be fitted as an approximation to any single-valued spherical distribution. SH are important in many applications, particularly in the computation of atomic orbital electron configurations, representation of gravitational fields, satellite navigation and the magnetic fields of planetary bodies and stars. They have been used in computer graphics to represent angular variations of an object or environment in properties such as BRDFs (Westin *et al*, 1992), to solve ambiguity problems in computer vision (D’Zmura, 1991), and to represent irradiance for rendering (Ramamoorthi and Hanrahan, 2001). They have also been applied to spatially varying transport effects like shadows (Sloan *et al*, 2002).

For distributions that apply only to a hemisphere, greater efficiency in the fitting may be achieved by limiting the domain to a hemisphere instead of the full sphere. Makhotkin (1996) showed that this could be done for radiative transfer between surfaces by an expansion of Jacobian polynomials analogous to spherical harmonics. An improved method of fitting directional distributions of reflectance was introduced by Gautron *et al* (2004) by limiting the domain to the top half of the sphere, from the north pole down to the equator, with co-latitude in the range $[0, \pi/2]$. Subsequently Huang *et al* (2006) made a similar transformation to an upturned hemisphere, in effect the lower half of the sphere with co-latitude in the range $[\pi/2, \pi]$, and showed how it could be used for 3D shape modelling of organs in medical imaging.

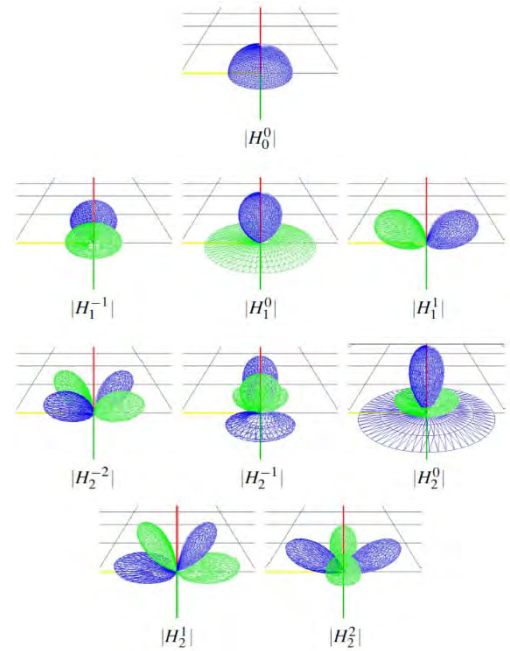


Figure 2.8: The first nine basis functions of hemispherical harmonics (Gautron *et al*, 2004).

Hemispherical harmonic (HSH) functions provide a more compact and accurate way of representing hemispherical distributions than the biquadratic function used in PTM. They have since been widely adopted for a variety of computer graphic applications where only half of the spherical distribution needs to be modelled, such as the representation of anisotropic BRDFs, environment map rendering of non-diffuse surfaces and global-illumination computation. HSH components are expressed as functions of the angles for azimuth φ and co-latitude θ over the hemisphere in terms of l_u and l_v :

$$\varphi = \text{atan2}(l_v, l_u) \text{ and } \theta = \arccos\left(\sqrt{1 - l_u^2 - l_v^2}\right) \quad (2-11)$$

The basis functions $H_l^m(\theta, \varphi)$, which are orthogonal over the two-dimensional support $\left[0, \frac{\pi}{2}\right] \times [0, 2\pi]$ with respect to both l and m , are defined as:

$$H_l^m(\theta, \varphi) = \begin{cases} \sqrt{2\widetilde{K}_l^m} \cos(m\varphi) \widetilde{P}_l^m(\cos \theta) & \text{if } m > 0 \\ \sqrt{2\widetilde{K}_l^m} \sin(-m\varphi) \widetilde{P}_l^m(\cos \theta) & \text{if } m < 0 \\ \widetilde{K}_l^0 \widetilde{P}_l^0(\cos \theta) & \text{if } m = 0 \end{cases} \quad (2-12)$$

where \widetilde{P}_l^m are the associated Legendre polynomials P_l^m shifted from the sphere to the hemisphere:

$$\widetilde{P}_l^m(\cos \theta) = P_l^m(2 \cos \theta - 1) \quad (2-13)$$

and the normalisation value is:

$$\widetilde{K}_l^m = \sqrt{\frac{(2l+1)(l-|m|)!}{2\pi(l+|m|)!}} = \sqrt{2} K_l^m \quad (2-14)$$

Table 2.2: The first three orders of hemispherical harmonics

FIRST-ORDER	$h_0 = H_0^0 = \sqrt{1/2\pi}$
	$h_1 = H_1^1 = \sqrt{3/2\pi} x' = \sqrt{6/\pi} \cos \varphi \sqrt{\cos \theta - \cos^2 \theta}$
	$h_2 = H_1^0 = \sqrt{3/2\pi} z' = \sqrt{3/2\pi} (2 \cos \theta - 1)$
	$h_3 = H_1^{-1} = \sqrt{3/2\pi} y' = \sqrt{6/\pi} \sin \varphi \sqrt{\cos \theta - \cos^2 \theta}$
SECOND-ORDER	$h_4 = H_2^2 = \sqrt{15/8\pi} (x'^2 - y'^2) = \sqrt{30/\pi} \cos 2\varphi (\cos^2 \theta - \cos \theta)$
	$h_5 = H_2^1 = \sqrt{15/2\pi} x'z' = \sqrt{30/\pi} \cos \varphi (2 \cos \theta - 1) \sqrt{\cos \theta - \cos^2 \theta}$
	$h_6 = H_2^0 = \sqrt{5/8\pi} (-x'^2 - y'^2 + 2z'^2)$ $= \sqrt{5/2\pi} (1 - 6(\cos \theta - \cos^2 \theta) \sqrt{\cos \theta - \cos^2 \theta})$
	$h_7 = H_2^{-1} = \sqrt{15/2\pi} y'z' = \sqrt{30/\pi} \sin \varphi (2 \cos \theta - 1) \sqrt{\cos \theta - \cos^2 \theta}$
	$h_8 = H_2^{-2} = \sqrt{15/8\pi} x'y' = \sqrt{30/\pi} \sin 2\varphi (\cos^2 \theta - \cos \theta)$
THIRD-ORDER	$h_9 = H_3^3 = \sqrt{35/16\pi} y'(3x'^2 - y'^2) = \sqrt{140/\pi} \cos 3\varphi (\sqrt{\cos \theta - \cos^2 \theta})^3$
	$h_{10} = H_3^2 = \sqrt{105/8\pi} z'(x'^2 - y'^2) = \sqrt{210/\pi} \cos 2\varphi (1 - 2 \cos \theta) (\cos \theta - \cos^2 \theta)$
	$h_{11} = H_3^1 = \sqrt{21/16\pi} x'(4z'^2 - x'^2 - y'^2)$ $= \sqrt{84/\pi} \cos \varphi \sqrt{\cos \theta - \cos^2 \theta} (1 - 5(\cos \theta - \cos^2 \theta))$
	$h_{12} = H_3^0 = \sqrt{7/8\pi} z'(2z'^2 - 3x'^2 - 3y'^2)$ $= \sqrt{7/2\pi} (12 \cos \theta - 1 - 30 \cos^2 \theta + 20 \cos^3 \theta)$
	$h_{13} = H_3^{-1} = \sqrt{21/16\pi} y'(4z'^2 - x'^2 - y'^2)$ $= \sqrt{84/\pi} \sin \varphi \sqrt{\cos \theta - \cos^2 \theta} (1 - 5(\cos \theta - \cos^2 \theta))$
	$h_{14} = H_3^{-2} = \sqrt{105/2\pi} x'y'z' = \sqrt{210/\pi} \sin 2\varphi (1 - 2 \cos \theta) (\cos \theta - \cos^2 \theta)$
	$h_{15} = H_3^{-3} = \sqrt{35/16\pi} x'(x'^2 - 3y'^2) = \sqrt{140/\pi} \sin 3\varphi (\sqrt{\cos \theta - \cos^2 \theta})^3$

In the unit sphere the Cartesian coordinates can be defined in terms of θ and φ as:

$$x = \sin \theta \cos \varphi ; \quad y = \sin \theta \sin \varphi ; \quad z = \cos \theta \quad (2-15)$$

After change of variables from the spherical to the hemispherical system these become:

$$x' = 2 \cos \varphi \sqrt{\cos \theta - \cos^2 \theta} ; \quad y' = 2 \sin \varphi \sqrt{\cos \theta - \cos^2 \theta} ; \quad z' = 2 \cos \theta - 1 \quad (2-16)$$

Good results are obtained with sixteen components, which include a constant plus three first-order, five second-order and seven third-order terms. These are expressed in Table 2.2 as functions of either (θ, φ) or (x', y', z') . Fitting of the coefficients follows the same SVD-based procedure as for PTM, and can be applied to the image luminance (weighted sum of R,G,B channels) to provide the angular modulation at each pixel of a constant R,G,B colour value. Alternatively, at the expense of greater storage, the fitting can be applied to each of the R,G,B channels independently. Comparison of the HSH and PTM renderings (Fig. 2.9) shows that the HSH (2nd-order with nine coefficients) gives a better representation of the directionality of the surface, with higher contrast for local gradients. Open-source software for fitter and viewer that supports both PTM and HSH image formats has been developed, and is used by a world-wide community of practitioners through the Cultural Heritage Imaging (CHI) organisation in California.



Figure 2.9: Interleaved strips of PTM and HSH renderings of a gold Merovingian triens (from Mudge et al, 2008).

2.2.4 Highlight RTI

PTM and HSH can be regarded as two examples of a more general family of Reflectance Transform Imaging (RTI) methods, in which the distribution of light reflected from an object surface can be modelled as functions of space, angle, spectrum and time. Key to the broader adoption of RTI has been the development of Highlight-RTI, in which a glossy sphere is placed in the scene so that the direction of the incident illumination can later be inferred from the coordinates of the highlight in each image. This obviates the need for a dome system, and enables the photography to be done *in situ*. Such a technique is essential for field work where the objects are so large or impossible to move that there is no alternative but to do the imaging on site, such as monuments, caves and rock art. A good review of applications in archaeology is given by Earl et al (2010). All methods of PTM and RTI capture all rely on knowledge of the light positions used to illuminate the object. For a dome system this is determined by the physical placement of the lights. Alternatively a template might be used to determine the exact positions of one or more movable lights in pre-determined locations (Mudge et al, 2005). An equivalent would be to employ a robot arm to position a lamp successively at pre-programmed locations. Equipment with fixed lighting positions has many advantages, including speed of acquisition, accuracy and repeatability. But it also has limitations, in particular the maximum object size, cost, portability, and difficulty of adapting to the site topography.

The H-RTI method was introduced by Mudge *et al* (2006) as a technique for on-site acquisition of image sets for rock art. The feature is that the illumination source, such as a flash or spotlight, can be moved freely to any position above the surface for each image, with no predetermined constraints and also no specific recording of its position. To maintain the level of illuminance approximately constant, the distance of the light from the object centre is maintained by use of a ruler or string. When working in a team typically the photographer holds the lamp while an assistant holds the string (Fig. 2.10). Within the field of view of the camera one or two glossy spheres are placed and recorded in each image. In each image the position of the light in relation to the surface is inferred by analysis of the highlight position on the ball. The accuracy of the coordinates is limited by the modulation transfer function (MTF) of the lens, image resolution, and the sphericity of the mirror ball. The method has proved to be versatile and effective for objects of all sizes from millimetres to decametres, both outdoors and indoors.



Figure 2.10: (left) Image capture with a Canon camera on a tripod and hand-held flash, using the H-RTI method on a marble tablet in the storeroom of the Ashmolean Museum, Oxford; (right) setup with two billiard balls alongside granite shaft of the Philae obelisk on a scaffold platform in the grounds of Kingston Lacy, Dorset.

The algorithm for calculating the lamp vector from the highlight on the ball in the image, as implemented in the RTI Builder software, makes the assumption that both the light source and the camera are far enough from the ball that the incident rays from the light and the rays reflected to the viewpoint can be regarded as parallel (Fig. 2.11). First, via the graphic user interface (Fig. 2.12), the centre coordinates (C_x, C_y) and radius r of the sphere are determined from the image, then the centre of the highlight spot, (H_x, H_y), all in units of pixels. The result is the normalised triplet (L_x, L_y, L_z) defining the direction of the lamp vector.

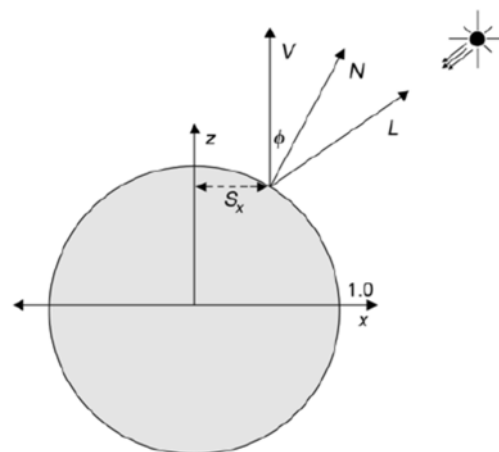


Figure 2.11: Vectors in the H-RTI calculation for determining the lamp vector from a highlight on a ball (from Mudge *et al*, 2006).

The highlight is normalised to the radial interval $[-1, +1]$ by:

$$S_x = \frac{(H_x - C_x)}{r} \quad S_y = \frac{(H_y - C_x)}{r} \quad (2-17)$$

The normal vector \mathbf{N} can then be expressed as:

$$\mathbf{N} = \left(S_x, S_y, \sqrt{1 - S_x^2 - S_y^2} \right) \quad (2-18)$$

Converting to polar coordinates, the inclination angle φ (co-latitude) of the surface normal is:

$$\varphi = \cos^{-1} \left(\sqrt{1 - S_x^2 - S_y^2} \right) \quad (2-19)$$

Assuming that the view vector \mathbf{V} is aligned with the Z axis, the inclination angle φ_L and azimuthal angle θ_L of the lamp vector \mathbf{L} are:

$$\varphi_L = 2\varphi = 2\cos^{-1} \left(\sqrt{1 - S_x^2 - S_y^2} \right) \text{ and } \theta_L = \sin^{-1} \left(\frac{y}{\sin\varphi} \right) \quad (2-20)$$

By the standard conversion from spherical to Cartesian coordinates, the components of the normalised lamp vector are thus:

$$L_x = \sin(\varphi_L) \cos(\theta_L) \quad (2-21)$$

$$L_y = \sin(\varphi_L) \sin(\theta_L)$$

$$L_z = \cos(\varphi_L)$$

When this process has been completed for every image, the region containing the ball(s) is trimmed off and the RTI file is constructed in the standard way with the calculated lamp vectors.

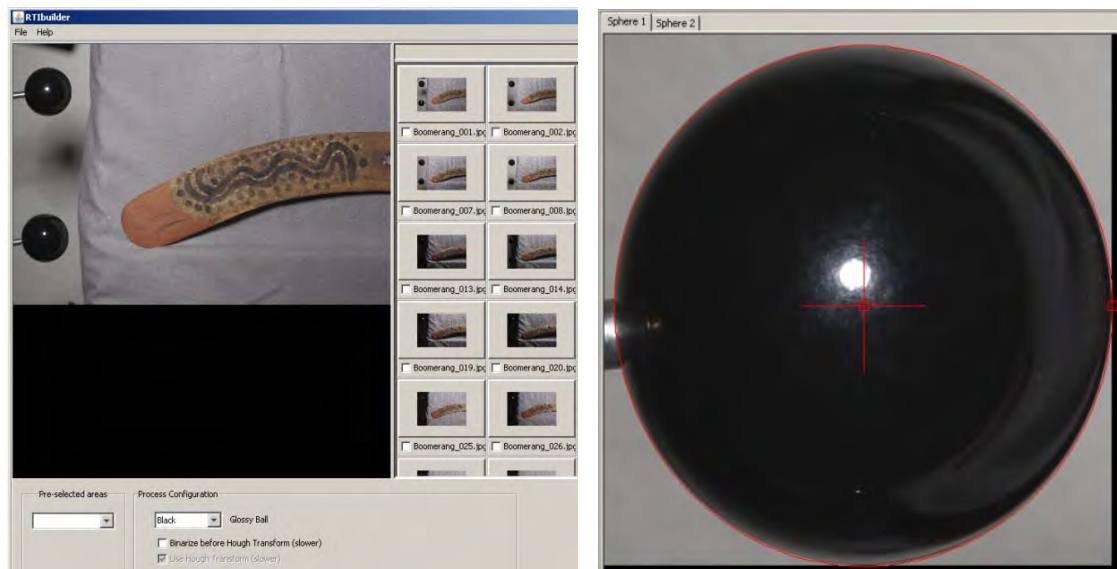


Figure 2.12: Two illustrations from the H-RTI Builder User Guide showing screenshots of: (left) image selection; (right) outline detection and selection of highlight on the billiard ball image (from Mudge et al, 2014).

Advantages claimed for the H-RTI method (Duffy *et al*, 2013) include:

- A non-invasive and non-contact recording method, which protects the heritage resource being recorded thus eliminating the need for destructive sampling.
- Cost effective, using a relatively inexpensive tool kit and freely available software.
- Transportable kit with a technique flexible enough to be used in many recording scenarios, such as at remote sites.

A disadvantage of H-RTI is that it requires an additional processing step, which may be quite time-consuming later when the RTI file is built. Also the method does not work well when the scene is illuminated by direct sunlight, and the precision of the incident light vector is limited by the size and quality of the highlight spot on the ball in each image. It should be considered as a qualitative rather than quantitative method.

2.2.5 Related Studies

How many images are needed to give an accurate rendering of the angular reflectance distribution in PTM or HSH? Ideally one should capture images with illumination from all necessary angles but no more. The challenge is to reduce the number of photographic image samples that need to be acquired while preserving the power of the digital model to represent the object realistically and to enable reliable analysis. A full BRDF analysis requires a systematic sampling of a four-dimensional space, with both illumination and view angles able to range over the complete hemisphere. With the PTM and RTI scenario, however, the view direction is always fixed (usually at the zenith, perpendicular to the centre of the object surface) and only the illumination direction is variable.

Gunawardane *et al* (2009) analysed the sampling of both view and lighting directions and whether methods for interpolation could be improved if both view and lighting information were available. They conducted a data-driven study in which a test object was illuminated by a hemispherical dome with 64 tungsten lights. The object sat on a turntable and a full set of 64 images was captured for 360 rotational angles of the object, at 1° intervals for a total of 23,040 images, taking 13 hours and producing 185 Gbytes of data.

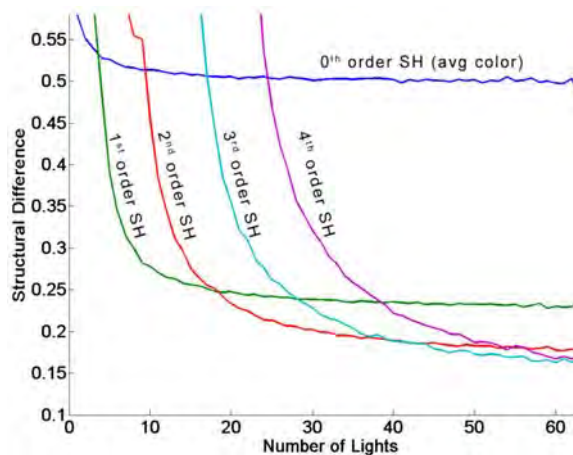


Figure 2.13: Error vs number of lights for four orders of hemispherical harmonics (Gunawardane, 2009).

The complete image set was then sub-sampled for intervals of both lighting and view angles and the errors calculated between HSH fittings of both the full image set and the subsampled set. The results (Fig. 2.13) indicated that the minimum numbers of lamps are approximately 10, 20, 36 and 56 for the 1st to 4th harmonic orders respectively. The pitfall in interpolation of images from different lighting directions was found to be that errors in flow vectors caused pixels to move to incorrect positions, causing visible tearing artifacts and structural discontinuities.

Happa *et al* (2009) reviewed aspects of historically authentic illumination, such as candles and gaslight, and the need to ensure that the illumination model adopted is valid. The removal of specularities and shadows can be regarded as the problem of extracting information contained in an image and transforming it into a meaningful representation of the intrinsic properties of the input image, such as interface and body reflections.

Devlin *et al* (2002) reviewed tone mapping methods for the rendering of realistic images. Conventional image capture and display technologies do not support the full dynamic range that the human eye is capable of processing. This High Dynamic Range (HDR) of luminance needs to be captured through multiple exposures and integrated to increase the reproduced level of realism in any virtual reconstruction. Kim and Kautz (2008) devised a technique for camera characterisation to improve colour accuracy in HDR imaging, and showed that it is very accurate even in unknown illumination conditions. Kuang *et al* (2007) introduced the novel colour appearance model iCAM06, which predicts a range of colour appearance phenomena (illumination level, field size, surround colour, etc.) that allow more precise simulation of how the HDR scene will appear to the user.

Drew *et al* (2012) observed that for non-Lambertian phenomena matrix factorisation methods can produce inaccurate surface normals and lighting directions. Because the basic PTM method relies on a matte reflectance model and linear regression, it fails to model higher-degree phenomena such as inter-reflections, specularities and shadows. Increasing the degree of the PTM polynomial model, for example by the use of HSH basis functions, may help to model these effects but at the expense of degrading the interpolated results at non-sampled light directions due to over-fitting. For computing the matte component of the photometric model, they used the Least Median of Squares (LMS) method, which provides automatic identification of outliers, both specular highlights and shadows. Knowledge of the inlier pixel values means that the recovered surface albedo, chromaticity and surface normals are robust, in the sense of ignoring outlier contributions and thus more accurately mapping surface reflectance, colour and shape (Fig. 2.14). They also modified the polynomial used in PTM so as to generate a subset of three regression coefficients that can be proved to be exactly correct in the case when the inliers are Lambertian.



Figure 2.14: (left to right) Two original images of a gold coin, an interpolated image, matte base image, and normal vectors in false colour (from Drew *et al* 2012).

Zhang and Drew (2014) have recently introduced a more efficient ‘mode-finder’ regression method to replace the computationally intensive multivariate LMS regression in the matte modelling stage. They also explored different combinations of basis functions for PTM, and found that 16 components for luminance plus 9 for chromaticity give a good balance between dimensionality and accuracy. They compared the performance of HSH against the polynomial basis of the same order and found that the polynomial model outperforms HSH in terms of the quality of appearance of the reconstruction, especially at large angles of incidence.

2.3 Photometric Stereo

2.3.1 Surface Normals

A normal \mathbf{N} to a surface S at point P is a vector perpendicular to the tangent plane touching the surface at P (Fig. 2.15). For a plane given by the equation $ax + by + cz = d$, the vector (a,b,c) is a normal. If the surface is defined as the set of points satisfying $S(x, y, z) = 0$, then a normal at a point (x, y, z) on the surface is given by the gradient formed by the partial first derivatives with respect to each of the three axes:

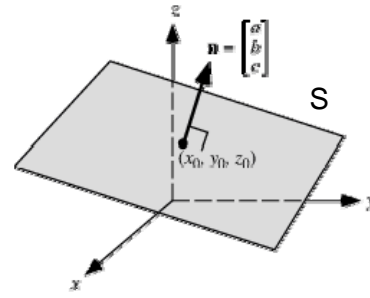


Figure 2.15: Normal to a surface

$$\mathbf{N} = \nabla f(x, y, z) = \left[\frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}, \frac{\partial S}{\partial z} \right]^T = \frac{\partial S}{\partial x} \hat{\mathbf{i}} + \frac{\partial S}{\partial y} \hat{\mathbf{j}} + \frac{\partial S}{\partial z} \hat{\mathbf{k}} \quad (2-22)$$

where $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ are the unit normal vectors along the x, y, z axes respectively. In the case where the surface is defined by $z = S(x, y)$ the normal is:

$$\mathbf{N} = \left[\frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}, -1 \right]^T = [p, q, -1]^T \quad (2-23)$$

where p, q are the first derivatives of $S(x, y)$ with respect to x and y , and the vector \mathbf{N} is an outward surface normal at (x, y) . The term outward refers to the direction of the normal with respect to the viewer. By convention, the viewer is assumed to be along the negative Z axis in relation to the surface being viewed. The surface normal $(0,0,-1)$ points directly at the viewer and is orthogonal to the image plane. The distance between the viewer and the object is assumed to be great enough in relation to the object's size that the projection onto the imaging plane can be approximated by an orthographic projection. The normalised (unitary) normal vector \mathbf{n} is defined as:

$$\mathbf{n} = \frac{\mathbf{N}}{\|\mathbf{N}\|} = \frac{1}{\sqrt{p^2 + q^2 + 1}} [p, q, -1]^T \quad (2-24)$$

Instead of gradients p, q the surface slant ϕ and tilt θ are commonly used, related to the surface normal by:

$$(n_x, n_y, n_z) = (\sin\phi \cos\theta, \sin\phi \sin\theta, \cos\phi) \quad (2-25)$$

For a plane given by the vector equation $\mathbf{r} = \mathbf{t} + \alpha\mathbf{u} + \beta\mathbf{v}$, where \mathbf{t} is a vector to get from the origin onto the plane and \mathbf{u}, \mathbf{v} are non-parallel vectors lying on the plane, the normal is given by $\mathbf{N} = \mathbf{u} \times \mathbf{v}$, i.e. the cross product of the vectors. There are two normals to every surface, one facing outward and the other inward, but for rendering the surfaces of objects the outward normal is usually chosen.

In a digital image the discrete pixels are represented in Matlab as P_{ji} where j is the row index (vertically down the column) corresponding to the $-y$ coordinate, and i is the column index (horizontally across the row) corresponding to the $+x$ coordinate (Fig. 2.16). The first pixel, with coordinates $(1,1)$, is at the top left of the array. The gradients p, q at pixel P_{ji} are defined as the average in the horizontal and vertical directions respectively of the differences between the value of the pixel and its immediate neighbour on either side:

$$\begin{aligned}
p &= \text{mean}(P_{j,i} - P_{j,i-1}, P_{j,i+1} - P_{j,i}) \\
&= \frac{(P_{j,i+1} - P_{j,i}) + (P_{j,i} - P_{j,i-1})}{2} = \frac{(P_{j,i+1} - P_{j,i-1})}{2} \\
q &= \text{mean}(P_{j,i} - P_{j-1,i}, P_{j+1,i} - P_{j,i}) \\
&= \frac{(P_{j+1,i} - P_{j,i}) + (P_{j,i} - P_{j-1,i})}{2} = \frac{(P_{j+1,i} - P_{j-1,i})}{2}
\end{aligned} \tag{2-26}$$

$P_{j-1,i-1}$	$P_{j-1,i}$	$P_{j-1,i+1}$
$P_{j,i-1}$	$P_{j,i}$	$P_{j,i+1}$
$P_{j+1,i-1}$	$P_{j+1,i}$	$P_{j+1,i+1}$

Figure 2.16: Immediate neighbours of pixel $P_{j,i}$

Second derivatives can similarly be approximated as:

$$\frac{\partial^2 S}{\partial^2 x} = \frac{\partial p}{\partial x} = (P_{j,i+1} - 2P_{j,i} + P_{j,i-1}) \quad \text{and} \quad \frac{\partial^2 S}{\partial^2 y} = \frac{\partial q}{\partial y} = (P_{j+1,i} - 2P_{j,i} + P_{j-1,i}) \tag{2-27}$$

2.3.2 Shape from Shading

Shape from shading is a member of the set of ‘Shape from X’ methods. These have been largely developed by computer vision researchers, and can be categorised according to the number of captured views into single-view, binocular, and multi-view methods. Single-view methods need one or more images of a particular scene captured by one camera in a fixed position. In order to achieve a depth map or 3D information, binocular (or stereo-pair) images can be captured with one camera moved successively to two specific positions or by two cameras fixed at those positions. Multi-view methods capture multiple images with a single camera, moving either the object or camera with respect to one another, or by multiple cameras situated around the target object.

Many algorithms have been developed to estimate depth map or 3D information from a single view. Methods include linear perspective, focus, atmosphere scattering, texture, occlusion (curvature or single transform), silhouette and statistical patterns. Almost all ‘Shape from X’ techniques that use a single view (except statistical pattern) can achieve a map of relative depth, but this may be nonlinear and may have a different scale from the real object. Moreover, the applicability of each method is limited by the assumptions pertaining to each specific situation.

In shape from shading (SFS) algorithms, given a grey level image, the aim is to recover the light source and surface shape at each pixel in the image. The surface reflectance can be used to infer the surface normal at each point and hence to determine the 3-D shape of an object. Christensen and Shapiro (1994) specified the assumptions about the scene and the object as:

- All parts of the surface under consideration have the same physical properties (colour, reflectance, transparency, roughness, etc.), which are known or estimated *a priori*;
- Every light source is distant, so the intensity fall-off across the surface can be neglected. The position, shape, and intensity of the light source is known;
- The directions to each light source must differ significantly between the images, otherwise the shape information in the two images is redundant. On the other hand, if the light source positions are too different, the part of the surface that is illuminated in both images will be small. Therefore the directions to the light sources should have been selected by some trade-off between these two contradictory requirements;
- The object does not receive light reflected from other objects (or other parts of the same object).

In the ideal case of a Lambertian surface, where the light falling on the surface is scattered equally in all directions, the apparent brightness of the surface at point P to an observer along vector \mathbf{V} is the same regardless of the observer's angle of view (Fig. 2.17), and the reflection is called diffuse. The luminance of the reflected light is given by the vector dot product:

$$L_r = \rho \mathbf{L} \cdot \mathbf{n} = \rho |\mathbf{L}| \cos \alpha \quad (2-28)$$

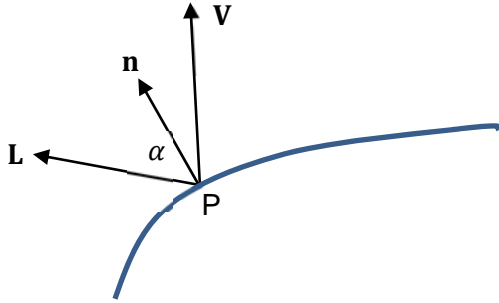


Figure 2.17: Geometry of reflection.

where L_r is the luminance of the diffusely reflected light (with no angular dependence), ρ is the surface reflectance (also with no angular dependence), \mathbf{L} is the incident light vector (by convention pointing from the surface toward the lamp), \mathbf{n} is the unit normal of the surface, and α is the angle between \mathbf{L} and \mathbf{n} . Because the normal vector has three components, at least three equations are needed to solve the system. This

can be achieved by illuminating the surface for successive images from three different lighting directions with incident light vectors \mathbf{L}^1 , \mathbf{L}^2 and \mathbf{L}^3 . Following Eq. 2-28 this system can be written as:

$$I^k = \rho \mathbf{L}^k \cdot \mathbf{n} \quad (2-29)$$

where $k=1,2,3$ are the three lighting directions. The three observed intensity values I^k of the reflected light can be stacked to form the 3×1 intensity vector $\mathbf{I} = (I^1, I^2, I^3)$, and the incident light vectors can be stacked row-wise to give the 3×3 lighting matrix $\mathbf{L} = (\mathbf{L}^1, \mathbf{L}^2, \mathbf{L}^3)$. Then Eq. 2-29 can be rewritten as:

$$\mathbf{I} = \rho \mathbf{L} \cdot \mathbf{n} \quad (2-30)$$

Provided the lighting directions \mathbf{L}^k are not coplanar, the matrix \mathbf{L} can be inverted, giving:

$$\rho \mathbf{n} = \mathbf{L}^{-1} \mathbf{I} \quad (2-31)$$

Since the normal vector \mathbf{n} is unitary, both its direction and albedo (modulus) can be recovered. Note that this simple method cannot be applied to outdoor images illuminated by sunlight, because the sun's linear trajectory across the sky means that the illumination vectors in

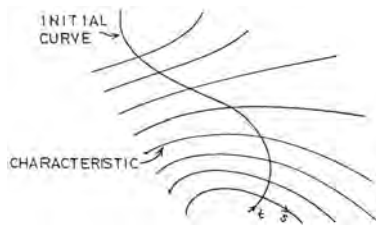


Figure 2.18: Characteristic strips sprouting from an initial curve (from Horn, 1970, p.50).

successive images would be coplanar. When there are more than three input images, the illumination matrix \mathbf{L} is not square and cannot be inverted directly. With the pseudo-inverse procedure, the albedo and the normal can then be recovered as:

$$\rho \mathbf{n} = (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T \mathbf{I} \quad (2-32)$$

Horn (1970) first showed how shape could be determined by solving nonlinear, first-order partial differential equations to find components of the surface gradient. While the gradient could not be determined locally, numerical integration of the equations permitted characteristic curves to be traced out on the object surface (Fig. 2.18). Woodham (1980) developed a method for determining surface shape without requiring the solution of a system of equations. Using one image illuminated by a single light source, the set of all possible surface orientations was limited to a set of iso-brightness

contours in gradient space. The actual gradient at each point was then determined by using relaxation-type algorithms. He further showed how multiple sources of illumination could be used to determine local surface gradients, a method he called photometric stereo. Where three images are used, illuminated from three independent light sources, the three contours intersect at a single point corresponding to the gradient of the surface. Woodham noted that photometric stereo is most accurate in regions of gradient space where the density of contours is great (i.e. rapid change in intensity across the image) and where the contours to be intersected are nearly perpendicular (i.e. where the light source vectors are orthogonal).

Woodham assumed the surfaces to be uniform, matte, convex and smooth with continuous first and second partial derivatives. The surfaces of real objects, however, have a microstructure that perturbs the angles of reflection, and also varying degrees of gloss that produces a specular component of reflectance. The effect is to introduce noise into the computation, manifested as a scattering of the directions of normals from adjacent pixels. Coleman and Jain (1982) adapted the technique to include surfaces varying in both specularity and visual texture, which they treated as resulting from non-uniform surface albedo. They showed how specular reflections can be detected and removed by including a fourth source of illumination. They calculated the overall deviation R_{dev} of the reflectances R_i from the mean μ by:

$$R_{dev} = \frac{\sum_{i=1}^4 (R_i - \mu)}{4 * R_{min}} \quad (2-33)$$

where R_{min} is the minimum of the four values. Before finalising the calculation of the surface normal at each point, a threshold value R_t is chosen representing the largest amount of reflectance deviation allowed before specularly occurs. To eliminate this specular component, the surface normal is averaged from the other three intensity values which have the smallest reflectance factor. This method could readily be extended to n sources, with the removal of multiple outliers.

Barsky and Petrou (2003) showed that the above method is susceptible to errors in the presence of shadows and they generalised the grey-scale photometric stereo technique for colour images, making use of spectral information as an additional cue for detecting highlights. Their combined method recovers local gradient and colour for three-dimensional non-Lambertian surfaces from four colour images, and copes with both shadows and highlights. Christensen and Shapiro (1994) observed that with Lambertian reflection the colour components are linearly dependent, which means that the information in a colour image is equivalent to that in three greyscale images. Hence there is no advantage in processing colour images for shape extraction, rather than an equivalent number of greyscale images. For non-Lambertian reflection, however, the information in the colour components is correlated. The specular and diffuse reflections usually have different colours, with peaks in different directions. Therefore a colour with contributions from both specular and diffuse reflection corresponds, in general, to a finite number of different orientations. In the dichromatic model (Shafer, 1985)

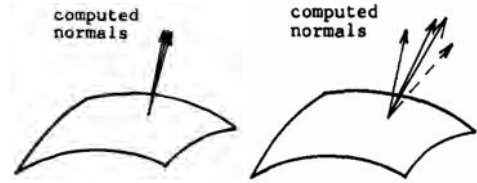


Figure 2.19: Computed normals from four light sources from (left) a matte surface, and (right) a surface exhibiting specularity (from Coleman and Jain, 1982).

and Torrance-Sparrow (1967) models, for example, at most two orientations can correspond to a colour with both specular and diffuse components. With a greyscale image, it is impossible to separate intensity into specular and diffuse components, and one intensity value can correspond to infinitely many orientations. For surfaces with specular reflection in one or more images, the orientation can be uniquely determined from the pixel colours. Using colour images also improves sensitivity in the presence of noise.

Kay and Caelli (1995) used an apparatus with a fixed video camera and rotating light source (Fig. 2.20) to capture successive images of a stationary object with different angles of illumination. The number of combinations of zenith and azimuth angles of illumination ranged from 55 to 75, although zenith angles were limited to a maximum of 30° from the camera axis. They applied non-linear least-squares regression over the set of images to estimate at each pixel the surface normals and parameters for a restricted version of the Torrance-Sparrow reflectance model.

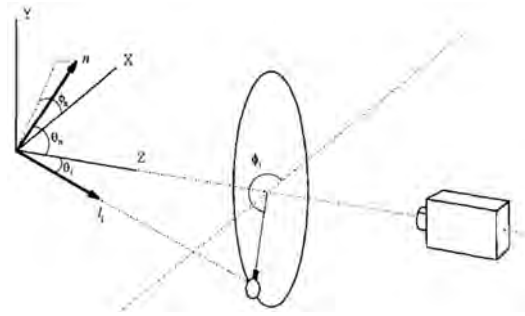


Figure 2.20: Setup for photometric stereo with a fixed camera and a rotating light source (from Kay and Caelli, 1995).

Plata *et al* (2009) used an alternative setup in which both camera and object rotated with respect to a stationary light source (Fig. 2.21). They showed that four-source photometric stereo could be applied effectively for recovering both shape and texture in RGB colour. Rushmeier and Bernadini (1999) employed a simple illumination rig with five tungsten lamps, one near the camera axis and four separated at approximately 45° from the axis. At each pixel, after correction for variation in lamp intensity, the five values were sorted into order and the highest and lowest discarded. If the three remaining values were all non-zero then the normal and relative reflectance were calculated. Debevec *et al* (2000) adopted an image-based approach with the photometric stereo technique. Instead of using reference objects, they recorded the object data under a dense sampling of light directions, then constructed a reflectance map for each point from 2,000 intensity values per pixel over the space of illumination directions, and recovered the surface geometry. Their method required dense image sets, gathered in real-time by several video cameras.

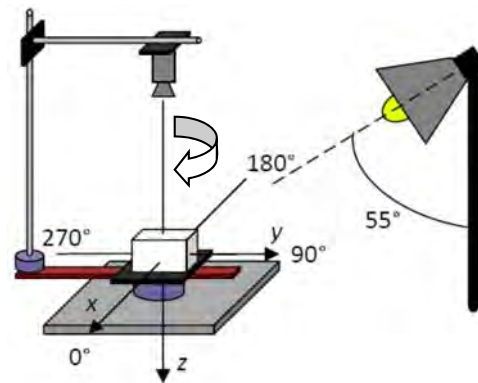


Figure 2.21: Setup for image capture in which both camera and object rotate wrt a stationary light source (Plata *et al*, 2009).

2.3.3 Surface Height Reconstruction

Zhang *et al* (1999) reviewed the literature up to that date and evaluated six algorithms for shape recovery. They noted that even if a Lambertian surface reflectance is assumed and the light source direction is known, and if the intensity can be described perfectly as a function of surface normal and light source direction, the problem is still not simple. This is because it is under-determined, i.e. a description of the surface shape at each point requires a linear equation

with three unknowns, whereas specification of the surface gradient gives a non-linear equation with only two unknowns. Therefore, finding a unique solution to SFS is difficult, and needs additional constraints. Zhang *et al* grouped the existing SFS techniques into four categories: (1) Minimisation approaches, which obtain the solution by minimising an energy function; (2) Propagation approaches, which propagate the shape information from a set of surface points (e.g. singular points) to the whole image; (3) Local approaches, which derive shape based on the assumption of surface type; and (4) Linear approaches, which compute the solution based on linearisation of the reflectance map. Zhang *et al* found that individual algorithms tended to work well for certain images but perform poorly for others, and none of the algorithms gave consistent performance for all images. In general, the minimisation approaches were more robust, while the other approaches were faster. Some reasons for errors were: the reflectance models used in SFS methods were too simplistic; perspective projection should be used instead of orthographic projection; and lights are generally not point sources at infinity. They also noted that the quality of the result can be improved by employing multiple images, by moving either the viewing position or the light source, in order to refine the shape successively. The advantage of moving the light source over moving the viewer is to eliminate the transformation of the depth map (warping) between image frames.

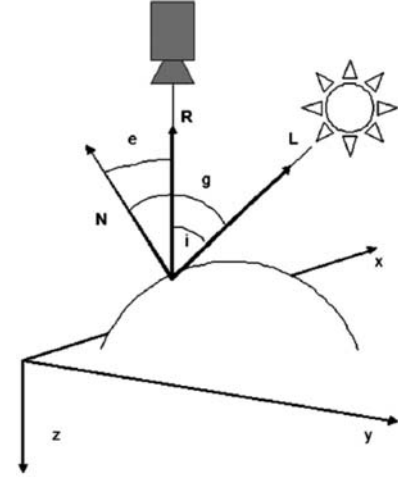


Figure 2.22: Vectors and reflectance angles in photometric stereo:
 R = viewer (camera) vector;
 L = illumination source vector;
 N = normal vector;
 i = angle of incidence;
 e = angle of emittance;
 g = phase angle.
 (from Plata *et al*, 2009)

In the case of a surface with constant albedo, when both observer and light source are far away, surface radiance depends only on surface orientation and not on position in space. Then the image projection can be considered as orthographic and the image equation simplified to:

$$E(x, y) = \mathcal{R}(p(x, y), q(x, y)) \quad (2-34)$$

where $E(x, y)$ is the image intensity at the point (x, y) and $\mathcal{R}(p, q)$ is the reflectance map, i.e. the (normalised) scene radiance of a surface patch with orientation specified by the partial derivatives $p = \frac{\partial z}{\partial x}$, $q = \frac{\partial z}{\partial y}$ of surface height $z(x, y)$ above a reference plane perpendicular to the optical axis. The task in SFS is to find $z(x, y)$, given the image $E(x, y)$ and the reflectance map $\mathcal{R}(p, q)$. Additional constraints, such as boundary conditions and singular points, are needed to ensure that there is a unique solution $\{p(x, y), q(x, y)\}$ that minimises the intensity error over the image:

$$\iint ((E(x, y) - \mathcal{R}(p, q))^2) dx dy \quad (2-35)$$

Horn (1989) proposed a further constraint on the solution to minimise a measure of ‘departure from smoothness’ by summing the squares of the second derivatives:

$$\iint \left(\left(\frac{\partial p}{\partial x} \right)^2 + \left(\frac{\partial p}{\partial y} \right)^2 + \left(\frac{\partial q}{\partial x} \right)^2 + \left(\frac{\partial q}{\partial y} \right)^2 \right) dx dy \quad (2-36)$$

The above metric assumes integrability, i.e. the existence of a continuous surface height function $z(x, y)$ such that $p(x, y) = \frac{\partial}{\partial x} z(x, y)$ and $q(x, y) = \frac{\partial}{\partial y} z(x, y)$. If this is not true, for example because of noise or discontinuities, one seeks at least a surface $z(x, y)$ for which the partial derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ come closest to matching the computed $p(x, y)$ and $q(x, y)$, by minimising:

$$\iint \left(\frac{\partial z}{\partial x} - p \right)^2 + \left(\frac{\partial z}{\partial y} - q \right)^2 dx dy \quad (2-37)$$

This leads to the Poisson equation, expressed in second derivatives, for every pixel:

$$\nabla^2 z = \text{div}(p, q) = \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} = p_x + q_y \quad (2-38)$$

These equations have to be solved iteratively because the system of equations is so large and the reflectance map $R(p, q)$ is typically nonlinear. An important constraint for integrability is that for a stable gradient field over the surface, the curl should everywhere be zero:

$$\nabla \times z = \text{curl}(p, q) = \frac{\partial p}{\partial y} - \frac{\partial q}{\partial x} = p_y - q_x = 0 \quad (2-39)$$

By Green's Theorem, the integral of (p, q) along a closed curve also equals zero:

$$\oint \{p dx + q dy\} = 0 \quad (2-40)$$

This means that in making a complete rotation about any point the increments should sum to zero, otherwise residual errors would propagate across the surface. Fig. 2.23 shows an example of the discrete version of the loop integral of (p, q) for a square circuit of 4x4 pixels. The total accumulated change in elevation around the full loop should be zero if the patches belong to a differentiable, single-valued surface. If a numerical method based on imperfect data is used to compute the orientation of the surface at every point, however, this integral will in general not be exactly equal to zero. One way of imposing a smoothness constraint is to find a solution that minimises the errors in such loop integrals while also satisfying the image irradiance equation (Ikeuchi and Horn, 1981).

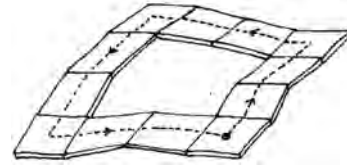


Figure 2.23: Integrability of a surface around a closed loop (from Ikeuchi and Horn, 1981)

Frankot and Chellappa (1988) found an effective way to enforce integrability by projecting the possibly non-integrable estimated surface slopes (p, q) onto a set of integrable slopes (\tilde{p}, \tilde{q}) that satisfy Eq. 2-40, while minimising the distance measure:

$$d\{(p, q), (\tilde{p}, \tilde{q})\} = \iint (p - \tilde{p})^2 + (q - \tilde{q})^2 dx dy \quad (2-41)$$

Fourier basis functions were chosen, so that the integrable surface could be represented by:

$$\tilde{z}(x, y) = \sum \tilde{C}(\omega_x, \omega_y) \exp(j\omega_x x + \omega_y y) \quad (2-42)$$

where $\{\tilde{C}\}$ are the coefficients of the Fourier series expansion of \tilde{z} . It follows that the distance measure in Eq. 2-42 is minimised when:

$$\tilde{C}(\omega) = \frac{-j\omega_x \tilde{C}_x(\omega) - j\omega_y \tilde{C}_y(\omega)}{\omega_x^2 + \omega_y^2} \quad (2-43)$$

Frankot and Chellappa observed an inherent difficulty in SFS, namely that the observed image intensity is a function of the first derivatives of the surface, and therefore low-frequency information for the surface is lost in the image formation process. Hence the reconstructed surface inevitably suffers from low-frequency distortion, the severity of which depends on observation noise characteristics. If additional low-frequency information could be obtained from another source, such as a digital terrain model (in remote sensing applications) or a laser scanner point cloud, then the accuracy of the surface reconstruction could be enhanced.

Epstein *et al* (1996) used the principles of superposition and principal component analysis (PCA) to analyse the space of images generated by one object at a fixed pose. The varying lighting conditions were sampled evenly over the view hemisphere, including extreme lighting angles. The experimental results showed that 5 ± 2 eigenvalues were sufficient, and for Lambertian surfaces the percentage of variance covered by the first five eigenvalues was approximately 90%. Specularities and shadows, though perceptually very salient, contributed little to the variance. They argued that the first three principal components of the linear space of eigenvectors correspond to Lambertian illumination of the object, and that higher-order principal components deal with specularities and sharp shadows. They went on to demonstrate for images of human faces that singular value decomposition (SVD) could be employed to minimise an energy function consisting of the difference between the set of images under multiple illumination sources and the weighted sum of a set of principal components.

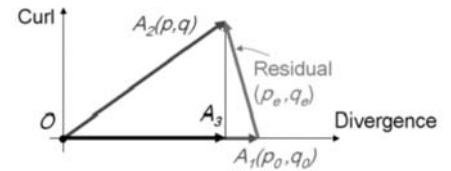


Figure 2.24: The curl-divergence space: a vector in this space represents a gradient field (from Agrawal, 2005).

In general Poisson solutions spread out the discrete non-integrable residual across the entire gradient field to ameliorate the inconsistency. Agrawal *et al* (2005) formulated a method to find the surface \hat{S} which minimised the least-square cost function J :

$$J(\hat{S}) = (\hat{S}_x - p)^2 + (\hat{S}_y - q)^2 \quad (2-44)$$

The method enforced integrability by finding a zero curl gradient field which has the same divergence as the observed non-integrable gradient field, derived from the SFS normals. They visualised a two dimensional curl-divergence space (Fig. 2.24), in which a vector represents a gradient field. All integrable (zero-curl, irrotational) gradient fields lie along the real (X) axis, whereas all divergence free (solenoidal) gradient fields lie along the imaginary (Y) axis. In algorithms that do not obey the zero-curl constraint, a non-integrable field OA_2 is given as an estimate of the true integrable field OA_1 . The residual gradient field is given by A_1A_2 . Enforcing integrability using Poisson reconstruction gives OA_3 as the solution. The Agrawal algorithm tried to estimate the residual gradient field A_1A_2 using the information in the curl to move from OA_2 to OA_1 . It corrected for the curl of the given non-integrable gradient field by solving a linear

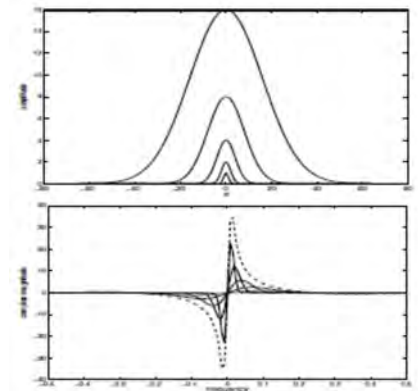


Figure 2.25: Five Gaussians as shapelet functions with height proportional to scale (top), and corresponding transfer functions of gradients (bottom) (Kovesi, 2005)

system. It was non-iterative and had the important property that the errors due to non-zero curl did not propagate across the surface.

Kovesi (2005) projected gradients onto a set of ‘shapelet’ basis functions. The correlation with the basis functions was formulated with respect to slant and tilt, rather than in terms of the surface gradient with respect to x and y , and proved to be highly robust to noise. The use of Gaussian basis functions (Fig. 2.25) implicitly imposed a continuity constraint in the reconstruction, yet also allowed sharp transitions to be represented in the surface via the finer shapelet scales. The method did not infer the locations of discontinuities in the surface or apply special conditions at these points. The correlation process treated slant and tilt separately and made their different roles explicit in the reconstruction process.



Figure 2.26: (left) Ramp test object; (right) normals map as a needle plot (from Wu and Tang, 2010).

Wu and Tang (2010) observe that it is inappropriate to apply the surface integrability constraint uniformly throughout the image domain in cases where surface discontinuities or abrupt changes in normal orientations should be preserved. A continuous surface is not necessarily differentiable every-where. For example, the ramp object in Fig. 2.26 has lines of discontinuity of slope at its sharp edges and around the base where the solid object meets the ground plane.

Wu and Tang propose a two-step optimisation by expectation maximisation (EM): in the first step normals and albedos are optimised (photometric stereo); and in the second step (visible surface reconstruction) a surface is inferred that preserves discontinuities inherent in the normal field.

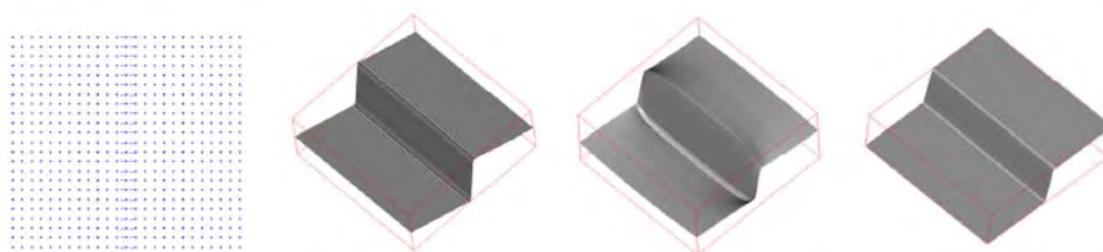


Figure 2.27: (left) Needle plot of normal directions of a step; (centre left) result of Frankot and Chellappa method with Fourier basis functions; (centre right) result of Kovesi method with Gaussian shapelets; (right) result of Ng method with kernel basis functions (from Ng et al, 2010).

Ng et al (2010) noted that discontinuity and height constraints cannot be imposed readily in methods that approximate the surface by sums of basis functions. As the occluding boundary of an object is approached, the gradients $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ approach infinity and the normals lie in the image plane. The same applies for any object with vertical planar surfaces such as a staircase or cuboid. Algorithms that were designed to handle height-fields fail to do so because such normals introduce severe numerical instability, leading to distortion of the reconstructed surface (Fig. 2.27). The observed distortion is in fact due to enforcement of the integrability constraint, which requires changing the orientations of the normals.

The method developed by Ng et al does not enforce discrete integrability and reconstructs a 3D continuous surface from a gradient field or height field, or a combination of both, which

can be dense or sparse. By the use of Gaussian kernel basis functions, they transfer the continuous surface reconstruction problem into high-dimensional space where a closed-form solution exists. An advantage is the avoidance of discretisation and finite approximation, both of which lead to surface distortion, typical of the Fourier or wavelet bases widely adopted by previous approaches. The overall energy function to be minimised is:

$$E(\mathbf{f}) = \|\mathbf{A}\mathbf{f} - \mathbf{b}\|^2 + \lambda\|\mathbf{S}\mathbf{f}\|^2 + \mu\|\mathbf{L}\mathbf{f}\|^2 \quad (2-45)$$

where λ and μ are weights to control the influence of terms for a smoothness constraint, which minimises the first derivative, and Laplacian regularisation, which minimises the second derivative.

Nehab *et al* (2005) developed a hybrid algorithm to produce a surface that optimally conforms to given surface positions from a triangulation scanner and normals from photometric stereo. They reduced the problem to solving a sparse linear system, enabling very efficient computation of optimal surfaces for large meshes.

2.4 Specular Reflectance

2.4.1 Reflectance Models

A key aspect of rendering a surface is modelling how an incident beam of light is reflected from the surface. One idealised case is the perfect mirror, in which all of the incident beam is reflected at an angle θ_r equal to the angle of incidence θ_i (Fig. 2.28). The angles are expressed with respect to the surface normal, which bisects the two beam angles, and all three vectors lie in the same plane. Such reflection is called specular. Another idealised case is where the light falling on the surface is scattered equally in all directions, so that the apparent brightness of the surface to an observer is the same regardless of the observer's angle of view. Such a surface is called Lambertian, and the reflection is called diffuse. The luminance of the reflected light is given by Eq. 2-28 in the previous section.

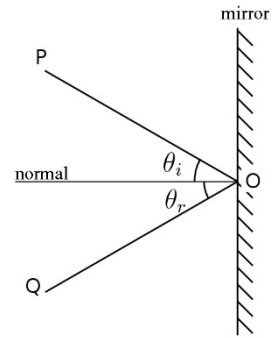


Figure 2.28: Specular reflection from a mirror surface.

In general real surfaces have neither ideal specular (mirror-like) nor ideal diffuse (Lambertian) reflection properties. The reflectance can be represented as a function of angle by the BRDF (Section 2.4.3). In the case of the Lambertian surface, the BRDF is constant in all directions:

$$R(\theta, \varphi) = \rho/\pi \quad (2-46)$$

Gouraud (1971) proposed a technique for shading curved surfaces, to reduce the visibility of Mach bands at the edges of planar facets in rendering solid objects. He noted that a discontinuity in the first derivative of the shading gives the illusion of a small bump along the edge between two differently shaded surfaces. His approach was to retain the polygon approximation of the surface but to use linear interpolation between the colour at the edges of each polygon to achieve continuity of shading. Phong (1975) developed a method of computing the normal for every point on the displayed surface, by linear interpolation between the normal at the edges of each polygon, i.e. applying Gouraud's technique to normals instead of shading. Phong also

introduced a reflectance model that combined ambient, diffuse and specular components (Fig. 2.29), with the specular component spread out in a lobe around the specular direction by using a cosine function raised to a high power:

$$L_p = \rho_a L_a + \rho_d \mathbf{L}_d \cdot \mathbf{n} + \rho_s(\theta_i)(\mathbf{R} \cdot \mathbf{V})^\alpha \quad (2-47)$$

where: ρ_a is the ambient reflectance (constant for the whole scene), L_a is the ambient illumination level (assumed to be equal from all directions), ρ_d is the diffuse (Lambertian) reflectance of the material, \mathbf{L}_d is the incident light vector, \mathbf{n} is the surface normal, $\rho_s(\theta_i)$ is the specular reflectance as a function of angle of incidence θ_i , \mathbf{R} is the direction vector of the reflected ray, \mathbf{V} is the direction vector towards the viewer, and α is a 'shininess constant' for the material. The diffuse term is not affected by the viewer direction. The specular term is large only when the viewer direction \mathbf{V} is aligned with the reflection direction \mathbf{R} .

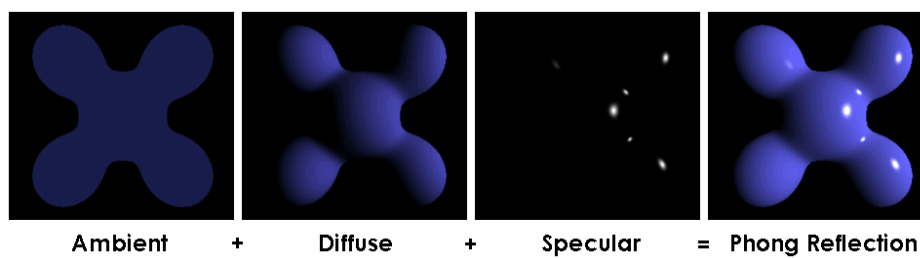


Figure 2.29: Three components summed in the Phong model of surface reflection (Wikipedia).

Blinn (1977) noted that the empirical Phong model could be improved by application of the Torrance-Sparrow (1967) model of light from real surfaces, which are assumed to be composed of a collection of randomly orientated mirror-like micro-facets. If the surface were a perfect mirror, light would only reach the eye of the viewer if the surface normal \mathbf{N} pointed halfway between the illumination direction \mathbf{L} and the viewer direction \mathbf{V} . He named this the direction of maximum highlights \mathbf{H} (Fig. 2.30).

The macroscopic specular component of the reflected light is assumed to arise from those facets oriented in the direction of \mathbf{H} , whereas the diffuse component comes from multiple reflections between the facets and from internal scattering. Blinn's model for the specular component combined factors for the directional distribution function of the micro-facets, the amount of self-masking by the micro-facets, and Fresnel reflection for grazing angles.

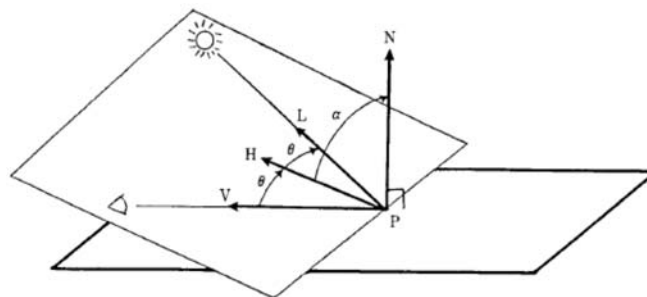


Figure 2.30: Principal vectors in the geometry of reflection (from Cook and Torrance, 1982).

These give a noticeable improvement over the Phong model for rendering non-metallic and edge-lit objects. Cook and Torrance (1982) developed a reflectance model for rough surfaces based on geometrical optics, applicable to a broad range of materials, surface conditions and lighting situations. The model predicted the directional distribution and spectral composition of the reflected light, with a procedure for calculating the R,G,B values from the spectral power distribution. They introduced the concept of splitting bidirectional reflectance into specular and diffuse components, and noted that these can have different colours if the material is not

homogeneous. They introduced a directional reflectance distribution function developed by Beckmann (1965), applicable to a wide range of materials and surface conditions:

$$D = \frac{1}{\alpha^2 \cos^4 \omega} e^{-((\tan \omega)/\alpha)^2} \quad (2-48)$$

where ω is the angle between \mathbf{N} and \mathbf{H} , and α is the rms slope of the micro-facets. The results give more realistic prediction of the spectral reflectance distribution from metallic surfaces, especially at grazing angles (Fig. 2.31).

Ward (1992) presented a BRDF reflectance model based on a Gaussian lobe, stressing its physical plausibility and ease of use. He successfully fitted the model to measurements of various surfaces and presented an equation for directional sampling of BRDF, which is important for Monte Carlo applications such as stochastic ray tracing. Ward (1994) presented the generalised RADIANCE framework for rendering computer graphic scenes, enabling various lighting and surface reflectance algorithms to be used for visualisation by ray-tracing. The algorithm makes a recursive evaluation of the following integral equation at each surface point:

$$L_r(\theta_r, \varphi_r) = L_e(\theta_r, \varphi_r) + \int_0^{2\pi} \int_0^\pi L_i(\theta_i, \varphi_i) \rho_{bd}(\theta_i, \varphi_i; \theta_r, \varphi_r) |\cos \theta_i| \sin \theta_i d\theta_i d\varphi_i \quad (2-49)$$

where: L_e is emitted radiance (for self-luminous objects), L_i is incident radiance, L_r is reflected radiance, and ρ_{bd} is the bidirectional reflectance-transmittance distribution. This formulation of the rendering problem is well-suited for ray tracing because it expresses outgoing radiance in terms of incoming radiance over the projected sphere without any explicit reference to the model geometry. No restrictions are placed on the number or shape of surface elements.

Kautz and McCool (1999) introduced a separable decomposition of BRDFs to implement arbitrary reflectances from point sources on existing graphics hardware. 2D texture mapping and compositing operations were used to reconstruct samples of the BRDF at every pixel at interactive rates. Lafortune et al (1997) generalised the specular cosine lobe of the Phong model to include non-Lambertian diffuse reflection, specularity at grazing angles, off-specular reflection, retro reflection and anisotropic reflection. The main drawback of the original cosine lobe model is that its behaviour contrasts sharply with the reflectance behaviour of most real surfaces, which appear more specular at grazing angles, because the apparent roughness decreases. The definition of the generalised cosine lobe model is:

$$f_r(\mathbf{L}, \mathbf{V}) = \frac{k_d}{\pi} + k_s (\mathbf{L}^T \mathbf{M} \mathbf{V})^n \quad (2-50)$$

where: \mathbf{L} is the unit vector to the light source, \mathbf{V} is the unit view vector, k_d is the diffuse coefficient, k_s is the specular coefficient, \mathbf{M} is a 3×3 symmetric matrix, and n is the specular exponent. Applying singular value decomposition, \mathbf{M} can be expressed as $\mathbf{Q}^T \mathbf{D} \mathbf{Q}$ where \mathbf{Q}

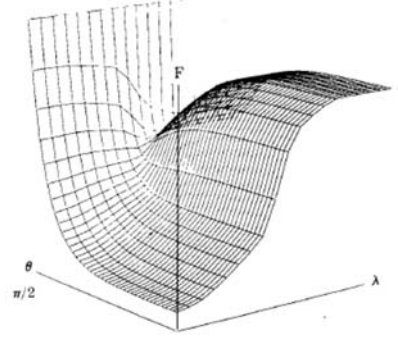


Figure 2.31: Reflectance of a copper mirror as a function of incident angle of illumination θ and wavelength λ (from Cook and Torrance, 1982).

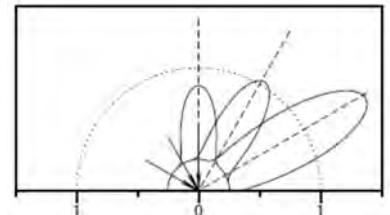


Figure 2.32: Polar plots of the generalised cosine lobe model (from Lafortune et al, 1997)

transforms the axes to a new local coordinate system aligned to the normal and to the principal directions of anisotropy, in which \mathbf{M} simplifies to the diagonal matrix \mathbf{D} . The authors claim that function has the advantages that: (a) it interpolates sparse data naturally; (b) it handles noise in the raw reflectance data gracefully; and (c) it can capture sharp reflectance lobes without suffering from small spurious errors in the data.

More complex BRDF models have continued to be developed for specialised materials, but in recent years attention has shifted, first to deriving parameters for the BRDF models from measured data (or extracting parameters from images) and then to using BRDF data measured

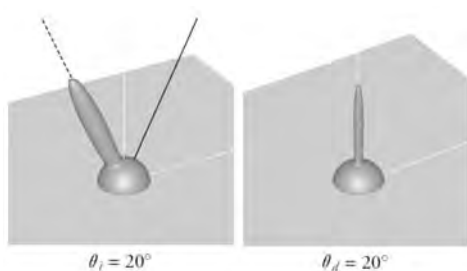


Figure 2.33: The Cook-Torrance-Sparrow BRDF in conventional axes (left) and transformed to difference about the half-angle (right) (from Rusinkiewicz, 1998).

from real materials. The latter approach has the advantage that measured data can capture a greater variety of real-world optical phenomena with greater accuracy than is possible with analytic models (Weyrich et al, 2008). A significant problem in using measured data is the high dimensionality of the BRDF function, with four dimensions for the incident and reflected angles and a fifth for wavelength, which produces massive datasets if each dimension is adequately sampled. Rusinkiewicz

(1998) proposed a method for making the decomposition of BRDFs into basis functions more efficient, by a reparametrisation of the BRDF as a function of the half-angle (i.e. the angle halfway between the directions of incidence and reflection) and a difference angle, instead of the usual angles of incidence and reflection. Because features in common BRDFs, including specular and retroreflective peaks, are aligned with the transformed coordinate axes, the change of basis simplifies the representation and reduces storage requirements for a large class of BRDFs.

Matusik et al (2003) tried to reduce the number of measurements by investigating: (1) What sampling frequency is required to measure BRDF adequately over its domain? (2) What set of basis functions is optimal to represent any isotropic BRDF? (3) Can new BRDFs be represented as linear combinations of these basis functions? They found that typical BRDFs exhibit high frequencies only in very specific regions of parameter space (e.g. near specular peaks), and hence only these regions require dense radiance sampling. Otherwise, BRDFs are smooth and slowly varying over most of their domain, and thus require fewer samples for accurate reconstruction away from the peak. This has enabled optimal BRDF sampling strategies for gonioreflectometers and scanning cameras.

Shafer (1985) introduced the dichromatic reflectance function in the context of a method for analysing a colour image to determine the amounts of interface (specular) and body (diffuse) reflection at each pixel (Fig. 2.34). The interface reflection represents the highlights from the original image, and the body reflection represents the original image with the highlights removed. The dichromatic model is defined by:

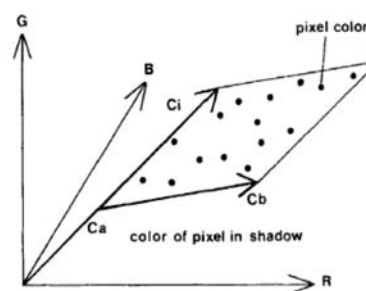


Figure 2.34: Distribution of pixel values lying on a parallelogram in a plane in RGB colour space (from Shafer, 1985)

$$L(\lambda, i, e, g) = L_i(\lambda, i, e, g) + L_b(\lambda, i, e, g) = m_i(i, e, g)c_i(\lambda) + m_b(i, e, g)c_b(\lambda) \quad (2-51)$$

where L_i is the light reflected at the interface and L_b is the light reflected from the surface body. Each of these components can be decomposed into two parts: a relative spectral power distribution $c_i(\lambda)$ or $c_b(\lambda)$, which depends only on wavelength λ and is independent of geometry, and a geometric scale factor m_i or m_b , which depends only on geometry and is independent of wavelength. The geometric parameters (i, e, g) are the angles of incidence, reflection and viewing. If the chromaticities of both components are known, a chromatic pixel can be decomposed into the two components, and the parameters can then be estimated. This model holds for inhomogeneous, dielectric, opaque surfaces.

Schlüns and Witting (1993) presented an implementation of the dichromatic model using sparse images based on the separation of the reflection components, in which they estimated the diffuse chromaticities from a histogram, requiring the surfaces to have uniform colour. They proposed a two-stage process to eliminate the highlights: first a refined specular separation, and second a conventional photometric stereo method for Lambertian surfaces. No restrictions were placed on the roughness of the material and, by using the measured values for each colour channel, an error minimisation was possible without additional light sources.

Sato *et al* (1997) proposed a colour analysis method with the dichromatic reflection model. They reconstructed an object's surface shape by merging multiple range images and a sequence of colour images of the object, then estimated parameters of a reflection model. First the diffuse and

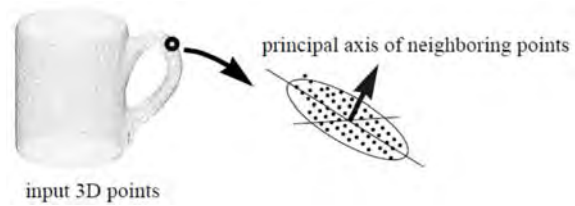


Figure 2.35: Surface normal estimation from a cloud of 3D points (from Sato *et al*, 1997).

specular reflection components were separated from the colour image sequence, and then reflectance parameters of each reflection component were estimated separately. The surface normal at each point was computed as an eigenvector of the covariance matrix of the neighbouring 3D points, specifically the eigenvector associated with the eigenvalue of smallest magnitude (Fig. 2.35). This approach enabled the estimation of reflectance properties of real objects with surfaces showing specularity as well as diffusely reflected highlights.

2.4.2 Gloss

Usually the appearance of coloured samples varies with the directions of illumination and viewing because most surfaces exhibit gloss, which results in some of the incident light being reflected directly from the top surface without passing through the material body. This is called the specular component and, for a smooth and uniform surface, is usually maximum when the angle of reflection is equal to the angle of incidence. In the case of a perfect mirror surface all of the incident ray of light would be reflected at this single angle. Conversely, an idealised perfectly matte surface with no gloss, a Lambertian surface, would reflect equally at all angles.

Harrison (1945) recognised that all real materials reflect both specular and diffuse components, and that the intensity is distributed over a range of angles of reflection. He used polar plots to show the intensity of the reflected radiation at all angles in the half-plane above the surface containing the incident and reflected rays (Fig. 2.36). The glossier the surface, the more sharply

defined is the specular 'spike' in the distribution around the primary reflection angle. A well rounded curve indicates a generally diffuse reflection, which in the case of a perfect Lambertian surface would be a circle. The complexity of gloss as a visual phenomenon is revealed by the many definitions in use:

- A surface shine or lustre; a smooth finish; a deceptively attractive appearance (OCED, 1996).
- Angular selectivity of reflectance, involving surface reflected light, responsible for the degree to which reflected highlights or images of objects may be seen as superimposed on a surface (ASTM, 2002).
- Sudden and often local increases in brightness of a sheet of paper when it is illuminated by a small source of light and viewed at certain definite angles (Harrison, 1945).
- The appearance attribute corresponding to the illuminance with which lights are seen to be specularly reflected (Christie, 1979).
- The relative amounts of diffuse reflectance to specular reflectance (Tilley, 1999).
- The property of a surface which involves specular reflection and is responsible for lustrous or mirror-like appearance (Hunter & Harold, 1987).
- Ability or capacity of a surface to direct reflected light (Trezza & Krochta, 2001).
- The power of a surface to reflect light specularly (Silvennoinen et al, 2008).

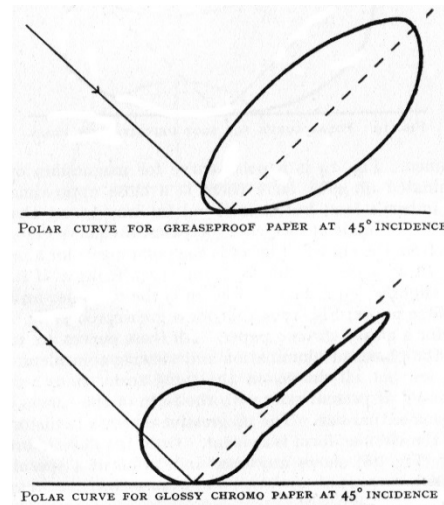


Figure 2.36: Polar reflection curves for two types of paper (from Harrison, 1945).

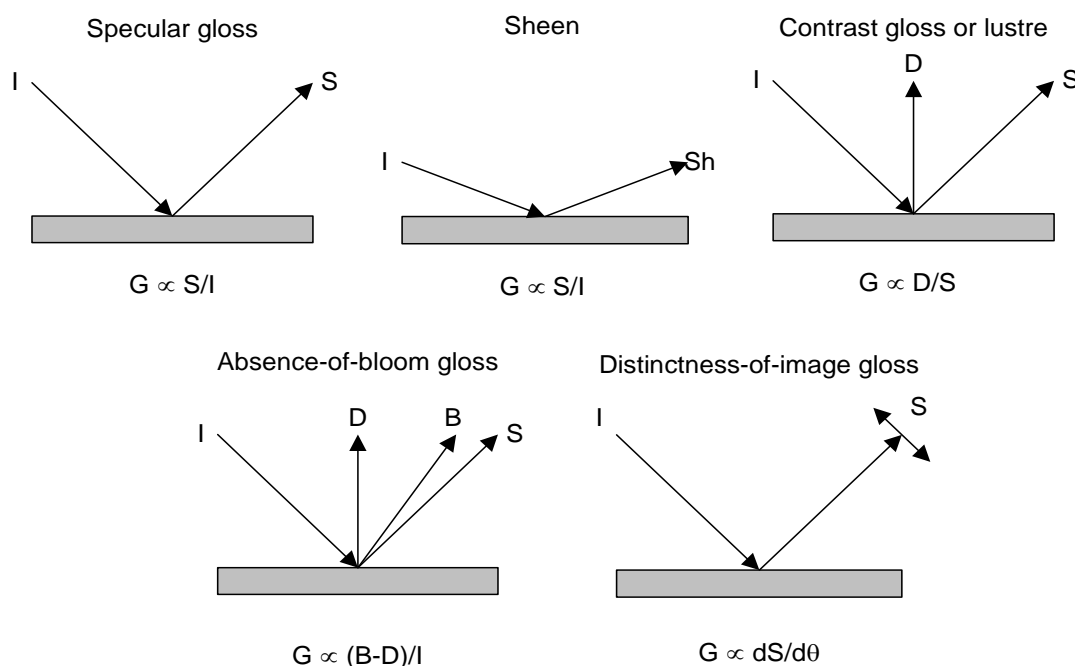


Figure 2.37: Hunter's five types of gloss (Hunter and Harold, 1987).

There are two reasons for this diversity of definition. The first reason is that gloss, like colour, is a physical process that produces a perceptual response in an observer. Hence there is a difference in approach between measurement of patterns of radiation and psychophysical experiment into how it appears in a given viewing environment. The second reason is that gloss is not a single attribute but a family of related appearance characteristics. Hunter and Judd (1939) recognised that the perception of gloss involves more than just specular reflection, and in addition to specular gloss they defined (Fig. 2.37):

- sheen as gloss at grazing angles of incidence and viewing;
- contrast gloss or lustre as the ratio of the specularly reflected light and that diffusely reflected normal to the surface;
- absence-of-bloom as a measure of the absence of haze or a milky appearance adjacent to the specularly reflected light;
- distinctness-of-image as the sharpness of specularly reflected light.

Ferwerda *et al* (2001) conducted psychophysical experiments and used multidimensional scaling (MDS) techniques to analyse the observer data. They found that gloss in painted surfaces could be well represented by two dimensions: apparent contrast and distinctness of reflected image.

2.4.3 Bidirectional Reflectance Distribution Function (BRDF)

In general the reflection from a surface at every angle in the hemisphere with axis normal to the surface depends on the angle of incidence of the ray of light. The distribution of intensity over all possible angles of incidence and reflection is known as the bidirectional reflectance distribution function (BRDF). This gives the ratio of reflected radiance exiting along angle ω_r to the irradiance incident on the surface from direction ω_i with respect to the surface normal n (Fig. 2.38). Each direction ω is itself parameterised by an azimuth angle φ and a zenith angle θ , therefore the BRDF as a whole is four-dimensional, with units of sr^{-1} . The BRDF was first defined by Nicodemus (1965), who expressed the relationship between incident radiance N_i and reflected radiance N_r by the fundamental equation:

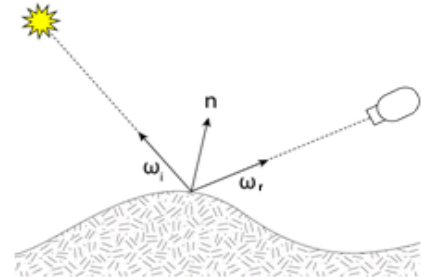


Figure 2.38: Scenario for definition of BRDF. All vectors are unit length. ω_i points toward the light source. ω_r points toward the viewer (camera). n is the surface normal (Wikipedia).

$$N_r(\theta_r, \varphi_r) = \int \rho'(\theta_i, \varphi_i, \theta_r, \varphi_r) N_i(\theta_i, \varphi_i) d\Omega'_i \quad (2-52)$$

where the integral is taken for small elements of solid angle $d\Omega$ over all angles of the hemisphere (Fig. 2.39). Ward (1992) restated the above equation as:

$$L_r(\theta_r, \varphi_r) = \int_0^{2\pi} \int_0^{\pi/2} L_i(\theta_i, \varphi_i) \rho_{\text{bd}}(\theta_i, \varphi_i, \theta_r, \varphi_r) \cos \theta_i \sin \theta_i d\theta_i d\varphi_i \quad (2-53)$$

using the conventional symbols L_i and L_r for the incident and reflected luminous intensity and making the integral explicit over the two angles. He labelled the BRDF function ρ_{bd} to

emphasise that it is bidirectional, because the physics of light dictates that the function will return the same value if the incident and reflected directions are reversed.

In reality, BRDF is also wavelength dependent (see Fig. 2.31), and to account for effects such as iridescence or luminescence the dependence on wavelength λ must be made explicit. Sun (2002) developed a physical reflection model based on a micro-facet scattering geometry, with both angle and wavelength as parameters, which successfully demonstrated the glossy appearance of both pure and enamelled copper plates. Other models allow the reflectance function to vary spatially over the surface of the object, abbreviated as SVBRDF, and to include the internal subsurface scattering of the radiation. The Bidirectional Texture Function (BTF) includes non-local effects such as shadowing, masking and inter-reflections, defined at each point on the surface (Dana et al, 1999).

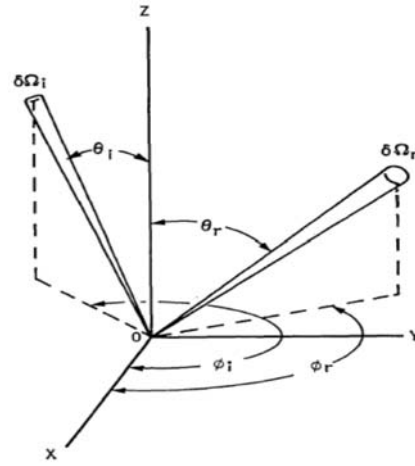


Figure 2.39: Geometry of incident and reflected beams relative to normal (from Nicodemus, 1965).

2.4.4 Specular Distribution Models

In the computer graphics community, motivated by the desire for realistic rendering from computational models, there have been many proposals for a specular distribution function of ω , the angle from the specular centre. A physically-based model can be derived by recognising that the surface of any real material has a microstructure, approximated by an arrangement of micro-facets (Fig. 2.40). These are assumed to be randomly orientated and their distribution can be modelled as a normal probability density function $D(\omega)$. This provides a statistical distribution of slopes for the mirror-like facets on the surface. Given a direction $\omega(\theta, \varphi)$ and an infinitesimal solid angle $d\omega$ centred around this direction, the probability for the normal \mathbf{n} of a micro-facet to be inside the cone $(\omega, d\omega)$ is $D(\omega)(\omega \cdot \mathbf{n})d\omega$. In order to be physically plausible, $D(\omega)$ should obey several constraints: it should be positive; the projected area of the micro-facets in any direction should be equal to the projected area of the macro-surface; and it should be normalised by $\int D(\omega)(\omega \cdot \mathbf{n})d\omega = 1$.

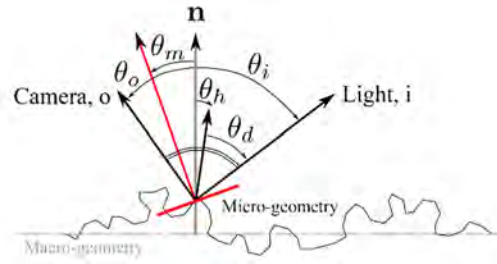


Figure 2.40: Micro-facet geometry for the Cook-Torrance model (from Bagher et al, 2012).

The simplest function is a Gaussian $e^{-(\omega/a)^2}$ which is symmetrical, continuous, integrable, and asymptotic to zero for increasing ω , but this does not resemble the distribution observed in real materials. An alternative was formulated by Beckmann and Spizzichino (1963) for surface scattering based on the Kirchhoff scalar diffraction theory, giving a function $e^{-((\tan \omega)/a)^2} / \cos^4 \omega$ (see Eq. 2-48). Blinn (1977) preferred the Trowbridge-Reitz (1975) form, based on $\cos^{-4}(\omega)$. Ward (1992) chose a variant of the Gaussian based on $e^{-\tan^2(\omega)}$. Watson and Raven (2001) described an empirical model called Mopaf, with a sum of Lorentzian functions to model BRDF,

based on $1/(1 + (\omega/a)^2)$. In the context of a transformation to a halfway specular coordinate system Edwards *et al* (2006) proposed a simpler ‘hump function’ based on $\left(1 - \left(\frac{\omega}{a}\right)^2\right)^n$.

More recently, Bagher *et al* (2012) studied the BRDF measurements of shiny materials and showed that they do not follow the Beckmann distribution. Instead the micro-facets are more likely to be aligned with the surface normal, resulting in a sharper peak. They also noted the broader flanks in real distributions and proposed a function based on $e^{-\omega/a}/\sqrt{\omega - a^2}$.

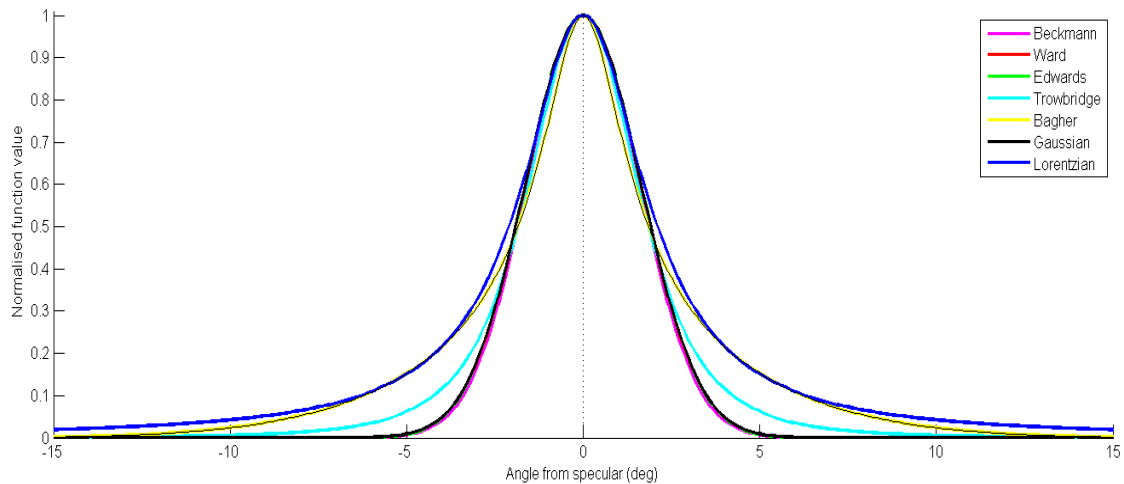


Figure 2.41: Various proposed specular functions vs angle from the specular peak.

Plotting these functions together, with parameters adjusted to give the same width at half maximum amplitude (Fig. 2.41), shows that for small angles the Beckmann, Ward and Edwards functions are almost identical to a Gaussian and that they fall quickly to zero with little amplitude in the flanks beyond 5° . The Trowbridge & Reitz function is broader, extending to about 10° , and the Bagher function is broader still, extending to about 15° . The Lorentzian function has the broadest flanks, extending out to beyond 45° which makes it a promising candidate for modelling the angular distribution of specular reflectance.

For evaluation of BRDF and specular distribution models, measurements of real materials are available. The MERL-MIT database has a hundred isotropic materials (Matusik *et al*, 2003), and the Cornell reflectance datasets include various types of paints, acrylic and felt (Greenberg *et al*, 1997). Measurements of anisotropic materials, including brushed aluminium, velvet and satin by Ngan *et al* (2005) showed that simple parametric reflectance models could not represent their appearances faithfully in many cases. They developed a method for a sampled micro-facet distribution computed from measurements. Filip and Vávra (2014) recently announced a database of 150 BRDFs for a wide range of materials, the majority exhibiting anisotropic behaviour, using an adaptive sampling method with template-based analysis of the measured BRDFs.

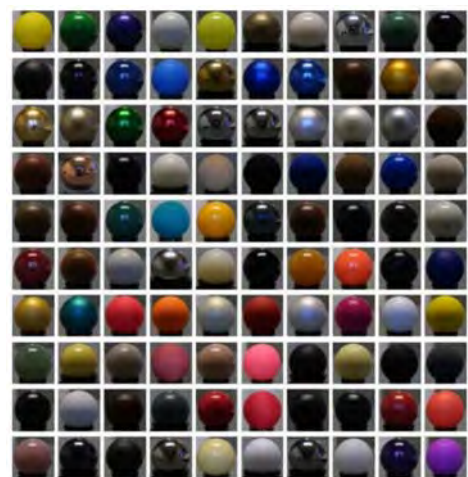


Figure 2.42: 100 BRDF samples in the MERL-MIT database (from Matusik *et al*, 2003)

2.5 Colour Specification

2.5.1 The CIE System

Three elements are needed to specify the colour of an object: the source of illumination, the reflection/transmission properties of the surface, and the sensitivity of the observer (Fig. 2.43). In 1931 the CIE (Commission Internationale de l'Éclairage) standardised a method for colour specification, including standard light sources and a standard observer. This has become the basis of all colour measurement (CIE, 2004a).

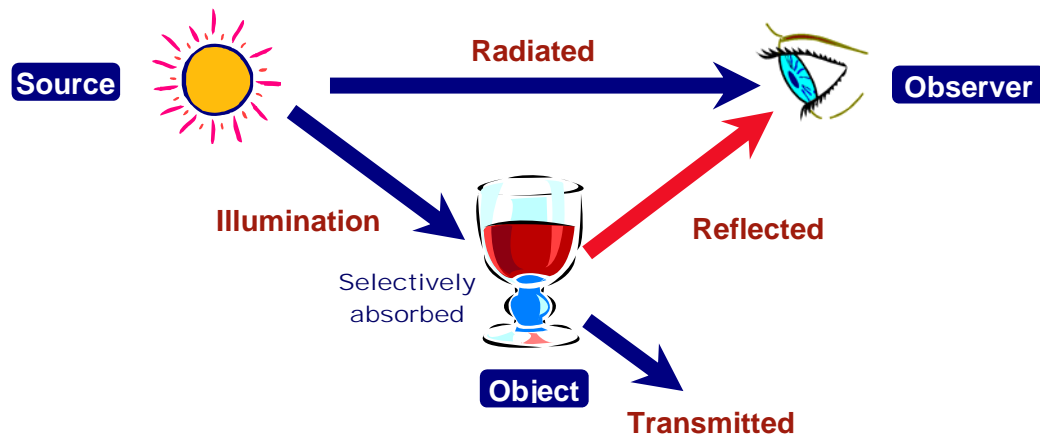


Figure 2.43: The three elements of colour specification: a light source, an object with a selectively reflecting or transmitting surface, and an observer.

All three elements (illumination, object and observer) must be characterised in spectral terms, i.e. their behaviour defined at each wavelength throughout the visible spectrum from 380 to 760 nm. For a source of illumination, this means specifying the power of the radiation incident on the surface as a function of wavelength through the spectral power distribution.

Reflection from a surface is characterised by the reflectance factor, i.e. the ratio of reflected to incident power, as a function of wavelength. For the great majority of real-world and synthetic colorants, the spectral reflectance distribution measured by instruments is smooth and well-behaved, as shown in Fig. 2.44. This may not mean that the spectrum itself is perfectly smooth, however, because of the double integrating effect of a spatial aperture (typically 6-10 mm diameter) and the wavelength interval, typically 5-10 nm.

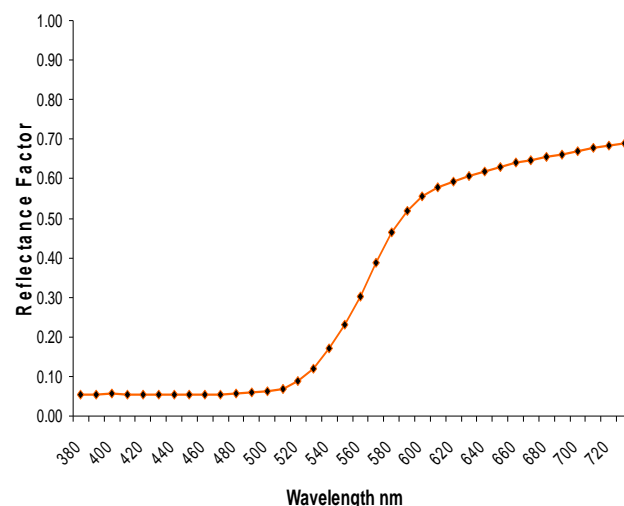


Figure 2.44: Reflectance factor of the orange patch in the Macbeth ColorChecker chart, measured with a GretagMacbeth i1Pro spectrophotometer at 10 nm wavelength intervals.

The perception of brightness as a function of wavelength is specified by the spectral luminous efficiency function, which was standardised by CIE in 1924. For daylight (photopic) vision, the function is denoted $V(\lambda)$ and peaks at $\lambda=555$ nm, whereas for night (scotopic) vision the alternative $V'(\lambda)$ peaks at $\lambda=507$ nm (Fig. 2.45). A separate function $V_{10}(\lambda)$ is used for large (nominally 10°) visual fields. A revised luminosity function $V^*(\lambda)$ has recently been proposed (Sharpe *et al*, 2005) to correct for the known deficiency in the original $V(\lambda)$ curve for wavelengths less than 480nm. Photopic luminance L_v is related to radiance L by:

$$L_v = 683 \int_{380}^{760} L \cdot V(\lambda) d\lambda \quad (2-54)$$

and has units $\text{lm} \cdot \text{m}^{-2} \cdot \text{sr}^{-1} = \text{cd} \cdot \text{m}^{-2}$. The factor 683 arises from the definition of the lumen as an SI unit, and the assignation of 683 lumens/watt for the spectral luminous efficiency (K_m) of monochromatic radiation having a wavelength of 555 nm. For scotopic luminance L'_v the above formula uses $V'(\lambda)$ and normalising constant $K'_m = 1700$.

Central to the CIE system of colorimetry is the specification of the responsivity functions of a 'standard observer', derived from the responses of only 17 observers in the classic colour matching experiments by Wright (1929) and Guild (1931). The three curves (Fig. 2.46) represent the mean sensitivity over the population of non-deficient observers, and collectively are known as the CIE Standard Observer.

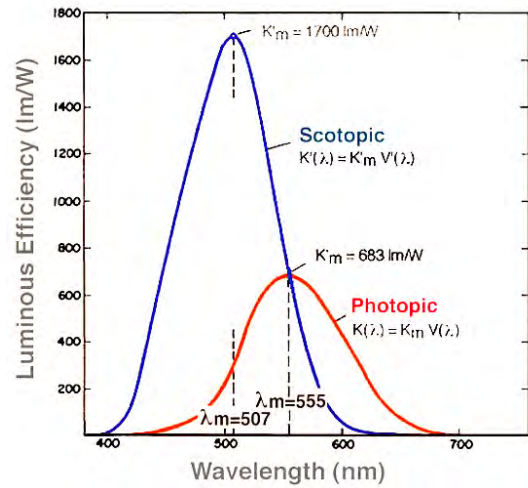


Figure 2.45: Spectral luminous efficiency functions for photopic and scotopic vision. (from <http://webvision.med.utah.edu>)

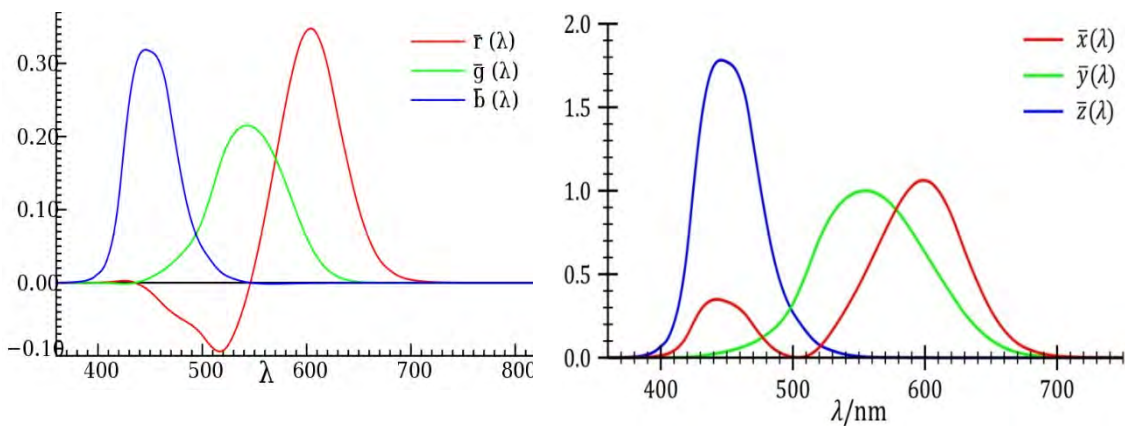


Figure 2.46: (left) Colour matching functions of RGB intensities required to match monochromatic test light; (right) Colour-matching functions for the CIE 2-degree 1931 Standard Colorimetric Observer.

The curves on the left represent the intensity value selected by the observer for each of three primary lights (red, green and blue) in a mixture to obtain a visual match with a monochromatic test light at each wavelength throughout the visible spectrum (Fairman *et al*, 1997). The curves on the right are a transformation to give non-negative functions where $\bar{y}(\lambda)$ is equal to $V(\lambda)$. For most imaging applications the data for the so-called 2-degree observer is used, because it

represents the foveal response of the eye to small-area stimuli in a scene. For applications in architecture and interior design the data for the 10-degree Supplementary Observer is used (Judd, 1968), particularly because of the greater sensitivity to blue in the peripheral visual field. The colour stimulus is quantified by multiplying together the power of the source by the reflectance of the object by the sensitivity of the observer at each wavelength and integrating over the visible spectrum. More formally, if the illumination source power is represented by $S(\lambda)$, the reflectance of the surface by $R(\lambda)$, and the observer's colour matching functions by $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ for each wavelength λ , then the visual stimulus of a colour is obtained by computing:

$$\begin{aligned} X &= k \int S(\lambda)R(\lambda)\bar{x}(\lambda)d\lambda \\ Y &= k \int S(\lambda)R(\lambda)\bar{y}(\lambda)d\lambda \\ Z &= k \int S(\lambda)R(\lambda)\bar{z}(\lambda)d\lambda \end{aligned} \quad (2-54)$$

$$\text{where } k = 100 / \int S(\lambda)\bar{y}(\lambda)d\lambda$$

The three scalar values defined by Eq. 2-54, known as the tristimulus values X, Y, Z , uniquely define the perceived colour as viewed in the given environment. Because the $\bar{y}(\lambda)$ colour matching function is identical to the photopic spectral luminous efficiency function $V(\lambda)$, the Y tristimulus value is equal to the luminance of the stimulus. The scaling factor k is chosen so that $Y=100$ for a perfectly reflecting diffuser. The X and Z tristimulus values give an indication of the relative intensities of long (red) and short (blue) wavelengths in the stimulus, but have no perceptual correlates. For many purposes it is useful to calculate the chromaticity coordinates, which normalise the tristimulus values with respect to the total power of the stimulus:

$$\begin{aligned} x &= X / (X + Y + Z) \\ y &= Y / (X + Y + Z) \end{aligned} \quad (2-55)$$

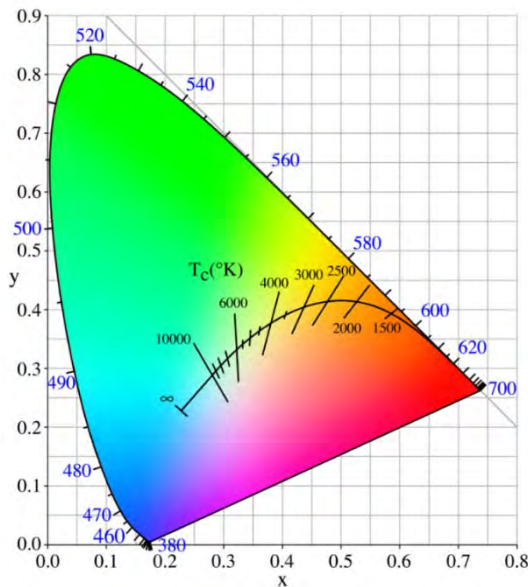


Figure 2.47: CIE x, y chromaticity diagram, showing the spectral locus with associated wavelengths (outer boundary) and the locus of blackbody radiators with correlated colour temperatures in Kelvin (centre).

The original CIE 1931 x, y chromaticity diagram provided a convenient way of mapping coloured samples relative to the 'white' light used to illuminate them. The boundary of this diagram, the spectrum locus, represents the chromaticity coordinates of monochromatic stimuli throughout the spectrum (Fig. 2.47). The chromaticity diagram is actually quite non-uniform perceptually, because the sensitivity of the eye to colour differences varies in different regions, particularly in the green where quite large coordinate changes are not very obvious visually. To correct this, the diagram can be altered in shape so that equal distances more nearly represent equal perceptual steps. This

led to the further recommendation of the CIE 1976 u',v' diagram as a more uniform chromaticity representation. A good description is given by Hunt (1998).

Colour is trichromatic because there are three classes of photoreceptors in the retina. Two-dimensional chromaticity diagrams are only strictly applicable to colours having the same luminance. In general, colour stimuli vary in both chromaticity and luminance, and some method of specifying the colours in a three-dimensional space is required. To meet this need, the CIE recommended in 1976 the use of two alternative uniform colour spaces, designated CIELUV and CIELAB, and the latter has found general favour for reflective surface colours.

$$\begin{aligned} L^* &= 116 f\left(\frac{Y}{Y_n}\right) - 16 \\ a^* &= 500 \left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right] \\ b^* &= 200 \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right] \end{aligned} \quad (2-56)$$

where:

$$\begin{aligned} f(x) &= x^{\frac{1}{3}} & x > \left(\frac{24}{116}\right)^3 \\ f(x) &= (841/108)x + \frac{16}{116} & x \leq \left(\frac{24}{116}\right)^3 \end{aligned}$$

Correlates of the perceptual quantities chroma C_{ab}^* and hue h_{ab} can be derived as polar coordinates from the Cartesian a^* and b^* . Together with L^* , these form a cylindrical 3D colour space (Fig. 2.48).

$$C_{ab}^* = \sqrt{(a^*)^2 + (b^*)^2} \quad (2-57)$$

$$h_{ab} = \arctan(b^*/a^*) \quad (2-58)$$

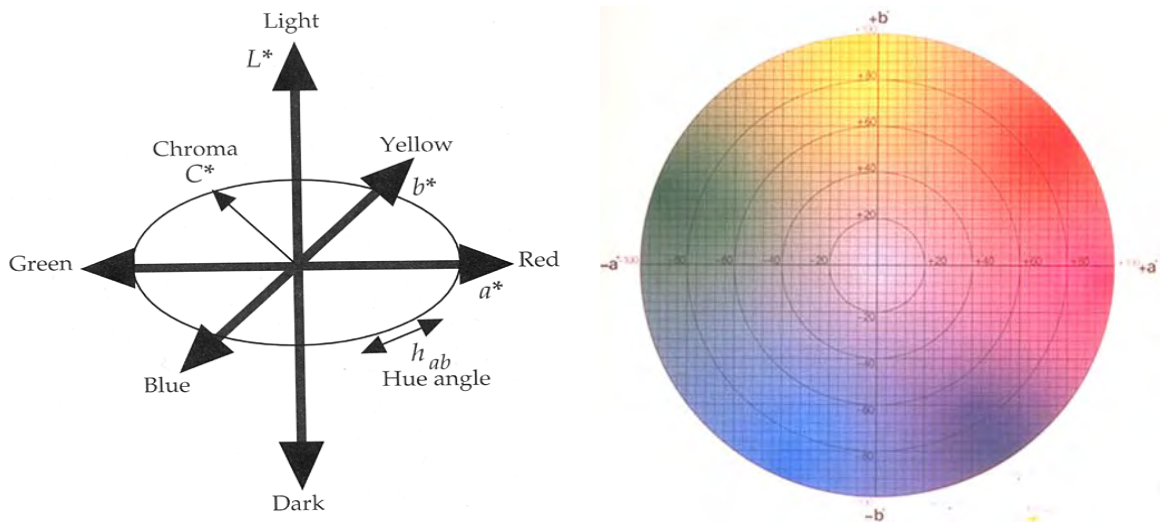


Figure 2.48: (left) Three dimensions of the CIELAB colour space: light-to-dark, red-to-green and yellow-to-blue; (right) the a^* - b^* plane, with achromatic grey at the centre.

Although the CIE system of colorimetry has been remarkably successful as a basis for industrial colour measurement and specification, it assumes the very simple scenario of an observer

performing a colour matching experiment under carefully controlled (and rather artificial) viewing conditions. The numerous inherent assumptions of the CIE system limit its applicability:

- There is a single source of illumination, uniform across the visual field;
- Tristimulus values are always referred to the light source. Sources with different spectral power distributions (SPDs) may produce different tristimuli from the same patch (the phenomenon of illuminant metamerism);
- There is no fluorescence, i.e. contribution to reflected power from stimulation of the object by wavelengths outside the visible spectrum;
- The observer is fully adapted to the viewing conditions and is not colour deficient;
- The scene's dynamic range is limited to approximately 100:1 so that no points are above the glare limit or below the black detection threshold for the observer's adaptation state;
- Illumination levels are above the upper mesopic threshold, i.e. there is no response from the rod cells in the retina of the observer;
- Individual variations between observers, leading to observer metamerism, are ignored;
- Colour patches subtend a visual angle of between 1 and 4 degrees – outside this range the standard 2-degree colour matching functions do not apply;
- Although the tristimulus values X, Y, Z may be absolute, their normalisation in the chromaticity and CIELAB formulae yield relative values from which it is not possible to deduce luminance;
- Background and surround effects are ignored, and the coloured patch is assumed to be viewed against a uniform achromatic field of similar luminance (avoiding simultaneous contrast effects);
- A unique white exists in the scene, or can be inferred, which is taken as the anchor point for vision and the reference value for computation;
- The stimuli are stationary and steady-state, so that temporal effects do not affect perception;
- Viewing geometry, including the distance and relative angles of the light source and observer to the sample, is controlled.

The CIE system works well for simple objects or static scenes with matte surfaces, viewed by a fully adapted observer in photopic conditions with a single source of illumination and a neutral surround. This is very appropriate for viewing a paint sample or a print in a viewing cabinet, or a television screen in normal room lighting. A mathematical model of colour vision, first proposed by Hunt (1982), has provided a basis for extending the CIE system to predict colour appearance using illuminants other than daylight. The input data includes the colour and absolute level of the illumination and the colour of the background to the sample. The output of the model includes correlates of appearance, such as hue, lightness and chroma of the sample. An extensively developed version of the model, CIECAM02, was recommended by the CIE

(Moroney *et al*, 2002). Colour appearance models are evolving into image appearance models, to augment traditional colour appearance with spatial vision attributes and colour difference metrics (Pointer, 2006).

Most colour appearance models, including CIELAB and CIECAM02, require the definition of a white point (X_n, Y_n, Z_n) in Eq. 2-56, to which it is assumed the eye is totally adapted. In many real viewing situations where there are multiple sources of illumination and patches of light and shade, however, this does not apply. Adaptation is never complete when the eye is viewing two scenes, objects or images alternately, by moving the focus of attention from one to the other. If these two scenes have different white points then the eye has to adapt dynamically to some intermediate state and will not be totally adapted to either.

An example of such a situation occurs in the graphic arts, where a soft-copy proof may be viewed on a computer display with a white point typically having a correlated colour temperature of 9300K, or perhaps 6500K if calibrated. A hard copy reproduction is typically viewed in an adjacent booth (Fig. 2.49) with a D50 light source having a correlated colour temperature of 5000K. Experimental studies have shown that observers in such a situation adapted approximately 40% to the monitor and 60% to the viewing booth (Kato *et al*, 1998) or the surrounding ambient field (Oskoui and Pirotta, 1999). Such mixed-adaptation situations need to be included in future models of colour appearance if they are to be successfully applied to colour image reproduction.



Figure 2.49: Mixed illumination environment for soft-copy proofing, matching display to print.

2.5.2 Colour Measurement

The CIE has recommended two basic systems of geometry for measuring instruments (CIE, 2004). The first is a bidirectional geometry for reflectance measurement, which requires the specimen to be illuminated at an angle of $45^\circ \pm 5^\circ$ from the normal and viewed in the direction normal to the surface (Fig. 2.50 left). This is known as 45:0 geometry and is the standard configuration for print viewing cabinets (Johnson and Scott-Taggart, 1993). The converse 0:45 geometry is also permitted, on the assumption that physical reflection of the light rays is identical in both directions.

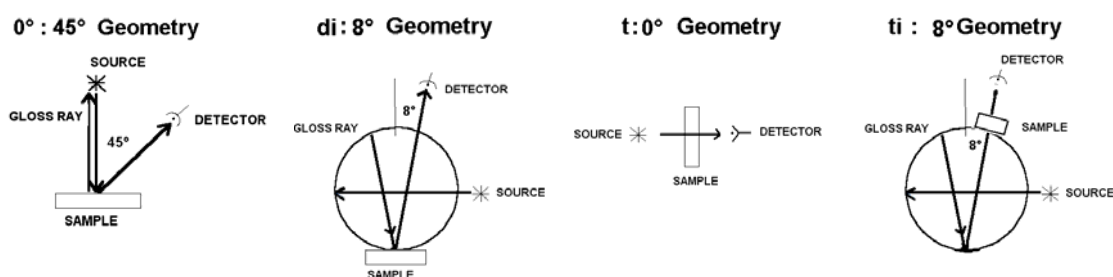


Figure 2.50: The configuration of standard illuminating/viewing geometries for the measurement of reflective and transmissive materials (from Rich, 2002).

In the second configuration the sample is illuminated diffusely by an integrating sphere and then viewed along a line at approximately 8° to the normal to the surface; this configuration, known as d:8 geometry, gives a good indication of diffuse reflectance while avoiding the direct specular component. The converse geometry, 8:d, illuminates the sample with a beam and uses an integrating sphere to collect reflected rays at all angles. By control of a port in the sphere at the specular angle, the specular component (the 'gloss ray') may be either included or excluded from the measurement, with corresponding nomenclature di:8 and de:8. Analogous geometries t:0 and t:8 are defined for transmissive materials.

Colour is specified in terms of the perception of a human observer, and is standardised through the CIE tristimulus functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$, described in Sec. 2.5.1. The colorimeter simulates the response of the eye with three channels having the sensitivities of the tristimulus functions. For each colour sample measured, the colorimeter generates three numbers for the tristimulus values X,Y,Z. Filter-based colorimeters have a source of light, a set of primary filters and a detector, and they compare the light reflected from the test colour with the light reflected from the reference colour (Fig. 2.51).

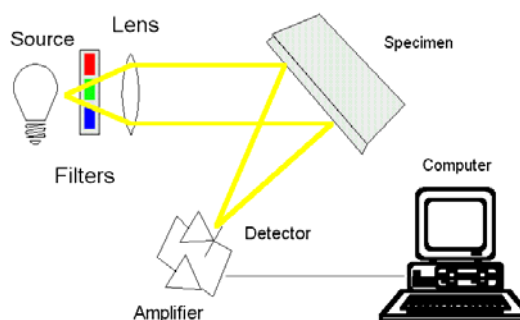


Figure 2.51: Schematic diagram of a tristimulus colorimeter (from Rich, 2002).

Colour separation is most simply achieved by three filters, with the short wavelength 'bump' in $\bar{x}(\lambda)$ (Fig. 2.46 right) approximated by a fraction of the blue filter. Better accuracy can be achieved by adding a separate indigo filter, thus forming a four-filter colorimeter, or by constructing mosaic arrays of small slivers of different coloured filters (Rich, 2002). A sophisticated method of fabricating filters involves thin-film deposition to form an interference stack with a precisely defined passband (Engelhardt and Seitz, 1993).

Photometry is the measurement of visible light, and is a subset of radiometry, the measurement of radiant power. The power captured by a photosensor is the product at each wavelength of the power of the incident illumination by the sensitivity of the sensor, summed over the wavelength range. By using a suitable filter to match the sensitivity of the instrument to the $V(\lambda)$ function a direct reading of

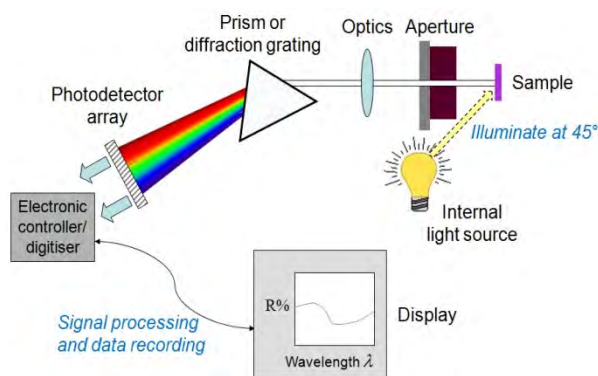


Figure 2.52: Schematic diagram of a spectrophotometer

luminance can be obtained. A spectro-photometer can be constructed by replacing the broadband coloured filters with a sequential series of narrowband interference filters. For example with 10 filters of width 30 nm covering the spectrum 400-700 nm, the signals derived from each can be summed to approximate tristimulus values. Use of a monochromator to disperse the wavelengths of the reflected beam enables a sampling slit to be scanned across the wavelength range, to take a measurement of the power in each wavelength interval Instruments

may divide the spectrum into intervals of 20 nm, producing 20 values over the range 380 to 760 nm, or 10 nm (39 values), 5 nm (77 values) or 1 nm (361 values). Instead of a scanning technique, most current spectrophotometers disperse the incoming light by means of a diffraction grating onto a fixed linear array of photo-detectors (Fig. 2.52). The geometry of the detector array determines the wavelength interval of each signal.

2.5.3 Measurement of Angular Reflectance

In the standard measurement geometries of Fig. 2.50 the angles of both illumination and collection of the light are considered together. But there are many surfaces that cannot be adequately measured using such limited conditions, such as 'gonio-apparent' colours (Johnston and Stanziola, 1969). A popular example is the coloured metallic finish applied to car bodies which changes its appearance according to the angles of illumination and viewing.

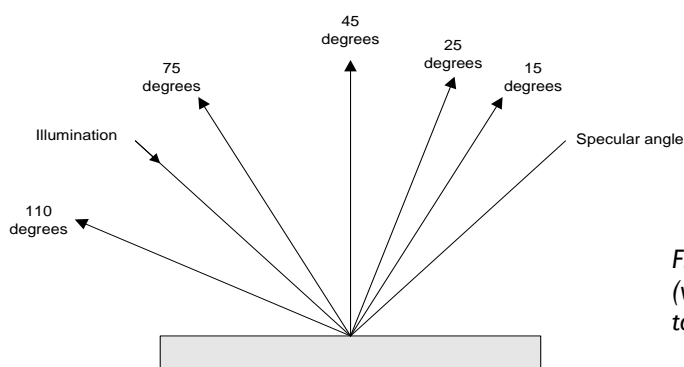


Figure 2.53: Multi-angle spectrophotometry (viewing angles are specified with respect to the specular angle).

For complete characterisation, the colour of such a surface should be measured at more than one illumination/viewing angle combination (McCamy, 1996). To use an integrating sphere is a compromise because it captures the average reflectance of the sample but not the detailed variation with angle. This limitation has been partly overcome by multi-angle spectrophotometers (Alman, 1987), which measure reflectance at three to five fixed angles (Fig. 2.53), as standardised in the automotive paint industry. These instruments use the concept of aspecular angle, i.e. difference of angle from the specular direction, rather than designating angles from the normal (90°) to the sample surface (Rössler, 1990). It should be noted that aspecular angles may be negative, indicating a detection angle further away from the normal than the specular angle.

In order to measure the distribution of reflectance as a continuous function of angle, a gonio-reflectometer is required (Fig. 2.54). By moving the source of illumination and/or sample and/or detector, the reflectance at every possible angle of incidence and reflection can be measured. By parameterising the azimuth and zenith angles with respect to the normal of the sample, the bidirectional reflectance distribution function (BRDF) is thus obtained (see Sec. 2.4.3).

Because the BRDF is a function of four angles, two incident and two reflected, the device must have four degrees of mechanical freedom to measure the complete function.

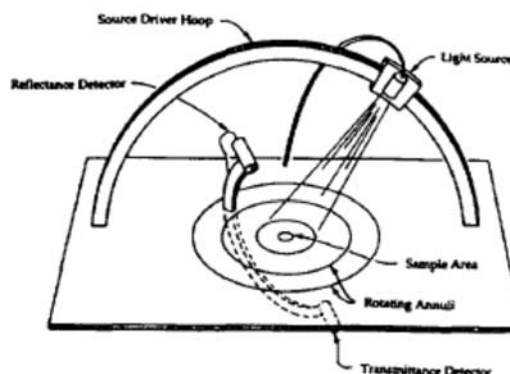


Figure 2.54: Schematic diagram of a gonio-reflectometer (from Ward, 1992).

In general, for materials with directional colour characteristics a goniospectrophotometer is necessary (Pointer *et al*, 2005) to measure the reflectance spectrum of the sample, when illuminated and viewed over a range of $\pm 80^\circ$ to the surface normal. Examples of such instruments are:

(1) the goniometric optical scatter instrument (GOSI) at NIST (Germer and Asmail, 1997), with three lasers at wavelengths of 633, 532 and 442 nm, and a broadband photomultiplier tube (PMT) detector giving high dynamic range;

(2) the automated gonireflectometer at Cornell (Li *et al*, 2005), with a white light source and a spectroradiometer with spectral resolution of 10 nm (Fig. 2.55);

(3) the bidirectional scatter distribution instrument in Gent (LeLoup *et al*, 2008), with a high-intensity xenon lamp and a CCD detector with spectral resolution of 4nm.

While such instruments are commercially available, their use is not widespread due to their relatively high cost, and the lack of a standard procedure for making multi-angle measurements. Nevertheless the colorimetric and appearance attributes derived from the rich multi-angle spectral dataset are of greater value than attributes derived from only a few isolated measurement angles (Mikula, 2003).

Dana *et al* (1998) developed a system at Columbia University with a robotic manipulator and CCD camera to allow simultaneous measurement of BRDF and BTF (Bidirectional Texture Function). They produced a database with image textures from over 60 different samples, each measured with over 200 different combinations of viewing and illumination directions. Ghosh *et al* (2008) described a high-speed approach for directly measuring BRDFs by projecting the incident light as a sequence of basis functions from a spherical zone of directions.

2.5.4 Imaging Techniques for Measurement of Gloss

It is advantageous to use a digital camera for the measurement of gloss because a large amount of spatial data may be captured in each image frame. The image of the surface includes the reflected profile of both the peak, i.e. the specular component, and the flank, representing the fall-off in gloss with angle of detection. Photometric calibration enables the slope of the flank to be obtained directly from the pixel values, from which the correlates of perceived gloss can be calculated.

Ji *et al* (2005) demonstrated an imaging technique to predict gloss. Several illumination/viewing geometries were tested and it was found that illumination at approximately 11° to the surface gave the most effective results in showing the gloss distribution. This angle agreed well with the subjective scaling of gloss by observers in earlier psychophysical experiments. Two models were proposed for the visual perception of gloss, one based on the luminance (Y) profile and the other on the largest gradient of the lightness (L^*) profile.

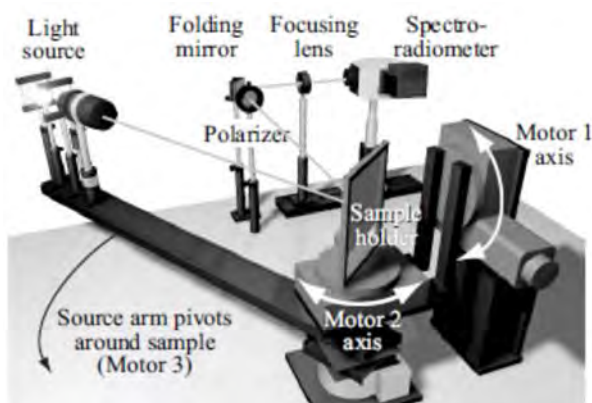


Figure 2.55: Automated three-axis gonireflectometer at Cornell University (from Li *et al*, 2005).

Marschner *et al* (2000) used a digital camera to record light reflected from the surface of test samples having a uniform BRDF (painted metal spheres and cylinders). The hardware setup involved a stable and uniform light source with a fixed 12-bit CCD camera to measure radiance reflected from the sample. Each sample was illuminated from a sequence of known positions, with a fixed camera angle, and the known geometry of the surface, light source and camera enabled the incident and exitant directions to be computed for each pixel. The camera sensitivity was adjusted by neutral density filters to control the exposure of bright samples without clipping. Flash illumination was modelled as a point source with uniform angular variation. The results were found to be close to true goniophotometric measurement. The advantages were that the BRDF was obtained with greater speed and higher angular resolution.

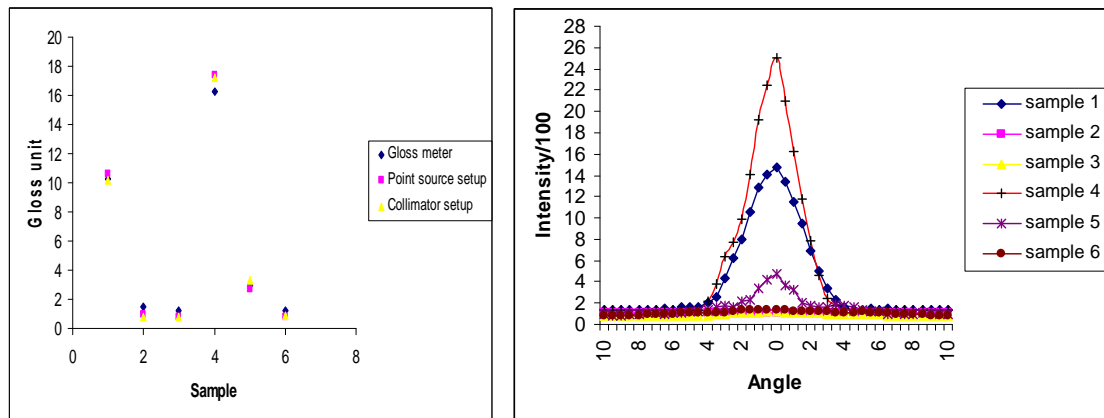


Figure 2.56: (left) Scatter plot of results obtained for six samples from three experiments; (right) Intensity vs. angle of tilt in increments of 0.5° (from Kumar and MacDonald, 2006).

Kumar and MacDonald (2006) showed that digital photography with flash illumination could be used as an alternative to multi-angle instrumental measurement for the determination of gloss. This was done in three phases: (1) convert digital camera response into gloss number; (2) evaluate the result on various substrates; and (3) verify the accuracy of the method with respect to a conventional gloss meter. Six white paper samples of different gloss values were selected. At each angle, the maximum intensity was determined from the centre of the image region for both the reference and sample. Each camera response was converted to a gloss number and evaluated against the reference gloss meter reading (Fig. 2.56).

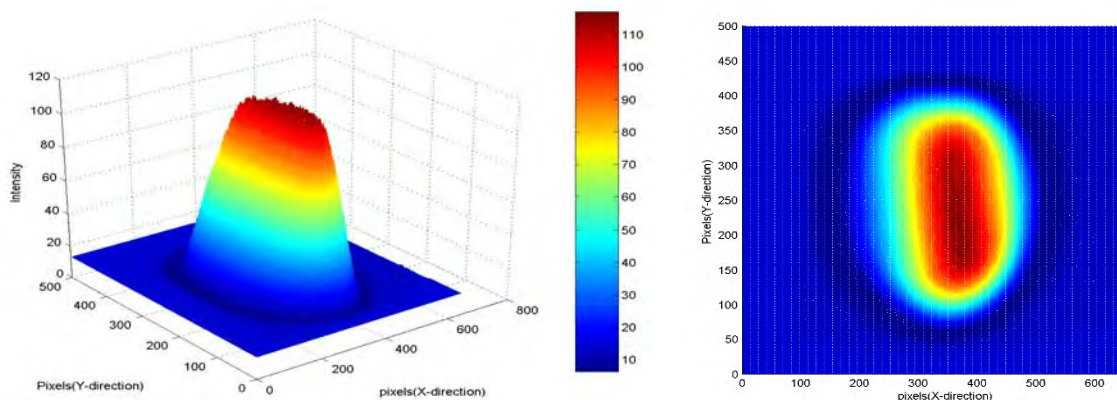
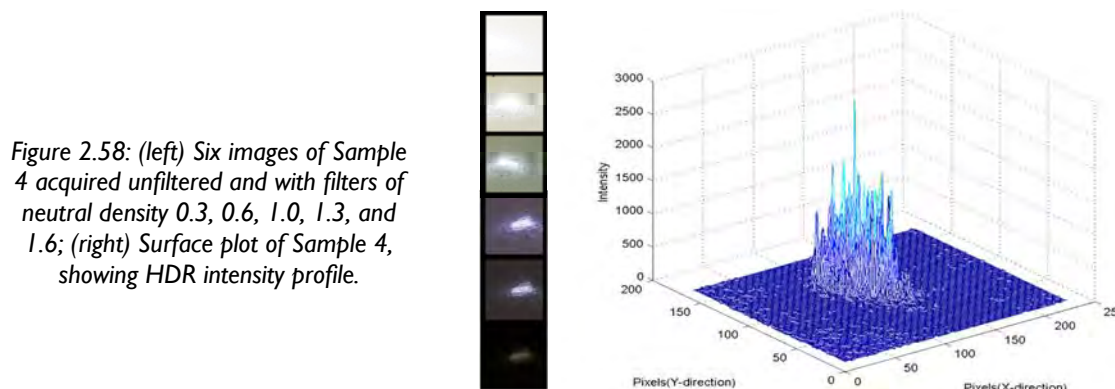


Figure 2.57: 3D and 2D plots of intensity over the illuminated area of the surface for matte Sample 3.

The illumination device collimated the light within a small region on the sample, and the intensity distribution of reflected light formed in the image depended strongly on the gloss and texture of the surface. The image of the matte Sample 3 was acquired through an ND filter of density 0.3 and the central region (Fig. 2.57) shows the uniformity of surface reflectance. A high dynamic range (HDR) distribution was constructed from the series of six images, taken first without any filter and then with five successively denser ND filters. The final intensity map of glossy Sample 4 had an extended dynamic range, with maximum value nearly 2,000 (Fig. 2.58).



2.5.5 Metrology

The conformance of a filter colorimeter to visual colour perception can be assessed by measuring a series of standard colours for which the tristimulus values (or a derivative such as CIE $L^*a^*b^*$) under a standard illumination source are known. Typical test objects are ceramic tiles, or chips from a colour atlas such as the Munsell Atlas or the Swedish Natural Color System. Care is required in the inter-comparison of values, because most colour order systems are specified by measurement with integrating sphere instruments (d:8 geometry), whereas most filter colorimeters are bi-directional instruments (45:0 geometry). Measurements are not transferable between the two geometries.

For a spectrophotometer the main operating parameters are the wavelength scale, radiometric scale (irradiance or radiance), spectral sampling interval and spectral pass band. For verifying the scale of reflectance or reflectance factor, there is a set of ceramic colour tiles (Fig. 2.59), which test the scale of reflectance factor, the linearity of the detector/amplifier/ digitiser system and the optical zero of the scale. The latter parameter is usually not an issue in transmittance readings, but reflection optics often produce homochromatic (similar wavelength) stray light that must be measured and subtracted from all subsequent readings. This correction should always be done at the time the white tile is used to standardise ('calibrate') the radiometric scale of the instrument (Clarke, 2006). For verifying the scale of transmittance, a series of neutral density filters is available from standards laboratories, such as the National Physical Laboratory (NPL), and from filter glass suppliers. These filters will test both the scale of transmittance and the linearity of the system.



Figure 2.59: BCRA Series II colour tiles for checking the accuracy and consistency of colour-measuring instruments.

For calibrating a radiometer a radiance standard is needed. This generally takes the form of a lamp with a continuous spectrum, such as tungsten halogen, powered by a well-regulated DC supply. It is enclosed in an integrating sphere with a white reflective inner surface, of some material such as barium sulphate, and a specified geometry for presentation to the instrument to be calibrated. An example is the SRS8 source from Bentham Instruments (Fig. 2.60), which has a luminance of 18,000 cd/m² at a correlated colour temperature of 3200K.

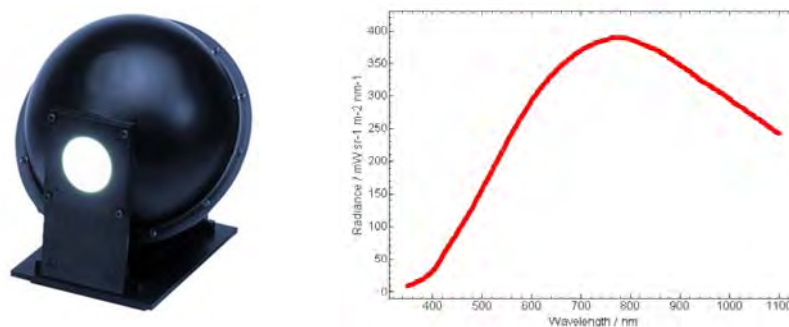


Figure 2.60: (left) Radiant calibration source in an integrating sphere; (right) spectral power distribution. (Bentham Instruments)

The key performance parameters for colour measuring instruments are:

- Repeatability, both short-term (within a minute) and long-term (typically an 8-hour period);
- Reproducibility, or inter-instrument agreement;
- Accuracy, the conformance to the correct or accepted value of test parameter.

All three parameters are frequently stated in terms of a colour difference between two instruments, or between a test instrument and a master instrument, or the average of a group of production instruments (often called the production aim point). The colour difference may be reported as a mean or median, as a maximum, as an RMS value, as a Coefficient of Variation (CV), or as the Mean Colour Difference from Mean (MCDM). Sometimes colour differences are plotted to show the relationships versus lightness, hue or chroma.

Typical results for short-term repeatability, expressed as ΔE^*_{ab} range from 0.01 for laboratory benchtop instruments to 0.1 for portable hand-held instruments. For long-term repeatability, these figures rise to 0.1 to 0.3 respectively (Wyble and Rich, 2007a). For reproducibility, the colour differences for measurements by ten instruments of the BCRA tiles (Fig. 2.59) expressed as differences from the grand mean (MCDM) for each tile range from 0.1 to 3.2 with an average of 1.69 (Wyble and Rich, 2007b).

A key requirement is traceability, defined as 'a property of a measurement result relating the result to a stated metrological reference through an unbroken chain of calibrations of a measuring system or comparisons, each contributing to the stated measurement uncertainty' (Gardner, 2006). In practice this means that it should be possible to relate any measurement back to a primary national standard, for example at NPL or NIST, via one or more artefacts as secondary (transfer) standards. For gloss scales, black glass with different degrees of etching is used as a calibration artefact. Samples of coated card are also available from Munsell and NCS in sets of up to three gloss levels and six lightness levels (Hanson, 1994).

2.5.6 Spectral Sampling and Principal Components

Any waveform within the wavelength range of the visible spectrum may be subjected to frequency analysis, in the same manner as time-varying and space-varying waveforms. The reflectance spectrum of a real material can be regarded as a waveform that is continuous, non-negative and frequency-limited. Buchsbaum and Gottschalk (1984) showed how it follows from the Sampling Theorem that the number of significant independent (basis) functions required to represent a collection of signals is:

$$n = \text{int}(2BT) + 1 \quad (2-59)$$

where $\text{int}(x)$ stands for the highest integer number smaller than x , B is the frequency limit in cycles per nm, and T is the signal extent (400–700 nm). The basis functions can be realised as projections on the orthogonal characteristic vectors (principal components) of the distribution of signals in the ensemble. Assuming a trichromatic system ($n=3$), the frequency limit for B is 0.005 cycles/nm. This means that all physical spectra having frequencies no greater than 1.5 cycles over the range 400–700 nm, with arbitrary phases and amplitudes (within physical and physiological limits), can be represented by three independent signals with minimal loss of information. They synthesised non-negative waveforms with sinusoidal modulation and plotted the locus of colours corresponding to all band-limited waveforms of frequency less than 0.005 cycles/nm. The colour gamut on the CIE 1931 chromaticity diagram (Fig. 2.61)

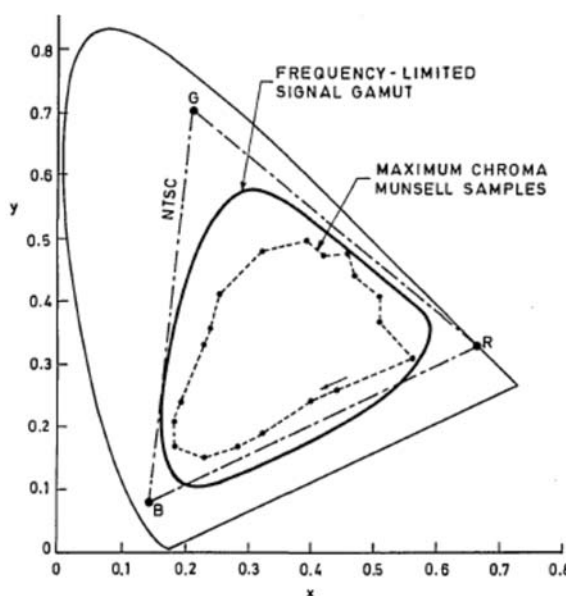


Figure 2.61: Chromaticity locus of spectra whose Fourier transforms are frequency limited to 0.005 cycle/nm (solid curve), the NTSC colour-television standard primaries (triangle), and the gamut of available Munsell samples (dashed curve) (from Buchsbaum and Gottschalk, 1984).

contains all colours on the surface of the Munsell colour solid, and it fits within the RGB triangle of NTSC phosphors, except for a small range of hues from indigo to magenta.

Principal component analysis (PCA) transforms a set of correlated variables into a number of uncorrelated variables, called principal components, related to the original variables by an orthogonal transformation. PCA can be performed by eigenvalue decomposition of a data covariance matrix, and reveals the internal structure of a multivariate dataset in a way that best explains its variance. It is defined so that the first principal component has the highest possible variance, and each successive component has the greatest remaining variance under the constraint that it be orthogonal to the preceding components. PCA is related to the discrete Karhunen–Loève transform (KLT) and the Hotelling transform.

Define a data matrix \mathbf{X}^T with zero mean (after subtracting the mean of the distribution from the dataset), where each of the n rows represents a different repetition of the experiment, and each of the m columns gives a particular kind of data value. In the present study, each row represents the spectral reflectance measurement of one sample, and the columns are the

reflectance factors at successive wavelengths. The PCA transformation is then expressed (Jolliffe, 2002) by:

$$\mathbf{Y}^T = \mathbf{X}^T \mathbf{W} = \mathbf{V} \mathbf{\Sigma}^T \quad (2-60)$$

where the matrices \mathbf{W} , $\mathbf{\Sigma}$, \mathbf{V} are defined by a singular value decomposition (SVD) of \mathbf{X} as $\mathbf{W} \mathbf{V} \mathbf{\Sigma}^T$. $\mathbf{\Sigma}$ is an m -by- n diagonal matrix with nonnegative real numbers on diagonal. Since \mathbf{W} (by definition of the SVD of a real matrix) is an orthogonal matrix, each row of \mathbf{Y}^T is a rotation of the corresponding row of \mathbf{X}^T . Each column of \mathbf{Y}^T is made up of the ‘scores’ of the cases with respect to the corresponding principal component. The matrix \mathbf{W} of singular vectors of \mathbf{X} is equivalently the matrix \mathbf{W} of eigenvectors of the matrix of observed covariances $\mathbf{C} = \mathbf{X} \mathbf{X}^T$,

$$\mathbf{X} \mathbf{X}^T = \mathbf{W} \mathbf{\Sigma} \mathbf{\Sigma}^T \mathbf{W}^T \quad (2-61)$$

Given a set of points in Euclidean space, the first principal component corresponds to a line that passes through the multidimensional mean and minimises the sum of squares of the distances of the points from the line. The second principal component corresponds to the same concept after all correlation with the first principal component has been subtracted out from the points. The singular values in $\mathbf{\Sigma}$ are the square roots of the eigenvalues of the matrix $\mathbf{X} \mathbf{X}^T$. Each eigenvalue is proportional to the variance (actually the sum of the squared distances of the points from their multidimensional mean) associated with each eigenvector. The sum of all the eigenvalues is equal to the sum of the squared distances of the points from their multidimensional mean.

PCA essentially rotates the set of points around their mean in order to align them with the successive principal components. This transfers as much of the variance as possible (using an orthogonal transformation) into the first few dimensions. The values in the remaining dimensions, therefore, tend to be small and may be dropped with minimal loss of information. PCA thus provides an effective method for reduction of dimensionality. In this study PCA is implemented by the Matlab function `princomp`.

Maloney (1986) analysed two large datasets of empirical surface spectral reflectances measured at 10 nm intervals and examined the conjecture that they could be reconstructed by a linear model composed of band-limited functions with a small number of parameters. He found that three basis functions could account for 0.9916 of the variance of a set of 462 Munsell samples, but that the number of parameters required to account for all of the variance in the spectral reflectances is five to seven, not three. Pakkinen *et al* (1989) repeated the analysis with a dataset of 1257 matte Munsell colour chips, measured at 5 nm intervals, and found that eight eigenvectors were required. Using the Karhunen-Loeve expansion, closely related to principal-component analysis (PCA), they showed that using the first four eigenvectors (Fig. 2.62) only 80.9% of reconstructed spectra had a mean error <0.03, but that with eight eigenvectors 100% of the spectra could be reconstructed with this accuracy. MacDonald *et al* (1999) proposed using CIE luminance and two chrominance values as the first three principal components, with higher-order components derived by PCA. Such studies indicate the number of channels required for multispectral image capture.

The set of reflectance values produced by a spectrophotometer represents a discretely sampled approximation to the continuous spectral reflectance distribution of the surface. The spectral data reported by a spectrophotometer may be abridged to a smaller number of values (corresponding to larger wavelength intervals). Conversely the spectral data may be interpolated to produce additional intermediate values (corresponding to smaller wavelength intervals), typically by using a cubic spline spanning the two values on either side of the interval CIE (2004)

recommended procedures for interpolation using cubic Lagrange polynomials. Grard and Trussell (1995) analysed the accuracy of these with respect to methods from digital signal processing using windowed sinc functions. They found that for interpolation, even in the extreme case of reducing from 20 nm to 2 nm intervals, there was little difference between the methods and in fact the Lagrange polynomials were generally better. For extrapolation, e.g. beyond 700 nm up to 760 nm, the effect on the spectrum depended strongly on the slope of the curve at the start of the extrapolation range, although it had little effect on the colorimetric accuracy of $L^*a^*b^*$ values because the colour matching functions approach zero at the extremities of the visible spectrum.

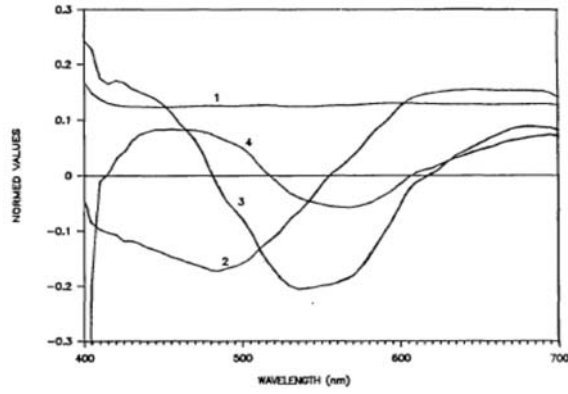


Figure 2.62: The four first eigenvectors of the set of 1257 Munsell reflectance spectra measured at 5 nm intervals (from Pakkinen et al, 1989).

2.5.7 Spectral Sensitivity

A key factor in the ability of a scanner to capture object colour or of a camera to capture scene colour is the response of its channels to wavelengths throughout the visible spectrum. The spectral responses of the colour channels, in general, do not match those of a typical human observer, as defined by the CIE Standard Colorimetric Observer. The colour sensitivity data is rarely available from the manufacturer of the scanner or camera, but is essential if the response is to be predicted for any surface reflectance under any illumination.

The relationship between scene radiance and output signal for a pixel in the k^{th} channel may be modelled as:

$$d_k = \mathcal{F} \left(e \int_{\lambda_l}^{\lambda_u} s_k(\lambda) i(\lambda) d\lambda + n_k \right) \quad (2-62)$$

where: d_k is the digital signal, \mathcal{F} is a monotonically increasing non-linear function, e is the exposure duration in seconds, λ_l and λ_u are the wavelengths of lower and upper bounds of integration (usually 380 to 760 nm), $s_k(\lambda)$ is the sensor response, $i(\lambda)$ is the incident power density per unit time, and n_k is a normal random variable. Typically there are three sensor types (red, green and blue) so that $k = 1, 2, 3$. The mean and variance of n_k specify the dark noise and response variability for the k^{th} sensor. Vora et al (1997a) described this as the static non-linearity model. Eq. 2-62 is an elaborated form of the tone response function, or opto-electronic conversion function (OECF). It will be assumed for the remainder of this section that \mathcal{F} has been determined and that the spectral processing is done in a linear signal domain.

The sensitivity metamerism index (SMI) is defined in ISO 17321 and describes the ability of a camera to reproduce accurate colours. The sensor can distinguish the same colours as the human observer if and only if the spectral responses of the sensor can be represented as a linear combination of the cone fundamentals (known as the Luther-Ives condition), which in practice never occurs (Berns, 2000). The index is defined as $SMI = 100 - 5.5\Delta E_{ab}^*$, where ΔE_{ab}^* is the average CIE $L^*a^*b^*$ error observed on a given set of 18 coloured patches in the Macbeth ColorChecker chart. The SMI varies depending on the SPD of the illuminant, and a value equal to 100 indicates perfect colour accuracy, only attained when the Luther-Ives conditions hold. Typical values of the SMI for digital cameras are in the range 75-85. Low-end cameras, such as camera phones, typically have a SMI of about 40 (Cao et al, 2008).

Camera sensitivity $s_k(\lambda)$ can be measured directly by using a narrowband (near monochromatic) light source, presented to the camera via one port of an integrating sphere or by reflection from a matte white tile (Fig. 2.63). The wavelength of the monochromatic source is scanned through the full visible spectrum, and the camera's RGB output signals recorded at each wavelength. Vora et al (1997b) measured the responses of two Kodak digital cameras to narrow-band illumination by passing light from a tungsten source through a monochromator and onto a white reflectance standard.

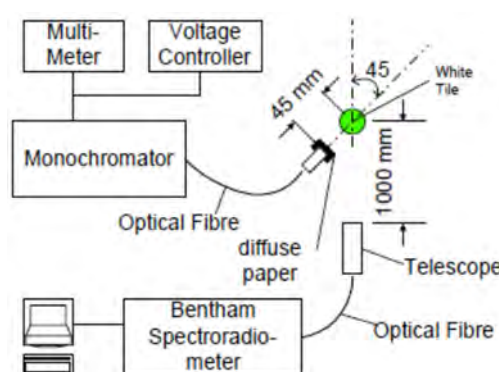


Figure 2.63: Measurement setup for camera calibration by monochromatic light source (from MacDonald and Ji, 2002).

To calibrate the radiant power they measured the integrated radiance of each narrow band stimulus using a spectroradiometer. The cameras' dynamic range was extended by varying the exposure duration across measurements. Wu and Allebach (2000) used a monochromator and tungsten-halogen light source to generate 31 colour stimuli of 10 nm bandwidth to measure the sensitivity of a Kodak DCS460c camera. MacDonald and Ji (2002) characterised a medium-format Rollei with a Jenoptik digital back by measurements at 5 nm wavelength intervals, and showed that the performance of the spectral measurement method was good enough to yield low mean-square errors when predicting the colours of a calibration target.

ISO 17321-1 (2006) defines a procedure for spectral characterisation of a digital still camera (DSC). Any automatic gain or adaptive tone reproduction should be disabled, and white balancing fixed. Fixed exposure settings are selected to generate peak output levels between 50% and 90% of full range (saturation level). The images are captured with the camera lens and any filters used for general picture taking (such as an infrared blocking filter) in place. The data output of each colour channel should remain independent, i.e. not be matrixed. The inverse OECF is applied to linearise the raw camera responses at each wavelength, and a 64×64 pixel block at the centre of each image is averaged to determine the linearised response of each channel at each wavelength.

IEC standard 61966-9 (2004) defines a similar procedure for consumer digital cameras, which may not have the facility to override the in-built automatic exposure and colour balance functions. Instead of an integrating sphere, the light output from the monochromator is passed

through a diffusing filter and presented to the camera through a small aperture in a flat test chart with background reflectance factor of 18%. The chart is front-illuminated by two or four tungsten-halogen lamps of illuminance 2000 lux and CCT 3100K. This illumination of the image field surrounding the test colour stabilises the camera's automatic exposure functions.

As an alternative to direct measurement, the spectral sensitivity of the channels of a scanner or camera may be estimated from images of a set of known reflectances captured under one or more sources of illumination. Indirect spectral characterisation of an image sensor is mathematically ill-posed because: (i) the sensors perform an input space reduction through integration over a range of wavelengths; and (ii) natural colours can be accurately approximated with only a

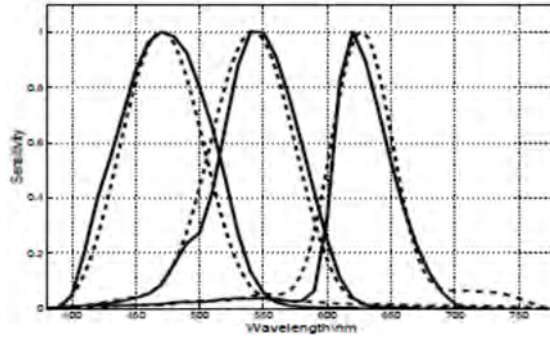


Figure 2.64: Measured (solid lines) and estimated (dotted lines) of the three sensors in a digital camera (from Finlayson *et al*, 1998).

few (typically between 3 and 9) basis functions. Therefore the same device output can be produced by an infinite number of stimuli, and only a limited set of linearly independent data may be collected for the estimation task. Therefore the solution space has to be constrained. Finlayson *et al* (1998) found that simple regression fails because of image noise and the limited dimensionality of the Macbeth reflectance set. They obtained good results by incorporating a number of natural constraints on the sensor spectra: positivity, modality and band-limitedness (Fig. 2.64). Carvalho *et al* (2004) assumed the positivity and smoothness of the sensor's spectral sensitivities, and applied a generalised cross-validation criterion for inequality constraint problems to estimate the error of a solution obtained under the least-squares framework.

Hardeberg (2001) found that the Principal Eigenvector (PE) method gave much better results than the Moore-Penrose pseudoinverse of the system of equations, and that the quality of estimation was almost as good when using 20 samples chosen with optimal reflectance spectra as when using the complete set of 1269 Munsell samples. Cheung *et al* (2005) compared four methods from multispectral imaging for a three-channel camera under a single light source. They found that the average colorimetric performance of all the methods was similar (median ΔE^*_{ab} in the range 2.6 to 3.9), but was significantly worse than conventional methods for camera characterisation such as polynomial modelling. Thomson and Westland (2001) were able to predict both the channel sensitivity and the non-linear response function (OECF) by fitting basis functions of the Gram-Charlier expansion, using Hermite polynomials to describe departures from the pure Gaussian shape to accommodate possible skew and kurtosis:

$$R(\lambda) = (1 + s_r\lambda + 2k_r\lambda^2)a_r\exp\left(-\left(\frac{\lambda-\lambda_r}{w_r}\right)^2\right) \quad (2-63)$$

The five parameters were: peak wavelength λ_r , amplitude a_r , width of Gaussian w_r , skewness s_r and kurtosis k_r . Nonlinear fitting was performed by the Levenberg-Marquardt technique.

2.5.8 Colorimetric Characterisation

For true colorimetric image capture the spectral sensitivities of the channels in the camera or scanner should satisfy the Luther-Ives condition that spectral sensitivities need not be exact duplicates of the CIE colour-matching functions but should be non-singular linear transformations of them. In practice this is never perfectly achieved because of difficulties in fabricating transmittance filter sets with the precise absorbance factor at every wavelength throughout the spectrum. Also the intrinsic noise in the imaging process limits their colour accuracy.

A common approach to obtain colorimetric data from a scanner or digital camera is to characterise the device in terms of CIE tristimulus values, i.e. to establish the transformation between the native signal space of the device (usually RGB) and a device-independent colour space (usually CIE XYZ). The device is presented with a series of colour patches of known colorimetric values (measured as CIE XYZ or $L^*a^*b^*$) and the averaged RGB output signals for each patch are recorded. Polynomial fitting techniques are then applied to interpolate the data over the full gamut to generate inverse transformations. In the case of a camera, the results are specific to the illumination under which the colour patches were captured.

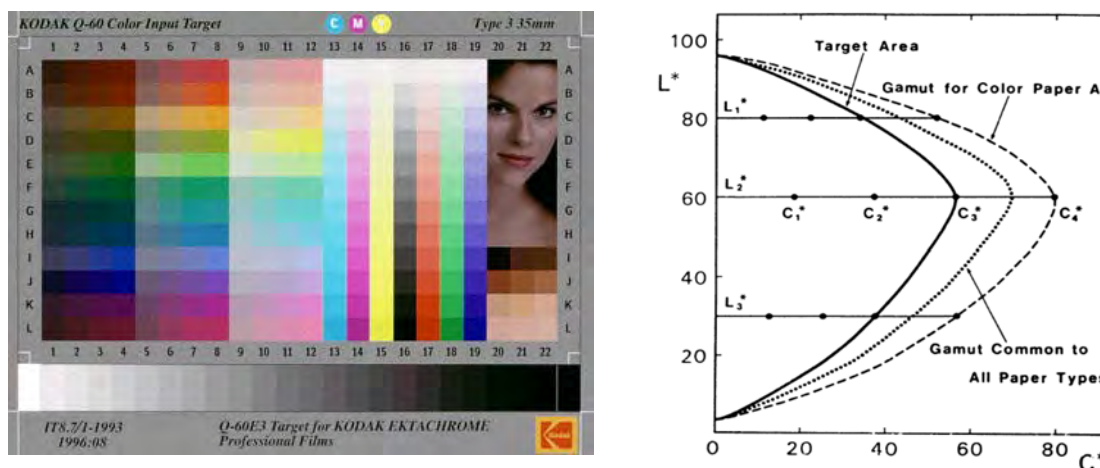


Figure 2.65: (left) Kodak Ektachrome version of the ANSI IT8.7/1 film target for scanner colour characterisation; (right) selection of four values of chroma for each of three values of lightness (ISO 12641).

Johnson (1996) described the development of photographic test targets for graphic arts scanners, evolving from the Kodak Q60A to the ANSI IT8.7/1 standard, which later became the ISO 12641 (1997) standard. The target (Fig. 2.65 left) has a 24-step neutral scale (bottom) from white (clear film or white paper) to black (D_{\max}), and 12-step density scales for each of the primary dyes (cyan, magenta and yellow) and secondary combinations (red, green and blue) and tertiary (three-colour black). It also has 144 colours sampled uniformly throughout the CIELAB gamut of the photographic film or paper, with 12 equally-spaced hues (h_{ab}), each at three levels of lightness (L^*) and four levels of chroma (C^*_{ab}). Additionally at the right there are 12 skin tones and a portrait, giving a total of 240 chromatic patches plus 24 achromatic patches. The target is also available as a photographic print for reflection scanners. The targets are supplied with measured colorimetric values for each patch. If the R, G, B signals from the scanner were truly a linear combination of the X, Y, Z values of the colour patches (i.e. if the Luther-Ives condition were obeyed) then the relationship for the averaged values in any colour patch could be expressed in just nine coefficients as:

$$\begin{aligned}
X &= c_{11}R + c_{12}G + c_{13}B \\
Y &= c_{21}R + c_{22}G + c_{23}B \\
Z &= c_{31}R + c_{32}G + c_{33}B
\end{aligned} \tag{2-64}$$

In fact the spectral absorption characteristics of the filters in most graphic arts scanners are designed not to be colorimetric but to optimise the colour separation of film dyes, e.g. as Status T densitometer filters (Giorgianni and Madden, 1998). To operate a scanner as a 'colour digitiser', Johnson therefore recommended a regression procedure to find coefficients that would minimise the error in the X, Y, Z domain over the full set of linear equations for X (and similarly for Y and Z) :

$$\begin{aligned}
X_1 &= c_{11}R_1 + c_{12}G_1 + c_{13}B_1 \\
X_2 &= c_{11}R_2 + c_{12}G_2 + c_{13}B_2 \\
&\dots \\
X_n &= c_{11}R_n + c_{12}G_n + c_{13}B_n
\end{aligned} \tag{2-65}$$

where in the case of the above target $n = 264$. Kang (2006) posed the problem as:

$$p_i = \mathbf{Q}_i^T \mathbf{C} = \mathbf{C}^T \mathbf{Q}_i \tag{2-66}$$

where the response p_i is X_i , and $\mathbf{Q}_i = [R_i \ G_i \ B_i]^T$ is the signal vector, and $\mathbf{C} = [c_1 \ c_2 \ c_3]^T$ is the coefficient vector. The whole system of equations for X can be written more compactly in matrix form as:

$$\mathbf{P} = \mathbf{Q}^T \mathbf{C} \tag{2-67}$$

where the $n \times 1$ response vector $\mathbf{P} = [X_1 \ \dots \ X_n]^T$, the $n \times 3$ matrix $\mathbf{Q} = [R_1 \ \dots \ R_n; G_1 \ \dots \ G_n; B_1 \ \dots \ B_n]$ is composed of all the signal vectors and $\mathbf{C} = [c_1 \ c_2 \ c_3]^T$ is the 3×1 coefficient vector.

If the measurement noise of the response is Gaussian then the system of equations can be solved to find values of coefficients \mathbf{C} that give the minimum overall error. The objective is to minimise the sum of squares of the differences Δp_i between each measured value p_i (X value of patch i) and the corresponding predicted value \hat{p}_i (weighted sum of R, G, B signals from the scanner):

$$\Delta p = \sum_{i=1}^n [p_i - \hat{p}_i]^2 = \sum_{i=1}^n [p_i - (c_1 q_{1i} + c_2 q_{2i} + c_3 q_{3i})]^2 \tag{2-68}$$

This can be rewritten in matrix form as:

$$\Delta p = (\mathbf{P} - \mathbf{Q}^T \mathbf{C})^T (\mathbf{P} - \mathbf{Q}^T \mathbf{C}) \tag{2-69}$$

The least-squares fit requires that the partial derivatives with respect to \mathbf{C} , i.e. $\frac{\partial p}{\partial c_1}, \frac{\partial p}{\partial c_2}, \frac{\partial p}{\partial c_3}$, are set to zero so that the minimum with respect to R, G and B is achieved when:

$$\mathbf{Q} \mathbf{Q}^T \mathbf{C} = \mathbf{Q} \mathbf{P} \tag{2-70}$$

Provided that there are $n \geq 3$ independent triplets of data, the determinant $\det|\mathbf{Q} \mathbf{Q}^T| \neq 0$ and the matrix $\mathbf{Q} \mathbf{Q}^T$ is invertible, then the equation can be solved for \mathbf{C} :

$$\mathbf{C} = (\mathbf{Q} \mathbf{Q}^T)^{-1} (\mathbf{Q} \mathbf{P}) \tag{2-71}$$

The outer product QQ^T is a symmetric 3x3 matrix and QP is a 1x3 vector as follows:

$$QQ^T = \begin{bmatrix} \sum R_i^2 & \sum R_i G_i & \sum R_i B_i \\ \sum R_i G_i & \sum G_i^2 & \sum G_i B_i \\ \sum R_i B_i & \sum G_i B_i & \sum B_i^2 \end{bmatrix} \quad QP = \begin{bmatrix} \sum R_i X_i \\ \sum G_i X_i \\ \sum B_i X_i \end{bmatrix} \quad (2-72)$$

Greater accuracy can be obtained by including additional terms with higher powers and cross-products of R, G, B in the transformation. For example, a second-order polynomial with ten terms would take the form: (2-73)

$$X_i = c_{10} + c_{11}R_i + c_{12}G_i + c_{13}B_i + c_{14}R_iG_i + c_{15}R_iB_i + c_{16}G_iB_i + c_{17}R_i^2 + c_{18}G_i^2 + c_{19}B_i^2$$

In this case the matrix Q would be $n \times 10$ and the matrix QQ^T would be 10×10 .

The above fitting procedure works with any continuous and monotonic functions of reflectance factor for R, G, B although in theory the least-squares minimisation should be performed in a linear signal domain. Variations on the above procedure include: (a) first linearising the R, G, B scanner signals by applying the inverse OECF function $R' = \mathcal{F}^{-1}(R)$ etc.; (b) taking logarithms of both sides, so that the fitting is performed in a density domain, i.e. from $\log_{10}(R/R_W)$ to $\log_{10}(X/X_W)$ etc., where the subscript W denotes the value for white – colorimetric density is preferred in photographic applications because of the close relationship between density and lightness; (c) transforming to measured CIE $L^*a^*b^*$ rather than the X, Y, Z values; (d) non-linear minimisation technique using other colour difference metrics, such as CMC or CIE ΔE_{94} or CIEDE2000, which are perceptually more uniform; (e) establishing the inverse transformation from X, Y, Z to R, G, B by interchanging the two datasets (device signals and colorimetric measurements) for the patches in the chart.

Herzog *et al* (1999) extended the polynomial regression procedure to multispectral reflectance data from a 6-channel sensor. Hong *et al* (2001) tested six different orders of polynomial for characterisation of a digital camera and found the best overall performance (i.e. lowest average error) with an 11-term polynomial, including a cubic term for RGB , based on 60 training samples. Pointer *et al* (2001) reviewed the major issues involved in the use of digital cameras to derive CIE X, Y, Z tristimulus values of objects in real scenes. They considered the spectral power and uniformity of the illumination, choice of test target and number of colours, camera signal processing, colour filter transmittance, characterisation method, and various quality measures. They showed that the choice of colour separation filtration is the most sensitive variable. In addition, the (non-)uniformity of the illumination of the test target used for characterisation was shown to be important.

The relationship between the original object colour and device signals can be expressed in an input profile (Fig. 2.66), within the framework defined by the International Color Consortium (ICC, 2004) system for colour management.

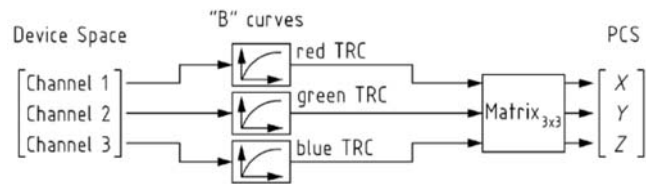


Figure 2.66: Structure of an ICC profile for an input device.

Signals from the device, typically R, G, B , are linearised by three tone reproduction curves (TRCs), in which the inverse OECF for each channel is loaded, then converted by a 3x3 matrix

into X, Y, Z tristimulus values for the device-independent Profile Connection Space (PCS). This enables the 3×3 matrix of Eq. 2-62 to be implemented. Higher-order polynomials and other transforms can be implemented in an alternative profile structure through a multi-dimensional lookup table. The conversion is not necessarily applied to the image at the time of capture because it is often preferable for archiving purposes to store the image data in its original form. By embedding the device profile in the image, the correct transformation can be applied later to generate a colorimetric image (MacDonald, 1996).

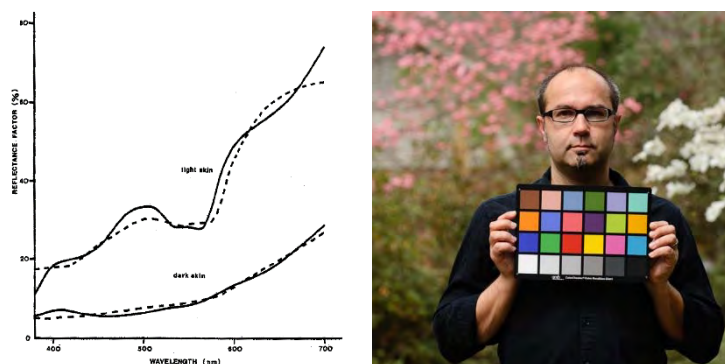


Figure 2.67: (left) Reflectance spectra of light skin dark skin and the corresponding patches in Macbeth ColorChecker (from McCamy 1976); (right) Chart in use for outdoor photography (www.photogtips.com).

For cameras the targets have colorants more typical of natural scenes and materials (Fig. 2.67 left) rather than the photographic dyes for film scanners. The traditional target is the Macbeth ColorChecker, developed for film and television cameras (McCamy et al, 1976). It has 24 patches in a 3:2 aspect ratio, including a 6-step grey scale. The size of the original target was 13 by 9 inches, with each patch 2x2 inches, so that it could be held at a distance from the camera in a studio or outdoor scene (Fig. 2.67 right). A new incarnation of the same target with identical colorants has been reduced to the size of a credit-card, 83 x 57 mm with each patch 10x10 mm, for macro photography on a copystand. For digital cameras the more recent GretagMacbeth DC and SG charts were developed (Fig. 2.68), with $12 \times 20 = 240$ and $10 \times 14 = 140$ sample patches respectively. The patches in the original ColorChecker and the DC chart are matte, whereas the SG chart is semi-gloss. The SG chart includes the same 24 colours of the original ColorChecker (top centre of chart) plus 16 skin tones, making it good for characterising cameras for portrait photography.

The SG chart includes the same 24 colours of the original ColorChecker (top centre of chart) plus 16 skin tones, making it good for characterising cameras for portrait photography.

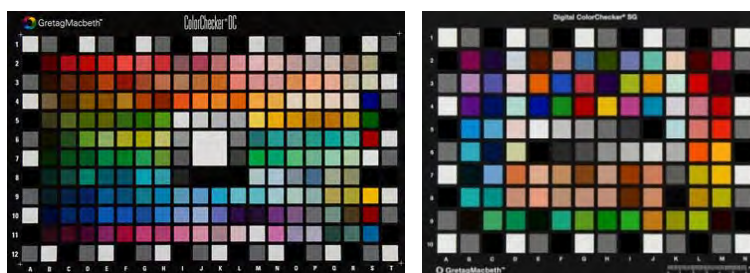


Figure 2.68: Two test charts for characterisation of digital cameras: (left) GretagMacbeth ColorChecker DC; (right) GretagMacbeth ColorChecker SG.

Mohammadi et al (2004) designed a colour target for spectral imaging with only 14 samples, consisting of a five-step grey scale and nine chromatic samples derived from statistical analyses of artists' paints. Different verification targets were reconstructed using a transformation matrix, and the spectral and colorimetric accuracy of the reconstruction was evaluated. It was shown that beyond a threshold number of samples in the calibration target, the performance of reconstruction became independent of the number of samples used in the calculation. The average spectral RMS errors for a calibration target consisting of 24 samples selected based on clustering were found to be less than 3.2% for the GretagMacbeth ColorChecker DC chart.

Chapter 3 – Dome Metrology and Calibration

Therefore geometry is founded in mechanical practice,
and is nothing but that part of universal mechanics
which accurately proposes and demonstrates the art of measuring.

– Isaac Newton (1687)
Preface to *Principia Mathematica*

In order to interpret images taken under directional illumination, two things need to be known: (1) the coordinates of the lamp; and (2) the relationship between points on the baseboard and pixels in the image. From these the geometry of incidence and reflection at each point on the object surface can be deduced. This chapter describes a series of studies in which five different methods were employed to determine the geometric coordinates of the lamp centroids.

3.1 The Illumination Dome

The UCL Dome is an acrylic hemisphere of nominal diameter 1030 mm, fitted with 64 flash lamps, each mounted on a separate circuit board (Fig. 3.1). Although not easily portable, this arrangement has proved to be stable and reliable for over six years in a laboratory environment. Because both the camera mounting point and the lamp positions are fixed, its geometry can be characterised precisely enough for the purposes of both PTM visualisation (Chapter 4) and 3D height reconstruction (Chapter 5), i.e. to half the diameter of the illuminated flash volume.

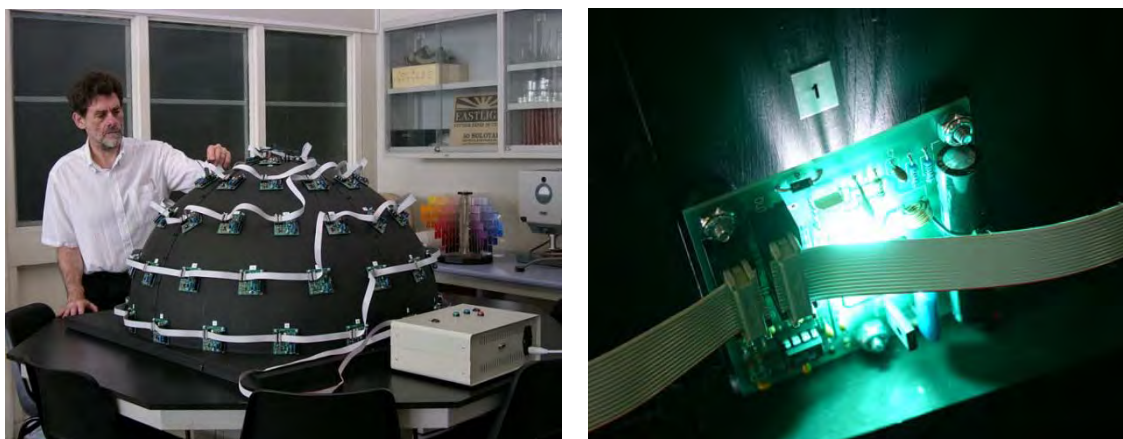


Figure 3.1: (left) Hemispherical dome in prototype before the camera mounting bar was fitted, with 64 flash lights on circuit boards, connected by 'daisy chain' ribbon cables to the control box; (right) flash lamp firing.

The 64 flash lamps are distributed around the hemisphere, arranged in five horizontal tiers: three tiers of 16, one tier of 12, and one tier of 4 lights at approximately equal intervals. The lowest tier (A) produces raking light on the equatorial plane, whereas the highest tier (E) is nearly polar. The electronic control unit and control logic on the circuit boards enable any combination of the flash lights to be selected and triggered, either singly or in groups. Although the original concept design called for the flash lights to be placed at regular intervals over the surface of the hemisphere (Fig. 3.2), the positions of the lamps in the actual dome, as

constructed, differ from the ideal. These irregularities arise from the manual construction method and from physical constraints caused by the mechanical supports.

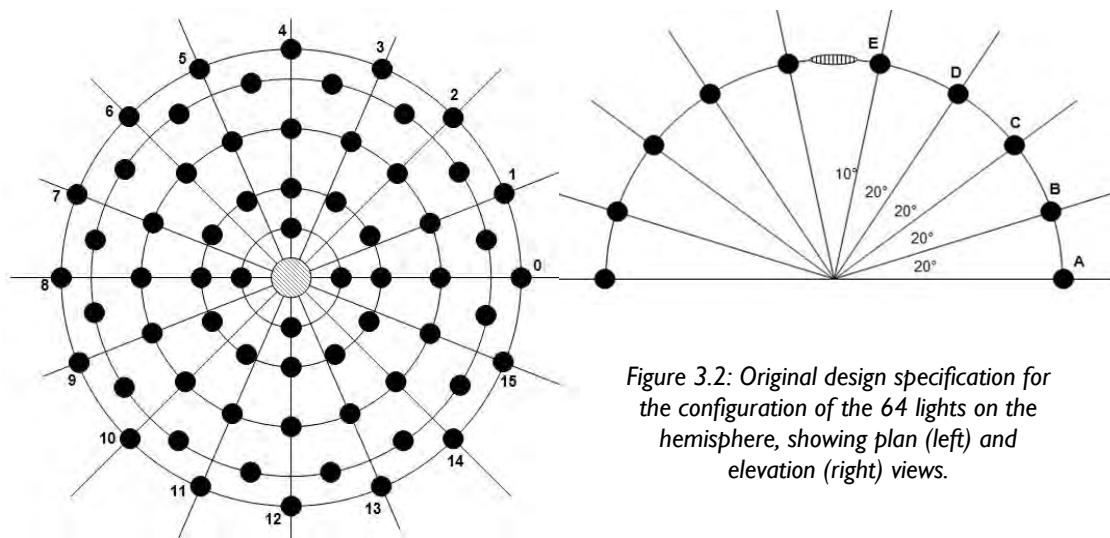


Figure 3.2: Original design specification for the configuration of the 64 lights on the hemisphere, showing plan (left) and elevation (right) views.

The flash lamp used in the dome is a small xenon arc unit, of the type used in consumer cameras, mounted on the back of a printed circuit board. It is triggered by applying a high voltage to ionise the gas inside the tube (Fig. 3.1 right). The boards are 'daisy-chained' together so that an 'increment' signal from the controller causes the 'enable' signal to shift along to the next board. This arrangement allows any combination of the lamps to be selected and fired together, although in normal operation single lamps are fired in sequence. Each lamp protrudes into the dome's interior through a lozenge-shaped hole cut in the acrylic shell. Both the inside and outside of the dome surface are painted with matte black paint to minimise stray light.

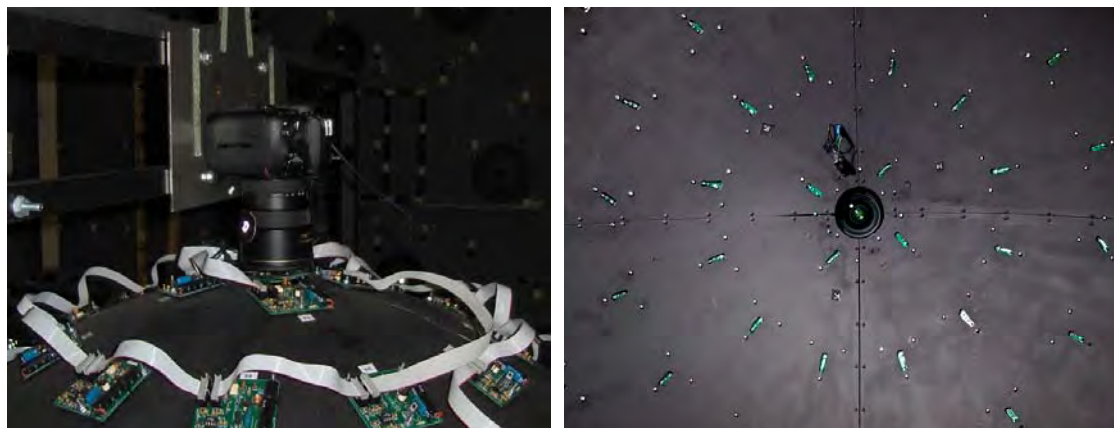


Figure 3.3: (left) Nikon D200 digital camera mounted above hole at top of dome; (right) View of inside of dome, looking upwards from baseboard, showing Tiers 3,4,5 of flash lamps.

A Nikon D200 digital camera is mounted on a rigid steel frame at the 'north pole' of the dome (Fig. 3 left) pointing vertically downwards with its optical axis normal to the horizontal baseboard in the 'equatorial' plane. The steel frame is braced to minimise movement of the camera. In effect the camera is held in a copystand configuration at a fixed height of approximately 700 mm. The view upward from an object placed on the baseboard (Fig. 3.3 right) is of the camera lens at the zenith, surrounded by concentric rings of flash lamps. By this means, multiple pixel-registered colour images of the object may be captured in sequence, each illuminated from a different direction.

3.2 Shadow Casting

The first method employed for determining the lamp coordinates was to take images of the shadow cast onto graph paper by a vertical steel pin when illuminated by each light (Fig. 3.4). The two-dimensional coordinates of the tip of the shadow can be determined simply by inspection of the image, and trigonometry applied. This is analogous to determining the position of the sun in the sky from the shadow of the gnomon on a sundial.

Let a pin of height h be placed perpendicular to the baseboard at the centre of a dome of radius R . A flash light mounted on the surface of the dome, assumed to be a point source, produces a shadow of length s on the baseboard, as shown in Fig. 3.5. The radial distance of the light from the base of the pin, along the baseboard in the plane of the shadow, is r and its height above the baseboard is z . Given R and h , and measuring s by observation, we seek to determine r and z from the relations:

$$s : h = (s + r) : z \quad (3-1)$$

$$r^2 + z^2 = R^2 \quad (3-2)$$

We assume that $z > h$, $r \leq R$ and $z \leq R$.

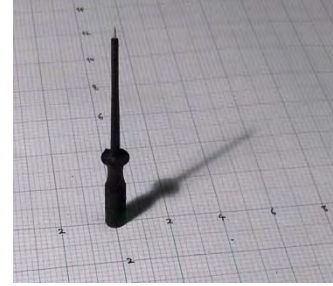


Figure 3.4: Shadow cast by a vertical steel calibration pin.

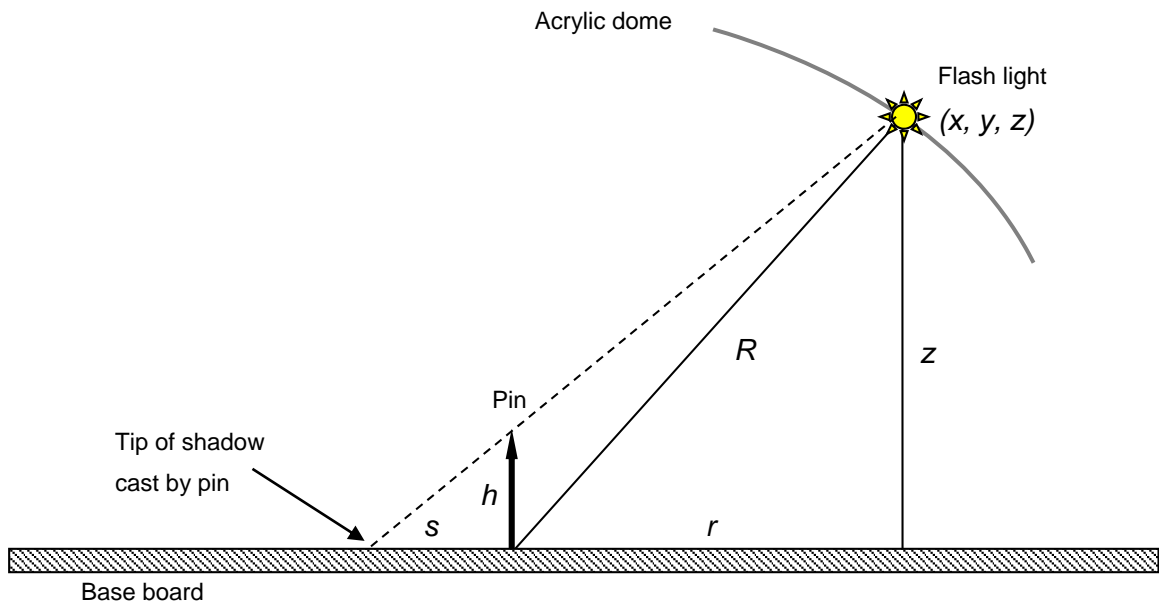


Figure 3.5: Elevation view of flash light and pin within the hemispherical dome.

Then from Eq. 3-1,

$$zs/h = s + r \Rightarrow r = zs/h - s = s(z/h - 1)$$

and from Eq. 3-2,

$$r = \sqrt{R^2 - z^2}$$

Hence

$$\begin{aligned} R^2 - z^2 &= (zs/h - s)^2 = z^2 s^2 / h^2 - 2s^2 z / h + s^2 \\ \therefore ((s/h)^2 + 1)z^2 - 2(s^2/h)z + (s^2 - R^2) &= 0 \end{aligned} \quad (3-3)$$

Solving Eq. 3-3 as a quadratic in z :

$$z = \frac{2(s^2/h) \pm \sqrt{4(s^2/h)^2 - 4((s/h)^2 + 1)(s^2 - R^2)}}{2((s/h)^2 + 1)}$$

or:

$$z = \frac{(s^2/h) \pm \sqrt{(s^2/h)^2 - ((s/h)^2 + 1)(s^2 - R^2)}}{(s/h)^2 + 1} \quad (3-4)$$

If $s = h$, Eq. 3-4 simplifies to:

$$z = \frac{s \pm \sqrt{s^2 - 2(s^2 - R^2)}}{2} = \frac{s \pm \sqrt{2R^2 - s^2}}{2} = \frac{h \pm \sqrt{2R^2 - h^2}}{2} \quad (3-5)$$

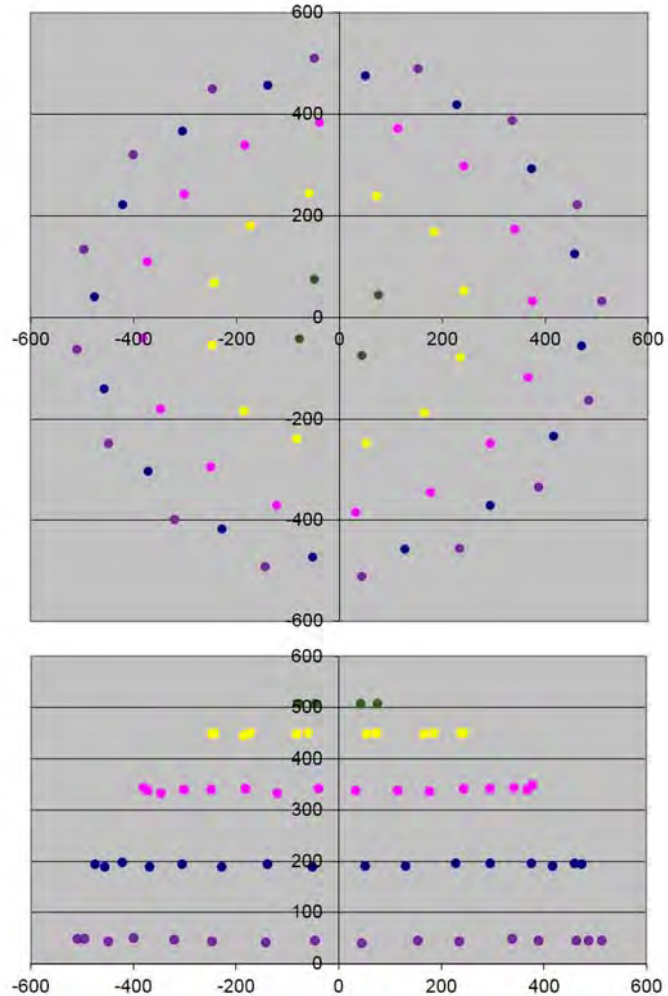
If $s = R$, Eq. 3-4 simplifies to:

$$z = \frac{(s^2/h) \pm (s^2/h)}{(s/h)^2 + 1} = \frac{2(s^2/h)}{(s/h)^2 + 1} = \frac{2(R^2/h)}{(R/h)^2 + 1} \quad (3-6)$$

In fact the acrylic dome is supported on a wooden frame that raises the equatorial plane by approximately 55 mm above the plane of the baseboard. It is hinged along the back edge so that it can be raised to place objects inside, and pin-registered so that the dome 'lid' always closes back into the same position relative to the base plane. The solution for this more complicated geometry is given in Appendix I.

The results of the pin-shadow experiment are shown in Fig. 3.6. The lamps are clearly arranged in five concentric rings, although departure of the actual lamp positions from the ideal is evident. The accuracy of the coordinates will be discussed in Sec. 3.5.

Figure 3.6: Lamp coordinates calculated by the pin-shadow method in plan (top) and elevation (bottom).



3.3 Photogrammetry of Retroreflective Targets

For more precise determination of the lamp coordinates, close-range photogrammetry was employed, with the aid of the Vision Measurement System (VMS). In order to apply VMS to measure the lamp coordinates in the dome, retro-reflective targets were stuck onto four well-defined positions around each of the 64 flash lamps, and a photogrammetric project was undertaken. The interesting problem in this case was that the locations of these targets on the inside of the hemispherical surface had to be estimated relative to the baseplane, but the photographs had to be taken with the dome raised on its hinges. Capturing images from inside while the dome is closed would be extremely difficult, so it was decided to partition the procedure into two stages.

Targets were placed in the central area and in locations around the supporting frame at the perimeter of the baseplane. The latter were carefully arranged in matched pairs, one on the base frame and one on the dome flange, so that they were in close contact when the dome was closed (Fig. 3.7). Two calibrated invar metrology scale bars were placed in suitable positions on the base plane. A set of overlapping images was captured from the plane using the Nikon D200, hand-held and moved freely with its integral flash. The 17-55 mm Nikkor zoom lens was set to 17 mm focal length to give the widest angle of view, and the aperture set to $f/16$ to give maximum depth of field and to minimise the over-exposure of the reflected flash light from the targets. Images were taken from both high and low viewpoints from five azimuthal angles, trying to ensure that all targets were visible and occupied a large area in each image, to achieve a robust imaging network. The location of each target was calculated in VMS by importing the images and running the photogrammetric bundle adjustment procedure.



Figure 3.7: Views of (left) the baseplane with retro reflective targets and scale bars; and (right) interior of the dome raised on its hinges, with retro-reflective targets around each lamp. The red circles in these images show one of the common point-pairs between the dome and plane.

In the first stage of the project, a target in the centre of the baseplane was arbitrarily designated as the origin of the coordinate system. In the second stage, 18 images were captured with the camera mounted in its usual position on the dome, and the dome raised on its hinges by varying angles. The locations and uncertainty estimates of common points between the plane and the dome achieved in the previous section were imported as control points. Finally, by running the bundle adjustment with external constraints in VMS, the positions of all targets on the dome were obtained.

After running the bundle adjustment, the mean precision of estimating target locations was 11.97 microns and the residuals of the estimated distances from the real distances of the two scale bars were -14.5 and 10.8 microns. In the second stage of the project, the match pair targets between the dome and the base plane were input into bundle adjustment process as control points with coordinates obtained from the first step. After running bundle adjustment, the target locations were estimated with a mean precision of 72.13 μm .

The primary objective of this study was to determine as accurately as possible the centre coordinates of the flash lights in the dome. For this purpose four circular target spots of 3.5 mm diameter were punched from a sheet of retro-reflective material and affixed to each lamp mounting, one on the head of each of the three mounting screws of the circuit board and one on the base of the flash lamp itself (Fig. 3.8). The origin (datum point) was chosen arbitrarily as the central target spot on the baseboard.

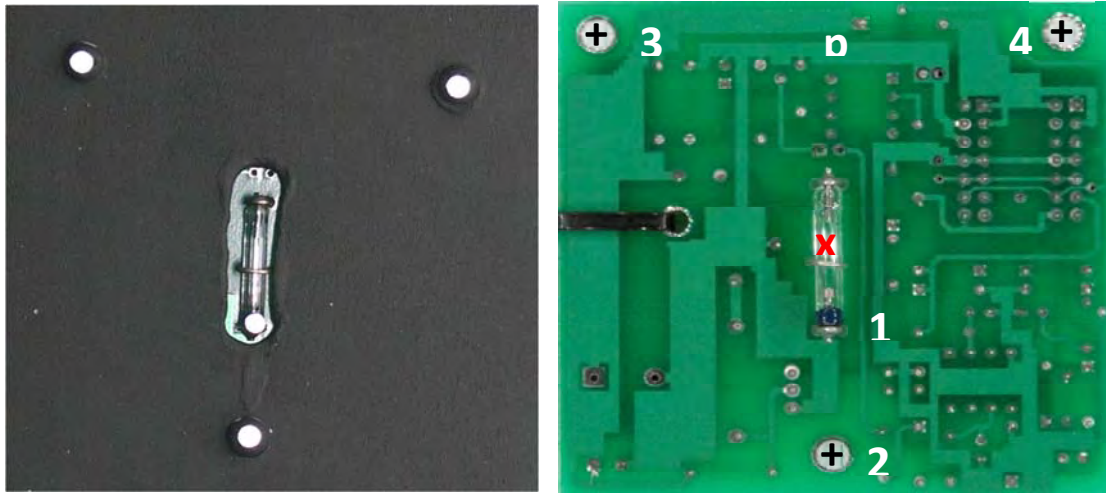


Figure 3.8: Locations of the four retro-reflective target spots on the mounting of each flash lamp: (left) targets in situ on the inside of dome, showing lamp 17 positioned in a lozenge-shaped hole cut in the acrylic shell of the dome; (right) underside of the flash board showing numbering of targets.

VMS produced very accurate estimates of the X,Y,Z coordinates of the geometric centre of each target, with a minimum error of less than 5 μm , mean error of 21.5 μm , and maximum error of 43 μm . Using this data, the centre of each lamp was then calculated as a weighted barycentre of the coordinates of the three screw positions, which are fixed in relation to the flash lamp by the geometry of the printed circuit board, Fig. 3.8b. The centre c of the flash (denoted by the red cross in Fig. 3.8b) lies on a line between point 2 and the midpoint p of the line between points 3 and 4, calculated by linear interpolation as:

$$x_c = \alpha x_p + (1 - \alpha)x_2 \quad (3-7)$$

where $\alpha = 0.522$, and similarly for the y and z coordinates.

The illuminated cylindrical volume of the flash discharge between the two electrodes is approximately 14 mm in length and 2.8 mm in diameter (Fig. 3.9). The outer diameter of the flash glass was measured as 3.5 mm, so the position of the true geometric centre of the flash illumination was moved outward along the normal by half this amount, i.e. the radius of the flash lamp cylinder.

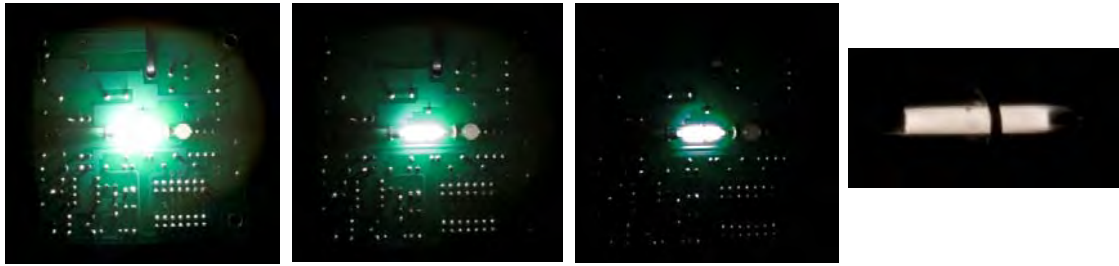


Figure 3.9: Illuminated area of flash light: (left) three renderings of the same raw image using different exposure settings in the Photoshop image processor; (right) detail from the image taken through a neutral density 3.0 filter.

The final coordinate positions of all target points and calculated lamp centres are shown in Fig. 3.10. Note that the normals to the boards do not intersect exactly at the centre of the dome because of variations in the mounting geometry.

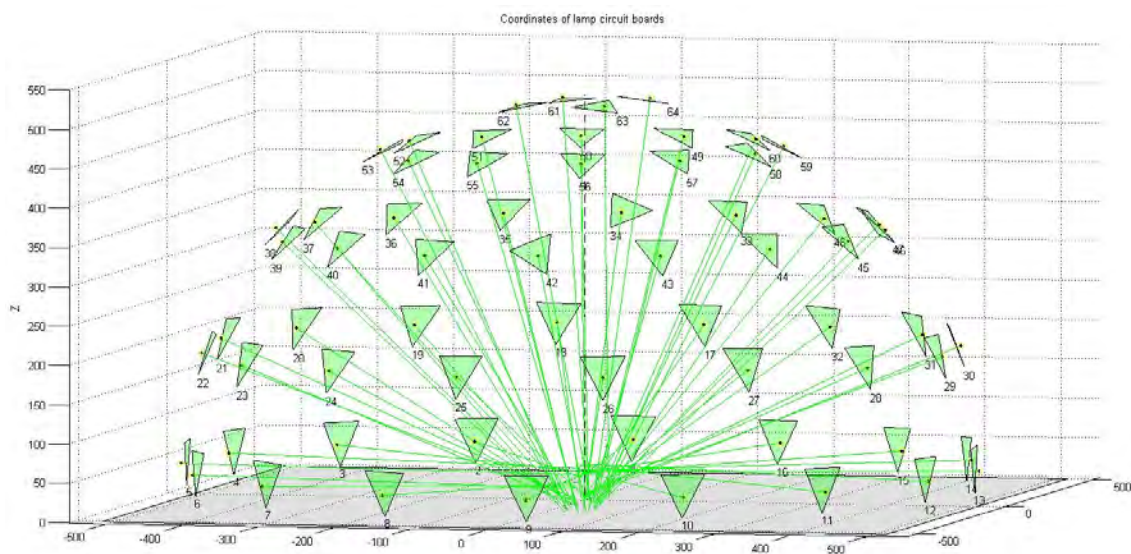


Figure 3.10: Triangles formed by coordinates of three mounting screws on each flash board, plotted in 3D. The normal to the plane of each board passing through the lamp centre is shown by a green line.

The thickness of the acrylic shell of the dome is not uniform, but tapers from the equator to the pole. This could be estimated because the flash boards are mounted on the outside of the dome, by calculating the normal distance of the target at point 1 (adhering to the bulb of the flash lamp glass) from the plane through points 2,3,4. The mean distances of the flash centres from the plane of the screw heads (outward from dome centre) for the five tiers were: 6.43, 5.52, 4.56, 3.67, 3.69 mm.

Table 3.1: Parameters (mm) of horizontal circles fitted to five tiers of lamp centre coordinates.

Tier	X centre	Y centre	Radius	Residual	X centroid	Y centroid
1	-2.87	-10.69	515.2	0.95	-2.43	-12.48
2	-1.39	-12.15	487.0	1.50	-2.04	-11.75
3	-2.31	-13.36	393.7	2.69	-2.01	-11.94
4	0.13	-13.99	257.6	2.91	-0.99	-13.56
5	1.52	-13.44	89.5	0.24	1.38	-12.80

To check the roundness of the dome, a horizontal circle was fitted to each tier of lamps (Fig. 3.11), ignoring the Z coordinates, using the Matlab function `nlinfit`. The parameters of the fitted circles, including the mean absolute residual are given in Table 3.1. The overall mean residual was 1.84 mm. For comparison, the centroids of the lamp coordinates in each tier (calculated as the mean of X and Y) are included.

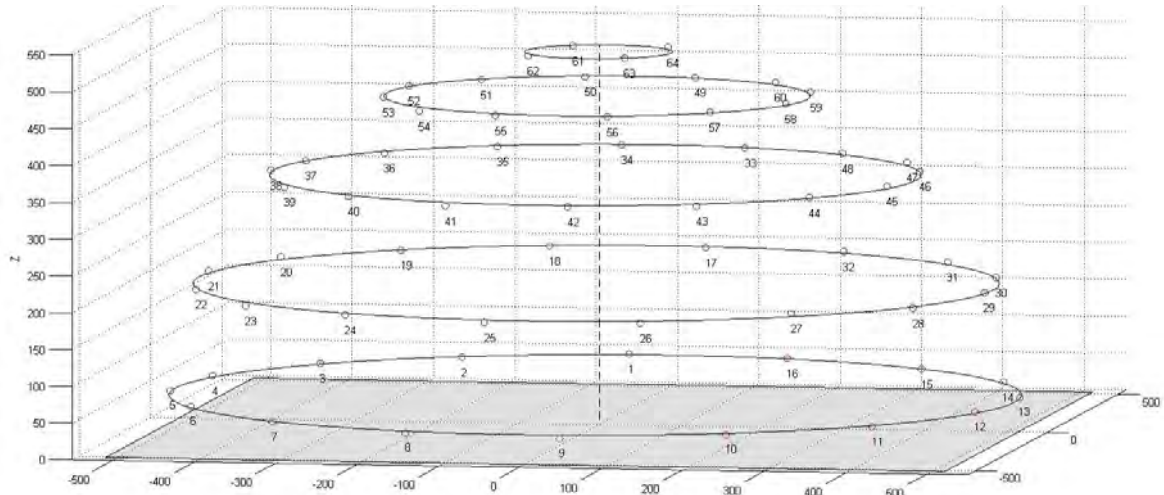


Figure 3.11: Circles parallel to base plane, fitted to the coordinates of the flash lamps in each tier.

The 'true' geometric centre of the dome was estimated by fitting a hemisphere to the 64 lamp coordinates (determined by VMS), again using the Matlab function `nlinfit`. The centre coordinates were $X_0 = -0.53$, $Y_0 = -12.18$, $Z_0 = 19.40$, and the radius was 513.5 (all in mm). Therefore the uplift of the best-fit hemisphere relative to the baseboard is $55 + 19.4 = 74.4$ mm. The residuals are shown graphically in Fig. 3.12, orientated along the vector to the centre, enlarged for clarity by a factor of x5. It can be seen that the dome is flatter at the top and wider at the base than a perfect hemisphere. The mean absolute residual was 12.61, with a maximum absolute value of 23.77 mm.

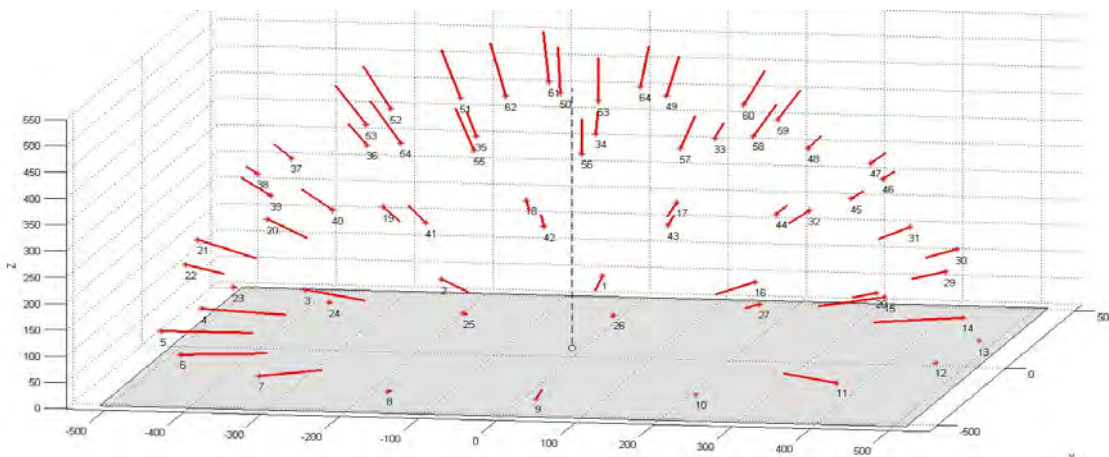


Figure 3.12: Residual values (length magnified by x5) of hemisphere fitted to coordinates of all 64 flash lamps.

The positions of the centre coordinates found from centroids, fitted circles and fitted hemisphere are plotted in Fig. 3.13. The overall mean was $X = -1.05$, $Y = -12.58$ mm. It is apparent from the graphs that there is a systematic variation of both X and Y with height above the baseboard, suggesting that the central axis of the dome is slightly tilted by about 4.5 mm in X and 3.5 mm in Y.

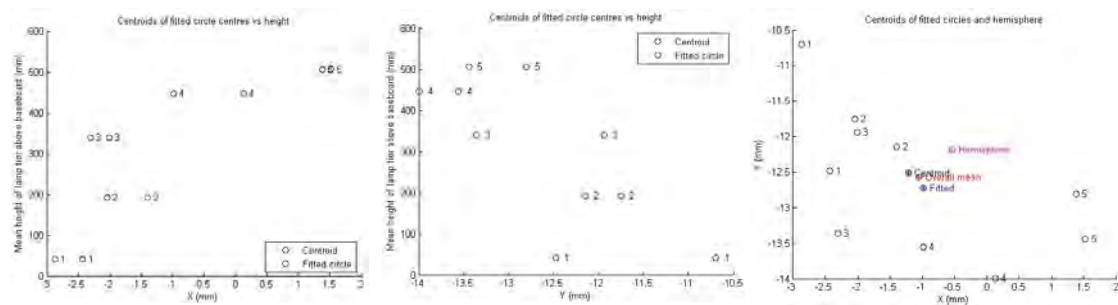


Figure 3.13: Mean height of lamp tier vs X (left) and Y (centre) offsets. (right) X vs Y coordinates of centroids (black), centres of fitted circles (blue), centre of fitted hemisphere (magenta), and overall mean (red).

With the Nikon D200 camera mounted on the plate at the top of the dome, the coordinates of the perspective centre of the 17-55 mm zoom lens (set at 28 mm), relative to the datum point on the baseboard (as determined by VMS) are: 7.50, -14.61, 577.63. Relative to the geometric dome centre the camera coordinates are therefore $X_c = 7.50 + 1.05 = 8.55$, $Y_c = -13.61 + 12.58 = -2.03$, and $Z_c = 577.64 - 19.40 = 558.24$ (mm). The height of the lens centre above the plane of the baseboard is $577.64 + 55 = 632.64$ mm.

3.4 Photogrammetry using Flash Lamps as Targets

In order to check the reliability of the previous method for estimating the flash lamp coordinates, a different approach was tried by using the flashes themselves as recognisable targets. In this way VMS was able to estimate the centroids directly, without the need for the post processing method presented in Section 3.3. To achieve this goal, images were taken of the open interior of the dome, synchronised with the flash controller. Since only 32 lamps could flash simultaneously, two sets of images were required. In the first set, the lamps in tiers 1 and 2 were illuminated, whereas in the second set, the lamps in tiers 3, 4 and 5 were illuminated, located closer to the centre of dome (Fig. 3.14).

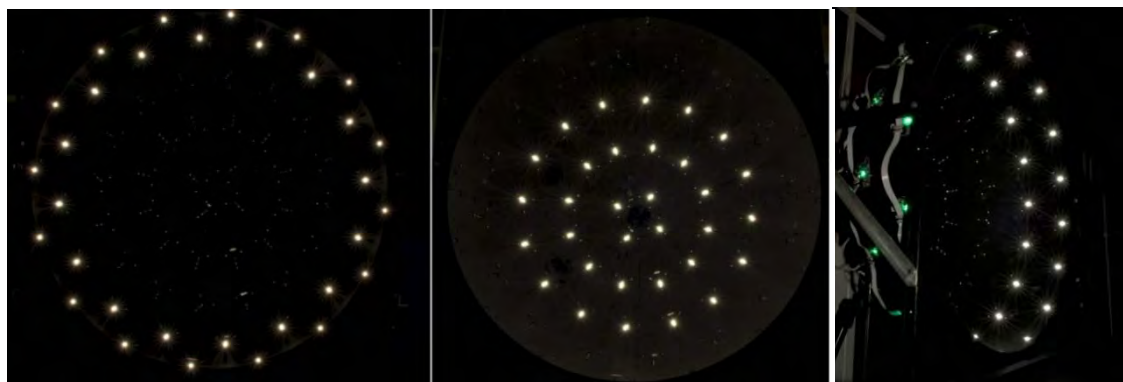


Figure 3.14: Images captured from lamp flashes, with the camera on the dome axis: (left) first set from tiers 1-2; (centre) second set from tiers 3-4-5. The oblique view (right) shows only eight of the lamps in tiers 1-2.

A set of overlapping images was captured using the Nikon D200, hand-held and moved freely. The zoom lens was again set to 17 mm focal length, with aperture $f/22$ and a shutter speed of $1/250$ second to minimise the over-exposure of the flash lights. Images were taken from many viewpoints, trying to ensure that all targets were visible in sufficiently many images to achieve a robust imaging network. The difficulty with visibility of the lamps, especially in tiers 1 and 2,

was clear when the camera was at an oblique angle (Fig. 3.14c). Some of the nearer flashes were visible through the green printed circuit boards, but these could not be used as targets.

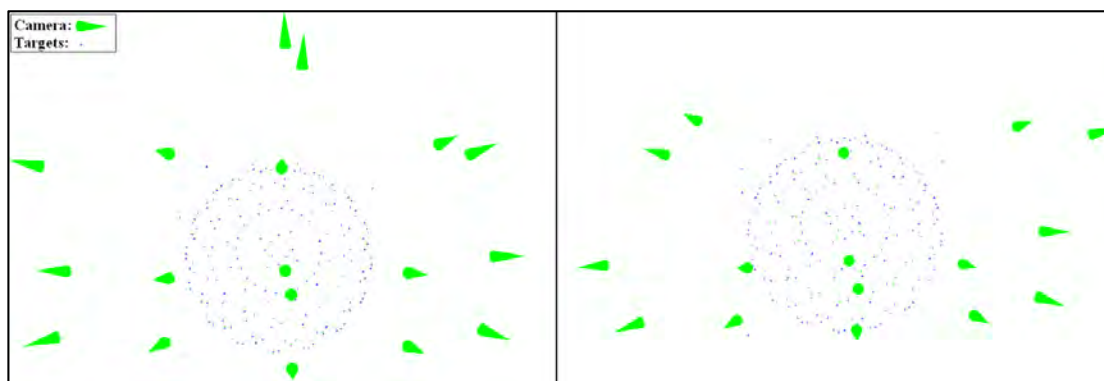


Figure 3.15: Imaging network of tier 1-2 image set (left) and tier 3-4-5 (right). Grey point cloud shows location of flashes and retro-reflective targets; green cones show estimated camera position and pose for each image.

On each set of images, after identifying locations of corresponding lamps, a photogrammetric bundle adjustment process with external constraints was run, in which the coordinates of the four targets around each lamp obtained from the previous study (Section 3.3) were considered as control points. The imaging network for each image sets is illustrated in Figure 3.15. The first set (left picture) included 18 images and the second set included 16 images (right). In both imaging networks, the targets and flashes nearer to the centre of the dome could be seen in more images than the targets and flashes nearer to the edge.

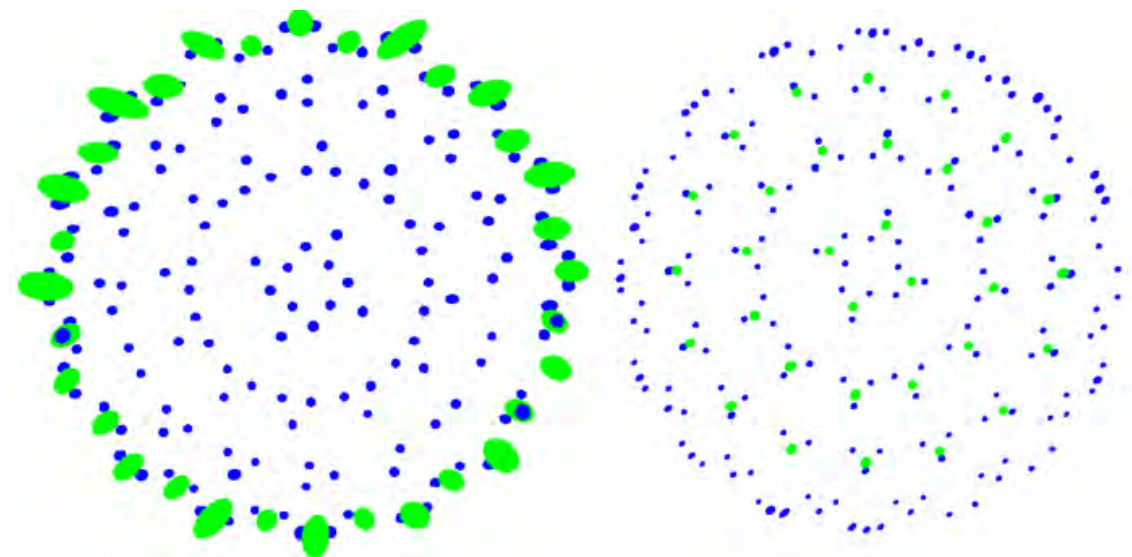


Figure 3.16: Error ellipsoids of retro-reflective targets (blue) and the lamp flashes (green) in tier 1-2 image set (left) and tier 3-4-5 image set (right).

After running the bundle adjustment, coordinates of lamps were obtained with different levels of precision: 80.91 microns for tier 1-2 image set and 43.60 microns for tier 3-4-5 image set. This difference in precision is due primarily to the geometry of the lamps relative to the camera locations. In tier 3-4-5 image sets the lamp flashes were close to the centre, where there is less obstruction to the rays of light from the targets to the camera; in tier 1-2 image set, however, the edge of the dome was difficult to avoid. This difference can also be seen in the error ellipsoids of the targets (Fig. 3.16) where the biggest errors occurred for the flashes in tier 1.

3.5 Comparison of Lamp Coordinate Methods

3.5.1 Pin Shadows vs VMS Targets

The coordinates determined by the procedures described in the previous three sections and their angles relative to the geometric dome centre are compared in Fig. 3.17 with the corresponding values previously calculated by the pin shadow technique. Tables 3.2, 3.3 and 3.4 show the data values. The figure shows a qualitative match between the two datasets, with vector differences between the corresponding X,Y,Z coordinates ranging from a minimum of 1.29 to a maximum of 20.77 mm. The mean difference was 11.93 mm.

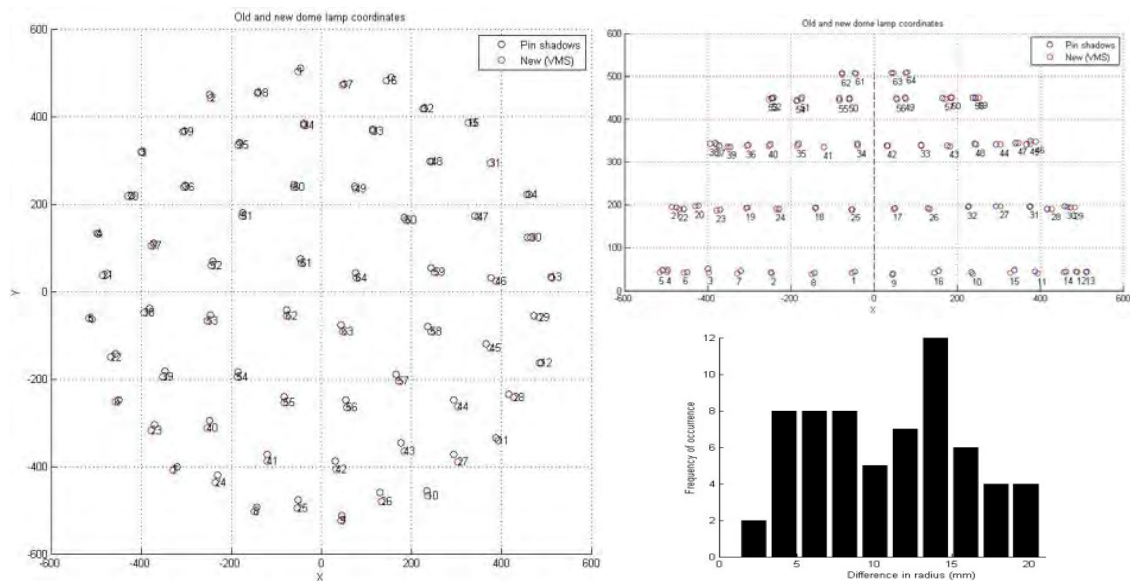


Figure 3.17: Plan (left) and elevation (right) of lamp coordinates determined by pin shadow (blue) and VMS (red) methods; (lower right) histogram of rms differences between corresponding points (3D vector lengths).

Histograms of the differences in radius, theta and phi angles (Fig. 3.18) show systematic discrepancies between the two datasets. Plotting the differences vs VMS values (Fig. 3.19) shows that in each tier of lights the radius changes as a function of azimuth angle (theta), indicating eccentricity. The differences in zenith angle (phi) are mostly negative, with a mean of -0.55° , indicating that the pin shadow method generally predicted a larger angle, i.e. higher lamp position, on the assumption that the structure was a perfect hemisphere.

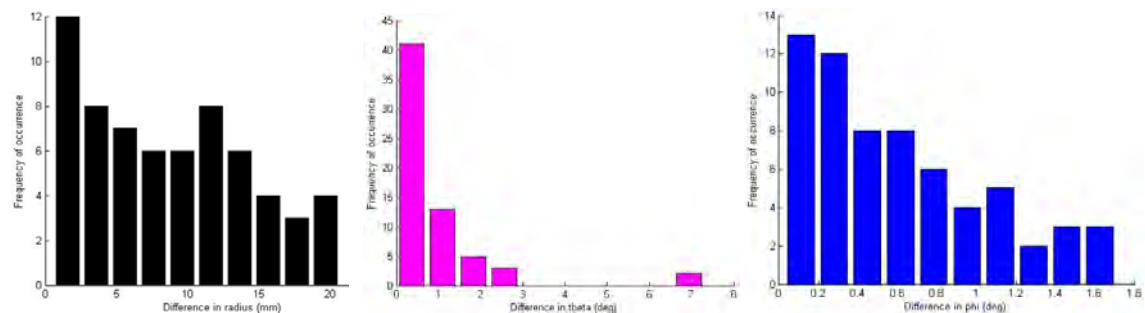


Figure 3.18: Histograms of absolute differences in radius (left), theta (centre) and phi (right) between the datasets for pin shadows and VMS targets.

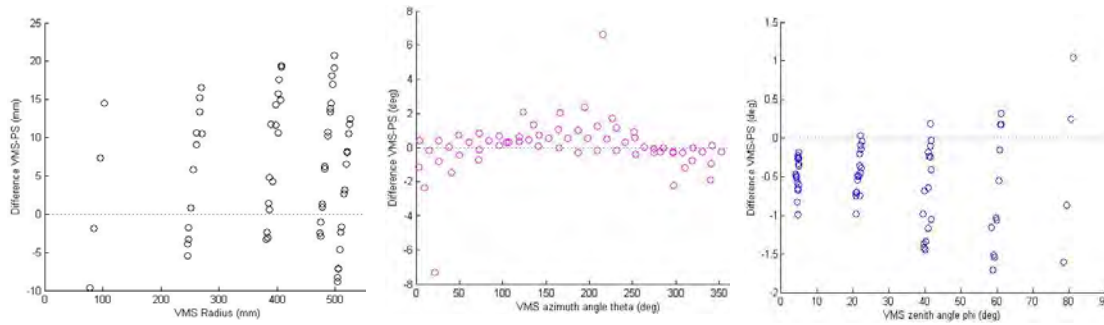


Figure 3.19: Difference between datasets vs VMS value for radius (left), theta (centre) and phi (right).

3.5.2 VMS Targets vs Flashes

Comparison of the estimates of the lamp centroids obtained by the two methods using VMS shows much closer correspondence than for the pin shadows. Plotting the 3D distance between target coordinates and flash coordinates vs lamp number (Fig. 20) shows values scattered mainly in the range 0.5 to 1.5 mm, with a mean of 1.13, stdev of 0.62, and maximum of 3.65 mm. Analysis of the component differences (target coordinates minus flash coordinates) in the X,Y,Z dimensions shows that the errors vary systematically over the dome geometry. The errors for X and Y oscillate between positive and negative values in groups of eight lamps (half of the dome circumference), suggesting a shift of origin in the X-Y plane. The means of the two datasets are: (-3.03, -13.38, 314.76) and (-3.21, -13.22, 314.22), with difference (0.18, -0.16, 0.54) and norm of 0.59 mm. After adjusting the means of the datasets to be equal (by subtracting their difference from the second dataset) the mean 3D error was reduced to 1.01, with stdev of 0.53 mm.

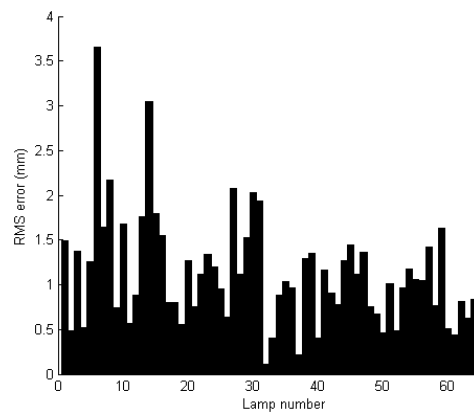


Figure 3.20: 3D distance between coordinates estimated by targets and by flashes, vs lamp number.

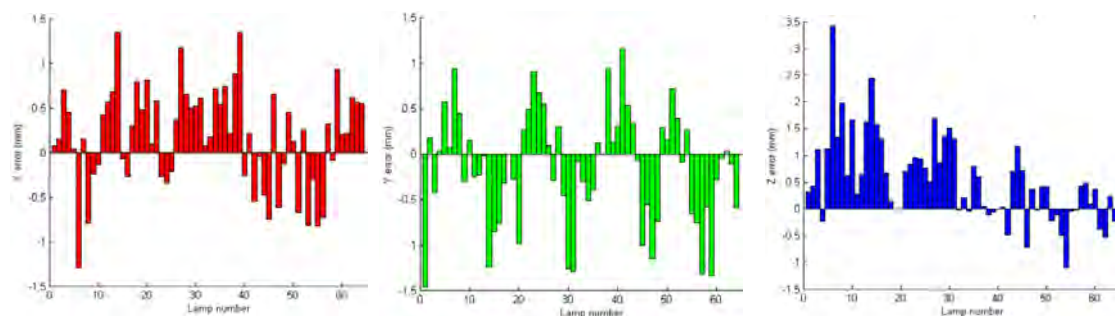


Figure 3.21: Differences in X, Y, Z dimensions between datasets for VMS target and flash vs lamp number.

After adjustment for the mean, the differences in radius, azimuth angle and zenith angle vs lamp number are plotted (Fig. 3.22), with rms values of 0.64, 0.10, 0.08 respectively. The azimuth θ error is mostly less than 0.1 degree, and the zenith ϕ error in the range ± 0.05 degree. The errors for Z, and hence also for ϕ , show a more uniform trend, suggesting a difference in the scaling of Z between the two methods.

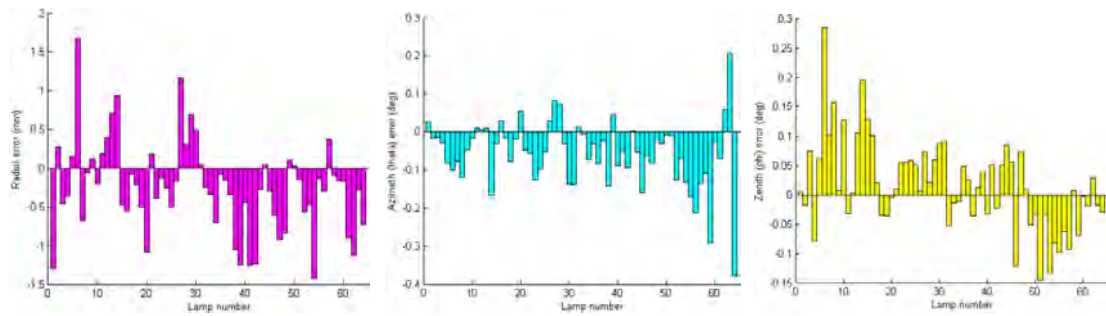


Figure 3.22: Differences in radius, azimuthal angle (θ), and zenith angle (ϕ) between datasets for VMS target and flash, after adjusting the means to be equal, vs lamp number.

3.5.3 Discussion

The objective of this study was to quantify the relative accuracy with which the lamp coordinates could be determined. A secondary objective was to determine the influence of errors in these coordinates on the process of estimating the normals of object surfaces, which requires repetitive solving of three simultaneous equations involving the cosines of the angles of incidence of rays from three lamps (MacDonald and Robson, 2010). This is critical if the overall aim is to be achieved of digitising cultural artefacts to resolve all of the man-made detail. A previous study showed that the main factors are the dimensions of tools and visual contrast sensitivity, and led to an estimate of finest detail of approximately $40\text{ }\mu\text{m}$ (MacDonald, 2010). By the Nyquist criterion, this requires a sampling pitch of $20\text{ }\mu\text{m}$, or 50 pixels/mm on the surface of the object, equivalent to a scanner resolution of 1200 dots per inch (dpi).

With close-range photogrammetry, i.e. at a range of less than 1 metre, it is expected to be able to locate all target positions to an accuracy of better than 10 microns. Although each individual image is limited in its pixel resolution, the use of relatively large retro-reflective target spots (diameter in range 2 to 5 mm) ensures that each spot has a diameter in the image of 6-15 pixels, and the centroid-finding algorithm can locate the geometric centre of each spot to an accuracy of 0.1 pixel. The redundancy inherent in many images (20 or more) from different viewpoints, combined with the iterative bundle adjustment procedure, produces a robust network. The scaling depends on the calibrated scale bar, traceable to national standards.

Table 3.1: Dimensions (mm) related to pin shadow technique.

The length of the pin, measured by vernier callipers to an accuracy of $\pm 0.05\text{ mm}$. Errors in this value are amplified by factors of up to 10x by the equations for Z and R (A.2 and A.3).	0.5 mm
The graph paper was ruled with a 2 mm square grid, and may also have had a scale error of up to 2% because of stretch. Also it was not lying perfectly flat on the table, causing some local distortion of the shadow.	2 mm
The length of the illuminated flash lamp, 14 mm, subtends an angle of 1.56° at the centre of the dome, which produces a penumbra around the shadow. This made it difficult to locate the precise location of the tip of the shadow, especially for low lamp angles (Fig. 3.23). Uncertainty in the shadow length, estimated at $\pm 1\text{ mm}$ is amplified by the equations A.2 and A.3, by a factor that for long shadows may exceed 10x that for short shadows, derived from the slope of the curves in Fig. A.2.	5 mm
The algorithm assumes a perfect hemisphere, with all lamps at an equal radius from the centre point, when in fact the dome is flattened at top by approximately 20 mm.	10 mm

Yet the differences between the three datasets were of the order of 10 mm for the pin shadows and 1 mm for the VMS methods. How can this be explained when individual targets can be located so accurately? The explanations lie in the mechanical movement of the structure and in the methodology. The poor accuracy of the first method, using pin shadows, was a consequence of the factors listed in Table 3.1. Taking the rms value of these three factors gives an overall uncertainty measure of 11.37 mm for the accuracy of lamp positions estimated by the pin shadow technique, similar to the mean 3D error between pin shadow and VMS datasets.

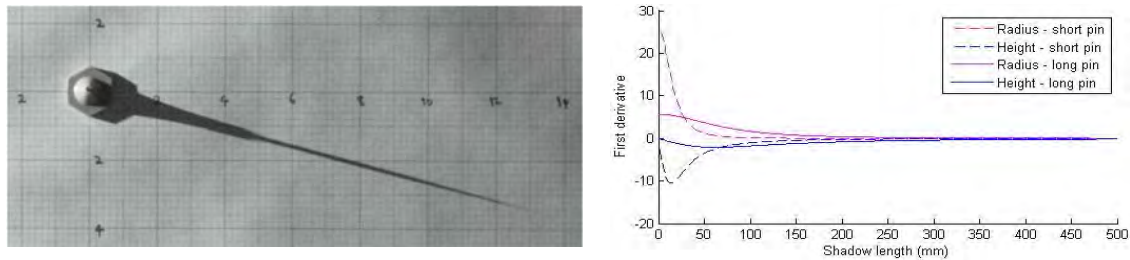


Figure 3.23: (left) Shadow cast by short pin from lamp 4 in tier 1 showing penumbra at tip; (right) first derivatives of the R and Z functions.

For the second method, using four reflective targets around each lamp, the uncertainty arose from the characteristics of the circuit boards and the processing steps applied to VMS results (Table 3.2). Taking the rms value of all these factors gives an overall uncertainty measure of 1.17 mm.

Table 3.2: Dimensions (mm) related to targets placed around lamps.

The transfer of the coordinate system from baseboard to dome by the pairs of juxtaposed targets assumes that the centroids of the upper and lower targets were identical. In practice the two may be in loose contact but still separated by 0.1 mm.	0.1 mm
No great care or consistency was employed in sticking the 256 target spots onto the three screw heads and flash lamp bulb of each of the 64 circuit boards. In some cases they were slightly angled to one side or partially folded. In three instances the spots were missing altogether and their positions had to be estimated.	0.3 mm
The circuit boards are mounted on the outside of the acrylic shell of the dome (on the opposite side to the target spots) with a rubber washer separating the board from the dome. The compressibility of each washer and the imprecise hand tightening of the nuts onto the screws at each of three positions on each board produces some variation in the angle of the board relative to the curvature of the dome (evidenced by the lack of convergence of the normals).	0.5 mm
The circuit boards were assembled and soldered by hand, and there is considerable variation in the length of the connecting wire at each end of the flash lamp. The uncertainty associated with this is estimated to be 0.5 mm in X, Y (plane of board) and 1 mm in Z.	1 mm
The algorithm for finding the barycentre of the flash from the screw head locations relies on the repeatability of the screw positions relative to the flash. In fact the holes in the flash board are larger than the screw shaft diameter, allowing a tolerance of ± 0.1 mm per screw.	0.2 mm

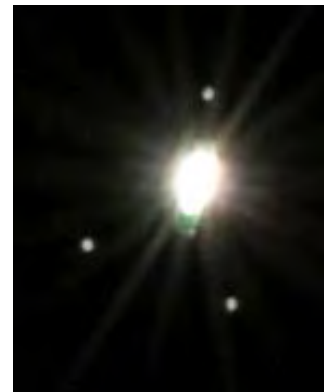
For the third method, using the flashes themselves as targets, other factors affected the accuracy of the results (Table 3.3). Taking the rms value of all these factors gives an overall uncertainty measure of 1.22 mm. This is consistent with the distribution of differences between the X,Y,Z coordinates estimated by the target and flash methods.

Table 3.3: Dimensions (mm) related to using flashes as targets.

The flash light is not a point source. The distance between the electrodes of the flash is approximately 8 mm, and the illuminated area of the interior of the xenon gas discharge within the tube is approximately rectangular of dimension 14 x 2.8 mm.	1 mm
The bright regions in the image corresponding to each flash are quite large (up to 50 pixels in diameter), elliptical in shape, and rotated. The central pixels are over-exposed and the whole shape is surrounded by lens flare, making it difficult to locate the centroid precisely (Fig. 3.24). With the 17 mm focal length of the zoom lens, the image resolution is approximately 2 pixels per mm on a surface at distance of 1 metre. Thus an error of 1 pixel in estimating the centroid of a flash corresponds to an error of approximately 0.5 mm.	0.5 mm
The lamps lie parallel to the surface of the hemisphere, so in the lower tiers they present more of an 'end view' to the camera. The flash may therefore be partially obscured by the internal structure of the flash tube, or by the dome wall around the cutout hole, or distorted by refraction through the glass, which may account for the larger errors observed in tiers 1-2.	0.5 mm

The results of the third method, using the flashes as targets, have the highest confidence level, because the processing by VMS is most direct. It should be noted, however, that the dome itself may not be entirely rigid. Although it is mounted around the base on a rigid steel frame, the top, where the acrylic shell is thinnest, is unsupported and may have 'slumped' under the force of gravity, when the dome was raised on its hinges for capture of the photogrammetric images.

Figure 3.24: Detail of Fig. 3.14c, showing rotated elliptical flash profile with lens flare.



3.6 Targets on Baseboard

Having established estimates of the lamp coordinates and some of the lens parameters by the Vision Measurement System (VMS) imaging methods were next applied to determine the transformation between points on the dome baseboard and pixels in the image. Information was derived from a series of images taken by the Nikon D200 camera in its usual mounting position with the 17-55 mm zoom lens. On the baseboard beneath the dome was a set of circular retro-reflective targets, arranged in three concentric rectangular patterns. When illuminated by the four flash lamps in Tier 5, which act as a ring light around the lens, the targets show up brightly against the dark background of the baseboard. At the high magnification of the 105 mm lens only the central rectangle lies within the image, whereas for the zoom lens at the 17 mm setting all three rectangles are visible.

The RAW image from the camera was converted via the DCRAW utility to 16-bit linear TIFF format, and the green channel extracted. All of the pixel values in the image were rearranged into a linear vector and sorted into descending order. The target regions were then located by scanning the list and noting the coordinates of each bright point not within a given radius (in the range 40-100 pixels) of all points already found. The centroid of each target in the image was then calculated as the weighted sum of all pixels in a region (of radius in the range 20-50 pixels) around the initial point:

$$\bar{i} = \frac{\sum_{ij} i \cdot P_{ij}}{\sum_{ij} i} \quad (3-8)$$

where P_{ij} is intensity value of the pixel at coordinates (i, j) . To reduce the effect of background noise, only pixel intensity values exceeding a threshold of $0.2 \times \text{maximum}$ within the target region were included in the summation. Although pixel coordinates and intensities are all integers, the position of the centroid could be determined by this method to sub-pixel accuracy. The target coordinates are sorted by radius and azimuthal angle (theta) so that they correspond to the numbering of the lamps in the dome.

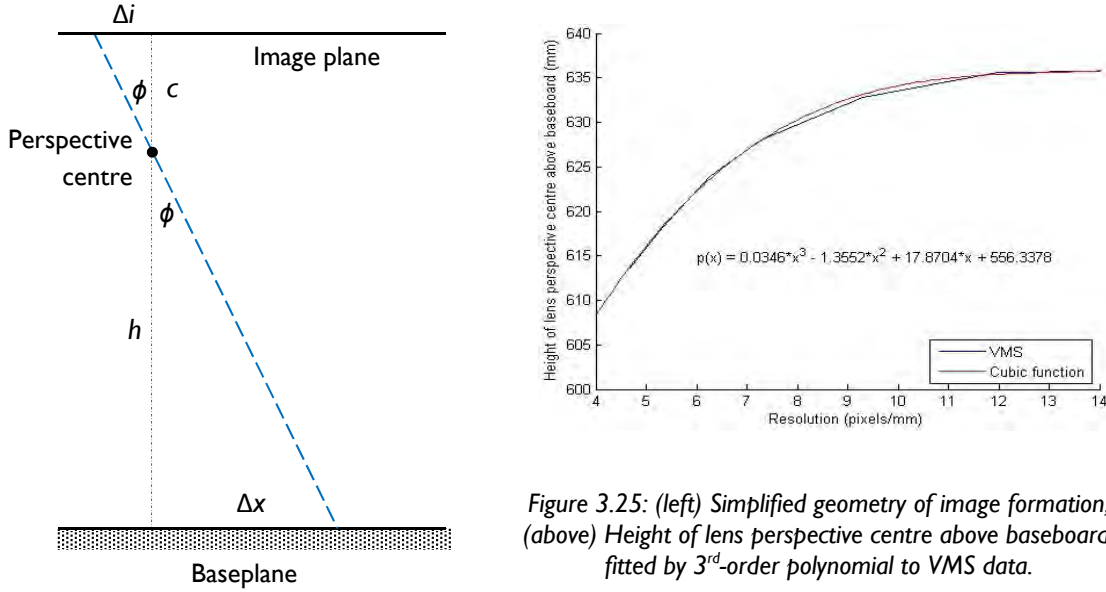


Figure 3.25: (left) Simplified geometry of image formation; (above) Height of lens perspective centre above baseboard, fitted by 3rd-order polynomial to VMS data.

A very simplified model of image formation was assumed (Fig. 3.25 left), in which all rays from positions in the baseplane pass through a single point in the lens, the perspective centre. A distance of Δx on the baseplane corresponds to a distance of Δi in the image plane, independent of position. It is further assumed that the angle ϕ of any ray relative to the optical axis is always preserved, so that by similar triangles the interior ratio $\Delta i : c$ is always equal to the exterior ratio $\Delta x : h$. The resolution \mathfrak{R} of the camera in pixels/mm is $\Delta i / \Delta x$, and the magnification \mathcal{M} is $c/h = p\Delta i / \Delta x$ where p is the dimension of one pixel (i.e. cell pitch in the sensor array). Using the known physical coordinates of the targets on the baseboard (previously determined by VMS), the image centroids relative to the image centre with a perfect lens could be related to the baseboard coordinates x_n, y_n by the linear relationship:

$$x_n = \frac{(i_n - i_c)}{\mathfrak{R}} + x_0, \quad y_n = \frac{-(j_n - j_c)}{\mathfrak{R}} + y_0 \quad (3-9)$$

where i_n, j_n are the pixel coordinates of the n^{th} centroid, i_c, j_c are the coordinates of the centre of the image plane, \mathfrak{R} is the resolution in pixels/mm on the baseboard, and x_0, y_0 are the coordinates of the central target. The Y axis is inverted because the camera's coordinate axes have the origin in the top left corner with the values increasing downwards. For the Nikon D200 raw image size of 3900x2616 pixels, the image centre coordinates are: $i_c = 1950.5, j_c = 1308.5$, and the pixel dimension p is $6.09 \mu\text{m}$. The VMS datum on the baseboard is taken as $x_0 = 0, y_0 = 0$. The results are plotted in Fig. 3.25 for the image from zoom lens set to its

widest field of view, nominally 17 mm. The resolution \mathfrak{R} can be found from Eq. 3-9 for any corresponding pair of image centroids and target coordinates spaced sufficiently far apart:

$$\mathfrak{R} = \frac{i_p - i_q}{x_p - x_q} \quad \text{for } (x_p - x_q) > d \quad (3-10)$$

Taking all pairs where the baseboard targets are spaced more than 10 mm, the mean value of \mathfrak{R} for the 17 mm lens is 3.62 mm in X and 3.68 mm in Y, with a standard deviation of 0.03 mm. This value can then be used as an index to obtain the height h_{pc} of the front perspective centre of the lens, using the third-order polynomial fitted to the values found by VMS (Fig. 3.25 right), in this case 613.39 mm. Given the magnification \mathcal{M} and the height h_{pc} the principal distance c of the lens can be calculated is:

$$c = h_{pc} \mathcal{M} = \frac{h_{pc} p \Delta i}{\Delta x} \quad (3-11)$$

In the present case the magnification \mathcal{M} is 0.028 from which the principal distance of the lens is 17.26 mm. The result of the simple linear projection of Eq. 3-11 shows a good correspondence between the reference (VMS) and predicted values, with a mean error over the 26 targets of 1.01 mm. It seems, however, that the errors increase with radius from the centre of the image plane (Fig. 3.26 right).

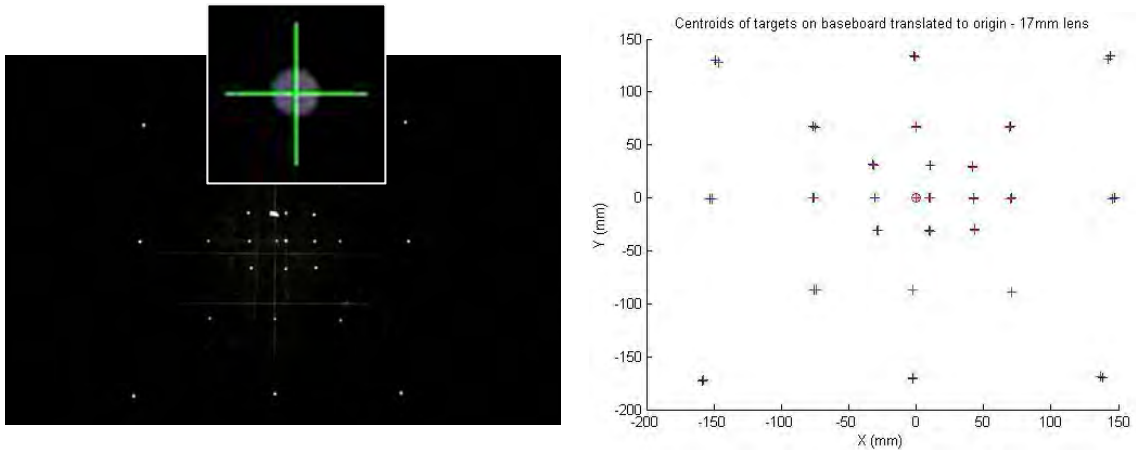


Figure 3.26: (left) Image of retro-reflective targets on the baseboard taken with zoom lens set at 24 mm under tier 5 flash lights; (inset) Cross-hairs show the location of centroid over the reflective target in image with diameter of 14 pixels and cross-hair length of 40 pixels; (right) coordinates of centroids of 26 targets in image scaled and translated to origin of physical targets on baseboard.

A better correspondence between coordinates in the image and coordinates on the baseboard can be achieved by fitting a 3x3 matrix \mathbf{M} to all target points, by the linear model:

$$- \begin{bmatrix} p(i_1 - i_c) & p(j_1 - j_c) & h_{pc} + c \\ \dots & \dots & \dots \\ p(i_n - i_c) & p(j_n - j_c) & h_{pc} + c \end{bmatrix} \mathbf{M} = \begin{bmatrix} x_{v1} - x_{v0} & y_{v1} - y_{v0} & z_{v1} - z_{v0} \\ \dots & \dots & \dots \\ x_{vn} - x_{v0} & y_{vn} - y_{v0} & z_{vn} - z_{v0} \end{bmatrix} \quad (3-12)$$

where $p = 0.00609$ mm is the spacing between pixels on the sensor, h_{pc} is the height of the front perspective centre of the lens above the baseplane, c is the principal distance of the image plane from the rear perspective centre of the lens, (x_{vn}, y_{vn}, z_{vn}) are the physical coordinates of the n^{th} target on the baseboard as measured by VMS, and (x_{v0}, y_{v0}, z_{v0}) are the coordinates of the VMS datum point. If Eq. 3-12 is written in vector form as $\mathbf{pM} = \mathbf{x}$ then regression fitting

over all n points is achieved by constructing the pseudo-inverse in Matlab by $\mathbf{M} = \mathbf{p} \backslash \mathbf{x}$. The errors between the estimated and VMS coordinates are smaller (mean of 0.28 mm) and show less dependence on radius (Fig. 3.28).

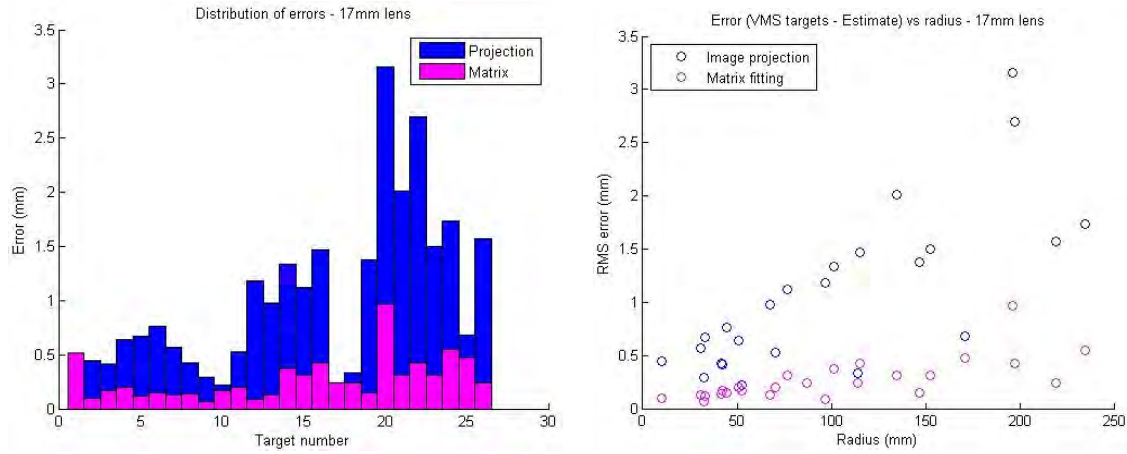


Figure 3.28: (left) Errors between VMS target coordinates and estimated coordinates in baseboard plane after fitting 3x3 matrix, using 26 points from an image with the 17mm lens; (right) Errors vs radius from VMS origin.

In the case of the data in Fig. 3.27, for $h_{pc} = 613.39$ and $c = 17.26$ mm, the matrix \mathbf{M} has coefficients:

$$\mathbf{M} = \begin{bmatrix} -35.5708 & -0.0708 & 0.0284 \\ 0.0262 & -35.5409 & 0.0013 \\ 0.0110 & 0.0020 & 0.0001 \end{bmatrix} \quad (3-13)$$

The two large values on the diagonal correspond to the inverse of the magnification factors for X and Y. They are slightly different, not because the pixels are not square but because the camera's optical axis is evidently not quite normal to the baseplane. The matrix can compensate for an offset of the origin (translation), different scales in X and Y, and rotation and tilt of the imaging plane.

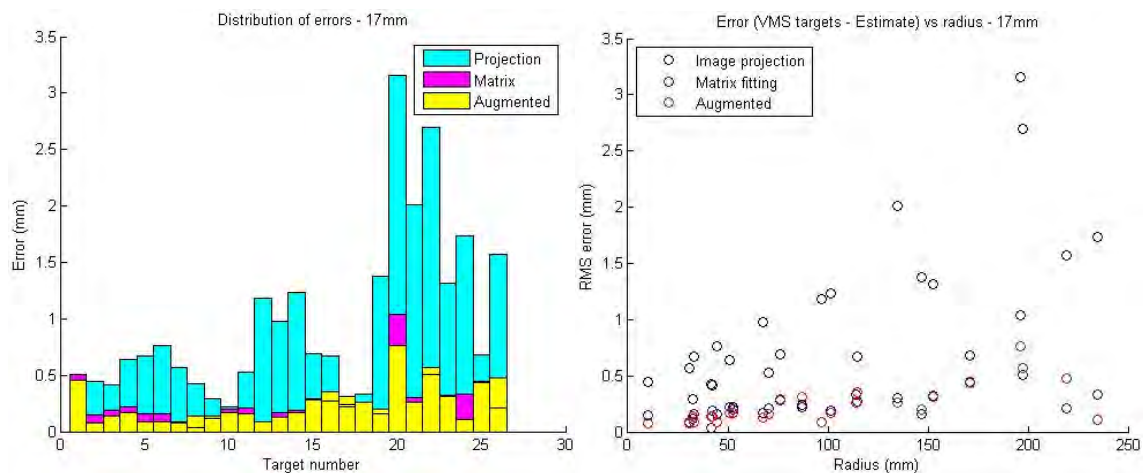


Figure 3.28: (left) Errors between VMS target coordinates and estimated coordinates on baseboard after fitting 4x3 augmented matrix, using 26 points from image with 17mm lens; (right) Errors vs radius from VMS origin.

The small values of the off-diagonal terms in \mathbf{M} in Eq. 3-13 indicate that these latter factors are small. Improvements to the accuracy of the transform can be made by adding terms to the matrix, for example by the cubic radial term:

$$- \begin{bmatrix} p(i_1 - i_c) & p(j_1 - j_c) & h_{pc} + c & r_1^3 \\ \dots & \dots & \dots & \dots \\ p(i_n - i_c) & p(j_n - j_c) & h_{pc} + c & r_n^3 \end{bmatrix} \mathbf{M}_A = \begin{bmatrix} x_{v1} - x_{v0} & y_{v1} - y_{v0} & z_{v1} - z_{v0} \\ \dots & \dots & \dots \\ x_{vn} - x_{v0} & y_{vn} - y_{v0} & z_{vn} - z_{v0} \end{bmatrix} \quad (3-14)$$

where $r_n = \sqrt{(p(i_n - i_c))^2 + (p(j_n - j_c))^2}$ and \mathbf{M}_A is the augmented 4x3 matrix. The small improvement in accuracy resulting from augmenting the matrix in this way is shown in Fig. 3.29. The mean error, which was reduced from 0.96 mm with stdev of 0.77 mm (projection) to 0.26 mm with stdev 0.20 mm (matrix), was further reduced to 0.25 mm with stdev of 0.17 (augmented matrix).

It is interesting to compare the linear formulation of Eq. 3-11 with the photogrammetric equation for single camera geometry (Cooper and Robson, 1996, page 17, Eq. 2.14a):

$$\mathbf{X}_A = \mathbf{X}_0 - \mu \mathbf{R}^t \mathbf{x}_a \quad (3-15)$$

In this form μ is a magnification factor and matrix \mathbf{M} corresponds to $\mu \mathbf{R}^t$, transposed for multiplication on the right. Note that Eq. 3-13 and 3-14 do not use the coordinates of the perspective centre (\mathbf{X}_0 in Eq. 3-15) because the collinearity condition (convergence of all rays through a single point) is not enforced in the present study. The off-diagonal terms represent rotations of the camera around each of the three axes. Because the camera is periodically removed from and replaced onto its mounting plate above the dome, there is inevitably some variation in the camera position (play of mounting screw in hole through the plate) and in the angles, especially in rotation about the Y axis. Likewise the uncertainty in the zoom and focus settings of the lens results in variation of the perspective centre and magnification factor.

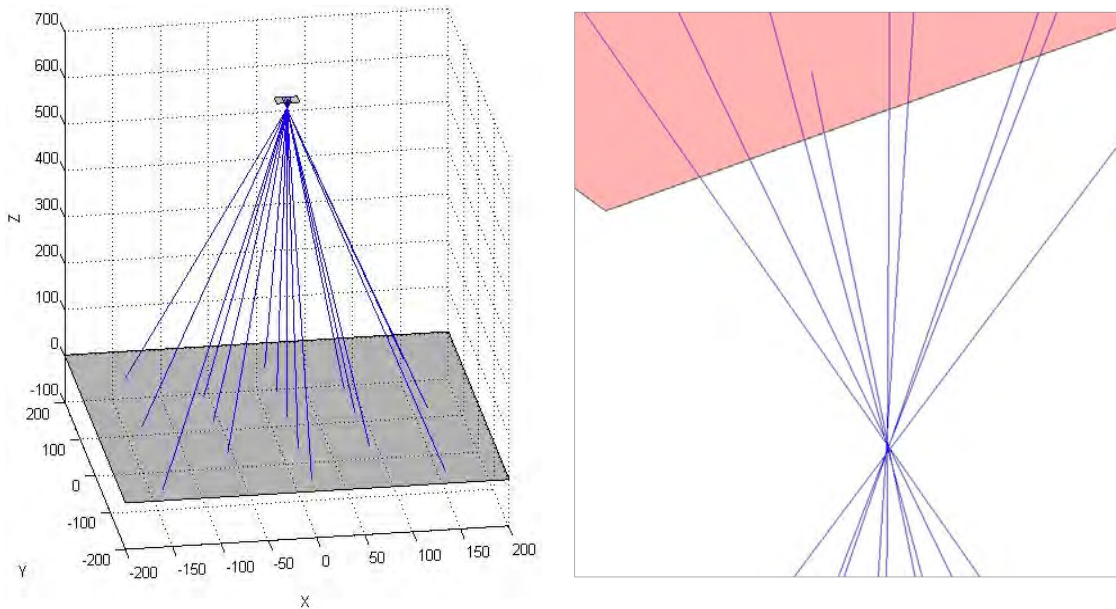


Figure 3.29: (left) Visualisation of rays from target centroids in image plane to corresponding target coordinates on baseboard; (right) Detail showing imperfect collinearity through perspective centre of lens.

The technique above provides a practical means of calibrating the camera geometry from a single image of the pattern on the baseboard, calculating the focal length c and resolution \mathfrak{R} from which the height h_{pc} is determined. From these values the matrix \mathbf{M} can be generated, to transform image pixel coordinates into physical coordinates on the baseboard. The

The zenith angle ϕ between the optical axis of the lens and the ray from C to the highlight point P is also the angle between the reflected ray and the normal \mathbf{n} :

$$\tan \phi = \frac{r_p}{h-t} = \frac{r_p - r_L}{t - z_L} \quad (3-16)$$

where $r_p = \sqrt{x_p^2 + y_p^2}$. The same angle ϕ is subtended between the normal and the ray from the lamp. Both rays and the normal lie in the same plane.

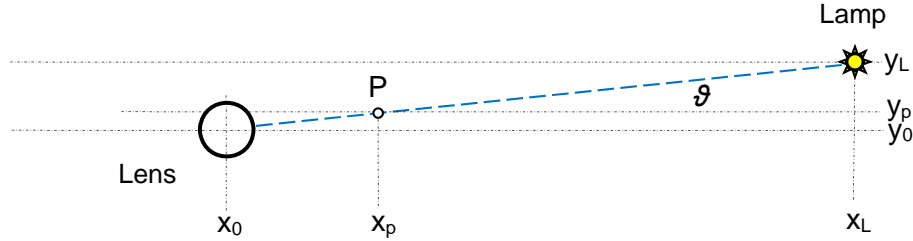


Figure 3.31: Plan view of reflection of a ray from the lamp to the glass sheet to the lens.

Operating the lamps in ‘ring flash’ mode with maximum f -number of 22, produces a dark image with a sharply-defined highlight on the glass for each flash lamp. These images are combined into a single composite image by taking the maximum value at each pixel (Fig. 3.31). The centroids are located for each highlight in the image, as described in the previous section for the baseboard targets, then sorted by radius and azimuthal angle into the lamp sequence.

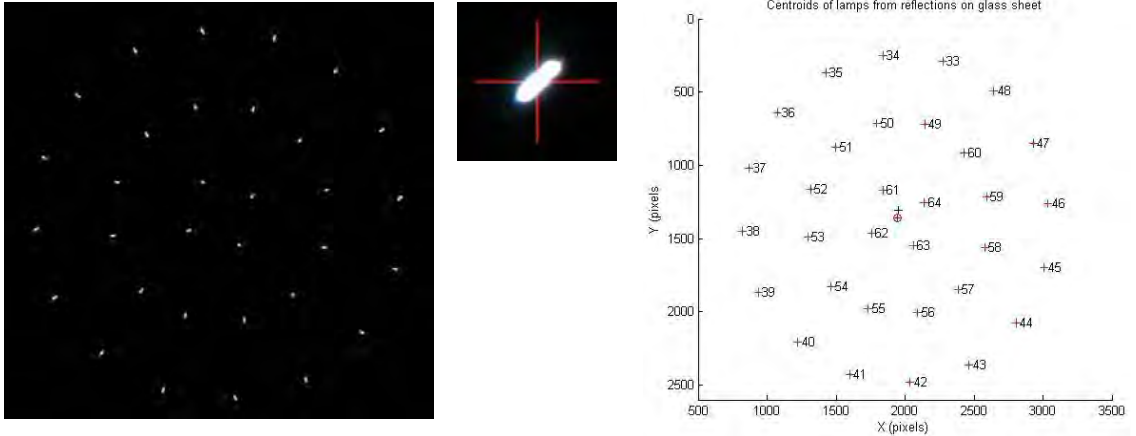


Figure 3.32: (left) Composite of three images of Tiers 3,4,5 taken with zoom lens set at 17 mm; (centre) enlarged detail of centroid of lamp reflection; (right) image centroid coordinates of lamps 33-63.

Looking down the optical axis from the camera (Z axis), the highlight P on the glass is located on a chord between the lamp L and the centre C (Fig. 3.31). Both rays have the same azimuthal angle θ :

$$\tan \theta = \frac{y_p - y_0}{x_p - x_0} \quad (3-17)$$

The projection of the ray $\vec{\ell} = \overrightarrow{PL}$ onto the baseplane is $\sin \phi$ and onto the Z axis is $\cos \phi$. Any point Q along the vector can therefore be represented as:

$$Q = P + \lambda \vec{\ell} = (x_p, y_p, z_p) + \lambda < \cos \theta \sin \phi, \sin \theta \sin \phi, \cos \phi > \quad (3-18)$$

where λ is a parametric scaling factor. Assuming that the dome can be approximated as a hemisphere, the lamp is located where the vector intersects the hemisphere:

$$(x - x_d)^2 + (y - y_d)^2 + (z - z_d)^2 = R^2 \quad (3-19)$$

where (x_d, y_d, z_d) are the coordinates of the geometric centre point D of the dome and R is the dome radius. Substituting Eq. (3.18) into (3.19) gives:

$$(x_p + \lambda \cos \theta \sin \phi - x_d)^2 + (y_p + \lambda \sin \theta \sin \phi - y_d)^2 + (z_p + \lambda \cos \phi - z_d)^2 = R^2 \quad (3-20)$$

The solution is a quadratic in λ :

$$a\lambda^2 + b\lambda + c = 0 \quad (3-21)$$

where:

$$\begin{aligned} a &= (\cos \theta \sin \phi)^2 + (\sin \theta \sin \phi)^2 + \cos^2 \phi \\ b &= 2((x_p - x_d)\cos \theta \sin \phi + (y_p - y_d)\sin \theta \sin \phi + (z_p - z_d)\cos \phi) \\ c &= (x_p - x_d)^2 + (y_p - y_d)^2 + (z_p - z_d)^2 - R^2 \end{aligned} \quad (3-22)$$

The calculated lamp positions can be visualised by plotting in 3D the ray for each lamp from the lens perspective centre to the centroid of the highlight on the glass plate and thence to the point of intersection with the flattened hemispherical surface of the dome (Fig. 3.33).

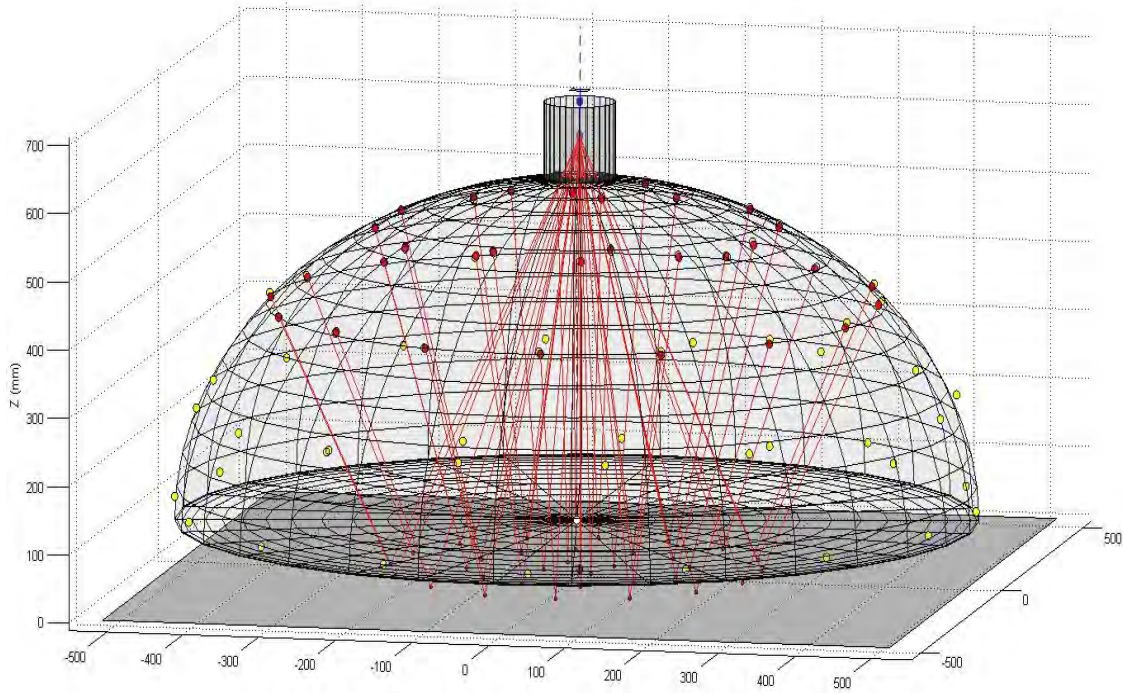


Figure 3.33: Rays from front principal point in lens reflecting off glass plate and intersecting the dome surface for lamps in tiers 3,4,5. The VMS lamp coordinates are represented by yellow dots and the calculated coordinates by red dots.

3.8 Reflections off Billiard Ball

A black billiard ball was placed on the baseboard of the dome, offset from the optical axis of the camera by a displacement (x_b, y_b) , small relative to the camera height h . A ray of light from the lamp strikes the ball at point P with coordinates (x_p, y_p, z_p) and is reflected toward the lens of the camera, where it appears at coordinates (i_p, j_p) in the image.



Figure 3.34: Two views of billiard ball on base board in north-west and south-east positions, taken with the Nikon zoom lens set at 55 mm, illuminated by the four flash lamps in Tier 5.

The image resolution on the baseboard was determined from observation of a scale ruler by the scaling factor \mathfrak{R} in pixels/mm and the ball's physical radius ρ is known from prior measurement with a micrometer. From each image of the ball can be observed (Fig. 3.34) the pixel coordinates of: (a) the retro-reflective targets on the baseboard; (b) the outline of the ball; and (c) the highlights on the ball's surface corresponding to each light source.

3.8.1 Outline Detection

The pixel coordinates of the outline of the ball can be determined from edge detection in the image as a point set $\wp = \{x, y\}$. Cross-sections of the image region enclosing the black ball (Fig. 3.35a) show that the dark inner area can be clearly discriminated from the lighter outer area (i.e. the baseboard in the background) by a simple threshold, in this case the pixel value of 400 in a 16-bit image. The location of the boundary is found by scanning each row and column inwards from the region boundary until the pixel value falls below the threshold.

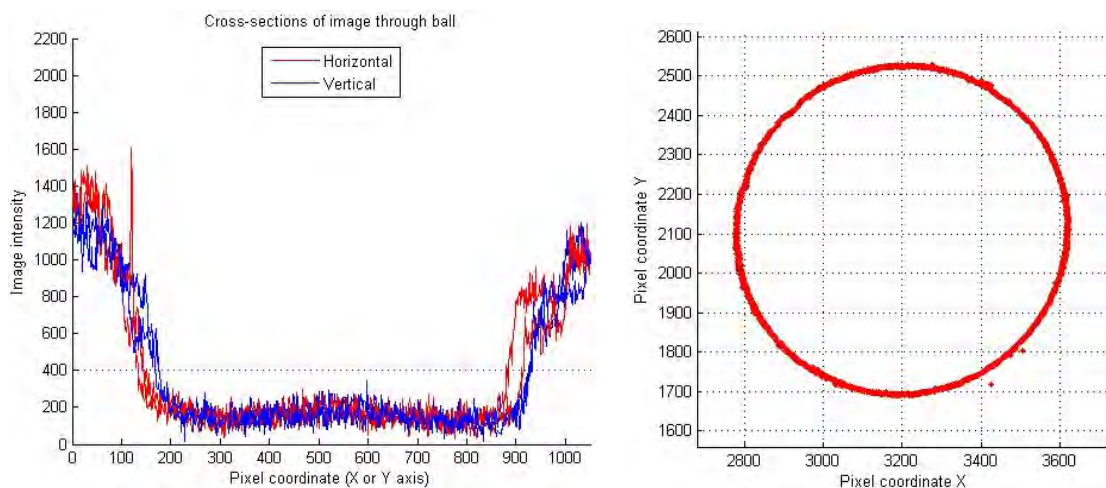


Figure 3.35: (left) Horizontal and vertical cross-sections through the billiard ball; (right) detected outline.

The resulting point set \wp is shown in Fig. 3.35 right, which for the case of the south-east ball position contains 2524 coordinate pairs. It is not exactly circular but elongated along the radius from the image centre.

3.8.2 Highlight Detection

Five successive images recorded the highlights on the surface of the ball from the five tiers of lights in the dome. These were combined into a single image by taking the maximum value at each pixel location, giving a composite image with 64 highlights (Fig. 3.36 left).

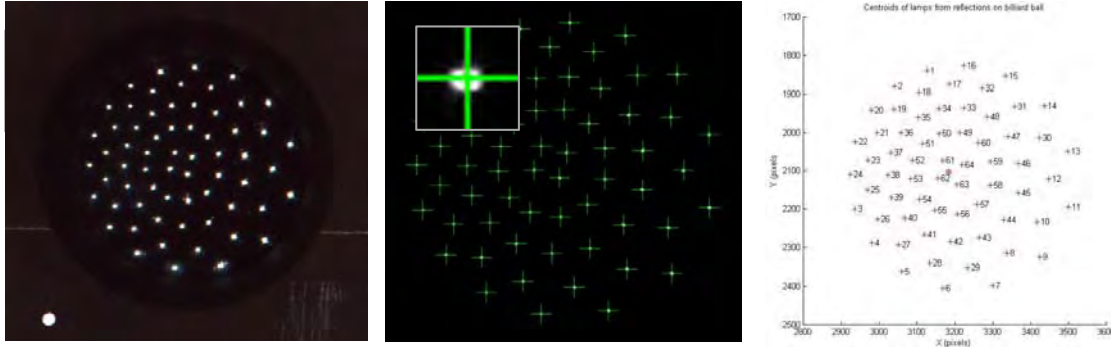


Figure 3.36: (left) Composite of five images showing highlights of 64 lamps on billiard ball in south-east position – the inset shows a magnified view of one highlight; (centre) green crosses indicating the calculated positions of centroids; (right) ordered and numbered centroid coordinates.

Each highlight covers an area of approximately 10x10 pixels but exhibits chromatic aberration caused by the lens. Therefore the centroid computation was based on only the green channel. All highlight clusters with normalised pixel values exceeding a threshold of 0.5 were identified and the centroid of each calculated as a weighted sum, using the pixel value as a weighting of the x and y coordinates (Fig. 3.36 centre). The centroid coordinate pairs were then grouped into tiers, based on the radial distance from the group centre, and sorted into order corresponding to the lamps in the dome (Fig. 3.36 right).

3.8.3 Horizon Effect

The critical problem is to determine the coordinates (x_b, y_b, z_b) of the ball centre B. Then, knowing the coordinates (x_p, y_p, z_p) of the highlight point P, the orientation of the normal vector \mathbf{n} can be determined, after which the vector \overrightarrow{PL} from P to the lamp L can be found in a similar way as for reflection off a plane glass surface. The geometry is projective, so what the camera ‘sees’ in the image plane are the pixel coordinates as if they were projected onto the baseboard with a resolution of \mathfrak{R} pixels/mm. All pixel coordinates (i_p, j_p) , $p = \{1..n\}$ relative to the image centre (i_c, j_c) can be transformed into physical coordinates $(x_{vp}, y_{vp}, 0)$ on the -baseboard (in mm), relative to the VMS datum $(x_{v0}, y_{v0}, 0)$, by Eq. 3-12 or 3-14. It follows that the observed pixel coordinates of each position (x, y, z) on the ball surface correspond to the projection along the ray to its intersection with the baseboard at coordinates $(x', y', 0)$. In the case where the ball centre is on the optical axis (Fig. 3.37a) the tangential rays at points T_1 and T_2 are projected symmetrically onto the baseplane at distances $r'_{t1} = r'_{t2}$ from optical axis, each subtending the half-angle ψ :

$$\tan \psi = r'_{t1}/h \quad \text{and} \quad \sin \psi = \rho/(h - \rho) \quad (3-23)$$

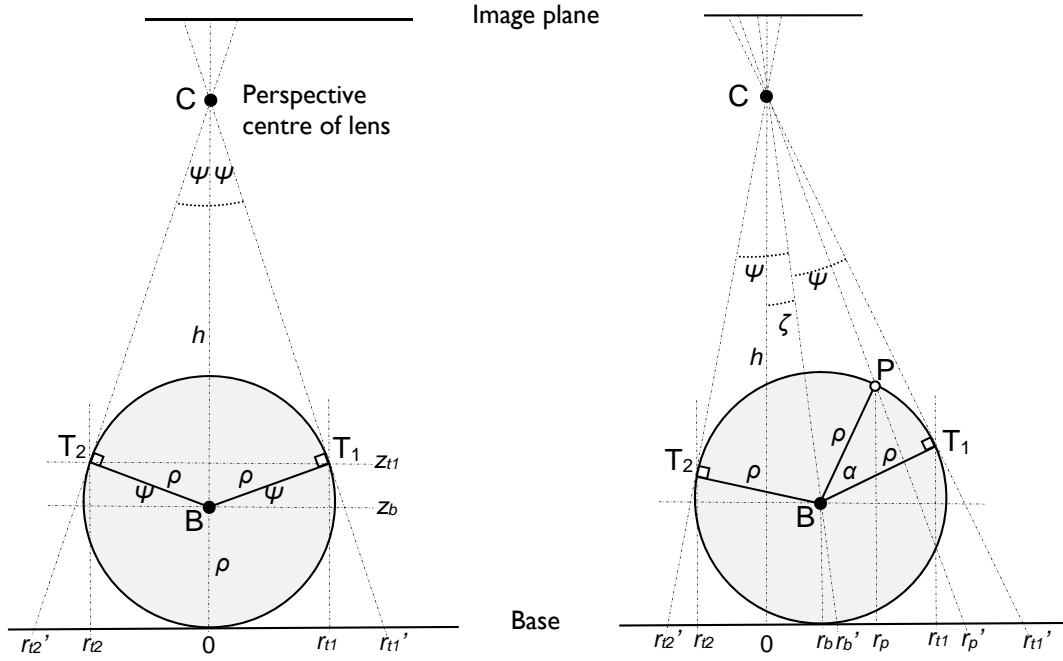


Figure 3.37: Elevation geometry of ball: (left) tangent rays in case where the ball is centred on optical axis; (right) tangent rays and highlight in general case where the ball is off axis.

The relationship between ψ and h is shown in Fig. 3.37 left. The height of T₁ and T₂ above baseplane is:

$$z_{t1} = z_{t2} = z_b + \rho \sin \psi = \rho(1 + \sin \psi) \quad (3-24)$$

and the actual radii of T₁ and T₂ are:

$$r_{t1} = r_{t2} = \rho \cos \psi \quad (3-25)$$

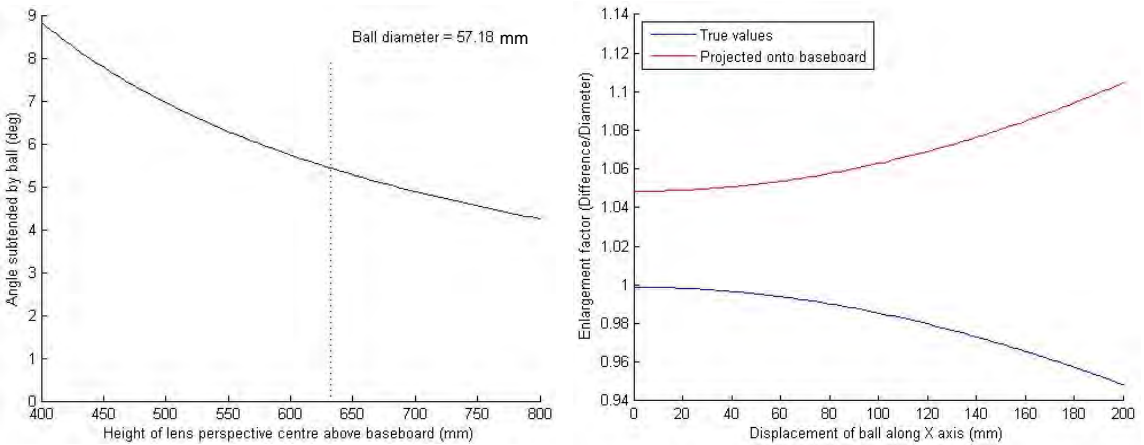


Figure 3.38: (left) Angle subtended by billiard ball vs height of perspective centre of lens above baseboard; (right) Ratio of difference between radial values divided by ball diameter vs displacement of ball from centre.

For the measured values of ball diameter $2\rho = 57.18$ mm and height of the perspective centre of the 55mm lens above the baseboard $h = 635.8$ mm, the angle subtended to the camera by the ball on the central axis is $2\psi = 5.40^\circ$. The height of the camera is approximately 12x the ball diameter and the apparent diameter of the ball outline is enlarged by a factor of $r'_{t1}/\rho = 29.97/28.59 = 1.0482$. In general the ball lies off to the side of the optical axis (Fig. 3.37 right),

so that the ray $\vec{c} = \overrightarrow{CB}$ to the ball centre subtends an angle ζ with the optical axis. Rays through the perspective centre of the lens C are tangential to the ball at points T_1 and T_2 , and meet the baseplane at radial distances r'_{t1} and r'_{t2} from the optical axis, making angles

$$\zeta - \psi = \text{atan}(r'_{t1}/h) \quad \text{and} \quad \zeta + \psi = \text{atan}(r'_{t2}/h) \quad (3-26)$$

with the optical axis, and producing a 'horizon effect'. When T_1 and T_2 are on opposite sides of the optical axis (i.e. r'_{t1} and r'_{t2} have opposite signs), both points are above the equator of the ball. When both points are on the same side then the nearer point is below the equator.

3.8.4 Locating Ball Centre

The centre of the ball lies on the bisector of the angle subtended at C . Hence the radial coordinate of the centre projected onto the baseplane is:

$$r'_b = h \tan \zeta = h \tan(((\zeta - \psi) + (\zeta + \psi))/2) = h \tan\left(\left(\text{atan}\left(\frac{r'_{t1}}{h}\right) + \text{atan}\left(\frac{r'_{t2}}{h}\right)\right)/2\right) \quad (3-27)$$

The true (unprojected) radial coordinate of the ball centre is therefore:

$$r_b = (h - \rho) \tan \zeta = (h - \rho) r'_b / h \quad (3-28)$$

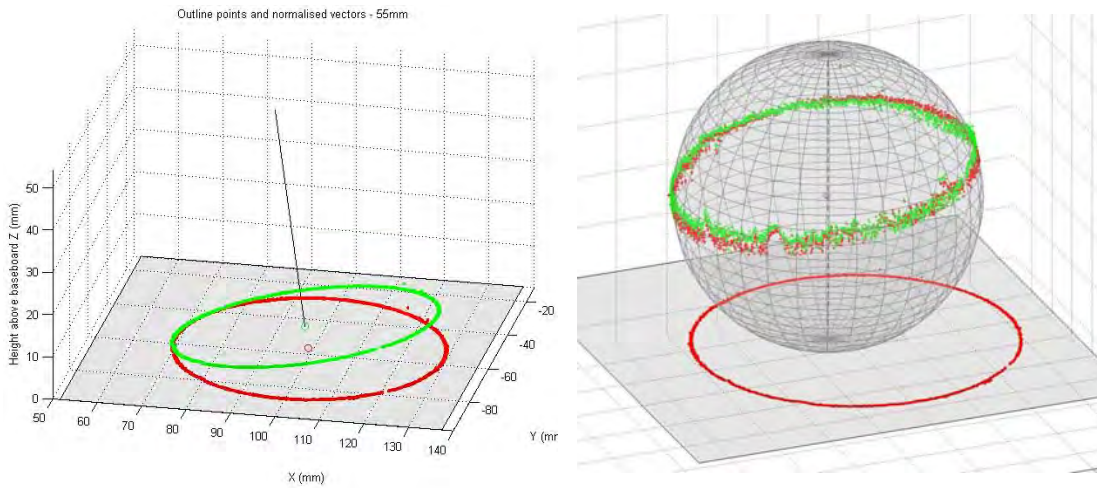


Figure 3.39: (left) Effect on outline of normalising the length of the vector from each point to perspective centre; (right) positions on surface of fitted sphere of normalised vectors (green) and tangents (red).

One way to estimate the ball centre $B = (x_b \ y_b \ \rho)$ is to normalise the length of the vector $\vec{v} = P - C$ from the perspective centre $C = (x_c \ y_c \ z_c)$ to every outline point in \wp projected onto the baseplane $P = (x_p \ y_p \ 0)$:

$$\|\vec{v}\| = \frac{P-C}{\|P-C\|} \quad (3-29)$$

The normalised vectors then lie on the surface of a unit cone with vertex at C , with its axis approximated by the mean of all \vec{v}_m over \wp :

$$\vec{v}_m = \text{mean}(\vec{v}) \quad (3-30)$$

The ball centre point B lies on this axis:

$$B = C + (h - \rho) \vec{v}_m / h \quad (3-31)$$

The effect of equalising all of the vectors is shown in Figure 3.39. The outline is tilted up from the baseplane to lie on a tilted plane normal to the vector \vec{v}_m .

Another way to estimate the position of the ball centre is to use regression to find the best-fitting sphere of radius ρ for which all vectors from the perspective centre to outline points on the baseplane would be tangential. This is conveniently done by the Matlab function `nlinfit`. The surface of the ball is represented as a sphere around the centre coordinates (x_b, y_b, ρ) :

$$(x - x_b)^2 + (y - y_b)^2 + (z - \rho)^2 = \rho^2 \quad (3-32)$$

Every point P on the baseplane can be mapped to a corresponding point P' tangential to the surface of the sphere along the vector $\vec{v} = P - C$:

$$P' = C + \lambda \vec{v} \quad (3-33)$$

where λ is a scalar. The value of λ corresponds to the intersection n of the vector with the surface of the sphere, and is found as a solution to the quadratic in λ :

$$a\lambda^2 + b\lambda + c = 0 \quad (3-34)$$

for which, by defining $\vec{w} = P' - B$, the coefficients can be written very compactly as:

$$a = \vec{v}\vec{v}^T \quad b = 2\vec{v}\vec{w}^T \quad c = \vec{w}\vec{w}^T - \rho^2 \quad (3-35)$$

The quadratic factor $\sqrt{b^2 - 4ac}$ is set to zero because the points are supposed to be tangential and so the roots of Eq. 3-35 must be equal. Hence

$$\lambda = -b/2a = -\vec{v}\vec{w}^T / \vec{v}\vec{v}^T \quad (3-36)$$

The intersection points form a close-fitting girdle around the sphere, shown in red in Fig. 3.38b and coincide well with the normalised vectors.

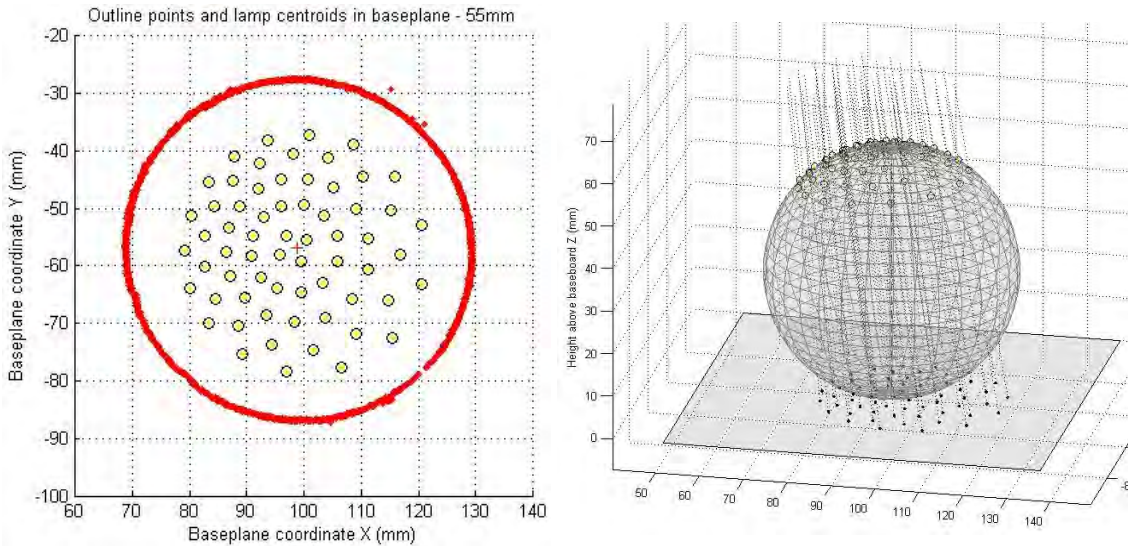


Figure 3.39: (left) Outline point set and highlight points projected onto baseplane in 2D; (right) Location of the highlight points on the baseplane and their location on the ball surface (yellow) at the intersection of each vector to the perspective centre of the lens.

3.8.5 Determining Lamp Coordinates

The coordinates of each highlight point on the ball surface can also be found by Eq. 3-32 to 3-35, with P being the image centroid of the highlight point projected onto the baseplane by Eq. 3-13. In this case the full form of the quadratic must be used, and the negative root is taken to give the shorter value of λ where the vector intersects with the upper surface of the sphere (Fig. 3.39b). The reflection geometry at the surface is shown in Fig. 3.40. The surface normal \vec{n} at point P can be written as a unit vector from the centre B:

$$\vec{n} = (P - B)/\rho \quad (3-37)$$

The angle ϕ between the vector \overrightarrow{CP} , denoted by \vec{c} , and the normal \vec{n} is:

$$\cos \phi = \frac{\vec{c} \cdot \vec{n}}{\|\vec{c}\| \|\vec{n}\|} \quad (3-38)$$

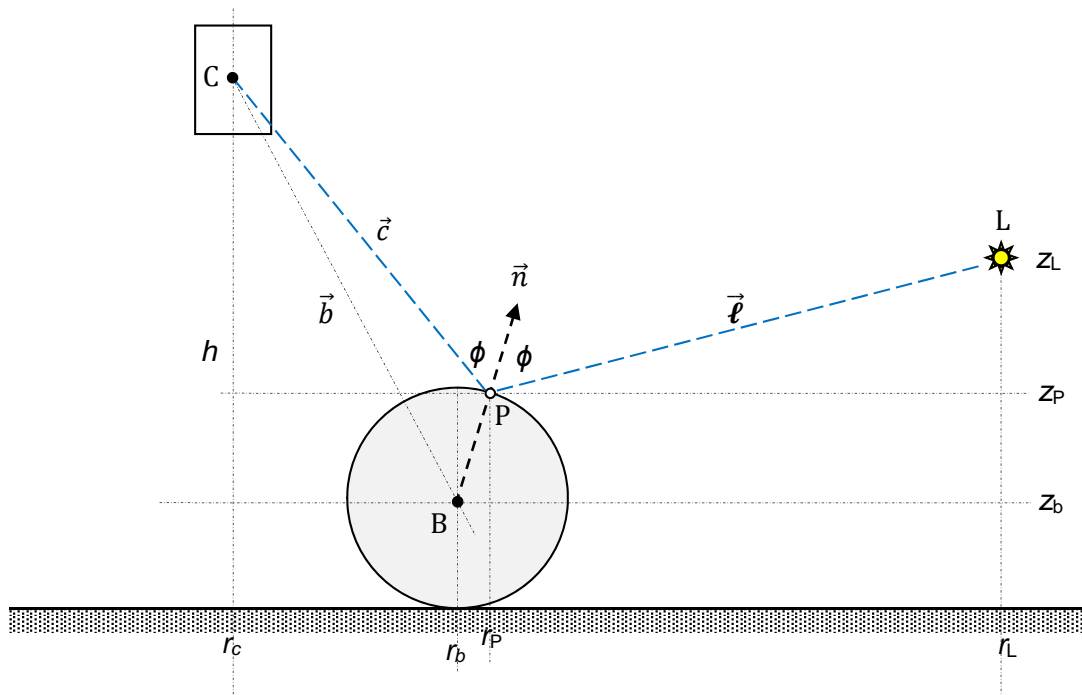


Figure 3.40: Elevation view of reflection of a ray from the lamp to the ball to the lens, in the plane of the rays.

Because the angle of reflection equals the angle of incidence ϕ , the vectors \overrightarrow{CP} and \overrightarrow{PL} , denoted by \vec{c} and $\vec{\ell}$, must both subtend the same angle ϕ with \vec{n} and lie in the same plane. It follows that the sum of the normalised vectors pointing outwards from P must lie on the normal:

$$-\frac{\vec{c}}{\|\vec{c}\|} + \frac{\vec{\ell}}{\|\vec{\ell}\|} = 2 \cos \phi \frac{\vec{n}}{\|\vec{n}\|} \quad (3-39)$$

The relationship of the vectors \overrightarrow{CP} and \overrightarrow{PL} is shown in Fig. 3.41, where the angle of reflection equals the angle of incidence on either side of the normal. L lies in the same plane as points C, P and B. The coordinates of the lamp L can then be written as:

$$L = P + \lambda \vec{\ell} \quad (3-40)$$

This is analogous to Eq. 3-17 for reflection off a glass sheet, and can be solved in conjunction with Eq. 3-18 to find the coordinates of the intersection of the vector $\vec{\ell}$ with the hemisphere and the corresponding value of λ .

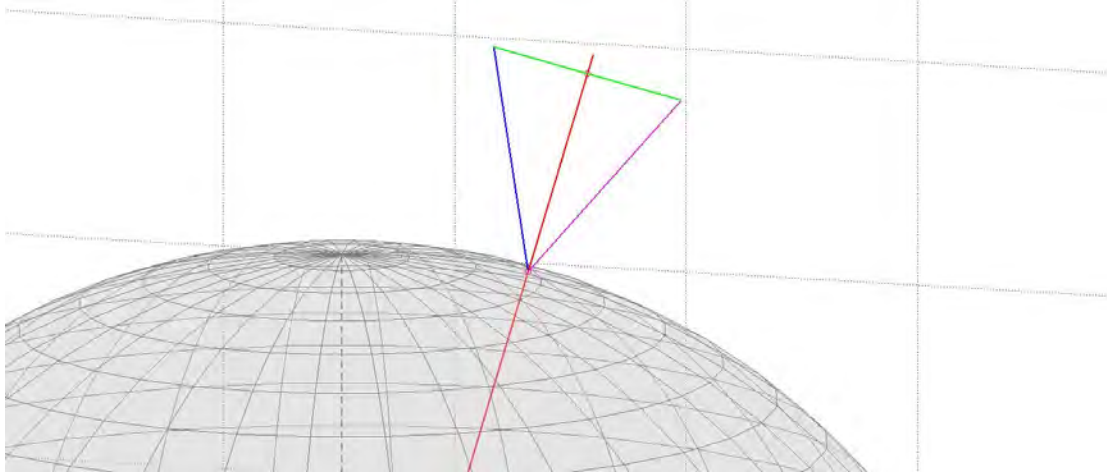


Figure 3.41: Geometry of reflection at the surface of the ball. The normal (red) from the ball centre bisects the angle between the incident ray (blue) and reflected ray (magenta). The displacements of both vectors from the normal are equal (green).

The plane in which lie points C, B, P, L and vectors \vec{c} , \vec{n} and $\vec{\ell}$ can be written in the form:

$$ax + by + cz + d = 0 \quad (3-41)$$

where a, b, c are defined by Cramer's Rule, for any value of $d \neq 0$:

$$a = \frac{-d}{D} \begin{vmatrix} 1 & y_1 & z_1 \\ 1 & y_2 & z_2 \\ 1 & y_3 & z_3 \end{vmatrix}, \quad b = \frac{-d}{D} \begin{vmatrix} x_1 & 1 & z_1 \\ x_2 & 1 & z_2 \\ x_3 & 1 & z_3 \end{vmatrix}, \quad c = \frac{-d}{D} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}, \quad D = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} \quad (3-42)$$

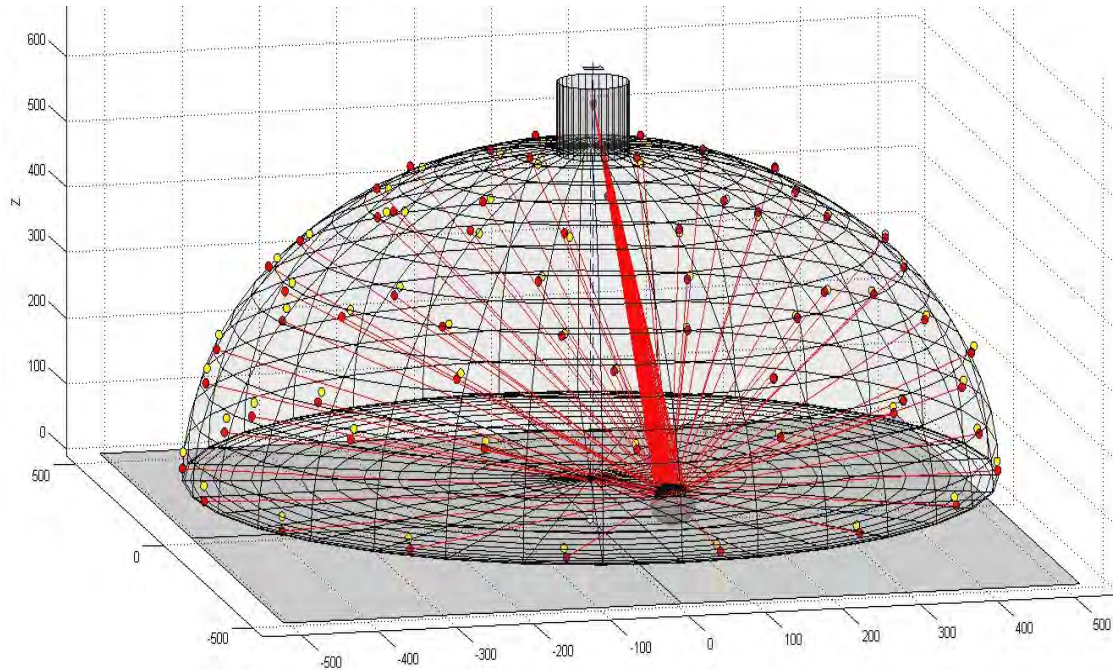


Figure 3.42: Rays from front principal point in lens reflecting off billiard ball and intersecting dome surface for all 64 lamps. VMS lamp coordinates are represented by yellow dots and calculated coordinates by red dots.

The calculated lamp positions can be visualised by plotting the ray for each lamp from the lens perspective centre to the highlight point on the surface of the ball and thence to the point of intersection with the flattened hemispherical surface of the dome (Fig. 3.42).

3.9 Analysis of Errors

3.9.1 Glass Sheet

The geometry of the constructions for both the glass sheet and the billiard ball relies on knowledge of the coordinates of point C, the front perspective centre of the lens (Figs. 3.29, 3.36 and 3.40), from a function interpolated through VMS estimates for different zoom focal lengths (Fig. 3.24). For the glass sheet, by adjusting the height of C in Fig. 3.32 the set of rays can be made to move outward and inward over the surface of the dome, like the opening and closing of the spokes of an umbrella.

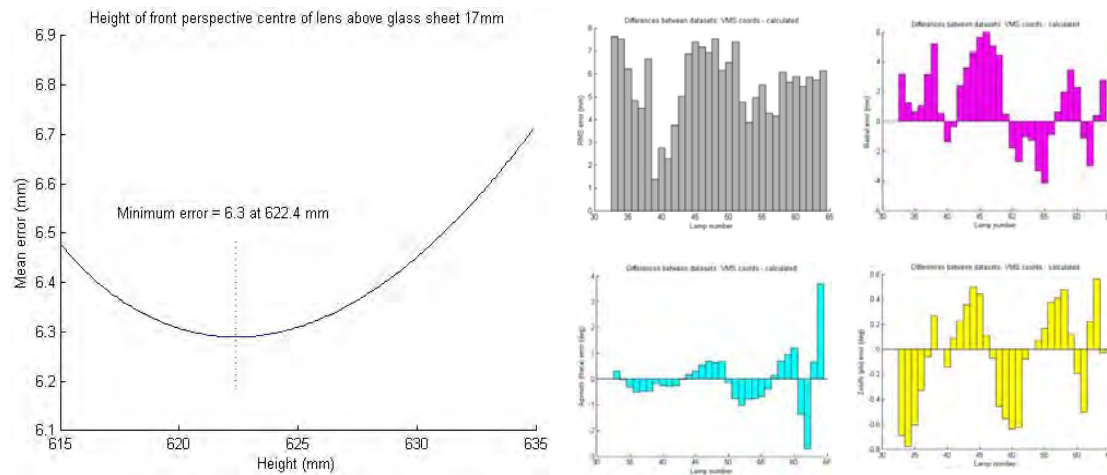


Figure 3.43: (left) Variation of mean error over 32 lamps with height of lens perspective centre; (right) Distribution of errors for RMS (grey), radius (magenta), azimuth angle theta (cyan) and zenith phi (yellow).

By repeating the computation for varying heights of C the minimum RMS error between the calculated and VMS reference coordinates for the 32 lamps in tiers 3, 4 and 5 was found to be 6.29 mm for a height of 622.4 mm (Fig. 3.43). At this optimum, the mean absolute errors in radius, theta and phi were 2.49 mm, 0.69° and 0.32° respectively. The distributions of errors show systematic variations in radius, theta and phi, suggesting an eccentricity in the shape of the dome. Repeating the error minimisation procedure for a series of focal lengths for the zoom lens over the full range 17 to 55 mm gave the corresponding optimum heights of C (Table 3.2 and Fig. 3.43).

Table 3.2: Calculated optimum height of lens perspective centre and VMS estimate vs focal length of lens.

Focal length (mm)	Num-ber of lamps	RMS fitting error	Calculated optimum height of perspective centre	VMS estimate
17	40	6.29	622.4	613.66
20	30	6.53	629.6	618.19
24	26	6.64	639.0	623.93
28	24	6.56	645.5	627.96
35	14	5.76	643.6	632.80
45	10	5.57	655.6	635.64
55	4	5.51	667.5	635.73

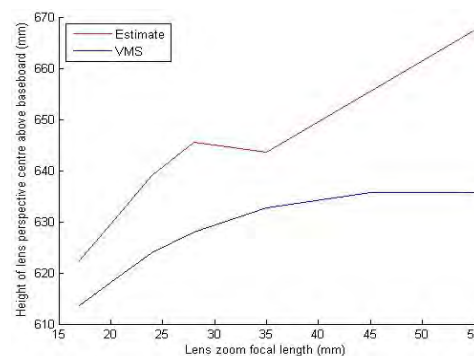


Figure 3.44: Calculated optimum vs VMS height.

Note that as the focal length increased and the field of view of the lens diminished, the number of lamps visible in the rectangular image of reflection from the glass sheet also diminished, until at 55mm focal length only the four lamps in Tier 5 were visible. The disparity between the calculated optimum and the VMS estimate may be explained by the absence of lamps in Tiers 1 and 2, forcing the calculation to prefer a steeper angle to match the lamp positions in the flattened upper regions of the dome.

3.9.2 Billiard Ball

For the billiard ball with the 55 mm lens, with the perspective centre fixed at the height reported by VMS, the mean error over all lamps between the VMS coordinates and the intersection of each reflected ray with a hemisphere was 23.14 mm. By moving the ball centre position systematically in X and Y, the mean error was reduced to a minimum of 9.23 at offsets $(\Delta x, \Delta y) = (-0.3, 0.6)$ mm, as shown in Fig. 22a. By allowing the perspective centre height also to vary, the mean error was further reduced to 6.86 at a height offset Δh of -73.5, and ball centre offset $(\Delta x, \Delta y) = (-1.1, 0.9)$, and the errors were well distributed over all lamps (Fig. 3.45b).

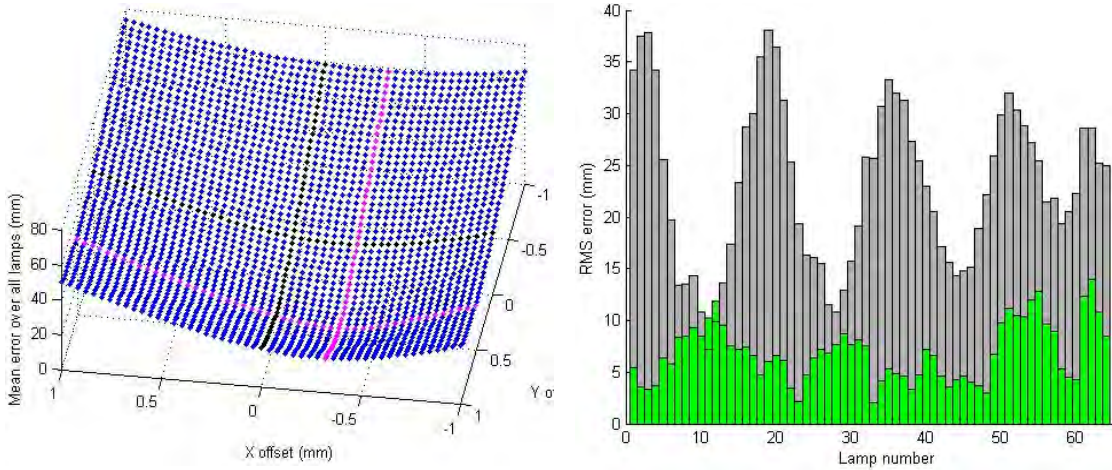


Figure 3.45: (left) Mean error over all lamps with variation in X and Y of ball centre; (right) Distance between calculated lamp position and VMS coordinates for each lamp, before (grey) and after (green) optimisation.

Because of its generally asymmetrical projection onto the baseplane, the outline of the ball is not circular but is stretched radially away from the axis (Fig. 3.24b), so the centre coordinates of the ball cannot be calculated simply as the centroid of the outline point set \mathcal{P} or as the mean of r'_{t1} and r'_{t2}

$$r'_b \neq h (\tan(\zeta - \psi) + \tan(\zeta + \psi))/2 \quad (3-43)$$

The true (unprojected) radial coordinate of the ball centre is given by Eq. 3-27. The error in calculation of the lamp position can be visualised in Fig. 3.46, exaggerated by a small height of the lens in relation to the ball diameter. The true ball centre B lies on the bisector of the angle 2ψ subtended by the ball at C, giving a radial coordinate r_b . If the ball centre is incorrectly taken as the centroid of the outline in the image then the false centre F with radial coordinate r_f is obtained by the projection of the outline midpoint $r'_f = (r'_{t1} + r'_{t2})/2$ back to the perspective centre C, at height ρ above the baseplane. Constructing the false normal \mathbf{n}_f from F through the highlight point P gives a lamp vector that is closer to the camera axis and hence a false lamp position L_f closer to the zenith than the true lamp position L.

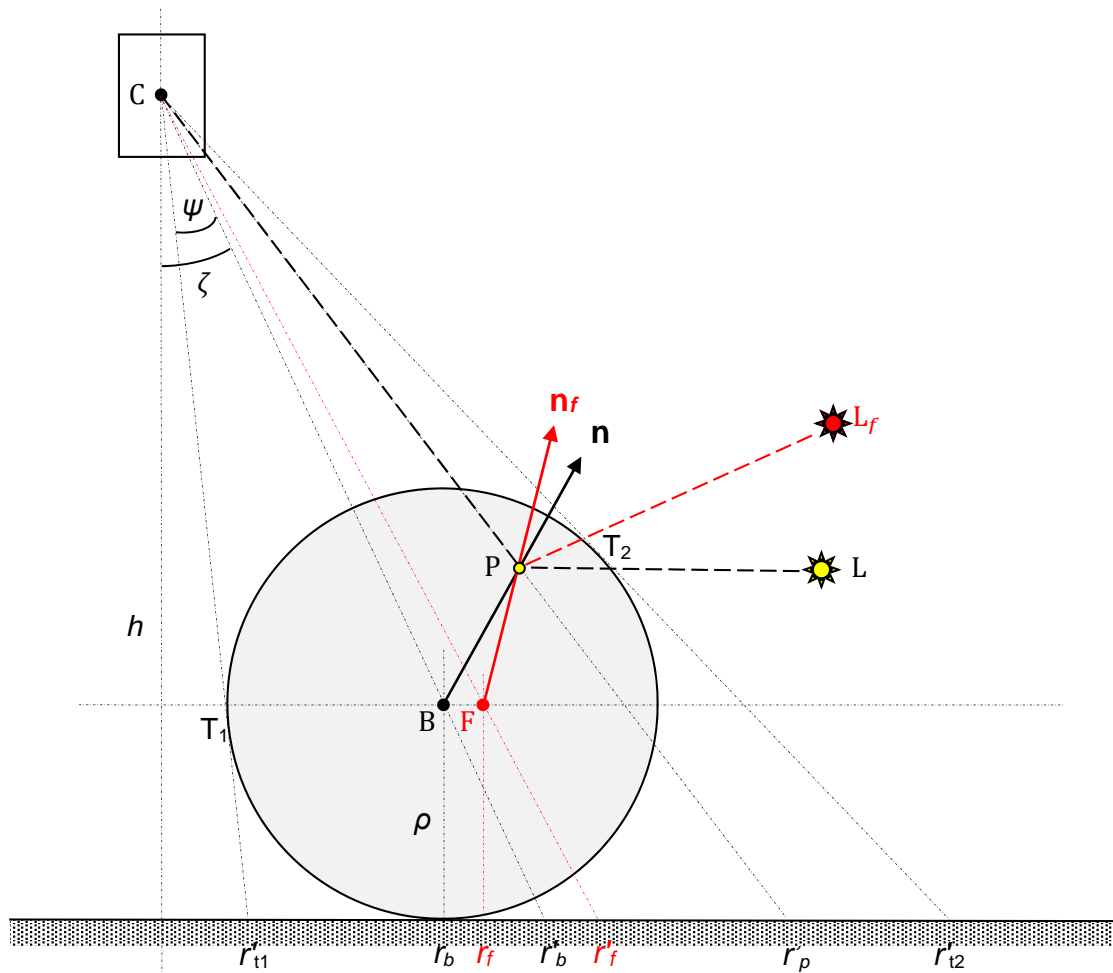


Figure 3.46: False lamp position (red) arising from assumption of ball centre at centroid of outline.

The angular error of the vector to the lamp is double the angular error of the normal. The error values calculated from a worked example are shown in Table 3.3, for data associated with the ball in the 'south-east' position (Fig. 3.33b). The positional error of the ball centre is 0.24 mm, which leads to an error of 0.47° in the normal vector, hence an error of 0.95° in the angle of the reflected ray, which produces an error of 1.34 mm in the lamp position on the hemispherical dome. The errors for all 64 lamps fluctuate from 0.1 to 1.8 mm (Fig. 3.47a).

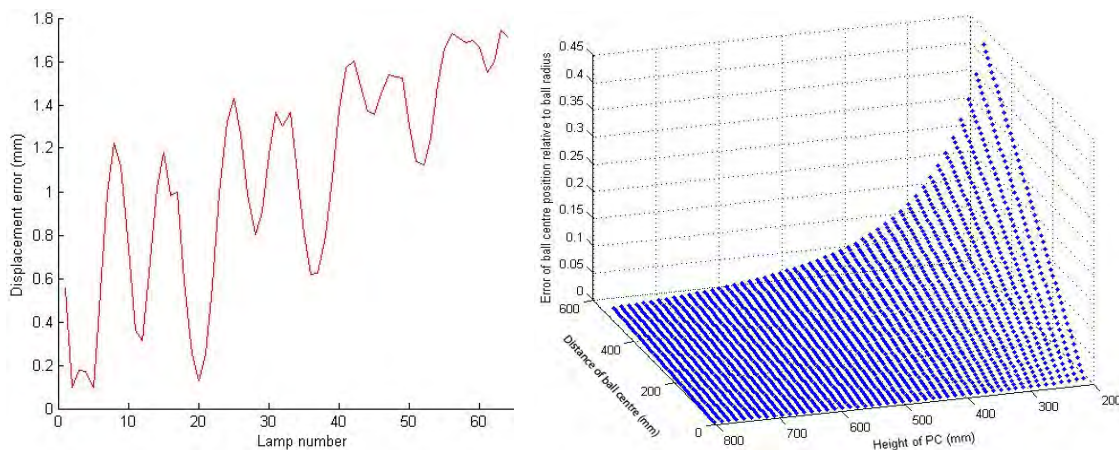


Figure 3.47: Errors arising from incorrect estimation of ball centre: (left) displacement of lamp position for all 64 lamps; (right) error in ball centre position relative to radius vs height and lateral position for $\theta = 135^\circ$.

The error in the position of the ball centre, relative to the ball radius, is plotted in Fig. 3.47b as a function of both the height of the lens perspective centre and the radial distance of the ball from the optical axis. For the values of the calculation in Table 3.3 (635 and 109 mm respectively), the error is low ($0.24/28.59 = 0.0084$), but in general it increases linearly with radial distance and exponentially with decreasing distance from the lens to the ball.

Table 3.3: Values of variables for configuration of the ball in Fig. 3.23b.

Quantity	Symbol	Value
Nominal focal length of lens	f	55 mm
Height of lens perspective centre	h	635.80 mm
Radius of ball	ρ	28.59 mm
True coordinates of ball centre	x_b, y_b, r_b	93.57, -53.72, 109.26 mm
False coordinates of ball centre	x_f, y_f, r_f	93.78, -53.84, 109.50 mm
Angle subtended by ball	2ψ	5.31 deg
Min/max radial coordinates of outline on baseboard	r'_{t1}, r'_{t2}	83.21, 145.11 mm
False radial centroid coordinate on baseboard	r'_f	113.66 mm
True radial centre coordinate on baseboard	r'_b	113.41 mm
False radial ball centre coordinate	r_f	109.50 mm
True radial ball centre coordinate	r_b	109.26 mm
Error between false and true ball centres	$r_f - r_b$	0.24 mm
True angle of ball centre	ζ_b	10.20 deg
False angle of ball centre	ζ_f	10.22 deg
Error between angles subtended at perspective centre by false and true ball centres	$\zeta_f - \zeta_b$	0.022 deg = 1.3 min arc
Highlight point of lamp #45 on ball surface	x_p, y_p, r_p	101.47, -56.11, 56.30 mm
Angle between true and false normals	$\text{acos}(\mathbf{n} \cdot \mathbf{n}_f)$	0.474 deg
Angle between true and false reflected rays		0.945 deg
Distance between true and false lamp coordinates		1.343 mm

3.9.3 Summary

The minimum values of mean error over all 64 lamps were similar for the glass sheet (mean 6.28, stdev 2.17) and the billiard ball (mean 6.86, stdev 2.83), indicating that both techniques can, under optimum conditions, estimate the lamp coordinates to an accuracy of approximately half the length of the illuminated area of the lamp (viz. 7 mm). The glass sheet method is more robust because the pattern of reflection highlights is much more spread out, and errors in estimating the position of the highlights are amplified less by the planar glass than by the convex ball surface. The optimum for the glass sheet was found by systematically adjusting the height of the perspective centre. To get to the minimum error value for the ball it was necessary to scan a three-dimensional parameter space, adjusting the height of the perspective centre of the lens and the X and Y offsets of the ball centre. Without this optimisation, the error (mean 23.14, stdev 7.96) was considerably worse. On the other hand, the glass sheet does not show the lamps in the lower tiers of the dome, as the angles are outside the field of view of the lens, whereas the billiard ball shows them all.

Likely sources of error in the estimation of lamp coordinates, in decreasing order of importance, include:

- (a) The flash lights are not point sources but glass cylinders of approximate length 18 mm by diameter 3.5 mm, within which the illuminated area is a cylinder of length 14 mm by diameter 2.8 mm. At the centre of the baseboard, at a mean distance of 520 mm from the lights, each lamp subtends a solid angle of approximately 1.5° by 0.3° . It is not entirely clear, however, that the centroid of the region of illumination always coincides with the geometric centre of the cylinder.
- (b) The fitting of the matrix to transform from pixel coordinates to physical baseboard coordinates is only an approximation and does not preserve the collinearity condition that all rays must pass through the front perspective centre of the lens.
- (c) The detection of centroids of highlights in the image, for both targets and lamps, depends on the distribution of light across the region. Despite thresholding and weighting, the position of the centroid may be distorted by specularity, noise, and saturation of the sensor.
- (d) The coordinates of the retro-reflective targets on the baseboard have an uncertainty arising from the estimation procedure of VMS.
- (e) The dome is not a perfect hemisphere but somewhat flattened and eccentric. This leads to positional errors when finding the intersection points between an assumed spherical surface and rays projected from the reflection points.
- (f) The glass sheet may not be perfectly planar and the billiard ball may not be perfectly spherical (these are mathematical assumptions in the reflection model).
- (g) The dome is made of acrylic, which has a high linear coefficient of thermal expansion (CTE) of $75 \mu\text{m}/\text{m}^\circ\text{K}$ ¹. A change of room temperature of 10° , therefore, could cause a change in a 1-metre length of acrylic of 0.75 mm. Because the dome structure is constrained at the bottom by its attachment to the frame, most of the expansion would be directed upwards along the Z axis.

¹ http://www.engineeringtoolbox.com/linear-expansion-coefficients-d_95.html

3.9.4 Finding Lamp Coordinates in RTI

The use of billiard balls to determine the direction of incident illumination is widespread in Reflectance Transform Imaging (RTI), implemented in the technique known as Highlight-RTI, as described in Section 2.2.4. The algorithm implemented in the H-RTI software assumes that all rays are parallel, and that the ball may appear anywhere within the field of view of the camera, and be processed by the same technique at any position in the image.

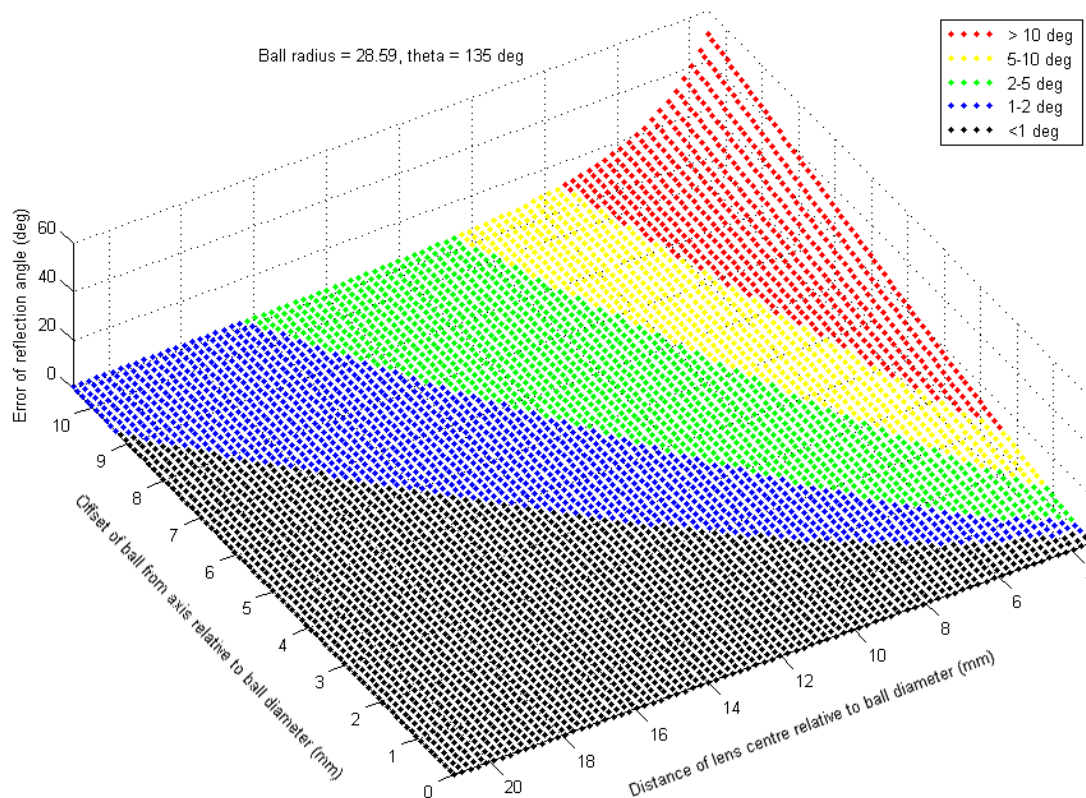


Figure 3.49: Error of angle to light source vs offset from optical axis and distance from camera, both relative to the ball diameter. The colour coding shows error zones.

The problem with this approach is that, except when the ball lies on the optical axis of the camera, the centre of the circular outline of the ball in the image is not the true centre of the ball because of parallax. The accurate method for calculating the mean was given above in Sec. 3.9.2. The angular error arising from the incorrect position of the ball centre within the image is more conveniently visualised in terms of the change in angle to the light source. Figure 3.49 plots these angular errors on normalised axes of distance relative to the ball diameter. Errors increase with lateral offset: even at a camera distance of 20 ball diameters, the angular error exceeds 1° when the ball is more than 8 ball diameters from the centre of the field. For a ball diameter of 50 mm (nominal 2 inches), this corresponds to a camera distance of 1 metre and a lateral ball distance of 400 mm, subtending an angle of 22° with the axis.

3.10 Conclusion

The objective of this study was to evaluate the precision of two image-based methods for determining the coordinates of flash lamps on the dome. The only *a priori* data were the coordinates of the retro-reflective targets on the baseboard and the coordinates of the perspective centre of the camera lens when mounted at the 'north pole' of the dome. The procedure was first to determine the transformation from points in the image plane to points on the baseboard by fitting a 4×3 matrix, then to use 3D vector geometry to find the points of intersection between the rays and the surfaces of objects placed on the baseboard. In the first method a sheet of glass gave a view of up to 40 lamps in linear projection on its planar surface. In the second method a billiard ball gave a view of all 64 lamps in a convex projection on its spherical surface. Both methods gave similar results, namely an uncertainty of approximately 7 mm, corresponding to half the length of the illuminated area of the flash lamp. It is clear from the results that these methods can never be more than qualitatively accurate, because they use a fitted approximation to the image formation geometry of the lens, and make various assumptions about the dome and the test objects.

In contrast the photogrammetric techniques, using the VMS software, were able to estimate the lamp centroids to an accuracy of approximately 1.5 mm, i.e. 0.3% of the dome radius. By employing a more precise model of lens distortion, the accuracy of the lamp coordinates could optimally be estimated with a mean error of approximately 1 mm. This result is considered satisfactory for the purposes of using the dome for reconstruction of surfaces by the photometric stereo technique, in which a system of three simultaneous equations is solved at each pixel to find the direction of the surface normal. Because the solution is implemented by regression over many angles, any errors in the positions of individual lamps are absorbed into the overall uncertainty. This is even truer for the PTM and RTI techniques described in Chapter 4, which fit an intensity distribution over the whole set of 64 lamps and are therefore quite insensitive to the direction vectors of individual lamps.

A significant result of this study concerns the accuracy of predicting lamp positions from the surface of a billiard ball, which is critically dependent on accurate knowledge of the coordinates of the ball centre. The conventional method used in highlight RTI relies on finding the centroid of the outline in the two-dimensional image, but this gives an incorrect result. Two methods are proposed as alternatives, using the 3D vectors from the lens perspective centre to the projections of the outline points from the image onto the baseplane: (a) normalising the vector lengths and taking their mean; and (b) fitting a sphere to the vectors as tangents.

The next step in the investigation was to use the dome for systematic capture of sets of 64 images of objects with surface relief. Knowledge of the precise coordinates of each lamp on the surface of the dome enabled the vectors for the incident rays of light to be quantified. This in turn enabled the intensity at each pixel to be modelled as a continuous function of angle, as described in Chapter 4.

Chapter 4 – Reflectance Transform Imaging

The simple rule for all illusions is this:
we always believe that we see such objects as would,
under conditions of normal vision,
produce the retinal image of which we are actually conscious.

– Hermann von Helmholtz (1867)
Handbook of Physiological Optics

A set of images contains more information about the scene than one image alone. If the image set is structured by systematic changes to exposure or wavelength or camera position or illumination direction then characteristics of objects in the scene may be inferred. A major focus of this research programme was how sets of images captured in the dome could be used for qualitative visualisation of cultural heritage objects, and how quantitative information about the objects could be derived through image processing. This chapter deals with the qualitative, creating from a set of images the illusion of dimensionality without any underlying 3D model. The following chapters deal with the more quantitative aspects of dimension, directionality and colour.



Figure 4.1: Test object used in this study, the 'Roman medallion'.

For this chapter a 'Roman medallion' is used as the test object (Fig. 4.1). It is modern, made in Rome, cast from brass and painted in glossy brown and powdery green to look like antique bronze. It is 10 cm in diameter, c.15 mm thick, and weighs 674 grams. The inscription DNHONORI VSPFAVG is an abbreviation of D(ominus) N(oster) HONORI-VS P(ius) F(elix) AVG(ustus). The design is probably based on a gold solidus coin, minted in Ravenna. It depicts the Emperor Honorius, who ruled from 395 to 423 AD in the declining years of the Roman Empire, and who is generally considered by historians as the least competent of all the Roman emperors. The most notable event during his reign was the assault and sack of Rome in 410 by the Visigoths.

The medallion is a good object for research because it is an ideal size and shape for photography in the dome. It has well-defined relief, a range of hues from green to reddish-brown, and a surface finish that includes both matte and specular areas. Moreover it is robust and can be handled freely without gloves. The image set illustrating this chapter was photographed with the Nikkor 17-55 mm zoom lens set to 55 mm focal length at aperture $f/8$. As the diameter of the medallion in the images is 1475 pixels, the spatial resolution on the surface is 14.75 pixels/mm (i.e. 1 pixel = 68 μm). Although the results described in this chapter are illustrated by this object, they are transferrable to a wide class of objects of any colour, texture and gloss. They are most applicable to flattish objects with relief, having opaque isotropic surfaces.

4.1 Image Acquisition

4.1.1 Image Processing Workflow

The procedure for capturing a set of images in the dome is given in Appendix I. The raw image (NEF) files from the Nikon camera are processed directly using the DCRAW software utility (with options -T -4), which generates linear TIFF files with 16 bits per RGB channel. Previous analysis has verified the linearity of the image data produced by DCRAW, i.e. the linear relationship between signal value in each channel and the measured reflectance factor of a greyscale. All subsequent image processing is done in Matlab, converting the 16-bit unsigned integer values into double-precision floats.

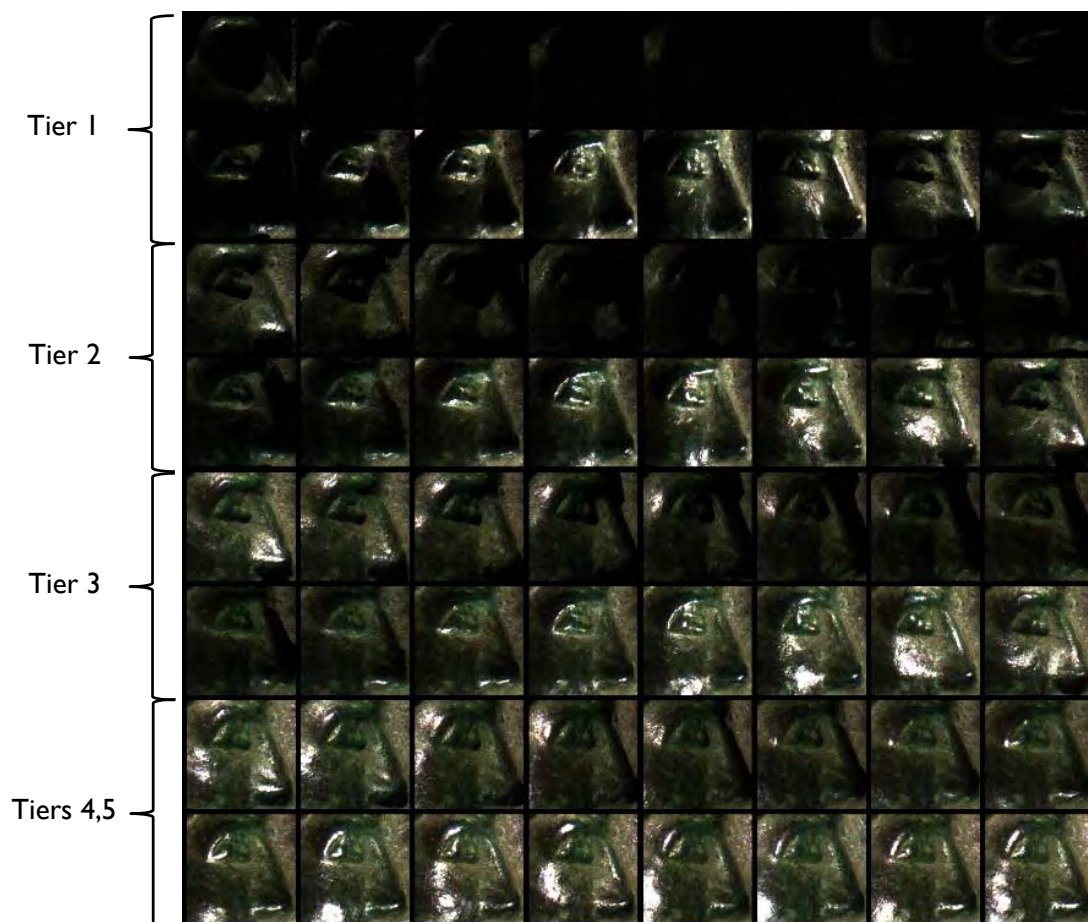


Figure 4.2: An 8x8 mosaic of a detail of the images from 64 lamps in dome.

The memory requirements increase substantially for this *modus operandi*: one full colour image of 3900x2616 pixels occupies 14.9 Mbytes in NEF format, 58.3 Mbytes in 16-bit TIFF format, and 233.3 Mbytes in double-precision floating-point format. The full set of 64 images in floating-point would require 14.9 Gbytes of RAM, which is beyond the capacity of most PCs. In the Matlab routines, therefore, the images are generally stored in the more compact 16-bit integer format, and converted to floating-point only for processing. The loss of precision is the rounding error of 1 part in 2^{16} or 0.0015% of the full range. In many cases the image area is reduced by cropping to a region of interest. When the full image is to be covered, the processing is done in tiles of approximately 1500x1500 pixels and the 64 source images are read again from disk to extract each tile.

The result of the dome acquisition procedure is a set of 64 images, all in pixel register but each illuminated from a different angle. The average lightness of the images increases with the increasing angle of incident illumination from higher tiers in the dome (Fig. 4.2).

4.1.2 Pixel Intensity Vectors

Each pixel in the 64-image set captured in the dome can be treated as a vector of 64 intensity values per channel. (The word ‘intensity’ is used here to mean the amplitude or strength of the digital signal.) The distribution depends on the gradient at the corresponding position on the object surface and is notably different for matte and glossy surfaces. Fig. 4.3 shows plots of intensity in the green channel vs lamp number for two different pixels. The position on the left is in the hollow to the side of the nose, where the surface consists primarily of a matte green powder, simulating the copper sulphate that would result from natural corrosion of the bronze. This has a granular powdery texture, giving a diffuse reflection that rises with increasing lamp number. The position on the right is on the tip of the nose which is shiny brown, simulating the copper oxide coating on bronze. This is smooth and glossy, giving a strongly directional reflection, with intensity peaks for just four lamps and much lower intensity for all the others. In both cases the periodicity of the pattern is clear, following the 16 lamps around the perimeter of the hemisphere (azimuthal angles).

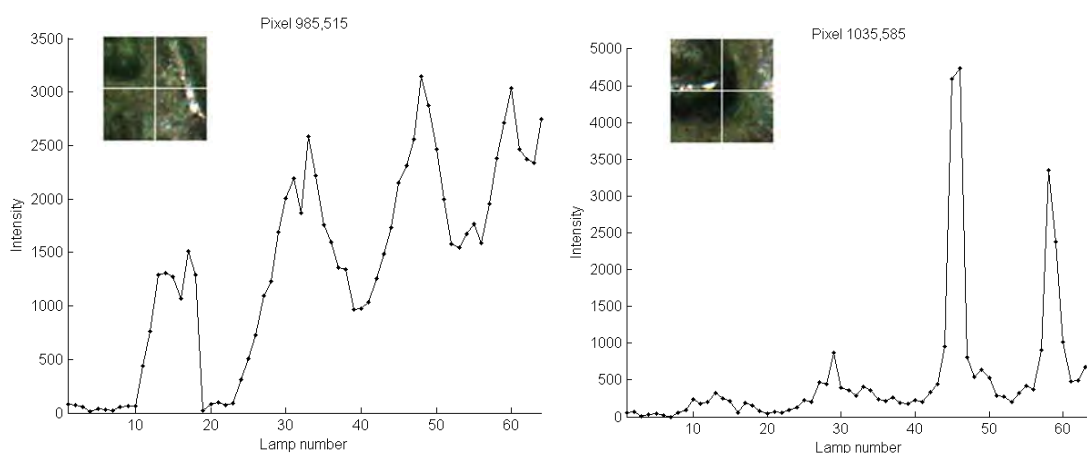


Figure 4.3: Intensity distributions of the green channel of 64 images at a single pixel: (left) matte surface; (right) glossy surface. Insets show local area of image.

4.1.3 Vector Projection

A more meaningful way to visualise the intensity distributions of Fig. 4.3 is in relation to the directions of the 64 lamps over the hemisphere. For the matte surface (Fig. 4.4 left) the vectors form a compact rounded shape, whereas for the glossy surface (Fig. 4.4 right) there is a long projection in the specular direction.

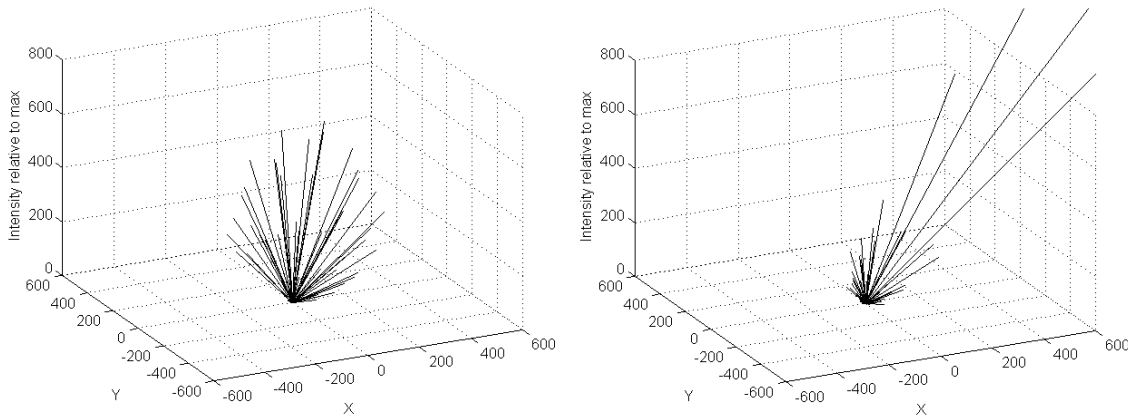


Figure 4.4: Intensity distributions of the green channel of 64 images at a single pixel, plotted along the lamp vectors: (left) matte surface; (right) glossy surface. (X and Y axes in mm).

4.1.4 Hemispherical Projection

Another way of visualising the intensity distribution is as a shell over the hemisphere. Fig. 4.5 shows the result of creating a triangular mesh over the convex hull of the vectors of Fig. 4.4, scaled as projections beyond the unit hemisphere. This projection exaggerates the volume of intensity in the specular direction.

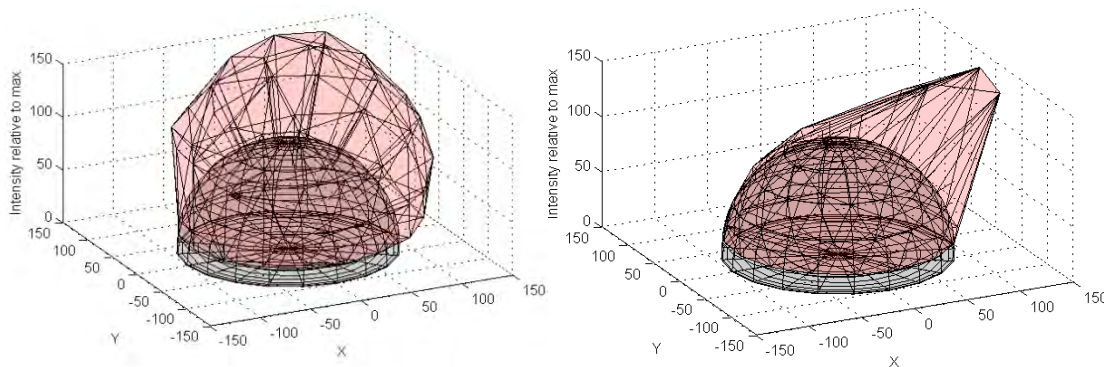


Figure 4.5: Intensity distributions of the green channel of 64 images at a single pixel, plotted as a mesh over the hemisphere: (left) matte surface; (right) glossy surface. (X and Y axes normalised).

4.1.5 Polar Projection

In cartography there are many projections of the surface of the sphere onto a plane. For this study, the azimuthal equidistant projection has been chosen, which places the north pole at the centre of a circle, with the equator at the circumference (Fig. 4.6). It has the convenient properties that: (a) all points on the map are at proportionately correct distances from the centre; and (b) all points on the map are at the correct azimuth (direction) from the centre point. In this projection all meridians (lines of longitude) are straight, with distances from the pole represented correctly as $r = \varphi$ where φ is the co-latitude ($= 90^\circ - \text{latitude}$). The disadvantage of this projection is that it is not isotropic (equal scale in all positions of the plane)

because distances in the azimuthal direction increase with r . Applying the azimuthal equidistant projection to the dome enables the intensity distribution to be plotted in 3D with the intensity on the Z axis over the polar plane. Fig. 4.7 shows the same distributions as in Figs. 4.3, 4.4 and 4.5, normalised to a maximum value of 100. The matte surface (left) has a conical form, whereas the glossy surface (right) has low values everywhere except in a small angular region corresponding to the specular direction.

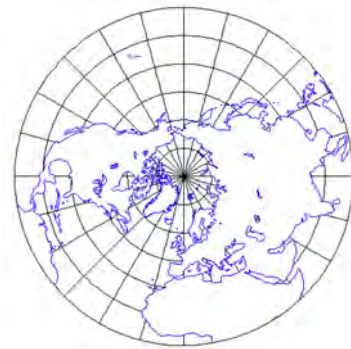


Figure 4.6: Azimuthal equidistant projection of the northern hemisphere of the earth, with north pole at centre.

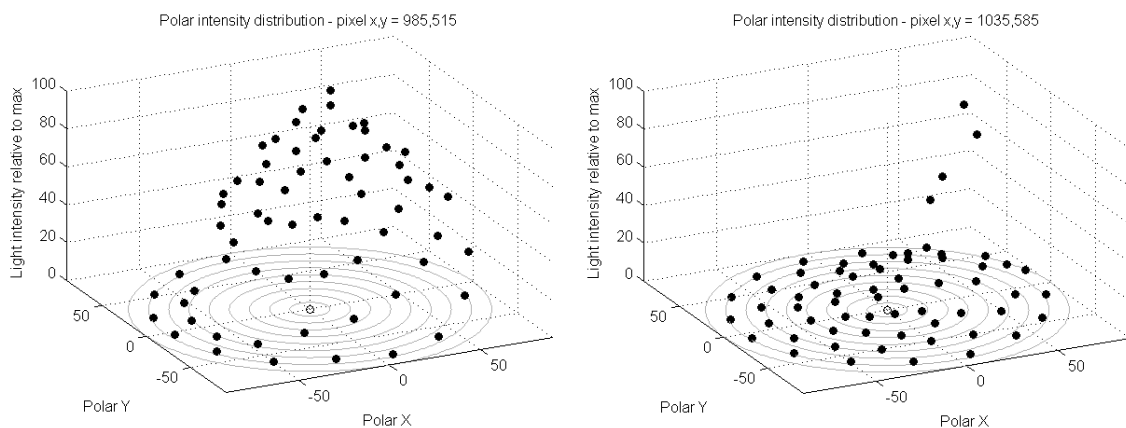


Figure 4.7: Normalised intensity distributions on the polar plane: (left) matte surface; (right) glossy surface.

4.1.6 Making an Outline Mask

For photography the object is placed in the dome on a black cardboard sheet to provide a dark image background free from texture. All subsequent image processing operations are limited to the object area within its outline. The outline is used as a pixel mask to limit the scope of operations. The algorithm for making the outline, as illustrated in Fig. 4.8, is:

- i. Compute the median image, i.e. each pixel is the median of the 64 values in intensity vector. For many objects it is sufficient (and quicker) to compute the mean of the 12 images from Tier 4 illumination.
- ii. Extract green channel and apply 5x5 spatial median filter to monochrome image.
- iii. Compute mean intensities in outside region (corner) and inside region (centre).
- iv. Calculate intensity threshold as: $t = 0.2 * (\text{mean}_{\text{outside}} + \text{mean}_{\text{inside}} - 2 * \text{stdev}_{\text{inside}})$.
- v. Plot one or more cross-sectional profiles. Check that the threshold level is above the background intensity and below the object intensity.
- vi. Make a binary mask at the same size as the original image, with 0 in the background (intensity < threshold) and 1 in the foreground (intensity > threshold).
- vii. Convert binary mask to floating point.

- viii. Apply a 21x21 median filter to mask image to remove noise (i.e. spots in the object below threshold level).
- ix. Apply a 3x3 smoothing filter to anti-alias the edge profile around outline mask.
- x. Convert floating point mask to 8-bit grey-scale image in range 0-255 and save as file.
- xi. Inspect the mask image and, if necessary, edit in Photoshop to tidy up.
- xii. Determine coordinate points around the image outline by scanning all rows from both left and right and all columns from both top and bottom. Filter to remove duplicate points and sort into order of angle around circumference.
- xiii. Apply 7-point linear smoothing filter to coordinate points in outline.
- xiv. Compute centroid, mean and maximum radius, length of perimeter and area.

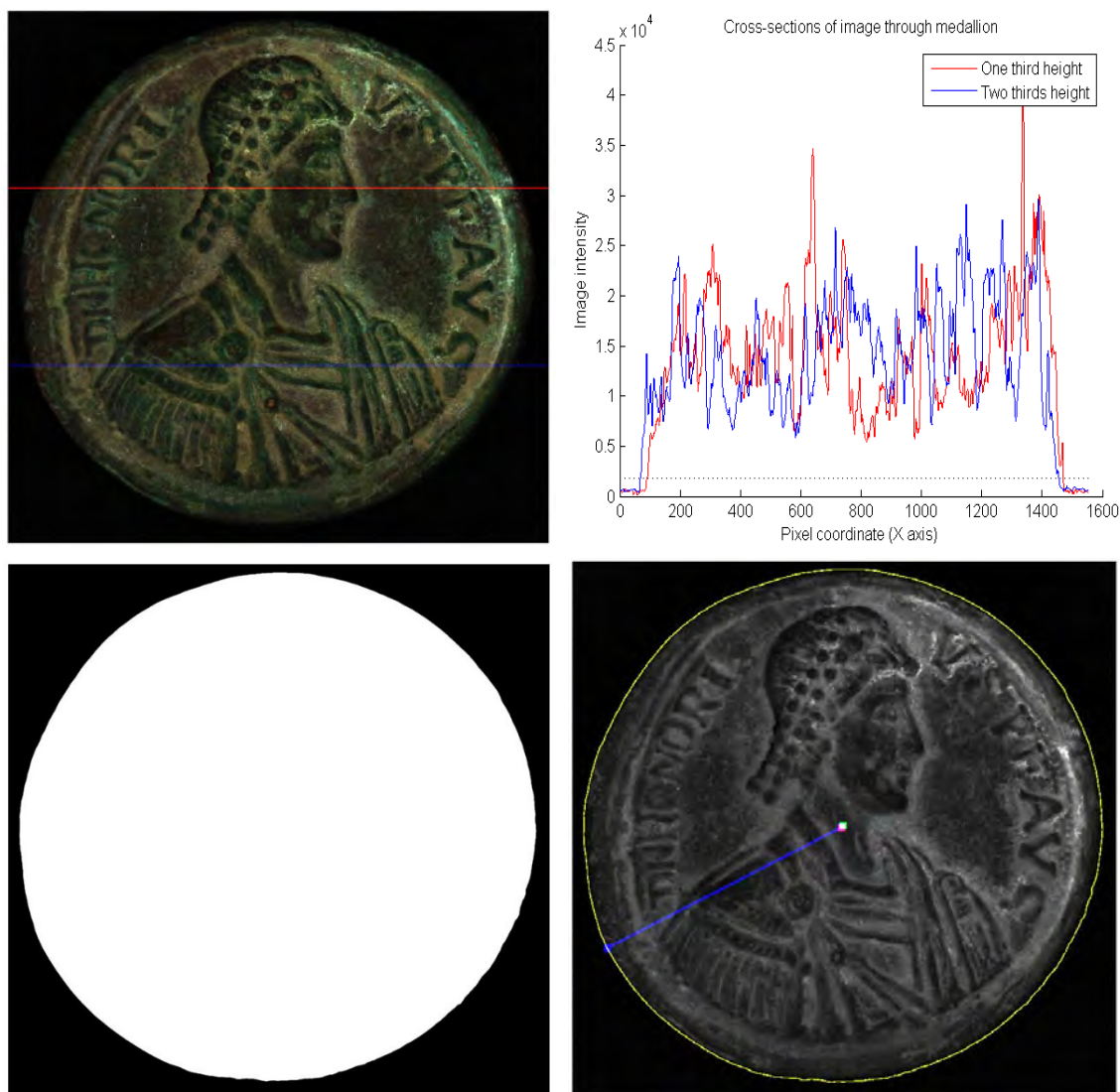


Figure 4.8: Two cross-sections through image: (upper left) median image with cross-sections; (upper right) profiles of intensity of green channel through cross-sections with threshold; (lower left) binary mask; (lower right) monochrome image showing outline, centroid and vector of maximum radius.

4.2 Polynomial Texture Mapping

The dependence of luminance on light direction in PTM is modelled in Eq. 2-7 by the biquadratic function of two variables l_u and l_v . A set of coefficients ($a_0 - a_5$) is fitted to the image data for each pixel and stored in the PTM file at the same spatial resolution as each of the original images. The six components can thus be represented as monochrome images, scaled so that mid-grey represents zero coefficient value, with darker values negative and lighter values positive (Fig. 4.9). The directionality of the components is clear, with the coefficients a_0 and a_3 for l_u^2 and l_u respectively showing strong contrast in the horizontal direction, emphasising vertical edges, and the coefficients a_1 and a_4 for l_v^2 and l_v respectively showing strong contrast in the vertical direction, emphasising horizontal edges.



Figure 4.9: Six coefficients of PTM luminance, shown as image planes.
Mid-grey represents zero, with positive values lighter and negative values darker.

Within the PTM file structure the data is organised as a short header followed by all the image components. The format most commonly used is the 'LRGB' format, in which each pixel is represented by six bytes for the six coefficients of the luminance biquadratic, plus three bytes for the RGB 'base image'. The RGB image with the PTM structure is not a standard RGB image, but the product of the albedo and the maximum luminance (Fig. 4.10), so that the effect of the luminance modulation by the biquadratic is to darken the image. The luminance coefficients are scaled to maximise the 8-bit integer range, as described in the original paper by Malzbender (2001). The data storage order for the image is [RGB, rows, columns] and the row order is inverted, with the origin at lower left. For display of the PTM components as images in Matlab, they have to be permuted to [columns, rows, RGB] and the row order flipped to place the origin at upper left.

4.3 Extensions of PTM

During this research programme several utilities have been developed in Matlab for manipulating PTM image files. All of these operations can be performed directly on the PTM files, without reference to the original image sets. Although not significant in terms of research, these functions have proved to be useful in many applications.



Figure 4.10: RGB component of PTM file.

4.3.1 Separating and Combining PTM Files

The PTM file structure for the LRGB format consists of a header followed by the six luminance components, followed by the RGB base image. Also the luminance coefficients are scaled by bias and offset values to fit the 8-bit byte. PTM files are therefore non-standard and cannot be processed directly by image utilities (such as Adobe *Photoshop*) in the same way as universal image file formats such as BMP, TIFF or JPG.

Rotating a PTM

The PTM file is decomposed into its constituent parts and the images corresponding to the luminance components and the RGB channels are rotated by a multiple of 90° . A new file is written in PTM format. There is no need to rescale the image (i.e. to recalculate the scale and bias factors) because none of the pixel data has been modified.

Extracting a section out of a PTM

The PTM file is decomposed into its constituent parts and the luminance scaling factors applied to each coefficient plane in floating point. The specified rectangle is cropped out of the luminance and RGB components, then scaling factors are recalculated. The cropped image is written in PTM file format. This function enables multiple objects to be photographed together and a PTM made of the ensemble, then the individual objects extracted as separate PTMs.

Combining two PTMs

The individual PTM files are decomposed into their constituent parts and the respective luminance scaling factors applied to each coefficient plane in floating point. The specified rectangular areas from each image, both RGB and luminance components, are composited together. The scaling factors are recalculated for the combined image and a new file written in PTM format. This function enables PTMs of two objects, or two views of the same object, to be combined and viewed together in the PTM viewer under a single interactive light source. This can be done even the original image sets had been captured by different cameras at different times in different domes under different lighting. It is particularly useful for making a mosaic of regions of a large object such as a marble slab or a wall.

Viewing PTM from lamp coordinates

The PTM file is decomposed into its constituent parts and the luminance scaling factors applied to each coefficient plane in floating point. An image is rendered for a virtual light source at the same coordinates as one of the lamps in the dome. This is useful for direct comparison between the rendering by the PTM model and the actual image captured by the camera under illumination from the specified lamp position. See discussion of viewer software in Chapter 8.

Making a movie

The PTM file is decomposed into its constituent parts and the luminance scaling factors applied to each coefficient plane in floating point. A series of images is rendered as above for the virtual light source traversing across the hemisphere in a scripted trajectory. The images can then be combined into a movie sequence by a utility program such as *MovieMaker* or *VirtualDub*. This enables the effect of the light source moving across the object surface to be simulated (Fig. 4.11). The continuous biquadratic function makes this possible, because the reconstruction can be done at any angle and is not limited to the specific coordinates of the dome lamps. For best results the increment of illumination angle between image frames should be less than 0.5° , to give the impression of smooth movement. At this rate, rendering the movie sequence at 24 fps would show a movement of the virtual light source by $12^\circ/\text{s}$, taking 15 seconds for a full transit from one horizon over the hemisphere to the other.



Figure 4.11: Four frames from a rendered image sequence for a movie, with the virtual light source at angles of 0° , 30° , 60° and 90° from the normal, moving towards the east.

4.3.2 Extracting Normals from PTM

Assume that the biquadratic in Eq. 2-7 attains its maximum value L_0 at $l_u = l_{u0}$ and $l_v = l_{v0}$, where (l_u, l_v) are projections of the normalised light vector into the local texture coordinate system (u, v) . The maximum can be found in terms of the coefficients a_0 to a_5 by setting the partial derivatives in u and v to zero:

$$\frac{\partial L}{\partial u} = 2a_0l_{u0} + a_2l_{v0} + a_3 = 0 \quad \text{and} \quad \frac{\partial L}{\partial v} = 2a_1l_{v0} + a_2l_{u0} + a_4 = 0 \quad (4-1)$$

Solving both equations in (4.3) for l_{u0} gives:

$$l_{u0} = \frac{-a_3 - a_2l_{v0}}{2a_0} \quad \text{and} \quad l_{u0} = \frac{-a_4 - 2a_1l_{v0}}{a_2} \quad (4-2)$$

This leads to an equation in l_{v0} :

$$2a_0(-a_4 - 2a_1l_{v0}) = a_2(-a_3 - a_2l_{v0})$$

And so by collecting terms:

$$l_{v0} = \frac{a_2 a_3 - 2a_0 a_4}{4a_0 a_1 - a_2^2} \quad (4-3)$$

Similarly:

$$l_{u0} = \frac{a_2 a_4 - 2a_1 a_3}{4a_0 a_1 - a_2^2} \quad (4-4)$$

In the original Malzbender (2001) paper, Eq. 4-3 and 4-4 were correctly presented as the equations for the maximum of the biquadratic. But then he claimed the surface normal \mathbf{N} as:

$$\mathbf{N} = \left(l_{u0}, l_{v0}, \sqrt{1 - l_{u0}^2 - l_{v0}^2} \right) \quad (4-5)$$

Eq. 4-5 is incorrect, because (l_{u0}, l_{v0}) represents the coordinates of the virtual light source in the local coordinate system at which the reflected luminance is maximum, which is the specular direction not the normal. The normal should be at half the specular angle (see Sec. 4.4.5).

But there is also a more subtle problem. The image of the surface normals computed for the medallion using Eq. 4-5 is shown in Fig. 4.12 (left), encoded in false colour where the N_x component is assigned to the red channel, N_y to green, and N_z to blue. The corresponding image of the normals generated by the PTM Viewer utility developed by Malzbender at HP is shown in Fig. 4.12 (centre). Comparing both with the normals calculated from HSH (Fig. 4.20), it is clear that something is amiss for many regions of the image. Plotting a map in Fig. 4.12 (right) of all pixels where the denominator of the terms in Eq. 4-3 and 4-4 is negative shows an exact correspondence to the incorrect areas in the normals images.



Figure 4.12: False-colour images of surface normals of medallion: (left) computed from coefficients; (centre) produced by the HP PTM viewer; (right) map of points with negative denominator.

The explanation can be found by plotting the biquadratic luminance distribution of two pixels in normalised $[l_u, l_v]$ space, in Fig. 4.13. On the left is a ‘good’ pixel for which the denominator is positive, and the biquadratic function shows a parabolic form in both axes, with a clearly defined maximum value (l_{u0}, l_{v0}) denoted by the red cross. On the right is a ‘bad’ pixel for which the denominator is negative and the biquadratic has no maximum. In this case the luminance continues to rise toward the perimeter of the hemisphere and the peak is therefore undefined. It doesn’t matter for the PTM viewer because the intensity is single-valued and continuous over the hemisphere, but it means that normals cannot be extracted at pixels where such distributions occur. Such cases correspond to the specular vector being at angle $> 90^\circ$ from the Z axis, effectively below the equator of the hemisphere.

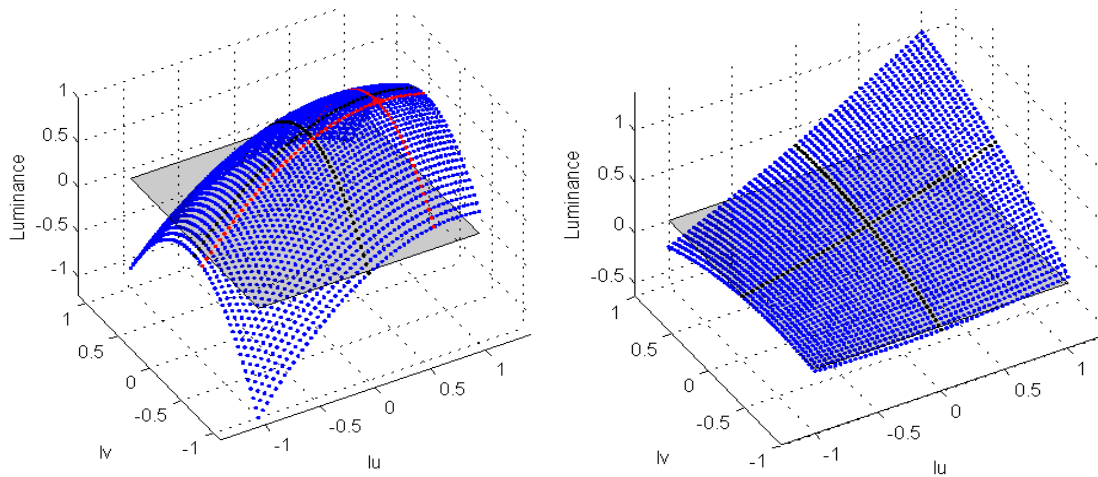


Figure 4.13: Two distributions of the PTM biquadratic luminance function: (left) well-defined maximum value denoted by red cross; (right) no maximum value.

4.3.3 Image Detail Enhancement

Once the surface normal vectors are known, the local image contrast can be enhanced by amplifying the normals (i.e. multiplying by a constant, typically in the range 1.2 to 1.5), which in effect increases the surface gradients. A set of 64 images is calculated from the modified normals by assuming a Lambertian surface with the same RGB colour as in the original PTM. Then a PTM is generated from the new image set. This method is effective in increasing the visibility of very shallow surface relief, and is helpful for palæography (Fig. 4.14).



Fig. 4.14: Detail of PTM of the Marmor Parium tablet in the Ashmolean Museum, Oxford, taken as screen shots from the RTI viewer with illumination from the north-west: (left) PTM as constructed by RTI fitter; (right) PTM after enhancement by scaling of normals by a factor of 1.5.

4.4 Hemispherical Harmonics

4.4.1 Basis Functions

The HSH basis functions in Table 2.2 can be visualised by projecting the hemisphere onto a plane by means of the azimuthal equidistant projection (Sec. 4.1.5). Fig. 4.15 shows the full set of 16 components in plan view. Apart from the first constant term, all functions are equally balanced between positive and negative values. Below the components h_3 , h_8 and h_{15} are also shown in 3D, scaled for visualisation to fit within a cube, indicating the increasing complexity of the first-order, second-order and third-order harmonics.

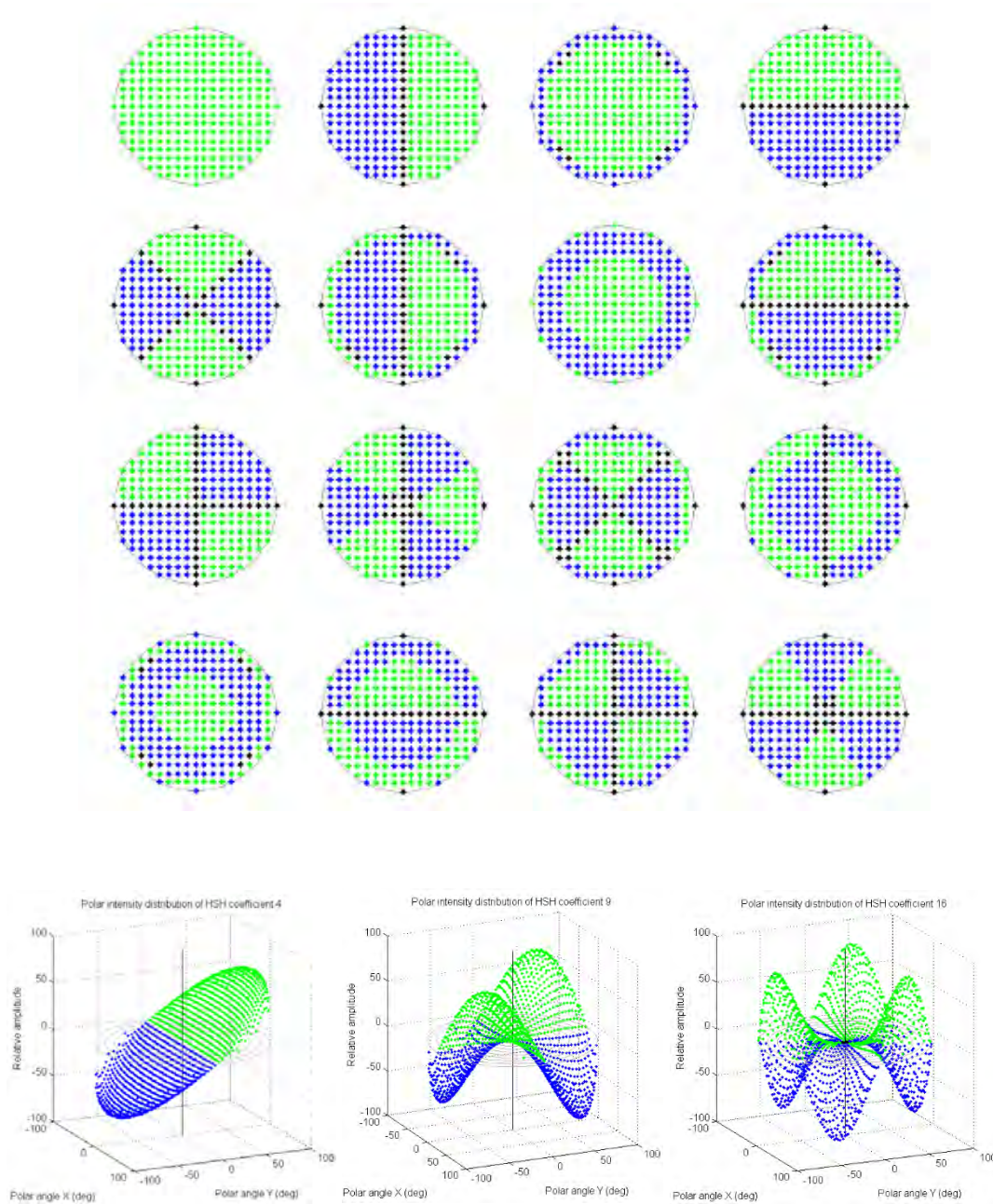


Figure 4.15: (top) The first sixteen modes of hemispherical harmonics, shown in azimuthal equidistant projection onto the plane. Positive values are indicated by green and negative by blue.
(bottom) 3D views of modes 3, 8 and 15.

4.4.2 Fitted HSH Components

Fitting of the coefficients to the image set follows the same procedure as for PTM, and can be applied to the luminance channel to provide the angular modulation at each pixel. These can be visualised as monochrome images (Fig. 4.16). Coefficients h_1 and h_3 in polar coordinates θ and φ are very similar to coefficients a_3 and a_4 in the PTM biquadratic (Fig. 4.9) in linear l_u and l_v coordinates, both representing linear image gradients in X and Y respectively.

The weighted sum of the HSH components approximates the angular distribution of intensity at each pixel of the 64 images captured by the camera in the dome. As the number of components increases the accuracy of the approximation improves. Fig. 4.17 compares the PTM and three orders of HSH in the azimuthal equidistant projection for one pixel, all normalised to the maximum amplitude of the lamp intensities. The PTM (six parameters) and first-order HSH (four parameters) give a similar performance, with a broad angular distribution. The second-order HSH (nine parameters) has a more narrow distribution but also rises opposite the peak, resulting in an alias, i.e. a false increase in brightness when the virtual light source is incident from angles opposite the peak. The third-order HSH (sixteen parameters) gives the narrowest peak, closest to the true intensity distribution.



Figure 4.16: Sixteen coefficients of HSH luminance, shown as image planes.

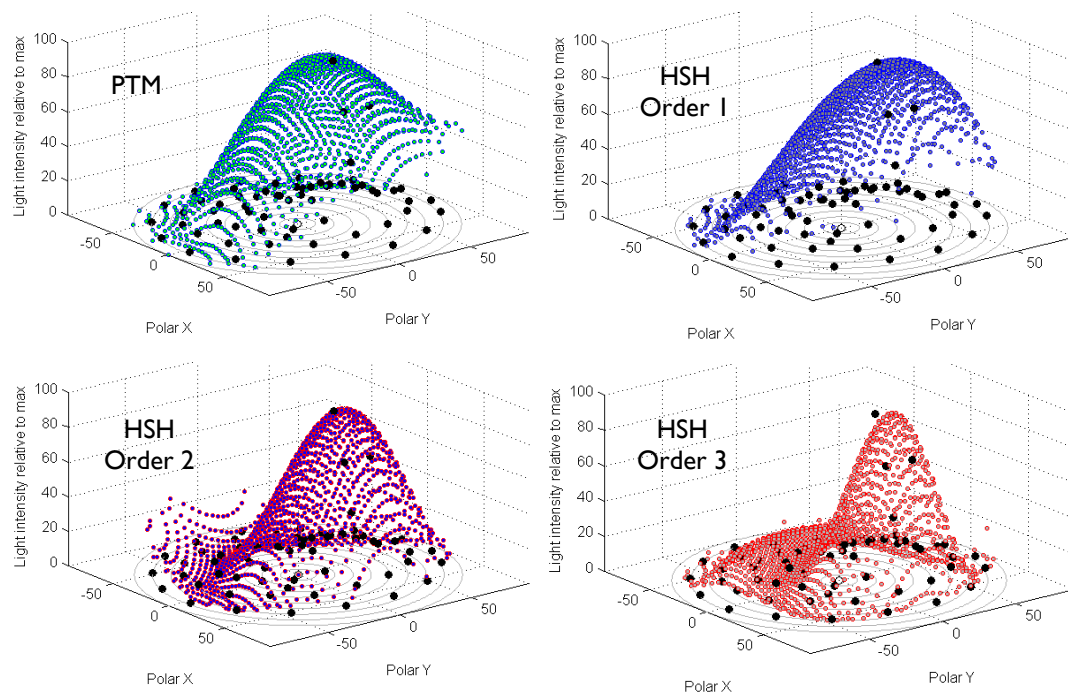


Figure 4.17: Angular distributions of intensity, plotted in the azimuthal equidistant projection, for four approximations to the measured intensity distribution. The black dots represent the actual reflected intensity of the 64 lamps. All intensities are normalised to a maximum of 100.

4.4.3 Quality of Fit

The improvement of the approximation with an increasing number of coefficients in HSH can be seen in Fig. 4.18, plotting intensity by lamp number (cf. Fig. 4.3) for a pixel on the cheekbone of Honorius on the medallion. The amplitudes of the modelled components are normalised for the sums over all 64 values (i.e. equal area under the graph \approx equal power over the hemisphere). Sorting the values into ascending order shows clearly the discrepancy between the model predictions and the actual peak intensities.

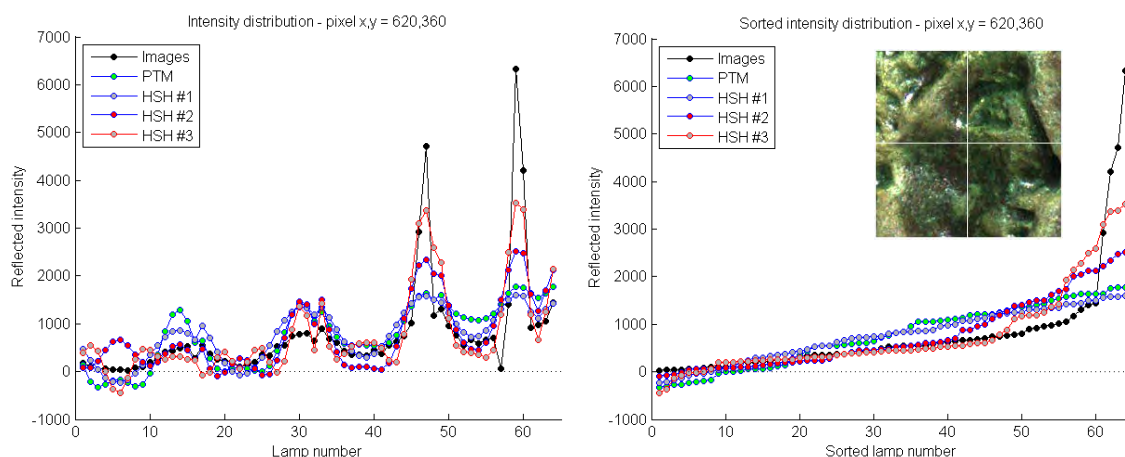


Figure 4.18: (left) Intensities at one pixel for 64 lamps of actual images and reconstructions by PTM and HSH models: (left) by lamp number; (right) sorted into ascending order.

The accuracy of the various reconstructions from the PTM and HSH models can be estimated taking the difference between the actual and modelled pixel intensity for each lamp, summed over all 64 lamps and divided by the maximum. Table 4-1 shows the results, first for the single

pixel illustrated in Fig. 4.18 and then for an average of 10,000 pixels selected at random from the image set within the outline of the medallion. Negative values produced by the models were clipped to zero, and the model distribution scaled so that the total power was equal to that of the 64 values in the image vector, as described above. In general the mean error for PTM lies between that of the first-order and second-order HSH reconstructions. The third-order HSH gives the best results, as expected.

Table 4-1: Mean % errors of PTM and HSH models

	PTM	HSH 1 st order	HSH 2 nd order	HSH 3 rd order
Single pixel	6.134	6.469	5.900	5.666
Mean over 10,000 pixels	6.094	6.265	5.776	4.818

4.4.4 Angular Resolution

By the geometry of the dome, the viewpoint (camera position) remains fixed but the incident rays from the 64 lamps sample the reflectance distribution function of the object surface over the hemisphere. The ability of the system to resolve angular variations in reflectance is therefore limited by the angular spacing of the lamps. The cosine of the angle between any two lamps is simply the projection of one lamp vector (from the lamp to the centre of the baseboard) onto the other. Plotting the histogram of all pair-wise angles between neighbouring lamps (Fig. 4.19 right) shows a peak at 20°. By the Nyquist criterion, a 20° sampling interval would resolve a maximum angular frequency of one cycle in 40°, or 4.5 cycles over the 180° period of the hemisphere. In particular this limits the fineness of the specular peak that can be resolved by the system. The full range of neighbouring angles is 12.02° to 28.45°, corresponding to a range of angular frequencies of 7.49 to 3.16 cycles per 180°.

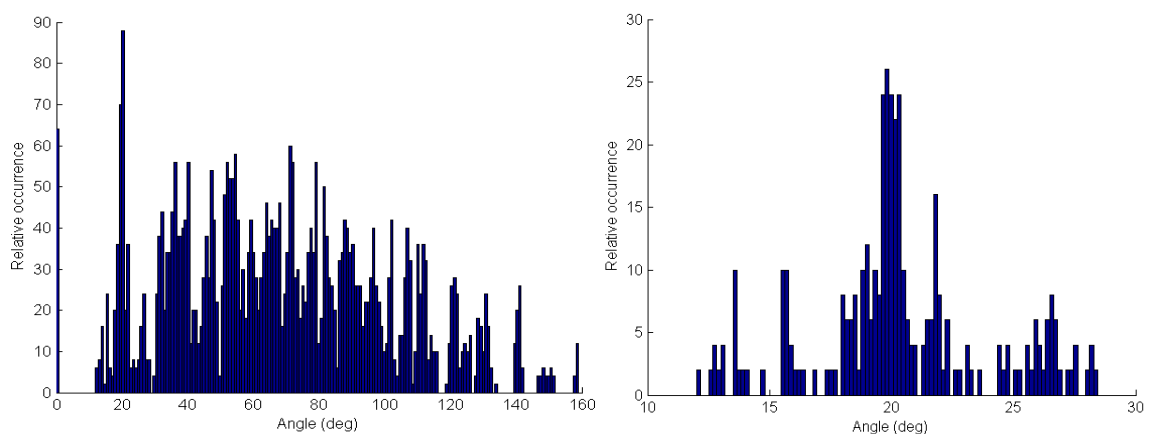


Figure 4.19: Histograms of angles between pairs of lamps: (left) all lamps; (right) neighbouring lamps.

The HSH basis functions are limited to one, two and three cycles per 180° for the first-order, second-order and third-order harmonics respectively, as is clear from examination of Eq. (4.6). Actually the maximum angular frequencies are 2 cycles per 180° latitude and 3 cycles per 360° azimuth, as is clear from Fig. 4.15, so the HSH representation with 16 basis functions is better able to resolve angular frequency in the latitude direction θ than in the azimuth direction φ .

4.4.5 Deriving Normals from HSH

Unlike the case for the simpler PTM biquadratic function (Sec. 4.3.2), there is no analytical solution for finding the maximum of the summation of hemispherical harmonics in Eq. (4.9). One way of solving this problem would be by numerical methods, by starting with an estimate then iteratively calculating local gradients in θ and ϕ and using the Jacobian matrix to update the estimate until a local maximum is found. An alternative method has been implemented in this study by direct calculation of the HSH value for every point in a grid in l_u, l_v space, transformed to polar angles θ, ϕ by Eq. 2-11. By choosing an increment of 0.01, the mesh over interval $[-1, +1]$ has 201 points on each axis, or 40401 points in all. The ability of Matlab to process the whole array of points in each operation makes this technique feasible. The maximum value of the distribution is found directly (Fig. 4.20), giving the coordinates (l_{u0}, l_{v0}) of the nearest point. The procedure can be iterated by taking the array patch $[l_{u0}-0.01: l_{u0}+0.01, l_{v0}-0.01: l_{v0}+0.01]$ and subdividing it into a further 201×201 points, calculating the HSH luminance value at each point, and finding the maximum. Each additional iteration adds two decimal places to the precision of the vector direction.

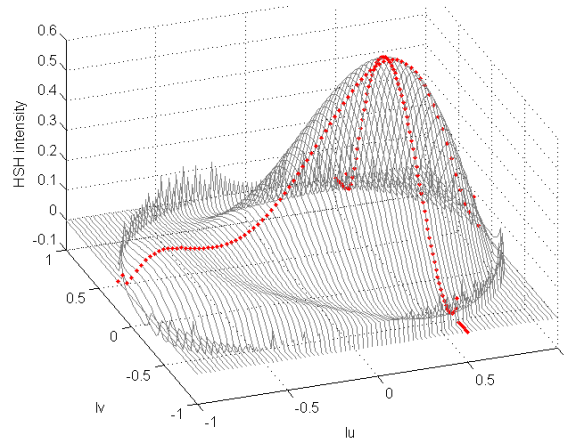


Figure 4.20: HSH intensity profile for one pixel. The red cross indicates the coordinates of the maximum.

The whole procedure has to be repeated for every pixel in the image. The coordinates found at each pixel correspond to the angle at which the reflected intensity, as modelled by the HSH

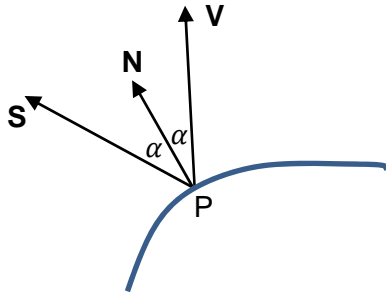


Figure 4.21: Geometry of specular reflection from a surface.

functions, is maximum. Consider the point P on the surface represented in Fig. 4.21, where \mathbf{V} is the view vector toward the camera at the 'north pole' of the dome. If \mathbf{N} is the normal vector at P, subtending an angle α with \mathbf{V} , then the specular vector \mathbf{S} lies in the same plane as \mathbf{N} and \mathbf{V} at an angle of 2α from \mathbf{V} . Thus \mathbf{N} and \mathbf{S} are inter-dependent, so if one is known then the other can always be calculated. The geometry is commutative, so a ray of light along \mathbf{S} incident at P would be reflected along \mathbf{V} , and vice versa. In the present case \mathbf{V} is assumed to coincide with the \mathbf{Z} axis,

and \mathbf{S} is the direction of maximum intensity of the HSH distribution.

The polar angles of the specular vector \mathbf{S} can be found analogously to Eq. 2-11:

$$\theta_S = \arccos\left(\sqrt{1 - l_{u0}^2 - l_{v0}^2}\right) \text{ and } \phi_S = \text{atan2}(l_{v0}, l_{u0}) \quad (4-6)$$

Then the polar angles of the normal vector \mathbf{N} are:

$$\theta_N = \theta_S/2 \text{ and } \phi_N = \phi_S \quad (4-7)$$

and the normal can be expressed as:

$$\mathbf{N} = (\cos \phi_N \sin \theta_N, \sin \phi_N \sin \theta_N, \cos \theta_N) \quad (4-8)$$



Figure 4.22: (left) Specular vectors for medallion calculated from HSH fitting; (right) derived normal vectors at half the specular angle.

Note that in the UCL dome the minimum elevation angle for the lamps in Tier 1 is approximately 4° , so the maximum angle θ_S of the specular vector \mathbf{S} from the view vector \mathbf{V} is 86° . It follows that the maximum co-latitude angle θ_N of the normal vector \mathbf{N} that can be determined by this method is 43° . The specular and normal vectors calculated from the medallion are shown as false colour images in Fig. 4.22. Because the angles of the speculars are greater, the X and Y components of the vectors are greater, and so the chromatic intensity of the red and green channels in the false colour representation is higher.

4.4.6 Comparing Normals with Arius 3D Geometry

The Arius 3D scanner at UCL has a moving laser head mounted on a massive coordinate measuring machine (CMM), which enables 3D surfaces to be digitised to a spatial resolution of 0.1 mm. Each pass over the object generates a 3D point cloud, including a normal vector at each point, of high geometric accuracy. It therefore provides a good source of reference data, against which to check the normals derived from the PTM and HSH fitting of the intensity data. Although the Arius laser scanner is a 3D digitising device, the point cloud it produces can be transformed into a 2D image by projecting the points onto a plane. In the simplest case the image is generated in the X-Y plane by ‘flattening’ all points down the Z axis.

The medallion was scanned in nine passes of the Arius, with the laser head orientated at varying angles to the surface to capture all faces of the relief and to avoid, as far as possible, the specular highlights from the glossy regions of the surface. Each file was saved in *a3d* format and imported into *Pointstream*. All nine files were merged into a single point cloud containing 3.42 million points and some editing done to remove extraneous points and fill holes. The composite point cloud was saved as a file in *asc* format, in which each point is represented by one line of 9 numerical fields encoded as ASCII text. Table 4.2 shows the fields in five text lines, for five successive points along one vertical laser scan line. Fields 1-3 represent the X,Y,Z coordinates in mm, relative to an arbitrary origin, expressed with a surprising number of decimal places (three would have been more than sufficient). It can be seen that the Y coordinate was decremented by about 0.117 mm (117 μm) for each successive point, as the laser beam swept

vertically, while the X coordinate remained almost constant, changing by only about 1.5 μm per point. The Z value was also nearly constant, indicating a flat horizontal surface. Fields 4-6 represent the 8-bit RGB reflectance values of the surface, which in this case was rather dark with a red cast. Fields 7-9 are the three components of the surface normal at the point, calculated by Pointstream from the coordinates of the neighbouring points.

Table 4.2: Five lines from the asc file for the medallion, showing nine fields per line.

1	2	3	4	5	6	7	8	9
X	Y	Z	R	G	B	Nx	Ny	Nz
-331.307251	-299.935699	-80.195625	73	50	33	-0.118710	0.026869	0.992565
-331.308777	-300.034088	-80.190849	69	50	33	-0.126527	0.056180	0.990371
-331.311218	-300.152771	-80.183632	66	50	32	-0.089399	0.042501	0.995089
-331.312775	-300.274506	-80.180893	67	48	33	-0.057157	0.010259	0.998312
-331.313538	-300.395660	-80.181236	62	46	37	-0.086468	0.006351	0.996234

The asc file was read via Matlab and each point mapped onto the nearest location in a two-dimensional image array with a resolution of 20 pixels/mm (grid spacing of 50 μm) using X,Y coordinates from the scanner, and ignoring the Z coordinate. Fig. 4.23 shows a detail of a flattened image. The spacing of points along the X direction is very regular, at 100 μm intervals, interrupted only once in every 46 lines approximately. In the Y direction the point spacing is much less regular, interrupted on every fourth or fifth pixel. These perturbations from the regular grid spacing lead to two distinct Moiré line patterns across the image: the diagonal white lines caused by X resampling and the nearly-horizontal black lines caused by Y resampling.



Figure 4.23:

Detail from an image generated from a3d scan data on a 50 μm grid. The vertical (Y) axis is the laser line scan direction (fast) and the horizontal (X) axis is the CMM traverse direction (slow).

The above image construction was repeated with each point being mapped onto the nearest location in a two-dimensional image array with a resolution of 10 pixels/mm (grid spacing of 100 μm). The image was further processed to fill the black lines, replacing each black pixel by the mean of the non-zero adjacent pixels in a 3x3 neighbourhood. In comparison with the normals calculated by the bounded regression technique (Fig. 4.24 right), as described in Chapter 5, they are slightly larger in value, manifested by greater chromatic intensity.

4.5 Conclusion

By capturing a set of images in pixel register, each illuminated from a different direction, a much richer representation of an object is obtained than by a single image in diffuse illumination. The

vector of intensity data for each pixel describes how the intensity varies according to the angle of the illuminating source. By fitting with suitable basis functions, a continuous function can be derived, approximating the variation of reflected intensity at that point for any angle within the hemisphere. The polynomial texture map (PTM) method implements this approximation by a biquadratic function with six coefficients, and the hemispherical harmonics (HSH) method by a variant of spherical harmonics with up to sixteen coefficients.



Figure 4.24: (left) Normal vectors from Arius 3D scan; (right) Normals calculated by bounded regression.

The objective of the investigation described in this chapter was to understand the characteristics and limitations of both PTM and HSH basis functions. The standard algorithms for both fitting and viewing were implemented in Matlab, using 16-bit linear TIFF images as input rather than the 8-bit JPG images used by the public-domain software utilities. In the course of the research the following additions to knowledge were made:

- Ability to composite two or more PTM files together, and so generate a combined PTM representation under the control of a single virtual light source.
- Explanation for the undefined normal vectors derived from PTM where the biquadratic function has no maximum (turning point).
- Enhancement of perceived surface relief by increasing the angle of the normal vector and re-rendering the PTM.
- Visualisation of the angular variation of intensity profiles and image components as surfaces on the polar plane, using the azimuthal equidistant projection.
- Quantitative estimation of the quality of approximation to the photographic image set by a summation of weighted basis functions.

To achieve a more accurate modelling of surface reflectance than is possible by the PTM and HSH methods, it is necessary to determine the normal and specular vectors at each point more precisely. This will open the way for 3D reconstruction by photometric stereo (Chapter 5) and realistic modelling of the specular component (Chapter 6).

Chapter 5 – 3D Surface Reconstruction

Indeed the more intelligible and comprehensive a theoretical system is,
The more obstinately it will resist all attempts at reconstruction or expansion.

– Max Planck (1933)
Where is Science Going? tr. J. Murphy, Ch. 1, p. 40
George Allen & Unwin, London

5.1 Lambertian Surfaces

The key to 3D reconstruction or rendering of an object is to find the surface normal vector at each point in the image set. Because all images are taken from the same viewpoint (neither the camera nor the object moves during acquisition of the image set), there is no baseline and so the methods of stereo photogrammetry cannot be applied. Instead the relationship between incident illumination angle and reflected intensity, taken over a subset of the images, is used to infer the slope of the surface and hence the direction of the normal vector.

The test object used to illustrate this chapter is a bas relief of the composer Chopin, of a design possibly based on his death mask in 1849. It is moulded from terracotta and is circular, of diameter 114 mm, weighs 250 g and is painted with a semi-gloss varnish (Fig. 5.1). The figure rests on an almost planar but slightly inclined clay disc of thickness varying from 9.5 mm at the upper right to 12.5 mm at the lower left. The letters FREDERIC CHOPIN are in raised relief around the curved edge on the left side, and the sculptor's name FBLACK is incised into the section at the base of the neck. The object is nearly monochrome (i.e. clay coloured) and is ideal for finding surface normals and reconstructing the 3D shape because of its smooth curves and absence of texture or occlusions. It was photographed in the dome with the Nikkor 15-55 mm zoom lens, set to 55 mm at aperture $f/8$, producing an image of 1626 pixels in diameter, i.e. a scale of 14.26 pixels/mm, equivalent to a pixel size of 70 μm .



Figure 5.1: Two views of terracotta bas relief of Frederic Chopin, dated 1871 on rear, illuminated by: (left) sixteen flash lamps in Tier 1 of dome; (right) four flash lamps in Tier 5 of dome.

The photometric stereo technique was explained at length in Section 2.3. It enables the normal at each point of the object surface to be determined for a single viewpoint, using the principle that the intensity of the reflected light depends on the angle of incidence of the light onto the surface and the reflectance factor. With a perfectly Lambertian surface and in the absence of noise, only three intensity values $[I_1, I_2, I_3]$ from non-coplanar light sources with unit direction vectors $[\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3]$ would be sufficient to solve for both the normal direction \mathbf{N} and the surface albedo ρ :

$$I_i = \rho \mathbf{L}_i \cdot \mathbf{N} = \rho |\mathbf{L}_i| \cos \alpha_i \quad (5-1)$$

where α_i is the angle between the lamp vector and the normal. In practice, normals calculated in this way from three light directions exhibit a high level of noise and vary widely according to the particular lamps selected. Fig. 5.2 shows the normals in a small region computed from the green channel of the image, using linear intensity data from the 16-bit TIFF file. There are numerous reasons for the variability seen in the computed normals:

- Non-uniform reflectance of the surface as a function of angle (i.e. non-Lambertian);
- Errors in the XYZ coordinate positions of the lamps;
- Non-point source (a flash lamp of length 9 mm at the mean dome radius of 515 mm subtends an angle of approximately 1° at the object surface);
- Variable intensity of the flashes (different illuminance from the three sources);
- Different lengths of the light paths from the three lamps (inverse square law);
- Flare light (scattered from other areas) adds to the direct illumination;
- Non-uniform profile of illumination across the image area;
- Physical movement of the camera relative to object between exposures;
- Non-linear signal from the camera (pixel value not proportional to luminance);
- Noise in the camera sensor (amplified by the matrix inversion);
- Quantising errors in the RGB image (3x16 bits per pixel);
- Quantising error in the representation of normals by 16-bit integers.

Better results can be obtained for noisy image data by calculating normals for many triplets of light sources (MacDonald and Robson, 2010). By selecting suitable combinations of three lamps, candidates for the normal can be calculated for every pixel. Figure 5.2 shows the normals for a 200x200 pixel detail, presented as false-colour images, for three combinations of lamp triplets from Tiers 2, 3 and 4 of the dome. There is clearly a good deal of variance among the three images, and also the noise is evident.

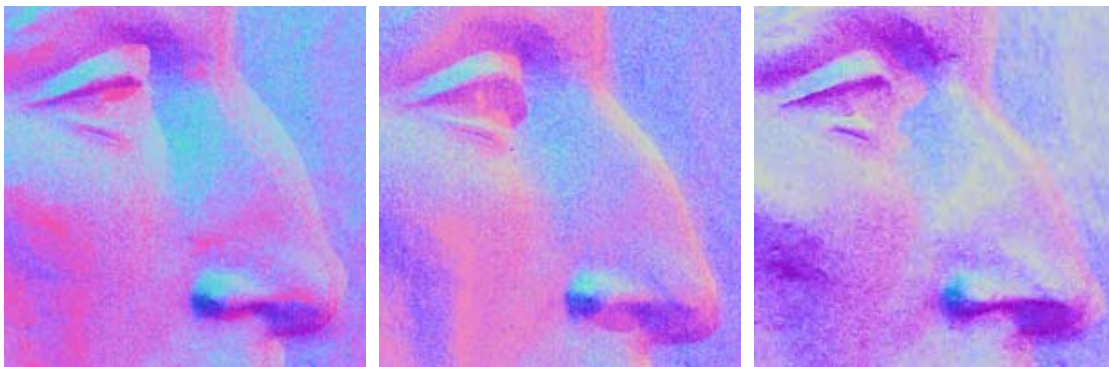


Figure 5.2: Detail of size 200x200 pixels of normals computed from the green image channel, for three dome lamp combinations: (left) 40,50,58; (centre) 35,45,54; (right) 37,43,60.

The extent of scatter in the normals calculated from triplets can be seen by plotting the values at a single pixel for all $16 \times 16 \times 12 = 3072$ triplet combinations of the lamps in Tiers 2, 3 and 4 of the dome (Fig. 5.3). Both the N_x and N_y components of the normal cover the full range of values from -1.0 to +1.0, with respective means of 0.202 and 0.249 and standard deviations of 0.328 and 0.309. Taking the median of each distribution for every pixel yields the normals shown in Fig. 5.4 (left).

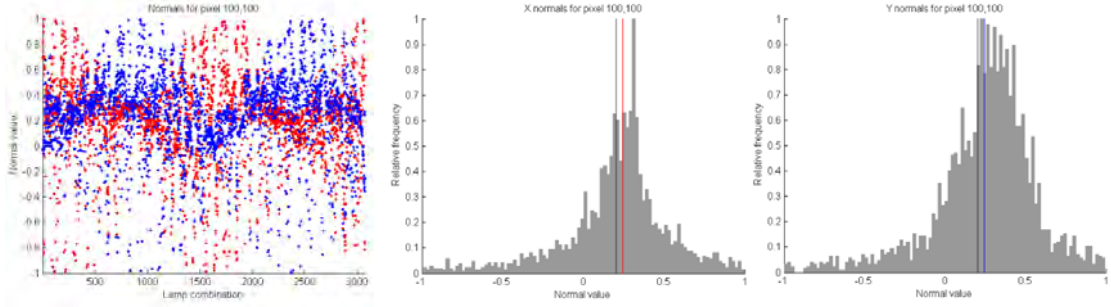


Figure 5.3: (left) Scatter plot of N_x (red) and N_y (blue) components of normals calculated for all 3072 triplet combinations of lamps for one pixel; (centre) N_x components, with mean and median shown by black and red lines; (right) N_y components, with mean and median shown by black and blue lines.

Better results can be achieved by including intensities from more than three lamps in the regression, by inversion of Eq. 5-1:

$$\rho \mathbf{N} = \mathbf{L} \backslash \mathbf{I} \quad (5-2)$$

where \backslash is the pseudo-inverse operator in Matlab, \mathbf{L} is an $n \times 3$ matrix of unit lamp vectors for n lamps, and \mathbf{I} is an $n \times 1$ vector of the corresponding intensity values. By including all 64 lamps in the dome, the result shown in Fig. 5.4 (centre) is obtained, which is quite close to the median of triplets in Fig. 4 (left) as the horizontal cross-section through the N_x normal values shows.



Figure 5.4: (left) Normals calculated as median of distributions of N_x and N_y for 3072 lamp triplets; (centre) Normals calculated by regression of intensity values over all 64 lamp vectors; (right) cross-sections of N_x values across horizontal midline.

5.2 Effect of Specularity

For a non-Lambertian surface the above method gives incorrect results. Consider the geometry of Fig. 5.5 (left) at point P on a curved shiny surface. The view vector V , passing through the perspective centre of the camera lens, subtends an angle θ with the normal N . Any lamp near the specular vector S , at an angle of θ from N and 2θ from V , will produce an intensity in the image greater than would be produced by a matte surface. In the solution of the photometric stereo equation Eq. 5-1 this is interpreted as if the surface normal were orientated closer to S .

than it actually is, producing the distorted normal N' . Thus the effect of surface gloss is to exaggerate the apparent gradient of the surface (dotted line).

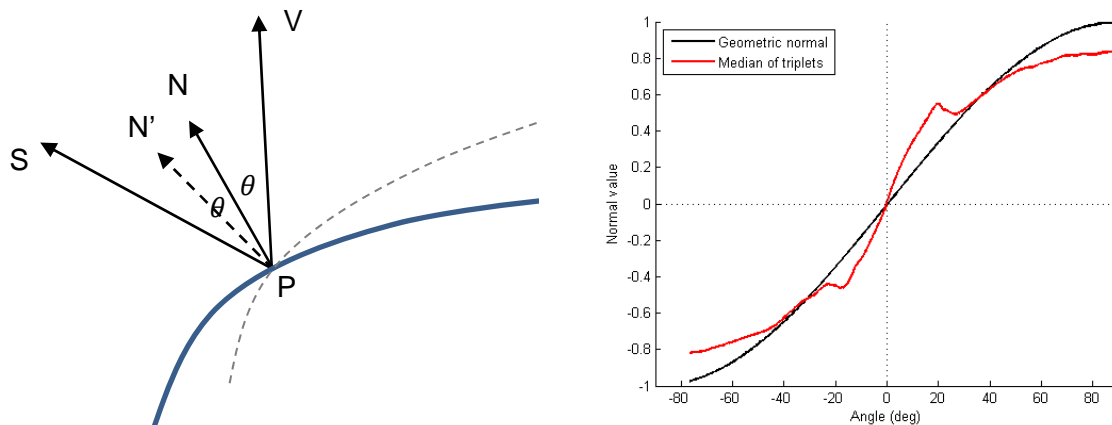


Figure 5.5: (left) Geometry of specular reflection, showing distortion of normal; (right) angle of normal computed from the median of triplets (red) vs true normal (black) for a spherical surface.

The effect was demonstrated by creating a digital test object (see Section 5.3), namely a sphere with a semi-gloss surface. Sixty-four images were rendered, each illuminated by a point source at successive lamp coordinates in the dome. These were then processed as if real images captured by the camera in the dome, and normals calculated by the median-of-triplets technique. The resulting normal angles can be compared everywhere with the true geometric angle on the surface of the sphere, and show significant over-prediction in the range $\pm 20^\circ$ from the view vector and under-prediction for angles greater than 45° (Fig. 5.5 right).

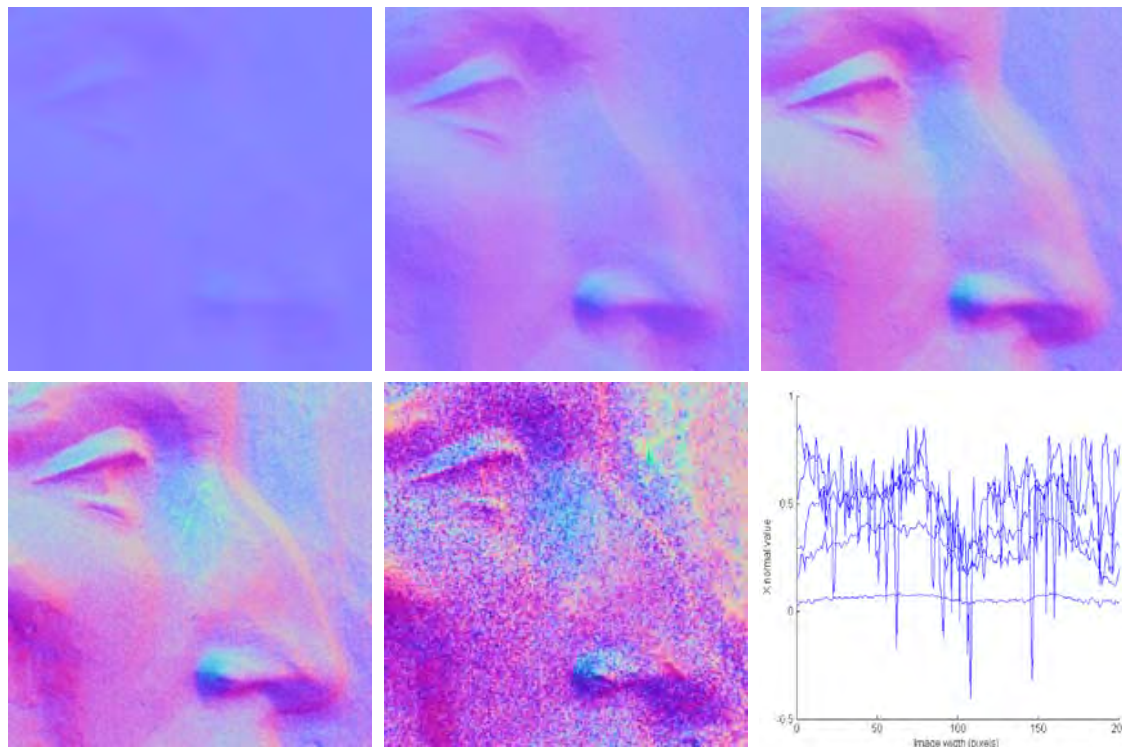


Figure 5.6: Normals calculated by regression over each of the five tiers of lamps in the dome; (lower right) cross-sections of five sets of N_x values across the horizontal midline.

Real surfaces are non-Lambertian and non-planar, and so exhibit both gloss and self-shadowing from oblique lighting, as shown in the image mosaic in Fig. 4.2 from all 64 images of the medallion, taken in the UCL dome under 64 different lighting directions. Although the mean brightness increases with illumination angle, the shadowing is severe at low incident light angles, especially for the lowest tier of lamps with an elevation less than 10° . The consequence is that calculating the normals by regression over different subsets of lamps gives different results, as shown in Fig. 5.6. At the lowest tier (the sixteen flash lamps nearest the baseboard) the normals are barely detected, whereas at the highest tier (the four flash lamps close to the camera) the slopes of the normals are extreme and very noisy.

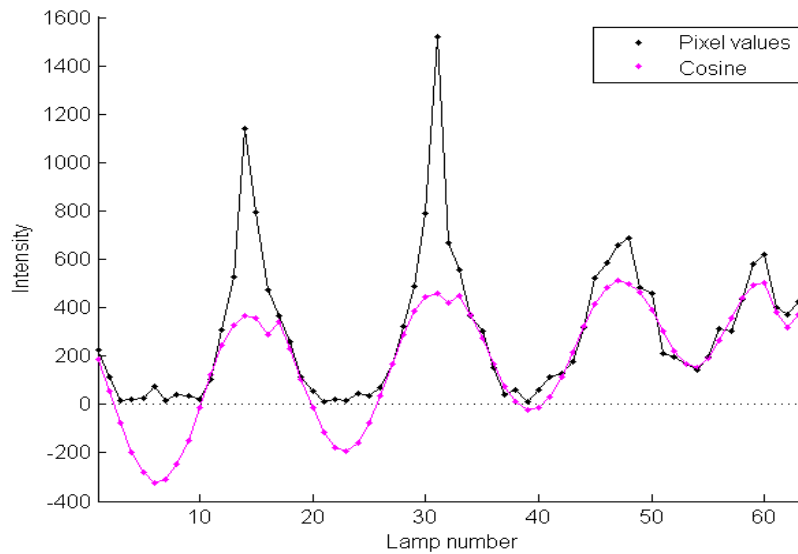


Figure 5.7: Intensity values at a single pixel for 64 illumination directions (black), compared with a similarly-illuminated perfect diffuser (magenta).

The problem is therefore how to select the best subset of all the intensity values at each pixel to avoid both shadows and specular highlights. A new method for estimating normals has been developed, which is robust and adapts to the presence of both shadows and surface gloss. First all the intensity values at a pixel are extracted from the image set and treated as a vector P (see Section 4.1.3). As an example a pixel is selected from the detail of Fig. 5.1 on the cheekbone just below the eye. Plotting the 64 values against lamp number (Fig. 5.7) shows the wide variation in intensity for different lamp angles, with peaks for lamp numbers 15, 31, 48 and 59.

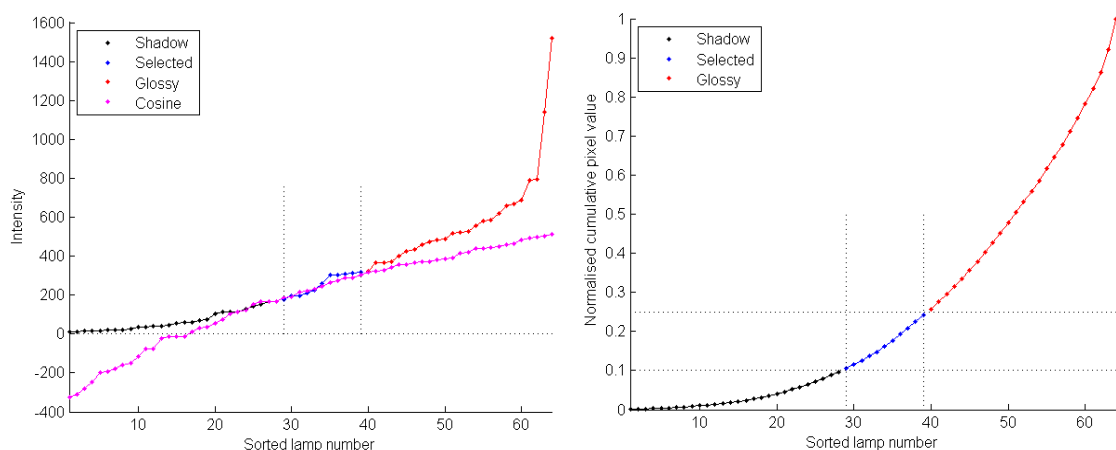


Figure 5.8: (left) Sorted intensity values for actual and cosine distributions; (right) Normalised cumulative distribution of sorted intensity values.

Also plotted for comparison (in magenta) is the intensity that would be expected from a perfect Lambertian surface of the same albedo illuminated by the same lamps, which follows the cosine of the angle between the lamp vector and the computed normal. This is negative for eight lamps in Tier 1, six lamps in Tier 2, and two lamps in Tier 3, corresponding to angles of incidence below the 'horizon' of the surface at that point. Note that for lamps with incident vectors close to the specular direction the intensity is higher than the cosine, whereas for others the intensity is close to that of the cosine.

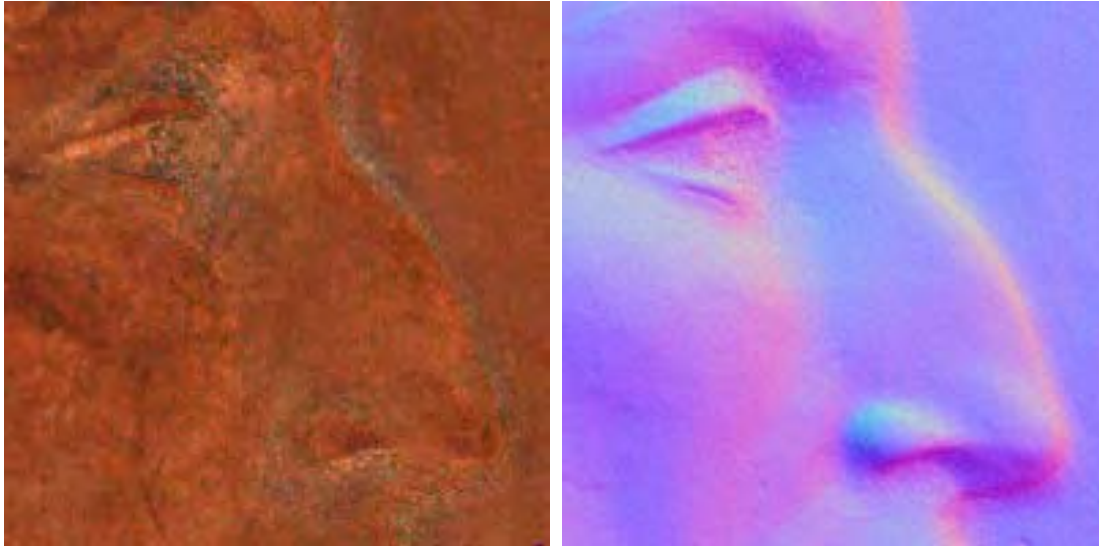


Figure 5.9: (left) Albedo and (right) normals calculated by the 'bounded regression' procedure.

The intensity values are then sorted into ascending order (Fig. 5.8 left) and the cumulative sum calculated (Fig. 5.8 right). The subset of lamps is selected for which the normalised cumulative values lie between two thresholds, set to 0.10 and 0.25. These thresholds are chosen somewhat arbitrarily to select a region of the distribution that follows the slope of the cumulative sorted cosine, i.e. related to the diffuse component of the reflection. In the example for this pixel, eleven lamps are selected (shown in blue in the figure). The lamps below (shown in black) descend into the shadow region and the lamps above (shown in red) ascend into the specular region. Regression is applied over the selected subset of lamp vectors to solve Eq. 5-2 to estimate the normal and albedo.



Figure 5.10: Enlarged detail of eye from normals generated by bounded regression procedure applied to: (left) single pixels showing the effect of noise; (centre) cells of 3x3 pixels.

The significance of this ‘bounded regression’ procedure can be seen by comparison of the sorted intensity values (Fig. 5.8 left) with the sorted cosine values (plotted in magenta), which show the baseline of the diffuse component. The selected subset coincides closely with the cosine. It lies above all cases where the pixel is in shadow, but below cases where there is an added component of reflectance from the surface gloss. The colour albedo arising from this procedure is a good approximation to the intrinsic diffuse component of surface reflectance (Fig. 5.9 left).

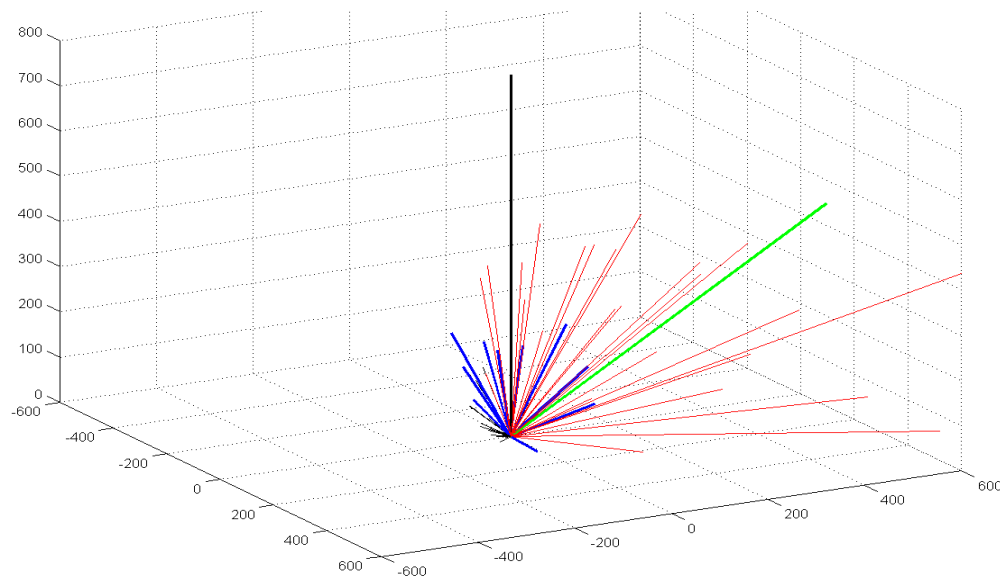


Figure 5.11: Visualisation of intensities at one pixel as 3D vectors toward lamp coordinates (mm).

The normals generated by this procedure (Fig. 5.9 right) are smooth and (as shown in the next section) accurate. When the algorithm is applied on a single pixel basis, however, the results are subject to perturbation by small blemishes on the surface and granularity and dust, as shown in Fig. 5.10 (left). Better results are achieved by calculating the normal at each pixel from the $9 \times 64 = 576$ values in a cell of 3×3 pixels, i.e. the central pixel and its eight nearest neighbours (Fig. 5.10 centre). This gives a smoother gradation of the normals (Fig. 5.10 right), but is more accurate than simply averaging the single-pixel normals by a 3×3 box filter because the bounded regression is applied to a larger number of values over the neighbourhood of the pixel.

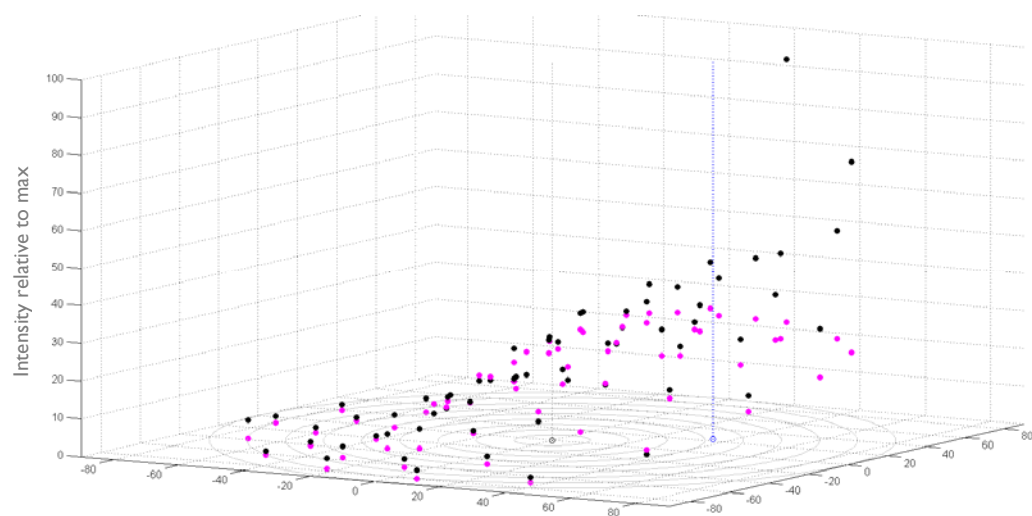


Figure 5.12: Visualisation of intensities at one pixel on an azimuthal equidistant projection. The dotted blue line represents the normal. The black points show pixel intensities and magenta points show the cosine distribution.

The subset of lamps selected by the procedure can also be visualized by plotting 3D vectors of length proportional to the pixel intensity corresponding to each lamp (Fig. 5.11). The heavy black line represents the view vector V and the green line the normal N . The black, blue and red lines correspond to the points in Fig. 5.8 for the 64 lamps. It is apparent that the selected lamp vectors (shown in blue) lie in an umbrella formation, well away from the normal where the influence of specularity is low. Plotting the 64 values of Fig. 5.7 for intensity and cosine in 3D, using the azimuthal equidistant projection (Sec. 4.1.6), gives the distribution in Fig. 5.12. The cosine distribution (magenta) lies on a smooth surface with maximum at the coordinates of the normal, whereas the intensity values (black) rise above the cosine to a peak at the specular angle, at approximately double the angle of the normal from the pole.

5.3 Verifying Normals with Digital Test Objects

To verify the accuracy of surface normal vectors generated by the procedure described in the previous section, an object with known gradients is needed. One way to achieve this is by creating a set of images of a digital test object which has the characteristics of a real object but for which all the dimensional values are specified.

1. Define the shape and key parameters in three dimensions X, Y, Z ;
2. Generate a 'height image' with a height value plus noise at each pixel on 1000x1000 grid;
3. Calculate the gradients in X and Y directions as first-order differences of height values;
4. Calculate ideal normal vectors from the gradients;
5. Generate an albedo image with an arbitrary pattern over the image grid;
6. Render 64 images of the object, as if illuminated by a virtual light source at each location of the real lamps in the dome, viewed from the camera's perspective centre;
7. Apply the 'bounded regression' method to estimate the normals and albedo;
8. Compare the two sets of normals.

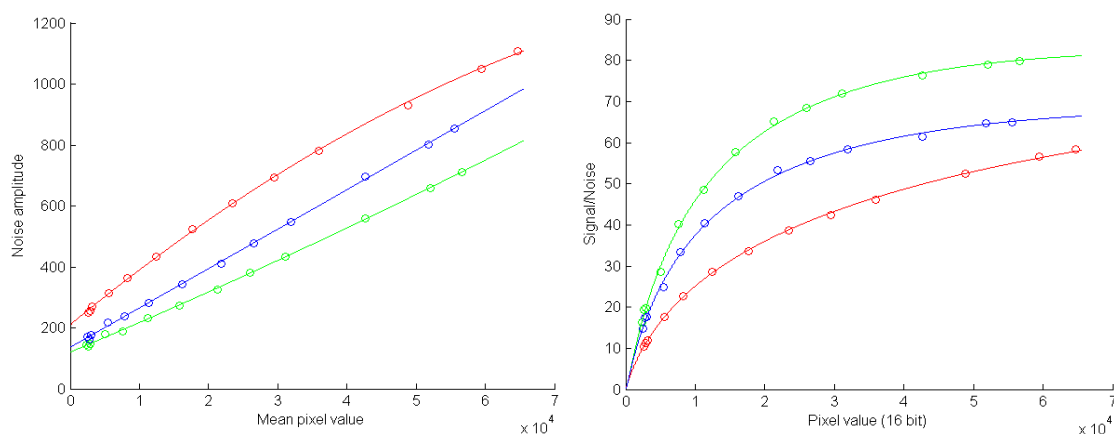


Figure 5.14: (left) 2nd-order polynomial (lines) fitted to mean noise (open circles) vs pixel value for R,G,B; (right) Result of fitting 2nd-order polynomials to SNR data vs pixel value in log-log domain.

5.3.1 Image Noise

To simulate real images from the Nikon D200 camera, random noise was added to each image channel. The signal-to-noise values determined from the greyscale of the GretagMacbeth target (see Section 7.1.5) were taken as representative of the typical performance of the camera. To

facilitate the generation of synthetic noise, the mean noise distribution in each channel was fitted with a second-order (quadratic) polynomial as a function of 16-bit pixel intensity value:

$$n_k = a_{2k}p_k^2 + a_{1k}p_k + a_{0k} \quad (5-3)$$

where: n_k is the mean noise level for channel k (one of R,G,B), p_k is the pixel value for channel k , and a_{jk} are the fitting coefficients.

The results shown in Fig. 5.14 (left) indicate that the noise increases monotonically with pixel value and is nearly linear. The corresponding functions for SNR need to be fitted on log-log data and transformed back to the linear domain, as shown in Fig. 5.14 (right).

Eq. 5-3 gives the mean standard deviation of the noise, equivalent to the root-mean-square (rms) value, so the amplitude of the zero-mean noise o_k added to the synthetic image value p_k at each pixel is calculated as:

$$o_k = 2\sqrt{2} n_k(\text{rand} - 0.5) \quad (5-4)$$

5.3.2 Square Pyramid

A simple four-sided peaked digital solid was generated using the Matlab code in Fig. 5.15. Strictly speaking, it is not a regular square pyramid but it is similar enough to justify the name.

```

imw = 1000; imh = 1000;           % image size (pixels)
bw = 50;                          % border width (pixels)
hmax = 200;                       % max height (pixels)
H = zeros(imh,imw,'double');      % image for height field
Z = peaks(3);                    % make a peak
[XP,YP] = meshgrid(1:3);          % coarse grid for peaks
[X,Y] = meshgrid(linspace(1,3,900)); % subdivide into 900
ZH = interp2(XP,YP,Z,X,Y,'linear'); % rescale peak to image
ZH = ZH-min(min(ZH));             % set min level to zero
H(bw+1:imh-bw,bw+1:imw-bw) = hmax*ZH; % insert inside outline

```

Figure 5.15: Fragment of Matlab code to generate square pyramidal test object.

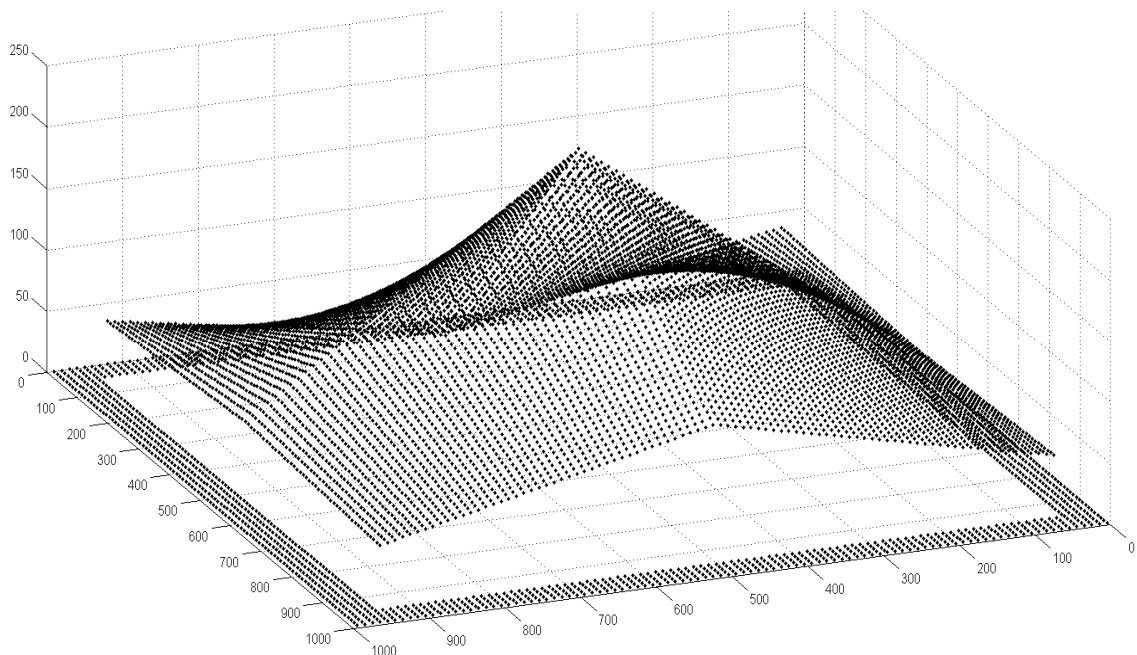


Figure 5.16: Visualisation of square pyramid test object, drawing a dot at every 10th pixel on surface.

The peak is at the centre of the image and there is a border of width 50 pixels all around with value zero, simulating the baseboard of the dome around a real object. The four edges of the upper surface, where there is an abrupt change of slope from one face to the next, are aligned with the X and Y axes (Fig. 5.16). The whole object is raised as if sitting on a planar surface, with no negative height values, and there are sheer vertical drops to the ground plane around the edges. Cross-sectional elevations make this clear (Fig. 5.17).

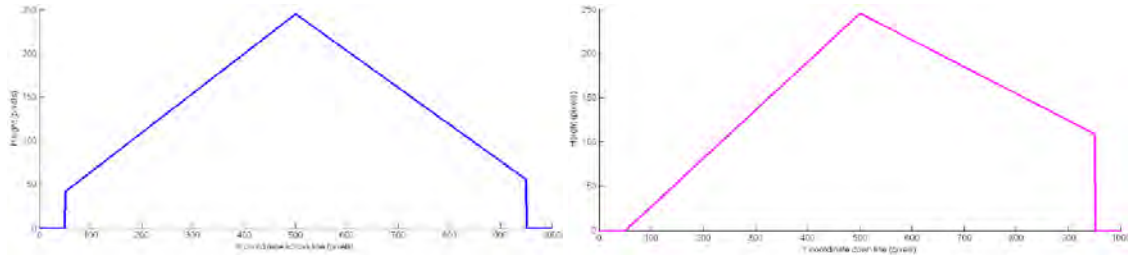


Figure 5.17: Horizontal (left) and vertical (right) midline cross-sectional elevations through pyramid.

The gradients were calculated as the average of the first-order difference of values on either side of the central pixel. For horizontal gradients (X direction) this means the pixels on right p_{i+1} and left p_{i-1} ; for vertical gradients (Y direction) it means pixels above p_{j+1} and below p_{j-1} , where i and j are the row and column indices respectively.

$$\left. \begin{aligned} z_x &= \frac{\partial z}{\partial x} = \frac{(p_{i+1} - p_i) + (p_i - p_{i-1})}{2} = \frac{(p_{i+1} - p_{i-1})}{2} \\ z_y &= \frac{\partial z}{\partial y} = \frac{(p_{j+1} - p_j) + (p_j - p_{j-1})}{2} = \frac{(p_{j+1} - p_{j-1})}{2} \end{aligned} \right\} (5-5)$$

In Matlab this is neatly done for the whole image by a few array operations, as shown in Fig. 5.18. The gradient at the pixel nearest the edge on all sides remains zero.

```

Zl = zeros(imh,imw,'double'); % neighbour left
Zr = zeros(imh,imw,'double'); % neighbour right
Zu = zeros(imh,imw,'double'); % neighbour up
Zd = zeros(imh,imw,'double'); % neighbour down

Zl(:,2:imw) = H(:,1:imw-1); % pixel to left
Zr(:,1:imw-1) = H(:,2:imw); % pixel to right
Zd(2:imh,:) = H(1:imh-1,:); % pixel below
Zu(1:imh-1,:) = H(2:imh,:); % pixel above
Zx = (Zr-Zl)/2; % gradient of Z in X direction
Zy = (Zu-Zd)/2; % gradient of Z in Y direction

```

Figure 5.18: Fragment of Matlab code to calculate first-order gradients in X and Y directions.

Image representations of the ideal heights, gradients and normals of the pyramid are shown in Fig. 5.19. The height map is a monochrome, scaled from black at zero (baseplane) to white at the peak. The gradients are in bichromatic false-colour with the X gradient in the red channel and the Y gradient in the green channel. The normals are in the conventional trichromatic false-colour for the three vector components N_x , N_y , N_z .

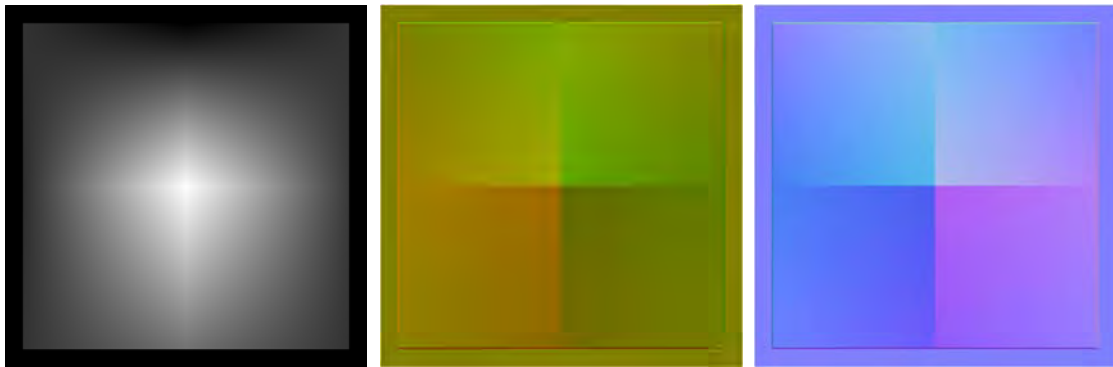


Figure 5.19: (left) Ideal height map, (centre) gradients and (right) normals of pyramid.

A set of 64 images was rendered by taking the coordinate position of each dome lamp in turn and at each pixel forming the product of Eq. 5.1, and adding random noise in each channel of each pixel as described in Sec. 5.3.1. To make the lightness gradations clearer, the albedo was set to a uniform light grey. The light intensity was assumed to be uniform over the whole area of the virtual object and the incident rays parallel, as if emanating from a point source at infinity. The images were saved in `uint16` format, i.e. as 16-bit positive integer values in range 0-65535.

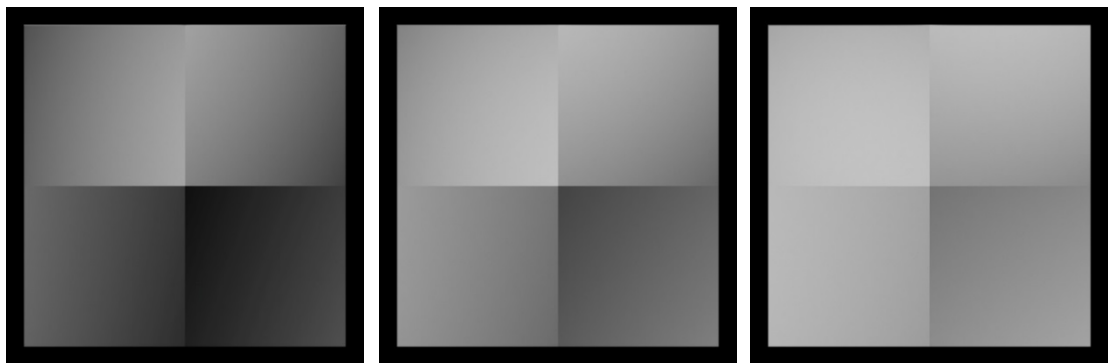


Figure 5.20: Rendered images of the pyramid for (left to right) lamps 18, 35 and 51.

Three of the rendered images are shown in Fig. 5.20, all illuminated from a similar azimuth angle to the north-west. The overall lightness increases as the virtual light source increases in zenith angle, i.e. for higher lamp tiers in the dome. Applying the ‘bounded regression’ procedure to the set of 64 rendered images gave the normals and albedo shown in Fig. 5.21.



Fig. 5.21: (left) Albedo and (right) normals computed from rendered images.

The albedo (left) is almost perfectly uniform grey, as it should be, except for a darkening at the four top edges of the pyramid, which appear as a cross at the centre of the image. These lines correspond to discontinuities in the slope of the surface. The normals (right) are smooth and well-modulated, and they appear very similar to the ideal normals of Fig. 5.19.

Plotting cross-sections shows a very close correspondence between the ideal and estimated normals (Fig. 5.22) at all positions except at the edges where the surface ends and the image reverts to the ground plane. The enlarged detail (right) shows that the mean estimated normal follows the ideal normal, but with a random error caused by the noise added to the rendered images. The rms value of this error is 0.000912. This cross-section is 10 pixels from the horizontal centre line to avoid the discontinuity at the edges of the pyramid faces.

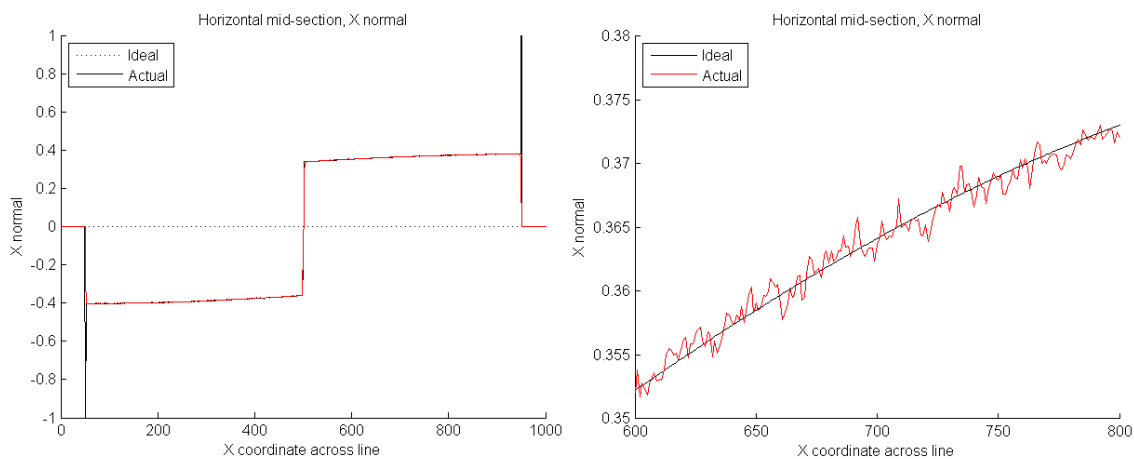


Figure 5.22: (left) Cross-section through N_x components of ideal and estimated normals; (right) detail.

Displaying as false-colour image maps the differences between the estimated and ideal normals for the N_x and N_y components (Fig. 5.23) shows that they are well distributed over the whole image, apart from the outer edges and the pyramid edges. To make the errors visible the values have been magnified by a factor of 10^4 . On the 8-bit image scale 128 represents zero difference, and it can be seen that the majority of errors lie in the range between 158 (yellow) and 98 (cyan), representing actual error values in the range ± 0.003 .

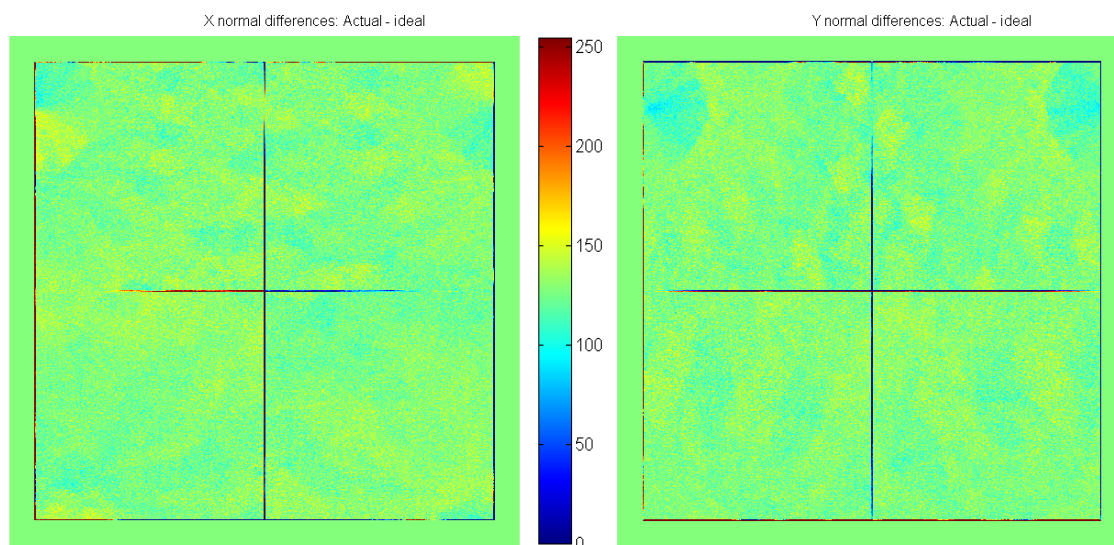


Figure 5.23: False-colour maps of magnified errors between: (left) N_x components and (right) N_y components. Green indicates zero difference, yellow indicates positive and cyan indicates negative.

5.3.3 Hemi-ellipsoid

A similar process was repeated to create a digital hemi-ellipsoidal solid, in effect a hemisphere with different scaling factors for the three semi-axes. The radii in X, Y and Z are 450, 350 and 250 pixels respectively. Image representations of the ideal heights, gradients and normals of the hemi-ellipsoid are shown in Fig. 5.25, encoded as in Fig. 5.19.

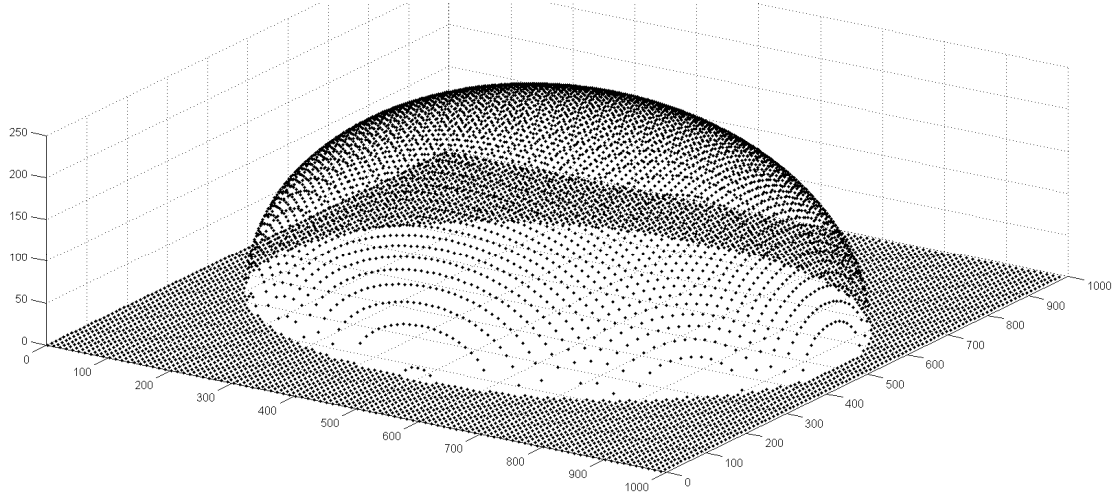


Figure 5.24: Visualisation of hemi-ellipsoidal test object, drawing a dot at every 10th pixel on surface.

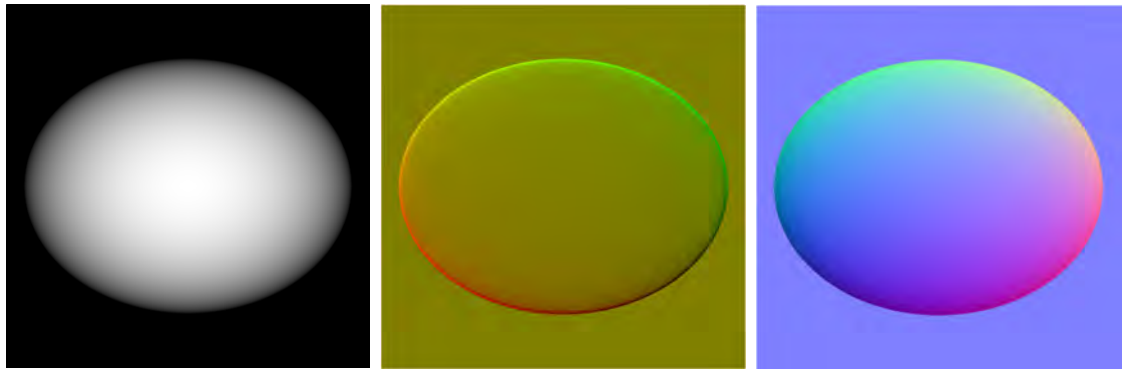


Figure 5.25: (left) Ideal height map, (centre) gradients and (right) normals of hemi-ellipsoid.

A set of 64 images was rendered by taking the coordinate position of each dome lamp in turn and at each pixel forming the product of Eq. 5.1, and adding random noise in each channel of each pixel as described in Sec. 5.3.1. To make the lightness gradations clearer, the albedo was set to a uniform light grey. Three of the rendered images are shown in Fig. 5.26, all illuminated from a similar azimuth angle to the north-west.

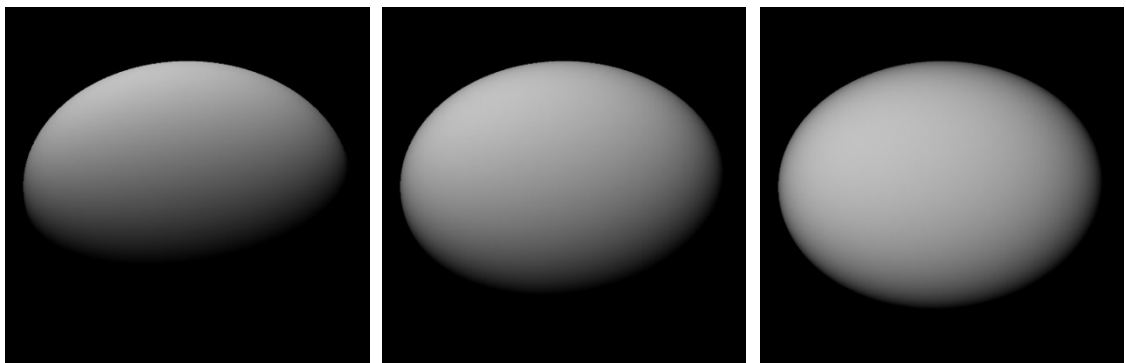


Figure 5.26: Rendered images of the hemi-ellipsoid for (left to right) lamps 18, 35 and 51.

Applying the ‘bounded regression’ procedure to the set of 64 rendered images gave the normals shown in Fig. 5.27 left, which appear very similar to the ideal normals of Fig. 5.25. Plotting a cross-section on the horizontal midline shows a very close correspondence between the ideal and estimated normals (Fig. 5.27 centre). The enlarged detail (Fig. 5.27 right) at the extreme edge of the curve shows that the mean estimated normal follows the ideal normal closely, but with a random error caused by the noise added to the rendered images. The estimated value is slightly less than the true value as the slope approaches vertical ($N_x = 1$). The rms value of the overall error is 0.0013.

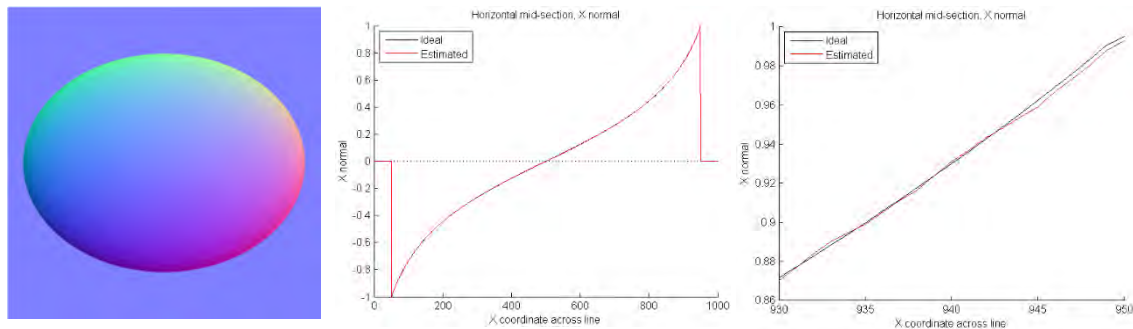


Figure 5.27: (left) Normals; (centre) cross-section through N_x components of ideal and estimated normals; (right) detail of peak of cross-section.

Displaying as false-colour image maps the differences between the estimated and ideal normals for the N_x and N_y components (Fig. 5.28) shows an asymmetrical distribution over the image. The differences are slightly greater on the negative side (left for N_x and bottom for N_y) and smaller on the positive side, particularly near the highest point of the surface (centre of the image), where the normal values change sign.

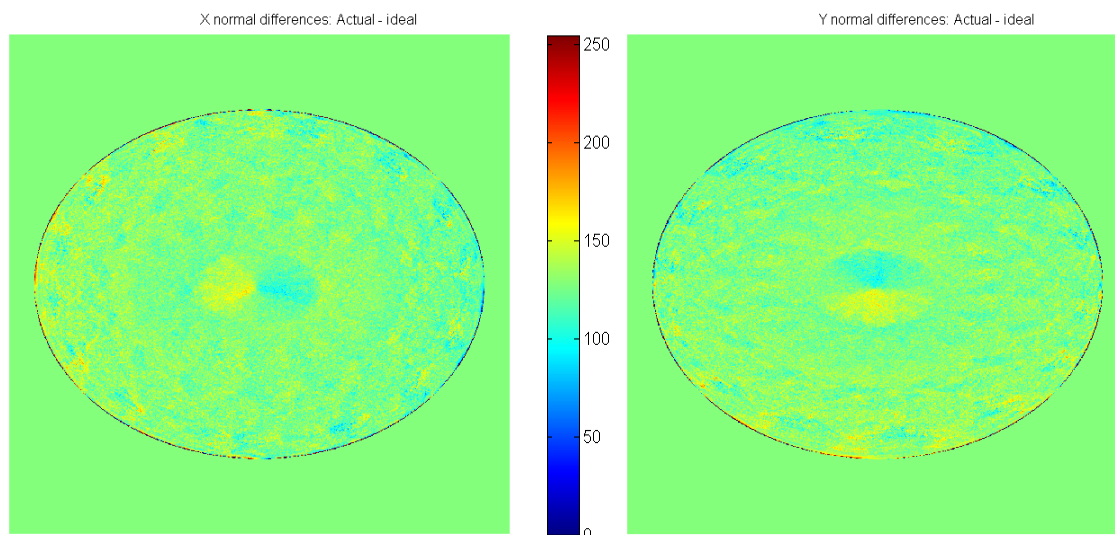


Figure 5.28: False-colour maps of magnified errors between: (left) N_x and (right) N_y components. To make the errors visible the difference values have been magnified by a factor of 10^4 .

5.4 Height Reconstruction by Integration of Gradients

In the examples in the previous section for digital objects, the gradients in X and Y at each pixel were calculated as the first derivatives of the known height, and then the surface normal vector was calculated from the gradients. The problem for photometric stereo is the opposite: given the surface normals, inferred from the gradations of intensity on a surface illuminated by a set of lights at known positions, then determine the (unknown) height of the surface at each pixel. The first step of calculating the surface gradients from the normals is easy, by taking the ratio at each pixel of the N_x and N_y components relative to N_z (Fig. 5.29). The Q gradient (in the Y direction) is inverted because the Matlab convention places the origin at the upper left corner of the image, rather than the lower left corner in geometric Cartesian axes.

```
Nx = N(:, :, 1); % components of normal vectors
Ny = N(:, :, 2);
Nz = N(:, :, 3);
P = -Nx ./ (Nz+eps); % dz/dx
Q = Ny ./ (Nz+eps); % dz/dy (inverted)
```

Figure 5.29: Fragment of Matlab code to calculate first-order gradients from normals.

The normals calculated by the bounded regression method for the digital pyramid are shown as a false colour image in Fig. 5.21. Applying the transformation of Fig. 5.29 to the normals gives the gradients, as shown in the false-colour image in Fig. 5.30. The values of the gradients P and Q , in the X and Y directions respectively, along horizontal and vertical cross-sections, are plotted in Fig. 5.31. In the horizontal direction the P gradient is constant and positive on the left side, where the surface is rising, and at the ridge line changes abruptly to a negative constant on the right side, where the surface is falling. The behaviour of the Q gradient is similar in the vertical direction, except that the upward gradient is greater than the downward gradient.

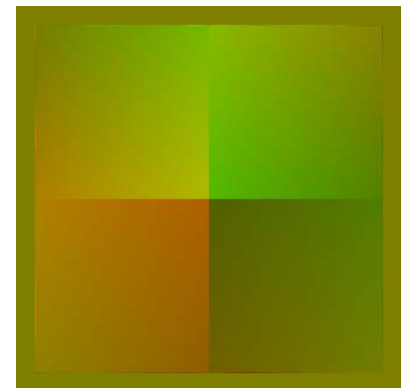


Figure 5.30: Gradients derived from normals.

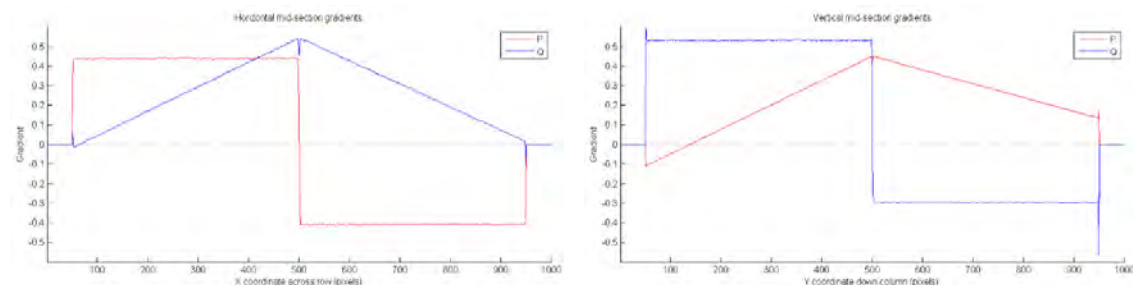


Figure 5.31: (left) Horizontal and (right) vertical midline cross-sections of P and Q gradients.

If the gradient were known precisely at every pixel then the height could be reconstructed simply by integrating the gradients across the surface. Unfortunately, even with a well-behaved surface, errors from various sources such as noise and quantisation accumulate to produce significant discrepancies. In particular, the gradients are not well defined at the edges of the

object where there is an unknown drop to the baseplane, and where the values of normals and gradients may be set to zero by the outline mask. This is illustrated in one dimension in Fig. 5.32, in which the Q gradient on the vertical midline (Fig. 5.31 right) is summed along the line, both top-to-bottom and bottom-to-top. Starting at zero (the assumed height of the baseplane) the gradients are added successively, pixel by pixel, in each direction. Although the resulting profiles of the pyramid are identical, their heights are very different, because of errors in encoding the large changes of gradient at the two edges. The upper profile (cyan) is actually the correct one, because the pyramid surface touches the baseplane at the top, but is raised substantially at the bottom (see Fig. 5.17 right).

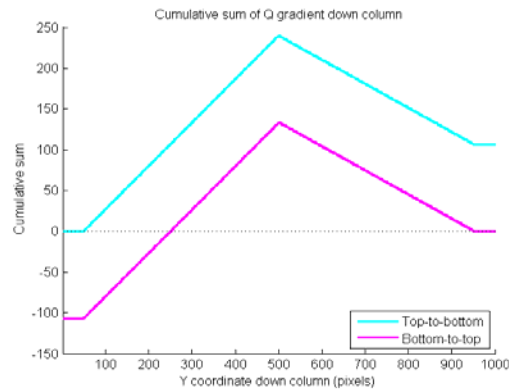


Figure 5.32: Cumulative sums of Q gradients in two directions up and down the vertical midline.

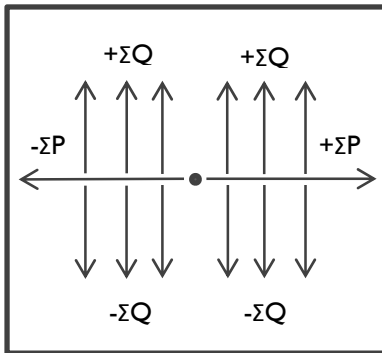


Figure 5.34: Method for integrating gradients from centre of image.

To avoid errors at the edges, the integration can commence at the centre of the image and work outward (code in Fig. 5.33). First the midpoint coordinates are calculated, and the height at this point taken as zero. Then the profile of the horizontal midline is calculated in two halves, summing the P gradients in the second half forward and the P gradients in the first half backward. The calculated values across the midline are then used as starting values for the same process applied to each vertical column, summing the Q gradients forwards down the lower half and backwards up

the upper half. The procedure is illustrated in Fig. 5.34. Finally the whole reconstructed surface is raised above the baseplane by subtracting the overall minimum value.

```

QI = zeros(imh,imw,'double'); % allocate image buffer
imid = floor(imw/2); jmid = floor(imh/2); % coordinates of centre point
Pm = zeros(imw,1,'double'); % allocate line buffer
Pm(imid:imw) = cumsum(P(jmid,imid:imw)); % sum across second half row
Ps = flip(P(jmid,1:imid-1)); % first half of line reversed
Pm(1:imid-1) = -flip(cumsum(Ps)); % sum backwards first half row
for I = 1:imw % for all columns across image
    QI(jmid:imh,I) = Pm(I)+cumsum(Q(jmid:imh,I)); % forward sum half column
    Qs = flip(Q(1:jmid-1,I)); % first half of line reversed
    QI(1:jmid-1,I) = Pm(I)-flip(cumsum(Qs)); % backward sum of half column
end
QI = QI-min(min(QI)); % raise all above baseplane

```

Figure 5.33: Fragment of Matlab code to integrate height from gradients.

The same procedure can be carried out by exchanging rows for columns, i.e. by calculating the profile of the vertical midline, summing the Q gradients in the second half downward and the Q gradients in the first half upward. The calculated values down the midline are then used as starting values for the process applied to each horizontal row, summing the P gradients forwards along the second half and backwards along the first half. The result can be averaged with the first to produce an overall estimate of the height at each pixel.

Applying the procedure to the pyramid gave a result close to the original ideal shape, with the maximum reconstructed height at the apex of 244.36 pixels, compared with a height of 244.78 pixels for the idealised original. The differences between the reconstruction and the ideal are shown in false colour in Fig. 5.35 (left) where the scale is in units of 0.01% of relative height (i.e. $0.0001 \times \text{difference} / \text{maximum height}$). Although the differences along the top edges of the pyramid are positive, the differences on the four faces are negative in the approximate range 30-60 units = 0.2-0.3% of height. The horizontal cross-section of the difference map in Fig. 5.35 (right) confirms that except for the edge (centre of graph) the estimated height is always less than the ideal height by 0.45 to 0.67 pixel.

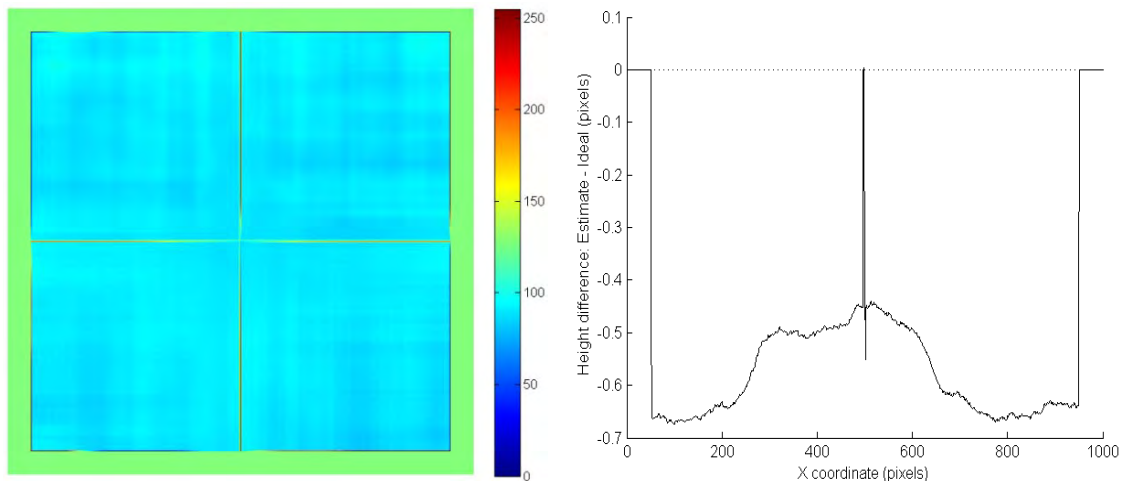


Figure 5.35: (left) False-colour map of height differences for pyramid between reconstructed and ideal heights; (right) horizontal cross-section through difference map.

Applying the process to the digital hemi-ellipsoid gives further insights. Fig. 5.36 (left) shows the gradients calculated from normals per the code in Fig. 5.29. They follow the tangent to the curved surface, with positive slopes to left (P) and below (Q), passing through zero at the top of the surface (centre), and negative slopes to right (P) and above (Q). Comparing the two alternative reconstructions of the surface (first P horizontal or first Q vertical) shows that the maximum height at the centre is very similar in both cases: 248.124 and 248.129 respectively. A false-colour map of the difference between the two reconstructions (Fig. 5.36 right), again coded in units of 0.01% of height, shows that the interior errors are less than 0.002% but that there are significant errors around the periphery where the gradients become large.

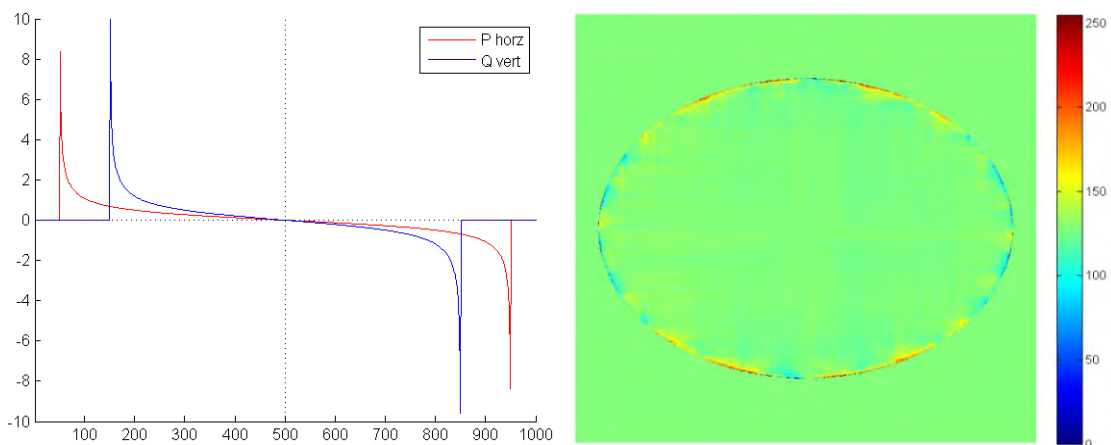


Figure 5.36: (left) Gradients of P on horizontal midline and Q on vertical midline; (right) False-colour map of differences between height reconstructions from P and Q.

The height at the centre of the averaged two reconstructions (P first and Q first) is 248.126, compared with 250.0 for the original. This loss of height can be attributed to errors in the gradients around the edge, resulting in the overall summation being too small. For analysis, the whole surface is therefore raised by 1.874 pixels, so that its height at the centre is exactly 250.0. Plotting the differences between the adjusted and ideal heights (Fig. 5.37) shows that the reconstruction is very close in the central region, perhaps +0.1 pixel higher (positive difference), whereas around the periphery it is up to 1.5 pixels lower (negative difference).

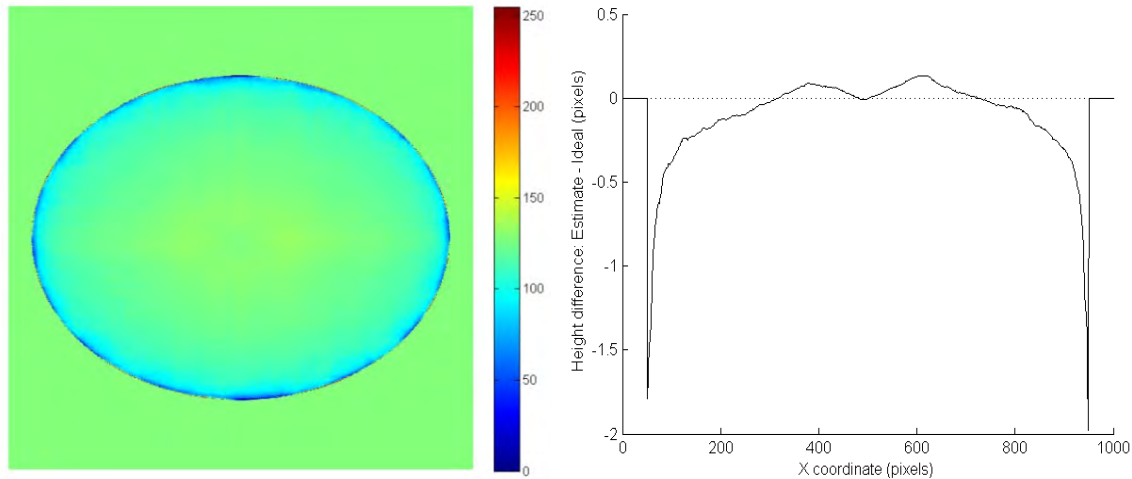


Figure 5.37: (left) False-colour map of height differences for hemi-ellipsoid between reconstructed and ideal heights; (right) horizontal cross-section through difference map.

Scatter-plotting the differences in height shows their dependence on angle (Fig. 5.38). Five thousand points were taken at random inside the outline and the difference values plotted against the angle of the normal vector with the Z axis, calculated as $\text{acosd}(N_z)$. The results show a systematic relationship, with the errors increasing to ~ 0.1 pixel at $\sim 5^\circ$ then decreasing steadily to ~ -0.5 pixel at $\sim 50^\circ$, then decreasing more rapidly for larger angles. This suggests that the gradients themselves become less reliable for angles $> 60^\circ$ and that this is a fundamental constraint on the accuracy of height reconstruction achievable from photometric stereo.

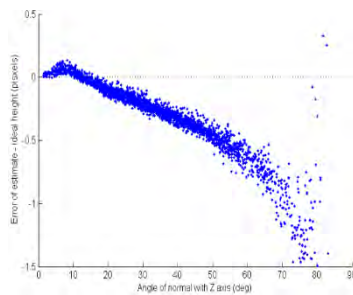


Figure 5.38: Scatter plot of height difference

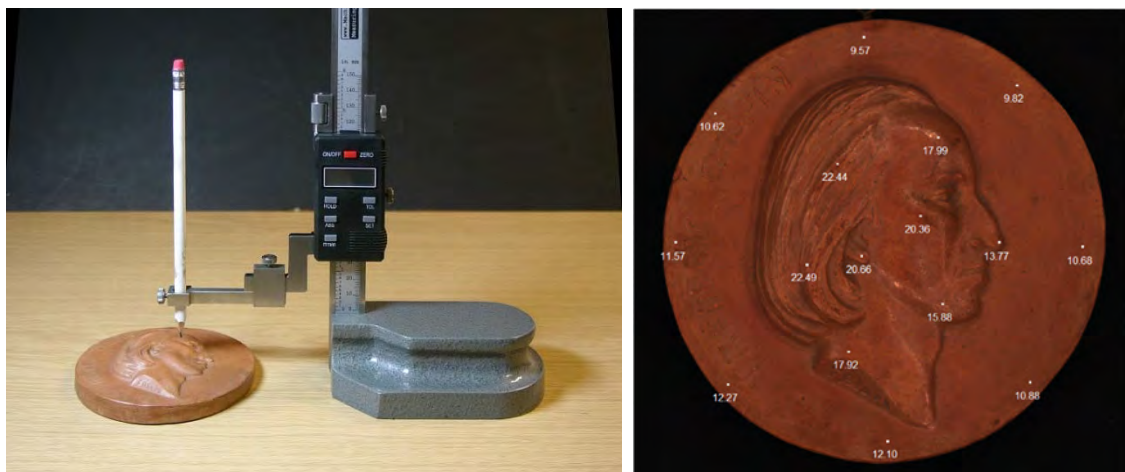


Figure 5.39: (left) Using a height measuring gauge; (right) Heights of selected points in mm, superimposed on an image of the Chopin terracotta taken under all lights in Tier 3.

5.5 Height Reconstruction with Reference Points

For physical test objects there is no idealised generating surface that can be used for comparison of the accuracy of the reconstruction. Instead measurements are needed. The simplest method is to use a manual height gauge at selected points on the object surface, with a nominal precision of 0.01 mm (10 μm). Fig. 5.39 shows such an instrument in use on the Chopin terracotta roundel, and at right the measured values in mm at 16 points over surface.

The normals calculated by the bounded regression method and the resulting gradients for the Chopin terracotta are shown in Fig. 5.40. The P gradient ($\partial I / \partial x$) and Q gradient ($\partial I / \partial y$) are encoded so that zero is mid-grey, white is maximum positive and black is maximum negative. Positive values of P are on the left, where the surface height is increasing, and negative on the right where it is falling. Similarly Q gradients are positive at the top and negative at the bottom (note the Matlab convention that the image origin is at the top left corner, so the Y axis is positive downwards). The values of P and Q along horizontal and vertical cross-sections, are plotted in Fig. 5.41. The image dimensions, cropped from the full Nikon D200 frame, are 1836(W) \times 1728(H) pixels.

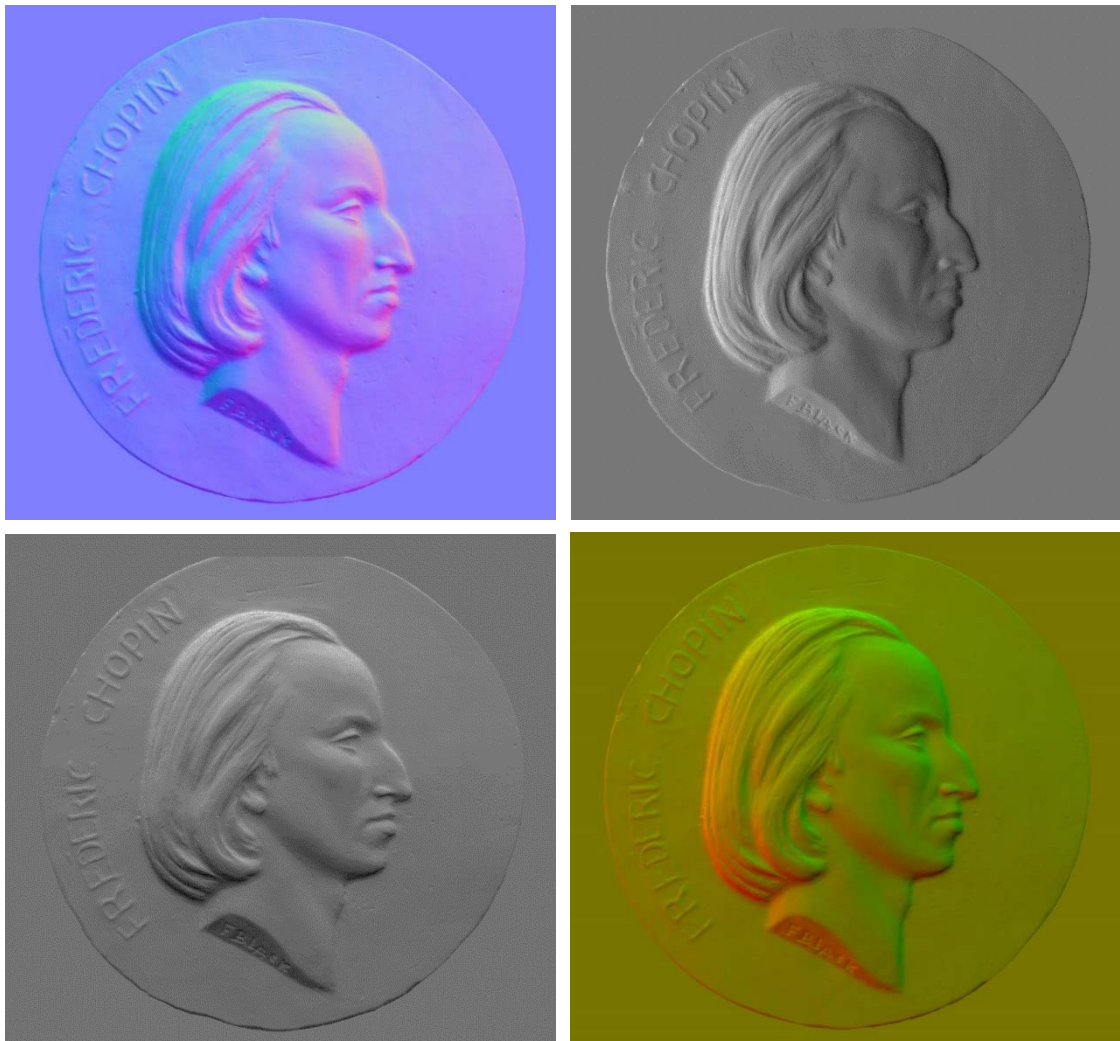


Figure 5.40: (top left) Photometric normal vectors in false colour (N_x red, N_y green, N_z blue);
 (top right) P gradient = $\partial I / \partial x$; (bottom left) Q gradient = $\partial I / \partial y$;
 (bottom right) False colour composite of gradients (P red, Q green).

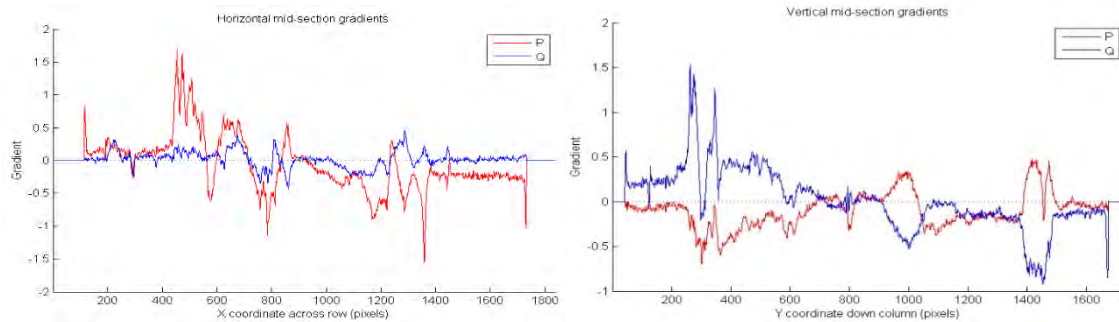


Figure 5.41: (left) Horizontal and (right) vertical midline cross-sections of P and Q gradients.

If the gradients were known at every point on a surface, one might suppose that it would be a simple matter to integrate them to reconstruct the height at every point. Unfortunately it is not so simple because errors from various sources accumulate over the surface to produce significant discrepancies. This is illustrated in one dimension in Fig. 5.42, in which the P gradient across the horizontal midline (Fig. 5.41 left) is summed both left-to-right and right-to-left. Starting at zero (the assumed height of the baseplane) gradients are added successively, pixel by pixel, in each direction. Although the resulting profiles are identical, their heights are very different, because of uncertainty about the large gradients at the edges of the object.

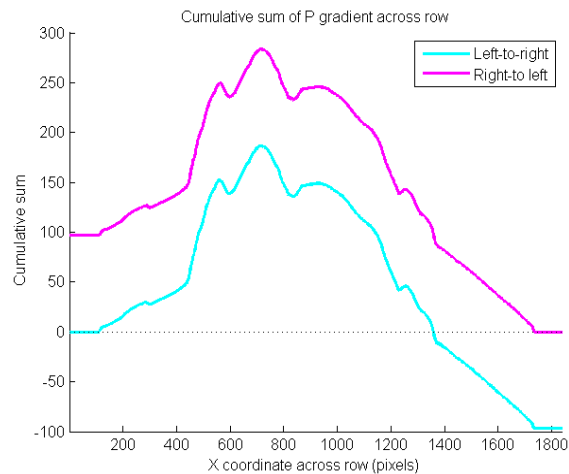


Figure 5.42: Cumulative sums of P gradients in two directions across the horizontal midline.

Performing the summation on the P gradients across every row and the Q gradients down every column, and raising each above the baseplane (i.e. minimum value of zero) gives height maps that are recognisable as Chopin (Fig. 5.43), but are heavily striated because of the variable height offset.

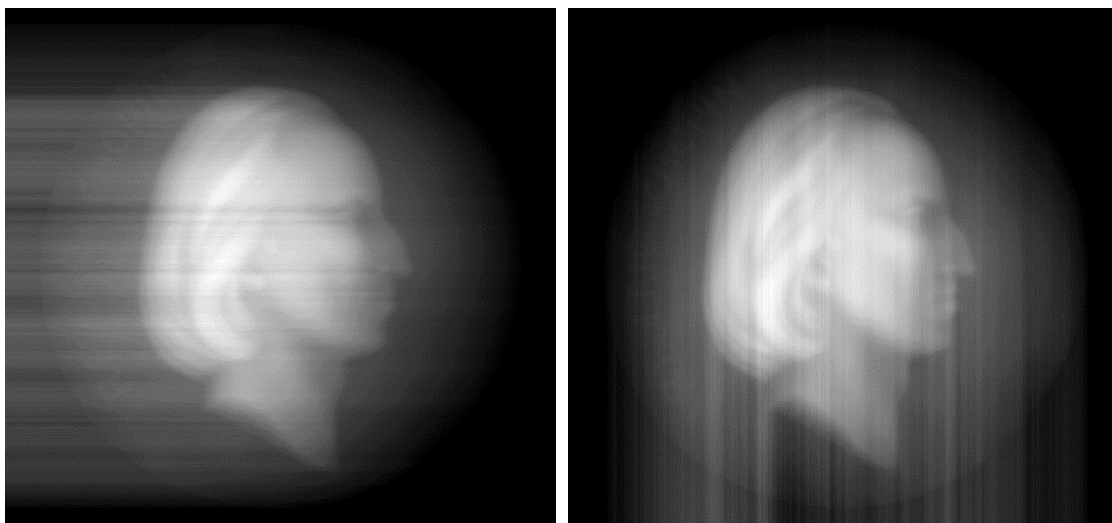


Fig. 5.43: (left) Horizontal and (right) vertical summation of gradients per line and column.

Applying instead the integration-from-centre procedure, as described in the previous section, gives heights that are better behaved in the centre, but still exhibit the striations in both directions (Fig. 5.44). The highest point of each height distribution is set equal, then the two distributions are averaged and raised so that the overall minimum value is zero. The resulting distribution (Fig. 5.45 left) is smoother but still shows striations.

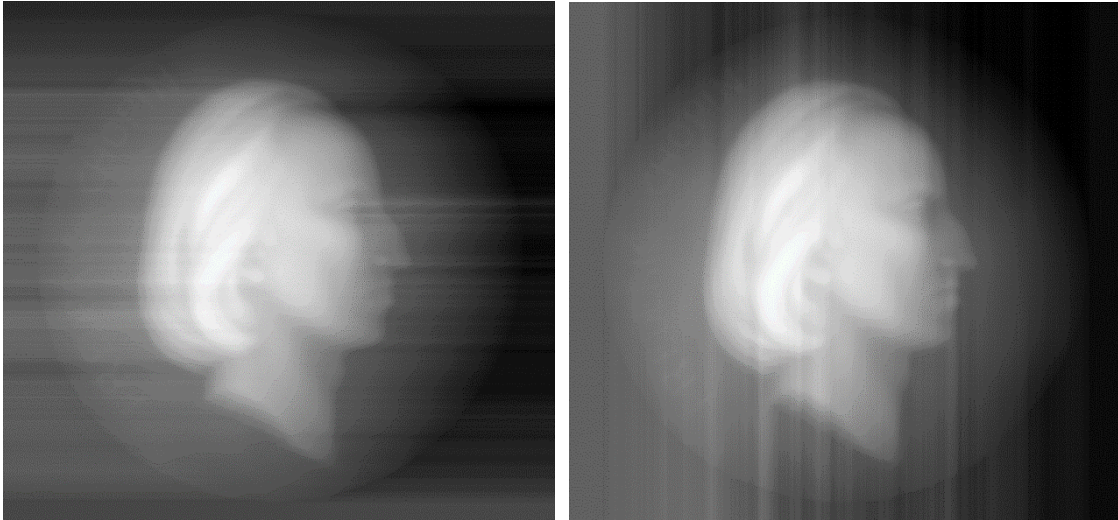


Fig. 5.44: Horizontal and (right) vertical summation of gradients from centre.

Comparing the reconstructed height with the measured values at each of 16 points (Fig. 5.45 right) shows that the match is better near the centre and worse around the periphery. In the bar chart the first eight bars correspond to the compass points around the edge, and the second eight to the central points. The height values have been converted from pixels to mm by the scaling factor of 13.92 pixels/mm calculated from analysis of the centroids of targets on the baseboard (note that this is less than the spatial resolution of the image of the top surface of the object, because it is closer to the camera).

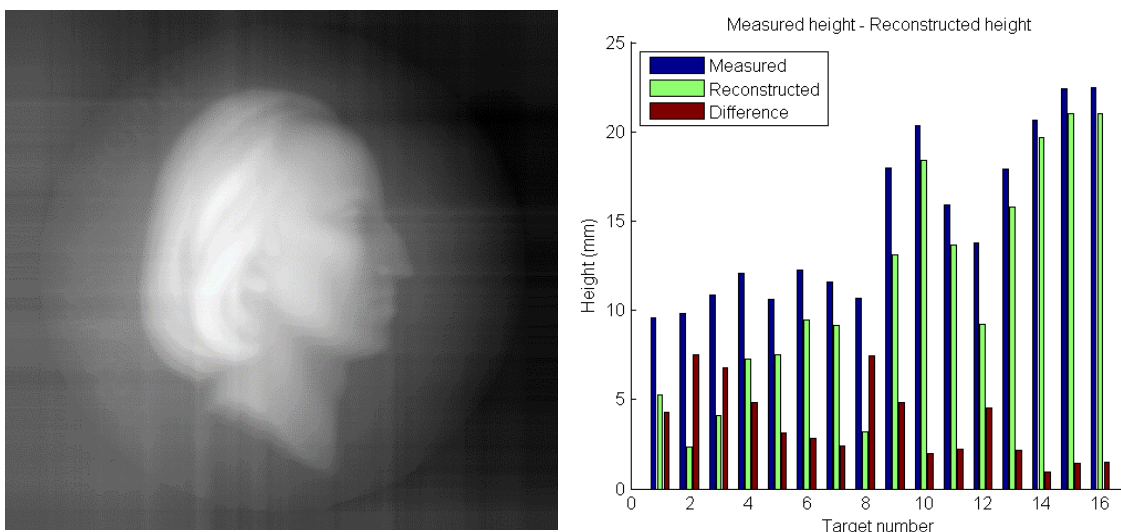


Fig. 5.45: (left) Averaged horizontal and vertical summation of gradients; (right) differences in height between reconstructed surface and measured values at 16 points.

The difference values (brown bars in the chart in Fig. 5.45) range from 0.96 to 7.48 mm, which is a disappointing result. The reason for the large errors can be seen in the reconstructed 3D

surface (Fig. 5.46), viewed from the south-east at zero elevation, i.e. parallel to the X-Y plane. The clay disc on which Chopin's head is moulded is tilted on a diagonal axis from upper left to lower right. Thus the disc is significantly higher at the south-west (bottom left) edge than at the north-east (top right) edge. The point measurements (Fig. 5.39) are 12.27 and 9.82 mm respectively, a difference of 2.45 mm, but the corresponding reconstructed heights are 6.98 and 2.71 mm, a difference of 4.27 mm. Thus the summation of gradients in the reconstruction has stretched the scale of the relief, while compressing the height of the disc.

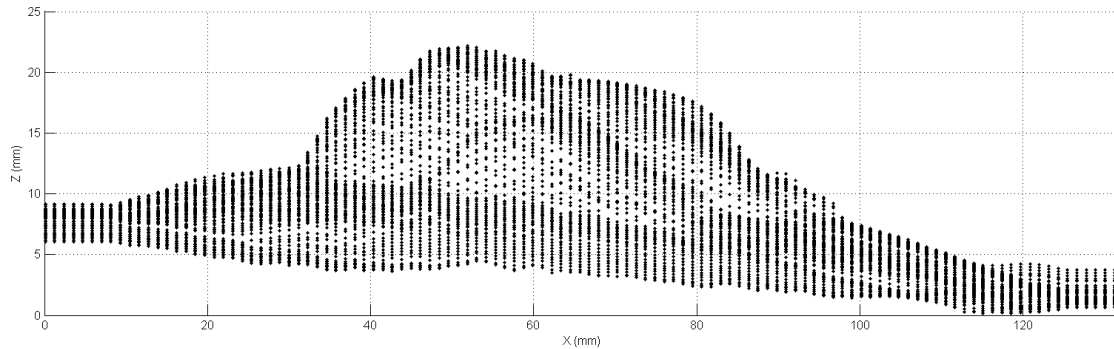


Fig. 5.46: Reconstructed 3D surface viewed from south-east parallel to X-Y plane.

5.6 Height Reconstruction by Fourier Transform

An alternative approach to integrating the gradients to reconstruct the height was introduced by Frankot & Chellappa, using the Fourier transform to regularise (i.e. enforce integrability of) the gradients in the frequency domain (see Sec. 2.3.3). This is neatly implemented in Matlab by a few lines of code, as shown in Fig. 5.47, taken from the library developed by Kovesi (2004).

```
dzdx = P; dzdy = Q; % gradients from normals
[rows,cols] = size(dzdx);
[wx,wy] = meshgrid(([1:cols]-(fix(cols/2)+1))/(cols-mod(cols,2)),...
    ([1:rows]-(fix(rows/2)+1))/(rows-mod(rows,2)));
wx = ifftshift(wx); wy = ifftshift(wy); % quadrant shift
DZDX = fft2(dzdx); DZDY = fft2(dzdy); % Fourier transform gradients
FC = (-j*wx.*DZDX-j*wy.*DZDY)./(wx.^2+wy.^2+eps); % F&C Equation 21
IFC = real(ifft2(FC)); % inverse transform
```

Figure 5.47: Fragment of Matlab code to enforce integrability in height reconstruction from gradients.

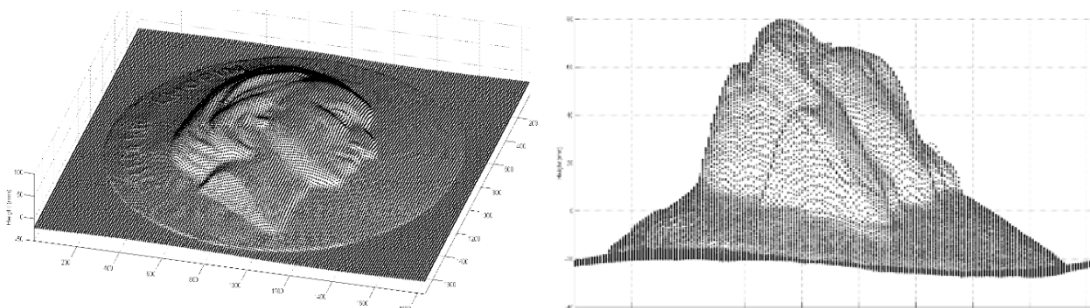


Figure 5.48: (left) Oblique view and (right) elevation of Chopin surface reconstructed by the basic Frankot & Chellappa integration of gradients in the frequency domain (all dimensions in mm).

Applying this technique to the Chopin gradients (Fig. 5.40 right) yields a 3D surface that is continuous and is recognisably Chopin, but is distorted over the whole area with the height greatly amplified. Fig. 5.48 (left) shows an oblique view, which looks very plausible, but when

the same structure is viewed in elevation (Fig. 5.48 right) it is seen that the height range is from -27.6 to +79.9, an overall maximum height of 107.5 mm, compared with the true maximum height above the baseplane of 22.5 mm. Also there is a false undulation of the base with a period of approximately one cycle over the whole width.

Adjusting the offset and rescaling the height range of the reconstructed surface to the height range of the measured points gives the error distribution of Fig. 5.49. The error values range from -2.72 to +2.34, mm with a mean over 16 points of -0.14, a mean of absolutes of 1.26 and a stdev of 1.52. The magnitudes of these errors are much lower than those of the simple summation (Fig. 5.45 right). The problem is that although the gradients give a good representation of the spatial frequencies in the surface, right up to the Nyquist frequency, they are not accurate

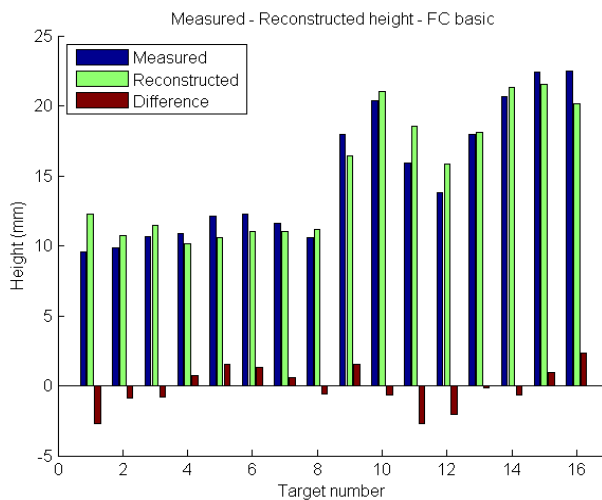


Figure 5.49: Errors between measured and reconstructed heights (mm).

for very low frequencies of a few cycles over the full object diameter. Such frequencies are represented in the Fourier plane by only a few sample points close to the (shifted) origin. Errors in these frequencies can result in 'curl' or 'heave' in the baseplane, even though the superimposed higher spatial frequencies may be accurate. Following a suggestion by Prof Shoji Tominaga, Chiba University, the solution is to replace the inaccurate low frequencies of the photometric normals by the more accurate low frequencies of a surface constructed from a few known heights (Tominaga *et al*, 2010). This can be conveniently achieved from the values measured by the digital height gauge (Fig. 5.39) by first interpolating them to produce a smooth 'hump' and then transforming into the frequency domain by an FFT.

```

XT = 1:imw;  XB = 1:imw;
XL = ones(imh-2,1,'double');
XR = imw*ones(imh-2,1,'double');
XE = [XT';XL';XR';XB'];
YT = ones(imw,1,'double');
YL = 2:imh-1;  YR = 2:imh-1;
YB = imh*ones(imw,1,'double');
YE = [YT';YL';YR';YB'];
mh = mean(XYZ(1:8,3)); % mean height of disc
ZE = mh*ones(2*(imw+imh)-4,1,'double'); % constant around perimeter
XP = [XE;XYZ(:,1)]; % column vector of X values
YP = [YE;XYZ(:,2)]; % column vector of Y values
ZP = [ZE;XYZ(:,3)]; % column vector of Z values
[X,Y] = meshgrid(1:imw,1:imh); % make image grid
ZS = griddata(XP,YP,ZP,X,Y,'cubic'); % construct smooth surface

```

Figure 5.50: Fragment of Matlab code to interpolate measured height values to a smooth surface.

In the code in Fig. 5.50 the function `griddata` interpolates a smooth surface constrained to pass through all given points. In order to make the surface coterminous with the gradient image,

points around the perimeter are included with the set of measured points, all of constant height equal to the mean height of the circular disc on which the figure is moulded. The resulting shape (Fig. 5.51) is a hump on a plane, extending to the edge of the image area and in 1-to-1 correspondence with the image pixels.

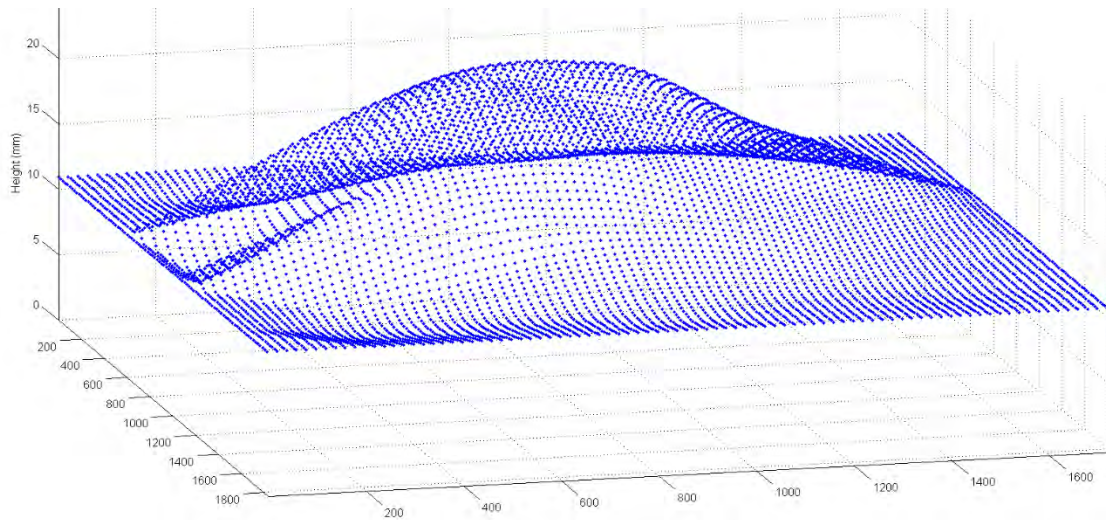


Figure 5.51: Smooth surface of hump produced by interpolation of measured points.

Taking the two-dimensional Fourier transform of the surface produces the $\log(\text{power})$ distribution shown in Fig. 5.52 (left). Almost all of the power is concentrated in a small region at the centre of the shifted frequency plane, apart from a few slanted lines across the plane caused by creases in the surface. Adding the power in concentric rings outward from the centre shows that 0.63 of the total power is at a radius of less than 1.5 pixels and 0.75 is at a radius of less than 4 pixels (Fig. 5.52 right).

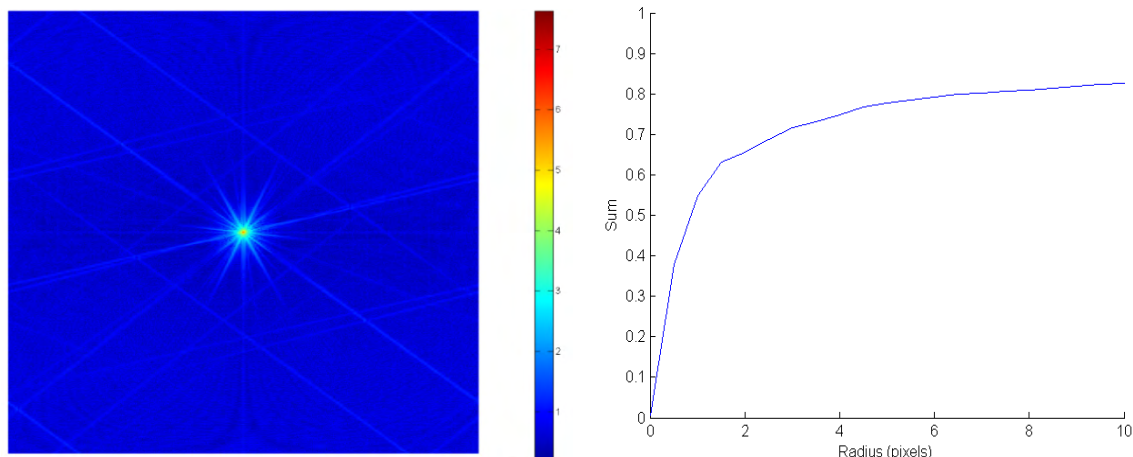


Figure 5.52: (left) $\log(\text{power})$ distribution of spatial frequencies of hump; (right) cumulative sum of power by radius from centre of spatial frequency plane.

The region of low frequencies is extracted by a circular mask, i.e. only pixels within a given radius of the centre point of the shifted frequency plane are selected. The effect of reconstruction by an inverse Fourier transform, using only the selected low frequencies, is shown in Fig. 5.53. On the left the nine pixels with radius less than or equal to 1.5 produce a very smooth flattened hump. On the right the 49 pixels with radius less than or equal to 4 produce a better defined shape, beginning to resemble the interpolated surface of Fig. 5.51.

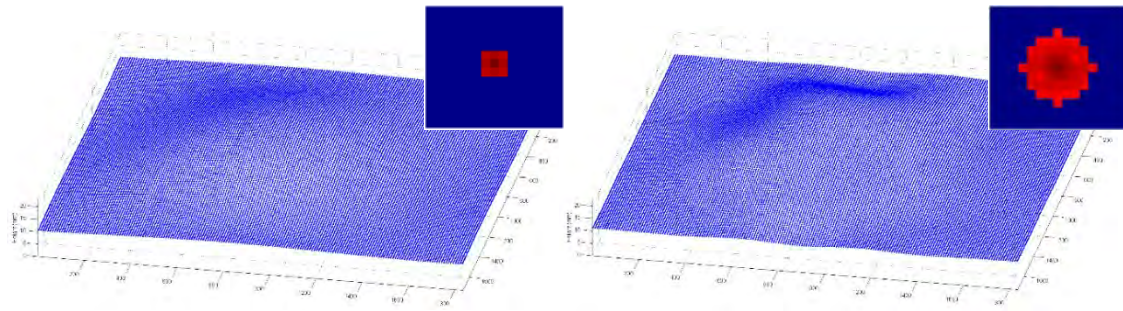


Figure 5.53: Reconstruction of hump from filtered low frequencies: (left) limit of 1.5 pixels; (right) limit of 4 pixels. Inset: region at centre of Fourier frequency plane showing pixels selected.

The high spatial frequencies of the surface are provided by the photometric gradients. Plotting the $\log(\text{power})$ distribution in false colour in the shifted frequency plane (Fig. 5.54 left) shows a concentration of power at all phases, extending out to approximately one third of diameter of the plane. Scatter-plotting all the points along the horizontal semi-axis from the centre to the right edge of the frequency plane (Fig. 5.54 right) shows that the distribution follows the typical $1/f$ form, with the horizontal scale expressed in cycles/pixel, normalised to the Nyquist frequency of 0.5.

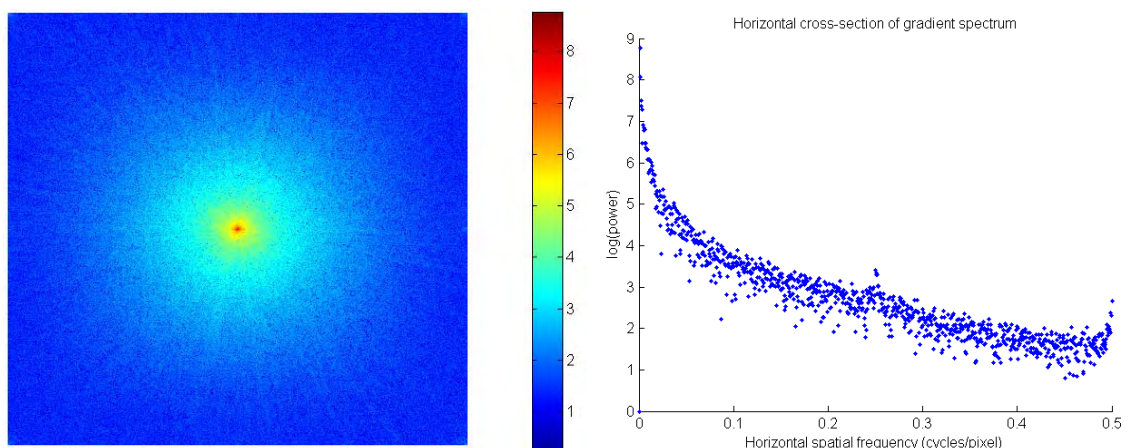


Figure 5.54: Log(power) spectrum of Chopin gradients in Fourier plane: (left) full spatial frequency plane with origin at centre; (right) horizontal cross-section of log(power) from centre along X axis.

The low spatial frequencies of the gradients from the Frankot-Chellappa integration are replaced by the corresponding frequencies from the hump. Rather than an abrupt change at a given threshold radius, they are blended over a radial distance in the range 1.5 to 4.0 pixels by a linear interpolation (lerp) function, as shown in Fig. 5.55. The power of the high frequencies from the gradients is scaled by the ratio of the low/high power in the region affected, in order to maintain the correct overall power distribution (code in Fig. 5.56).

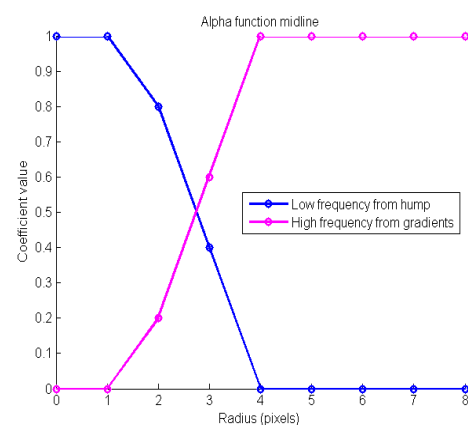


Figure 5.55: Cross-over of contributions from low and high frequencies in blend.

```

cx = 1+floor(imw/2); cy = 1+floor(imh/2); % centre point of plane
[X,Y] = meshgrid(1:imw,1:imh);           % make image grid
R = sqrt((X-cx).^2+(Y-cy).^2);           % unit radius from centre
rmin = 1.5; rmax = 4;                    % radial frequency limits
Alpha = 1-(R-rmin)/(rmax-rmin);          % linear weighting plane
T = R<rmin; Alpha(T) = 1;                % 1 for LF
T = R>rmax; Alpha(T) = 0;                % 0 for HF
TF = R<rmax;                             % select pixels affected
SumLF = sum(abs(LF(TF)));                 % total power in LF region
SumHF = sum(abs(HF(TF)));                 % total power in HF region
CF = Alpha.*LF+(1-Alpha).*HF*SumLF/SumHF;% lerp LF and HF
CFS = ifftshift(CF);                     % shift centre to corners
ZC = real(ifft2(CFS));                    % reconstruct surface

```

Figure 5.56: Fragment of Matlab code to reconstruct surface from combined low and high frequencies.

The reconstructed surface from the combined low and high spatial frequencies is shown in Fig. 5.57. It is scaled so that the mean height of the 16 reference points is equal to the mean of the 16 measured values (Fig. 5.39 right), in this case a factor of 0.937. Comparing the result with the reconstruction from the high frequencies only (Fig. 5.48), it is seen that the base is more planar with less low-frequency distortion. Also it is close to the ‘true’ height above the baseplane, rather than being distributed about a zero mean.

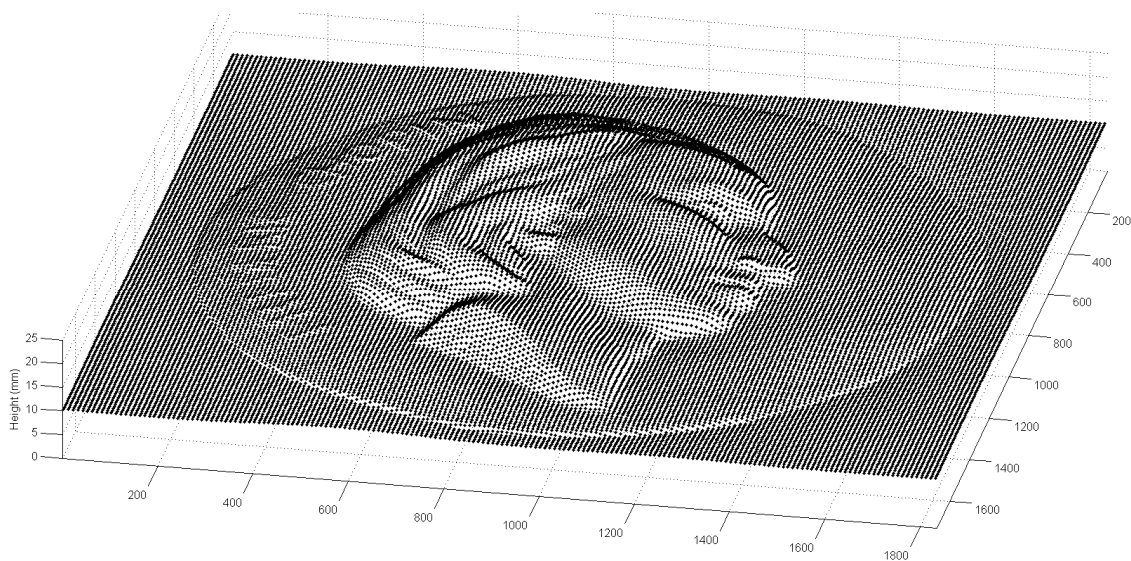


Figure 5.57: Reconstructed surface, after combining the low frequencies of the hump with the high frequencies of the photometric gradients.

Plotting the same surface in elevation, viewed from the ‘south’, shows a detailed profile set in a horizontal plane of height 10 mm (Fig. 5.58 left). The errors are well distributed (Fig. 5.58 right), with values ranging from -1.31 to +1.92 mm with zero mean, mean-of-absolutes 0.61 and stdev 0.83. These errors are lower than for uncorrected integration of the gradients (Fig. 5.49).

Better results still can be obtained by noting that the top surface of the circular disc that forms the base of the Chopin terracotta lies approximately on a slanted plane. By subtracting the height of this plane from every point on the surface, the moulded figure is transformed to lie on the XY plane (i.e. $Z = 0$), where it has minimum energy and hence the least amount of aliasing in the Fourier frequency domain.

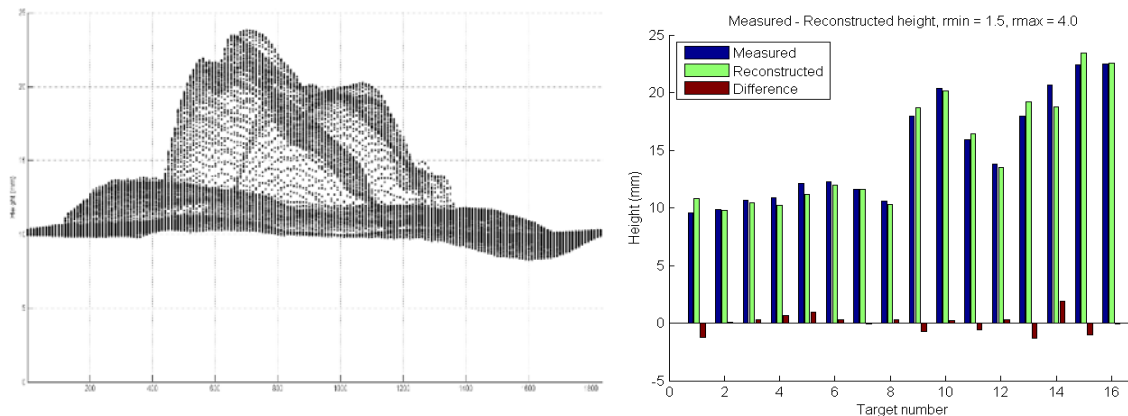


Figure 5.58: (left) Elevation of reconstructed surface, viewed from south; (right) errors between the measured and reconstructed heights (mm).

```
[Pcoeff,Pscore] = princomp(XYZ(1:8,:)); % principal components
Pnorm = Pcoeff(:,3); % normal to plane
Pmean = mean(XYZ(1:8,:)); % mean of points on disc
a = Pnorm(1); % plane: ax+by+cz = d
b = Pnorm(2); % so z = d-(ax+by)/c
c = Pnorm(3);
d = Pmean*Pnorm;
[X,Y] = meshgrid(1:imw,1:imh); % make image grid
Z = d-(a*X+b*Y)/c; % height at every point
```

Figure 5.59: Fragment of Matlab code to fit a plane to the disc around the base.

The code of Fig. 5.59 finds the three principal components (vectors in 3D space) of the eight points around the periphery of the basal disc on the Chopin terracotta. The first two components lie in the plane of best fit and the third component is the normal to the plane. The height of this plane can then be calculated for all points around the perimeter of the image for making the basic hump (Fig. 5.60).

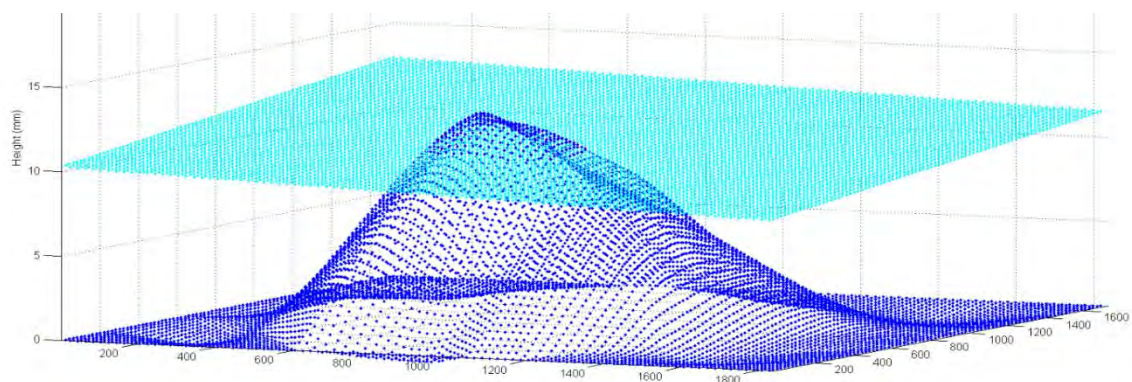


Figure 5.60: Slanted plane of best fit to disc (cyan) and hump constructed from residual (blue).

After following the same process as before — i.e. taking the Fourier transform of the basic hump, combining low frequency components of hump with high frequencies of the gradients, and transforming back to the spatial domain — and then adding back in the slanting base plane, the reconstruction of Fig. 5.61 is obtained. It is again scaled so that the mean height of the 16 reference points is equal to the mean of the 16 measured values (Fig. 5.39 right), in this case a factor of 0.952. The outline mask has also been applied to clear the surrounding region to zero.

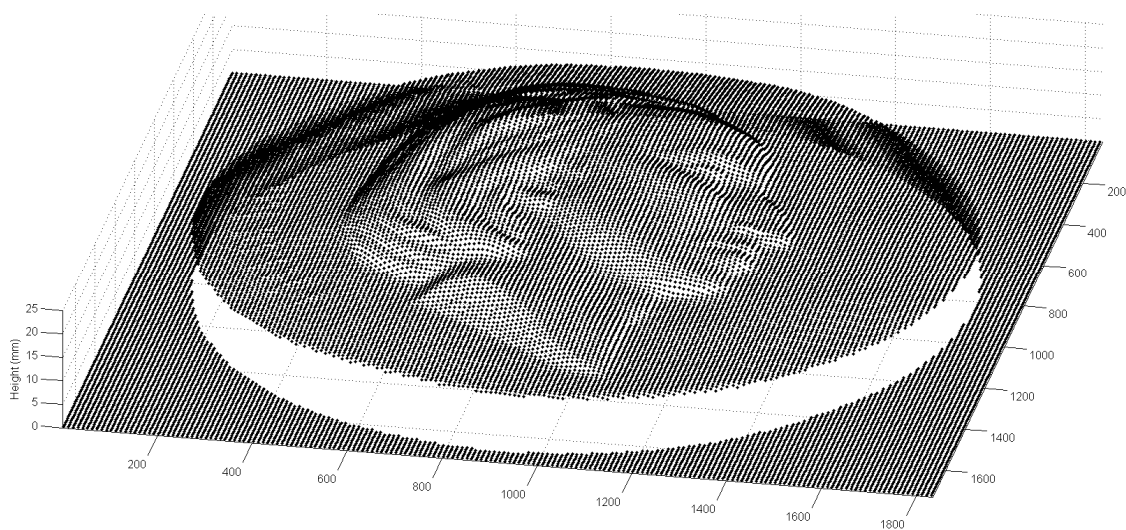


Figure 5.61: Reconstructed surface, after adding slanted plane and cutting around outline.

The height map can be scaled and visualised as a monochrome image (Fig. 5.62 left) and by comparison with Figs. 5.43–5.45 is seen to be free of striations and visible defects. The errors are well distributed (Fig. 5.62 right), with values ranging from -0.95 to $+0.77$ mm with zero mean, mean of absolutes of 0.32 and stdev of 0.42 . These errors are the lowest of all the methods tried.

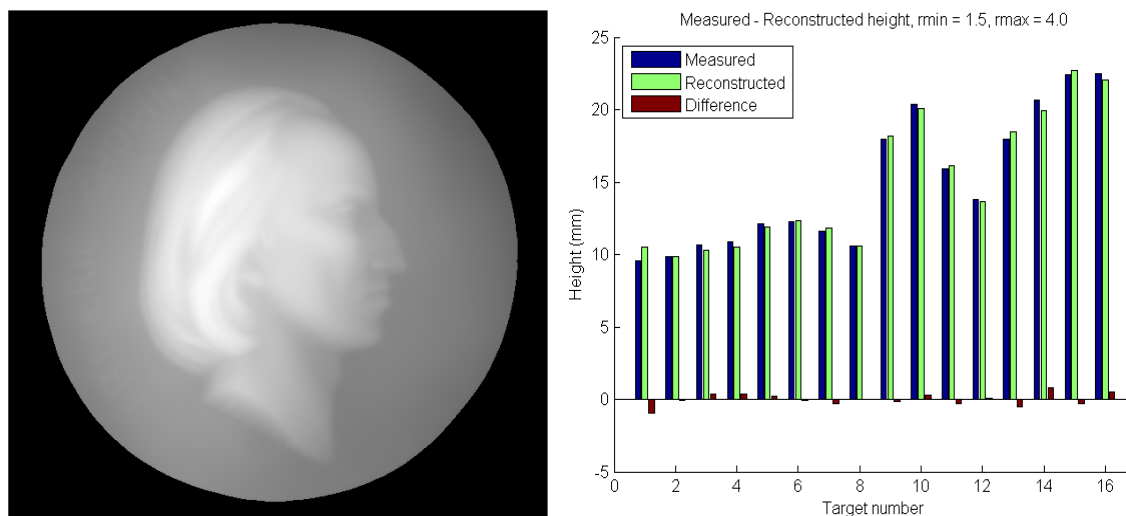


Figure 5.62: (left) Height map of reconstructed surface, scaled to fill the tonal range from black to white; (right) errors between measured and reconstructed heights (mm).

5.7 Height Reconstruction with Surface from a 3D Scanner

By using a 3D laser scanner an accurate digital terrain map of the whole object surface can be produced. Thus instead of manually measured heights at just a few points, it is possible to know the height to an accuracy of ~ 25 μm at every position on a grid over the surface at intervals of 0.1 mm. The problem is then to find the correspondence of features between the photometric and geometric surfaces to enable the one to be mapped to the other.

A previous study of the Arius 3D colour laser scanner characterised it as an imaging device (MacDonald, 2010b). The Chopin terracotta was digitised by the Arius scanner and the point

clouds from multiple scans were registered by the Pointstream software and decolourised (Fig. 5.63). The combined point cloud was written out as an ASCII file (.asc format) containing 1.45 million points, each represented by one line of 23 fields, including X,Y,Z coordinates, R,G,B colour values, and N_x, N_y, N_z normal components.



Figure 5.63: (left) Scanning Chopin on Arius 3D laser scanner; (right) point cloud without colour.

A two-dimensional image was constructed by ‘pancaking’ the point cloud, i.e. projecting all points of the top surface down onto the $Z=0$ plane and filling gaps, where necessary, by interpolating the nearest neighbours. The same technique was used to make images of the normals and a height map (Fig. 5.64). The diameter of the roundel in the generated images was 1135 pixels, corresponding to the physical diameter of 113.5 mm at 10.0 pixels/mm.

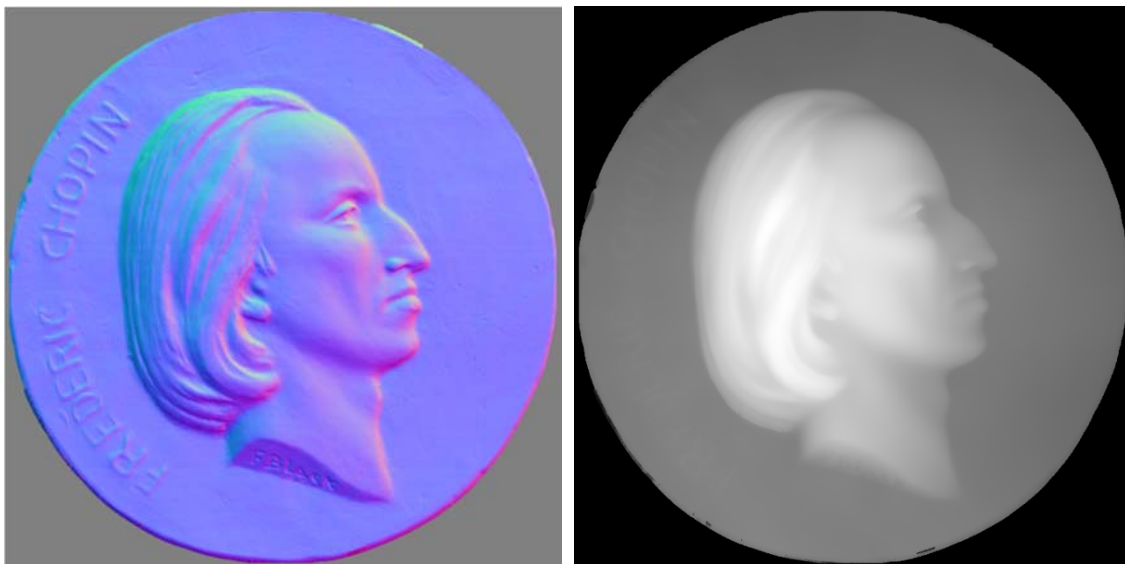


Figure 5.64: (left) Two-dimensional images of normal vectors and (right) height map, generated by flattening the Arius 3D point cloud.

The photographic image from the camera in the dome and the pseudo-image generated by flattening the 3D scanner point cloud are in general of different sizes and at different orientations. An efficient way is needed to determine both scale factor and rotation angle to bring them into alignment. Although this could be done on the 2D images by a search-and-correlate algorithm such as SIFT, the method preferred in this study uses the outline. Even for

an almost perfectly circular shape such as the disc of the Chopin terracotta, the slight irregularity is sufficient to provide a characteristic signature that facilitates the transformation. The technique for generating an outline was described in Section 4.1.7. Applying this to both versions of the Chopin roundel gives the results of Fig. 5.65. The detected outline is shown in yellow, the centroid by a white patch and the maximum radius by a blue line. It is surprising that slight differences in the centroid position have caused the line of maximum radius to be almost diametrically opposite in the two cases, because the terracotta disc is slightly elongated in the northeast-southwest direction.

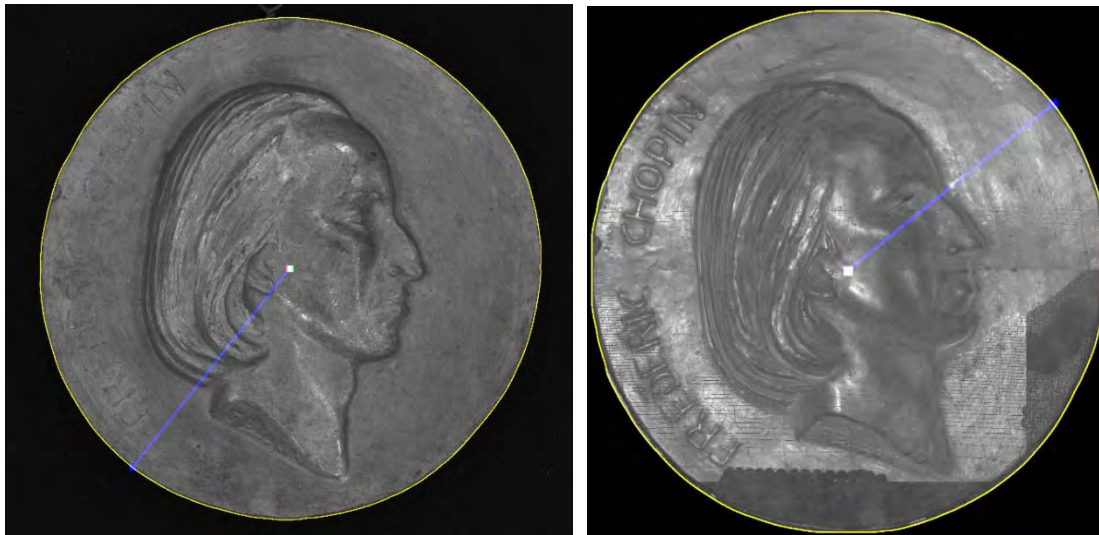


Figure 5.65: Outline, centroid and line of maximum radius for monochrome composite of: (left) camera image; (right) image generated by flattening Arius 3D scanner point cloud.

The computed measures for the two outlines are given in Table 5.1. Normalising the radius of each point to the mean radius and plotting against angle gives the graphs of Fig. 5.66 (left) showing $r(\theta)$ vs θ . Although the radius varies by less than 3% a similar pattern can be seen for both outlines. The number of points around the outline is a measure of the circumference of the shape, and hence the ratio of their lengths gives the relative linear scale factor, in this case $4593/3245 = 1.415$.

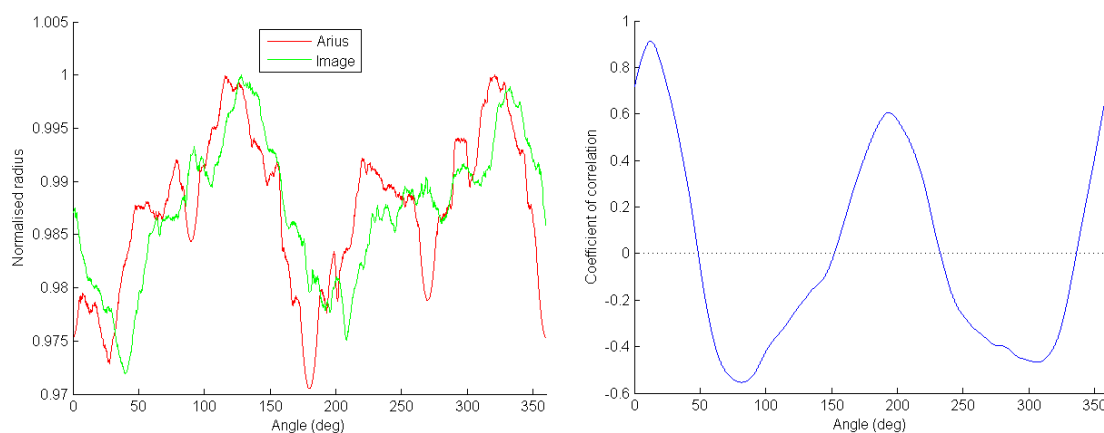


Figure 5.66: (left) Normalised radius vs angle; (right) Outline correlation coefficient vs angle.

Table 5.1: Measures of outlines from Arius and Nikon images

Measure	Arius outline	Image outline
Scale (pixels/mm)	10	14.26
Perimeter length (pixels mm)	3245 324.5	4593 322.1
Mean radius (pixels mm)	570.1 57.1	811.4 56.9
Maximum radius (pixels mm)	577.6 57.8	821.7 57.6
Angle of max radius (degree)	321.1	128.2
Area by polygon (mm ²)	10218.5	10170.3
Area by outline (mm ²)	10257.4	10197.6

By scaling the two outlines to equal length, then interpolating each to units of 0.01 degree (i.e. a vector of 36,000 values) and sliding one vector along the other, the position of best fit can be found. To avoid wraparound at 360° the reference outline is duplicated so that its vector of 72,000 points represents two full revolutions 0-720°.

```

an = 360*100; % expand range by factor of 100
A100 = 1:an; % angle scale in 1/100th degree
AR = (A100-1)/100; % angle range 0-359.99 degree
ARI = interp1(ASF,ARF,AR,'spline')'; % interpolate Arius outline
IRI = interp1(ISF,IRF,AR,'spline')'; % interpolate image outline
IRI2 = repmat(IRI,[2,1]); % double scale to range 0-720 deg
C = zeros(an,1,'double'); % allocate buffer for corr coeffs
for a = A100
    IRP = IRI2(A100+a-1); % extract section of image outline
    C(a) = corr(ARI,IRP); % correlate with Arius outline
end
[cmx cmi] = max(C); % find correlation peak
cang = cmi/100; % rotation angle in 1/100th deg

```

Figure 5.67: Fragment of Matlab code to determine best fit between outlines by cross-correlation.

Determination of the goodness-of-fit is achieved by cross-correlation between the two 36,000-element vectors, one fixed (the reference outline from the Arius) and the other extracted at successive positions from the duplicated image outline (code in Fig. 5.67). The maximum value of the coefficient of correlation indicates the angle of best fit (Fig. 5.66 right), with the peak at 12.56°. Note that the second peak at 193.44° is almost diametrically opposite, indicating a slightly elliptical outline shape.

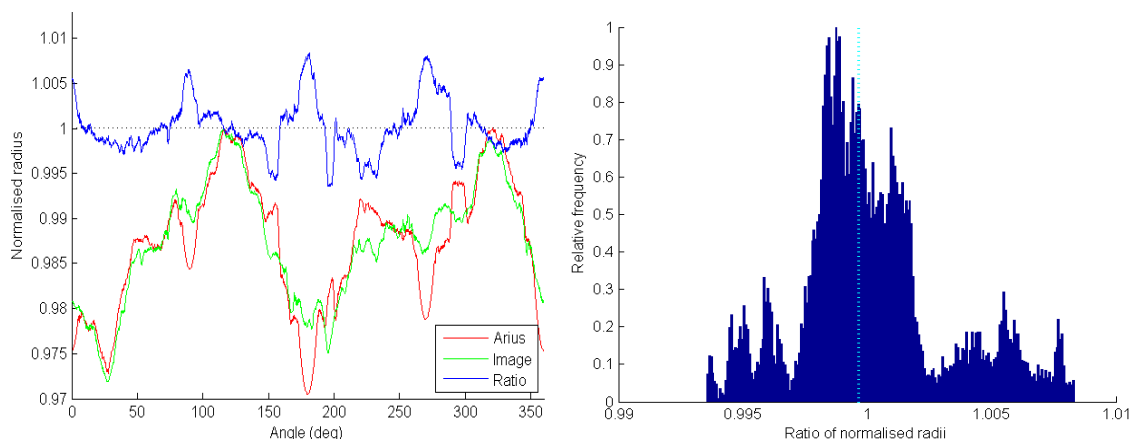


Figure 5.68: (left) Position of optimum fit; (right) Distribution of ratio of normalised outline radii.

Placing both outlines at the angle of optimum fit, the ratio of the normalised radii of the two outlines can be calculated at every angle (Fig. 5.68 left). The resulting distribution shows a narrow spread of values between 0.9935 (min) and 1.0083 (max), with a median value of 0.9997 (Fig. 5.68 right).

By the above method of outline correlation, both the scaling factor κ and rotation angle θ can be found to map one outline onto the other. In this case the scaling factor is 1.415 and the rotation angle is 12.56° (clockwise). This provides a means for every point $[x_r, y_r]$ in the reference images for albedo, normals and height, derived from the point cloud produced by the Arius 3D scanner, to be transformed into close registration with a corresponding point $[x_t, y_t]$ in the test images from the Nikon camera by a 2×2 matrix:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \mathbf{M} \begin{bmatrix} x_r \\ y_r \end{bmatrix}, \text{ where } \mathbf{M} = \kappa \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} = \begin{bmatrix} 1.3893 & 0.3095 \\ -0.3095 & 1.3893 \end{bmatrix} \quad (5-5)$$

The result is shown in Fig. 5.69, in which the red channel of the image contains the transformed Nz component of the Arius normals and the green channel contains the unmodified Nz component of the photometric image normals.

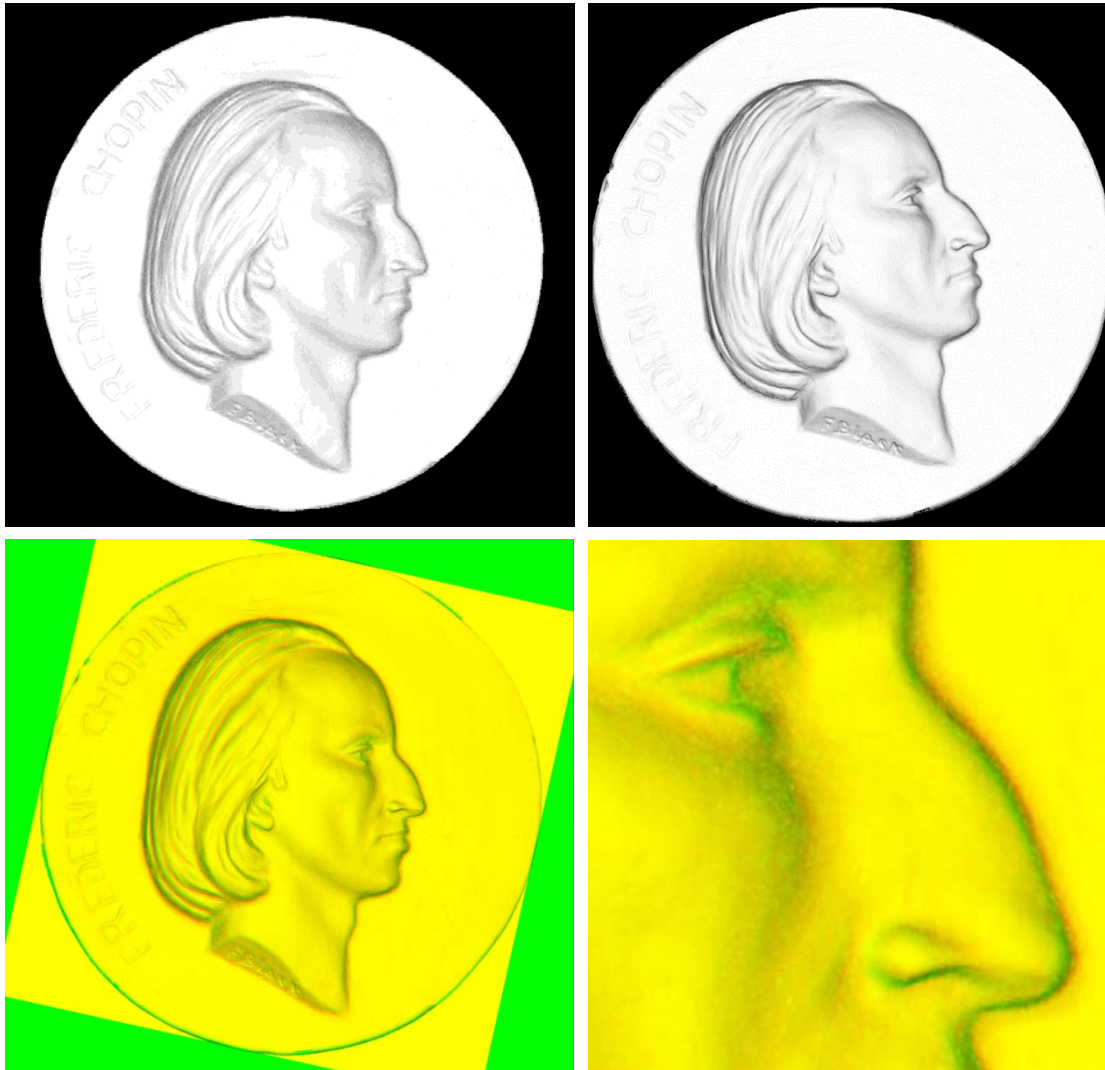


Figure 5.69: (top left) Nz component of photometric normals; (top right) Nz of Arius normals; (bottom left) Registered images from transformed Arius and original photometric; (right) detail.

The N_z component is preferred for correlation because it is well bounded in the range $[0,1]$ and gives good contrast at all angles, unlike N_x and N_y which are directional. It can be seen from Fig. 5.69 (lower left) that, after the transformation, the two images coincide closely, although when magnified some discrepancies are visible (right). In this step the photometric image is taken to be the reference image because it is at higher resolution and will determine the fine detail of the rendered surface; the Arius image is transformed globally to match it because it is at lower resolution and will be used to provide the height of the low frequencies. In the next step, however, the transformed Arius image is taken as the reference because the superior geometric accuracy of the scanner provides a better standard for X,Y,Z coordinate values than the projective geometry and lens distortion of the camera. The actual displacement between the photometric image and the transformed Arius image can be determined at any position by cross-correlation between surrounding regions. In effect a section of the test image A is taken as a template and correlated with a corresponding section of the reference image B , using the Matlab function `corr2` to calculate the correlation coefficient r :

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\left(\sum_m \sum_n (A_{mn} - \bar{A})^2 \right) \left(\sum_m \sum_n (B_{mn} - \bar{B})^2 \right)}} \quad (5-6)$$

This formula has the advantage of being independent of the brightness and contrast of the template A and image B , because it takes differences from their respective means \bar{A} and \bar{B} and normalises the intensity values of both. A small section of 51×51 pixels from the test image is moved over the reference image with successive displacements over ± 25 pixels in both axes. At each of the $51^2=2601$ positions the correlation coefficient is calculated. The resulting array of 51×51 values in the range $[-1,+1]$ can be visualised as a surface on which the maximum value represents the position of best fit (Fig. 5.70 left). Because the computation is too time-consuming to perform at every pixel in the image, it is restricted to points at intervals of 50 pixels in both X and Y (Fig. 5.70 right).

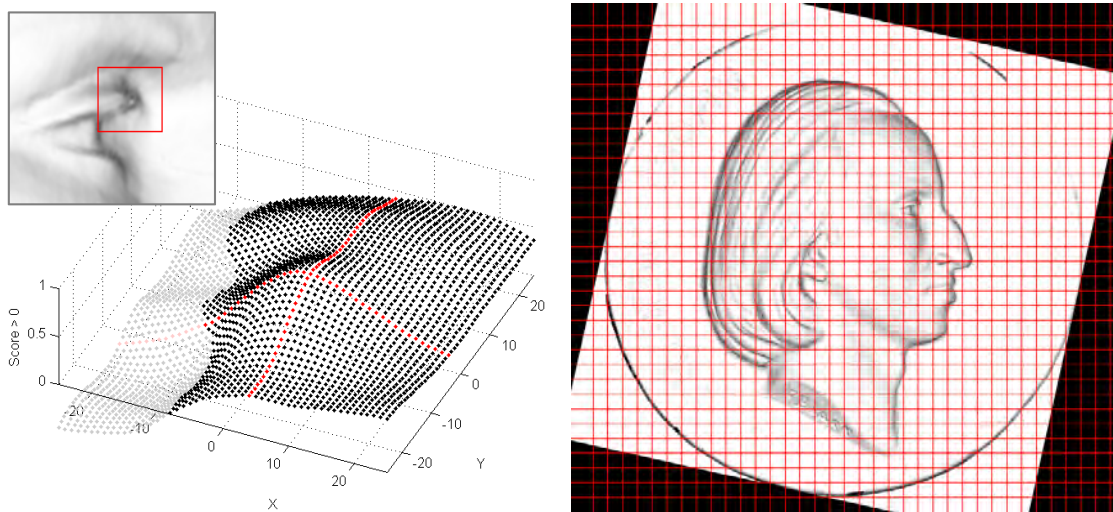


Figure 5.70: (left) Surface formed by the values of correlation coefficients over an area of 51×51 pixels. The red lines show the row and column of the peak value; (right) Grid of points over image at intervals of 50 pixels for analysis of the displacement vectors.

The resulting array of offsets shows a well-behaved distribution in X ($\mu=0.11$, $\sigma=4.2$) and Y ($\mu=-1.29$, $\sigma=4.6$) with the central values close to zero, indicating that the global registration and scaling by the correlation of outlines distributed the errors evenly (Fig. 5.71 left). After applying a 5x5 median filter to reduce noise, a clear trend can be seen when the offset values are plotted over the image area (Fig. 5.71 right). The direction of displacement is positive to the left and top and is negative to the right and bottom of the image.

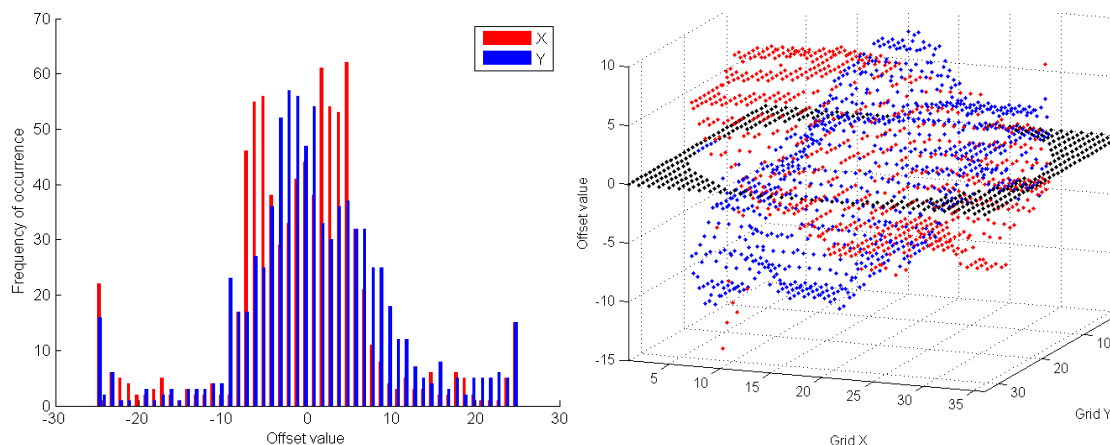


Figure 5.71: (left) Distributions of offset values (mm); (right) X and Y components of filtered displacement vectors over image area.

Plotting the composite displacement vectors (magnified by a factor of 5 to make them more visible) on the image (Fig. 5.72) makes it clear that the pattern is primarily radial, centred on a point near the ear lobe. As this corresponds to the centre of the original (uncropped) image frame from the camera, it is highly likely that the cause is geometric lens distortion.

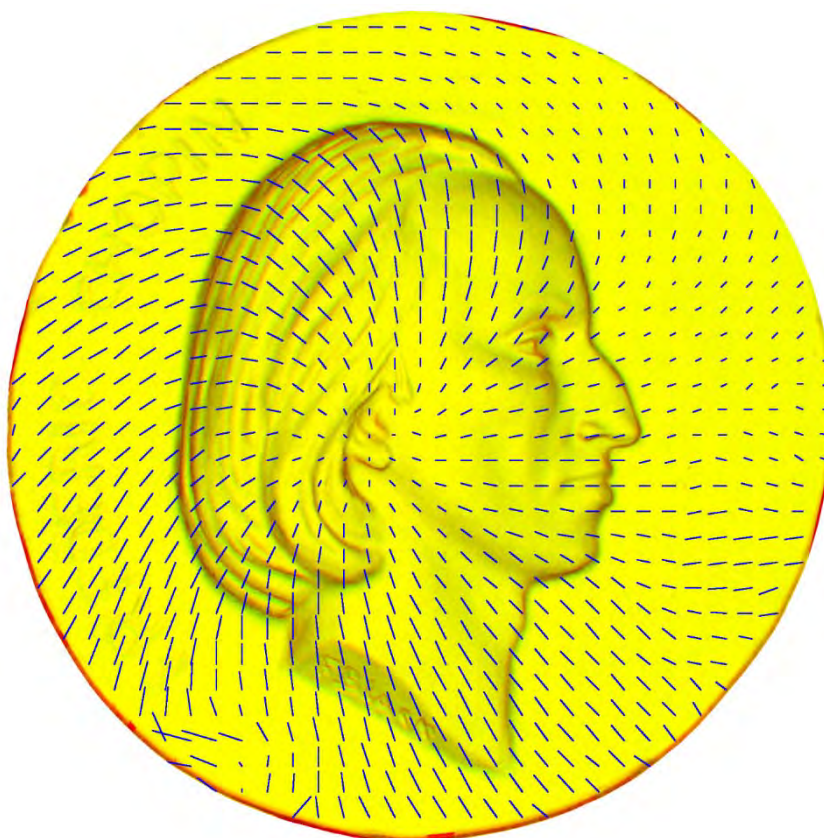


Figure 5.72: Displacement vectors magnified by a factor of 5 at each point on correlation grid.

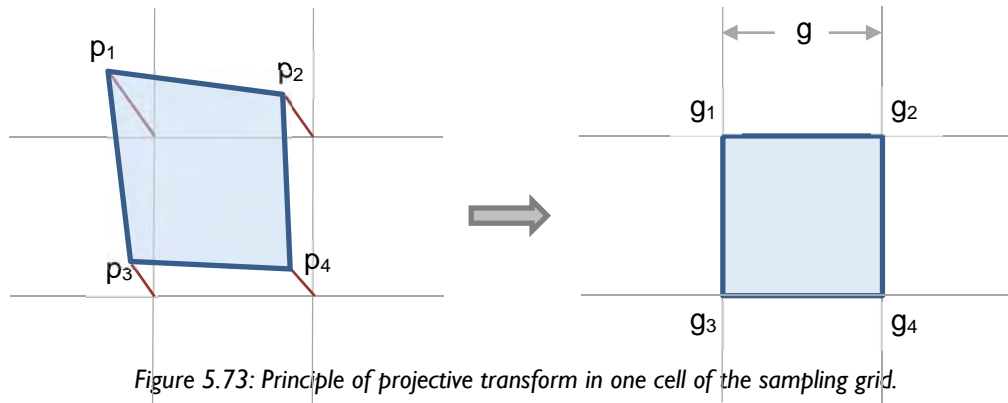


Figure 5.73: Principle of projective transform in one cell of the sampling grid.

The displacement vectors of Fig. 5.72 can be utilised in a ‘rubber sheet’ distortion transform of the photometric normals, to bring them into exact registration with the Arius normals. This is achieved by a four-point projective transform in each cell of the grid (Fig. 5.73), implemented by the code in Fig. 5.74. The displacement offsets $\{o_i\}$ when added to the four corner points $\{g_i\}$ of one grid square of size $g \times g$ form the vertices $\{p_i\}$ of a quadrilateral region of the input image. The variables k_i and k_j in the code are the index values to the grid array, g is the grid spacing, and g_b is the boundary offset.

```
kic = g*ki+gb;  kjc = g*kj+gb;           % grid point coords
p1 = g1+offset(kj-1,ki-1,:);             % add displacements ..
p2 = g2+offset(kj-1,ki,:);               % to grid points ..
p3 = g3+offset(kj,ki,:);                 % to give corners ..
p4 = g4+offset(kj,ki-1,:);               % of quadrilateral
grid_points = double([g1;g2;g3;g4]);     % output grid square
input_points = double([p1;p2;p3;p4]);    % input quadrilateral
TF = cp2tform(input_points,grid_points,'projective'); % set up transform
Ireg = Inormals(kjc-2*g+1:kjc+g,kic-2*g+1:kic+g,:); % get input region
Itile = imtransform(Ireg,TF,'XData',[g+1,2*g],'YData',[g+1,2*g]); % do it
i = g*(ki-1)+gb;  j = g*(kj-1)+gb;      % output coordinates
Inorm_tran(j+1:j+g,i+1:i+g,:) = Itile;  % save tile in image
```

Figure 5.74: Fragment of Matlab code to perform a 4-point projective transform.



Figure 5.75: Rotated and scaled Arius normals (red), and transformed photometric normals (cyan).

Superimposing the rotated and scaled Arius normals with the transformed photometric normals (i.e. the N_z component of each) in a false-colour image (Fig. 5.75) shows that an excellent registration has been achieved. Comparison of the detail with the untransformed photometric image (Fig. 5.69 right) shows reduced colour fringes (caused by misregistration) around the contours of the nose and eyelid.

The rotation and scaling transformation of Eq. 5-5 is applied to the height map from the Arius scanner, and the Fourier transform taken. The $\log(\text{power})$ spectrum (Fig. 5.76 right) shows that the spatial detail extends out to frequencies of approximately half of the Nyquist limit. The rotated cross pattern is the Fourier signature of the Arius sampling grid for the image, rotated by 12.56° (same as the image).

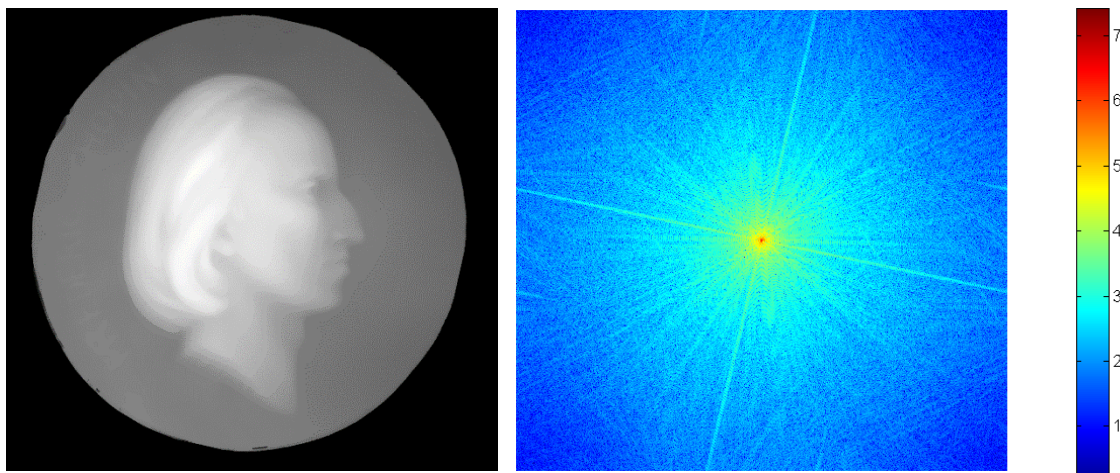


Figure 5.76: (left) Rotated and scaled Arius height map; (right) $\log(\text{power})$ of Fourier spectrum.

The rubber-sheet distortion transformation is applied to the three components of the photometric normals to bring them into register with the Arius height map. Then the X and Y gradients are calculated and the Frankot-Chellappa integration performed, as described in Section 5.6 (code in Fig. 5.47). Comparison of radial sections of the Arius and photometric $\log(\text{power})$ spectra (Fig. 5.77) shows that there is more power in the photometric spectrum, although the two converge into the noise floor for radius greater than 200 pixels. Also there is a periodic structure to the Arius spectrum, with a period of approximately 8 pixels, caused by the image raster of the scanner.

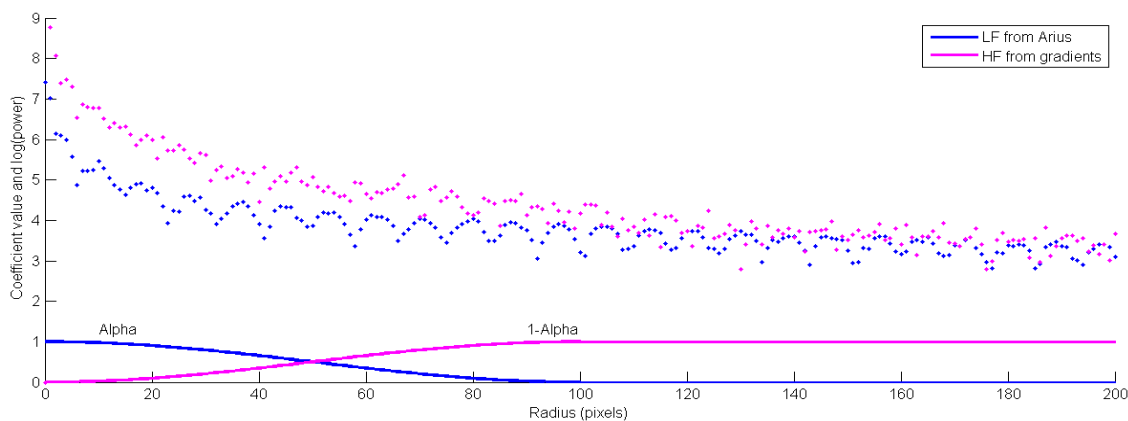


Figure 5.77: (Upper part) Radial sections of $\log(\text{power})$ spectrum for Arius and photometric gradients; (Lower part) Radial blending function (α) and its inverse.

The blending function (alpha) is chosen to be continuous (no discontinuities of value or slope) and to extend down to zero frequency. This avoids any artifacts, such as ringing patterns, in the reconstructed surface that would result from sudden changes in the power spectrum. The function selected is a cosine, as used in the Hann filter in signal processing:

$$\alpha(r) = \begin{cases} (1 + \cos(\pi r/r_{\max}))/2 & 0 \leq r \leq r_{\max} \\ 0 & r > r_{\max} \end{cases} \quad (5-7)$$

The limiting frequency r_{\max} is chosen to be high enough to include sufficient of the low frequency power from the Arius spectrum to preserve the geometry of the object, and also low enough to include sufficient fine detail from the photometric normals. In this case $r_{\max} = 100$ pixels, which is equivalent to 1.6 cycles/mm. The function $\alpha(r)$ and its inverse $1-\alpha(r)$, illustrated in Fig. 5.78, are used to blend the low and high frequency power spectra:

$$C(r) = \alpha(r)L(r) + \beta(1 - \alpha(r))H(r) \quad (5-8)$$

The power spectra $L(r)$ and $H(r)$ are of course two-dimensional arrays of a complex variable which includes both power and phase, but are combined by the one-dimensional variable, i.e. radius from the origin of the shifted Fourier frequency plane. The weighting factor β is the ratio of the total power in each spectrum up to the limiting frequency r_{\max} which balances the contribution of the photometric gradients to the composite spectrum:

$$\beta = \sum_0^{r_{\max}} \text{abs}(L(r)) / \sum_0^{r_{\max}} \text{abs}(H(r)) \quad (5-9)$$

Fig. 5.78 shows the relationship between the two weighted power distributions and their sum, all on the same scale of $\log(\text{intensity})$.

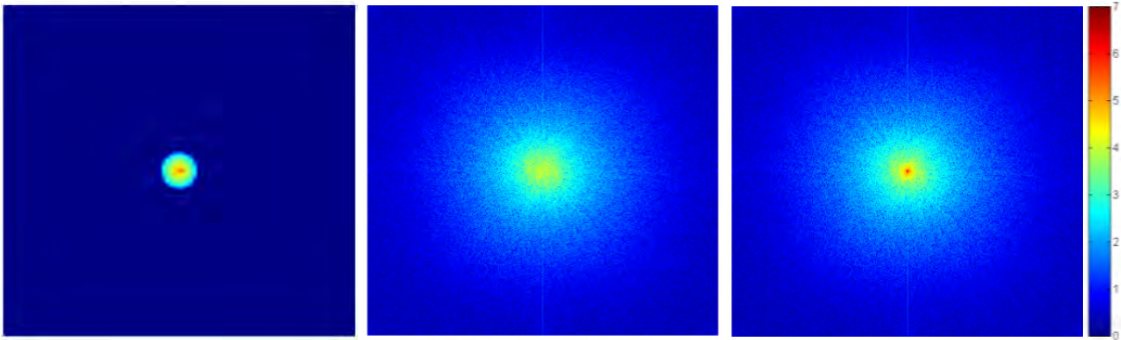


Figure 5.78: Fourier $\log(\text{power})$ spectra of: (left) Arius height map multiplied by $\alpha(r)$; (centre) photometric normals multiplied by $1-\alpha(r)$; (right) sum of two parts.

The high spatial frequencies from the photometric normals are then combined with the low spatial frequencies from the Arius height map and the inverse Fourier transform is applied to reconstruct the surface (code in Fig. 5.56). The components can be reconstructed separately, by applying the inverse Fourier transform, to visualise their relative contributions to the final result (Fig. 5.79). For clarity the edge of the roundel has been masked out. It is impressive to compare the high frequency components of the Arius and photometric spectra (top right and bottom left) and to see how much extra detail there is in the latter. The effect on the final composite height map (bottom right) is to enhance both the fine detail and also the edge contrast of the lettering, the hair and the profile of the face. The 3D reconstructed surface is shown in Fig. 5.79.

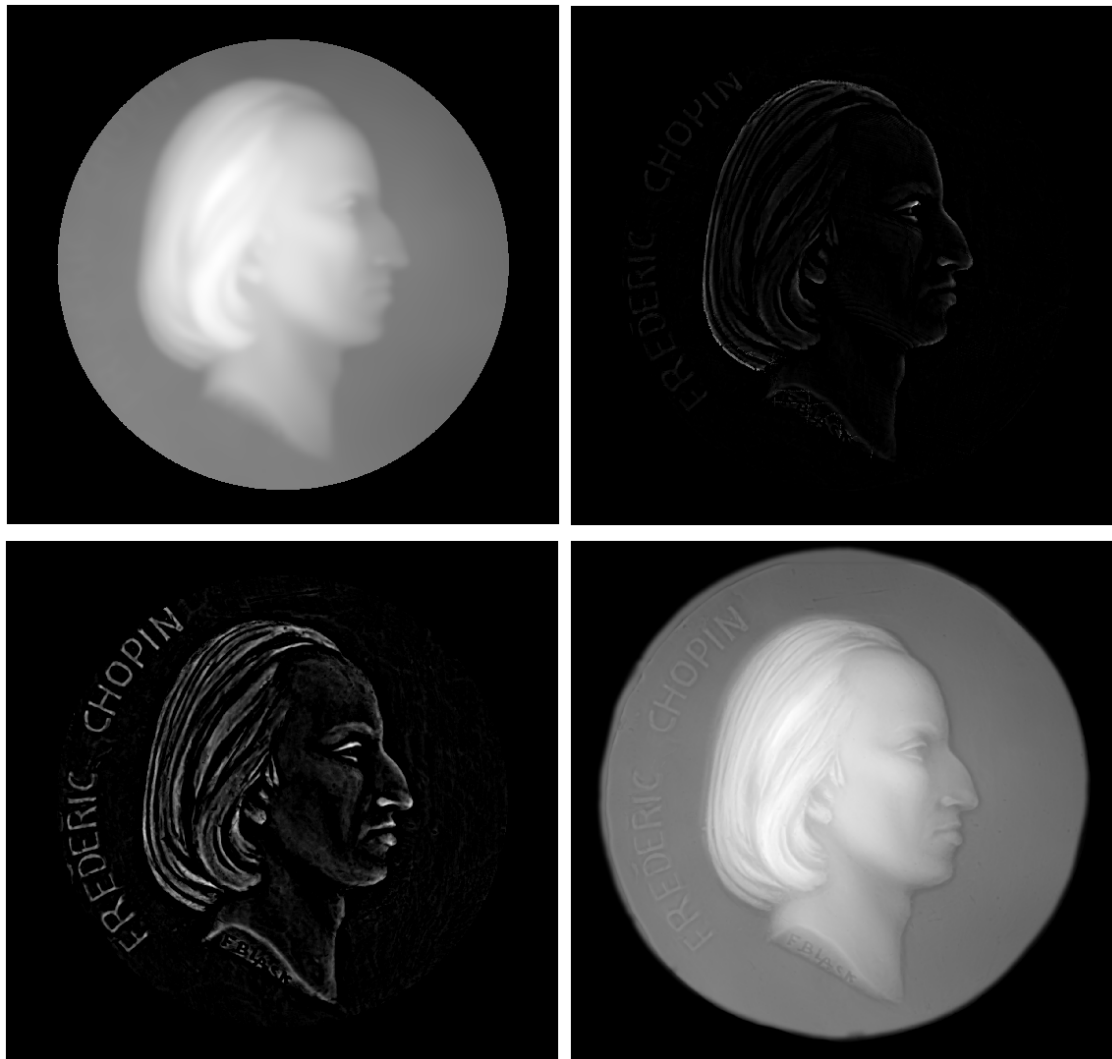


Figure 5.79: Components of height map reconstructed from weighted frequency components: (top left) tapered Arius low frequency $\alpha(r)L(r)$; (top right) residual Arius high frequency $(1 - \alpha(r))L(r)$; (bottom left) tapered photometric high frequency $(1 - \alpha(r))H(r)$; (bottom right) combined frequency $C(r)$.

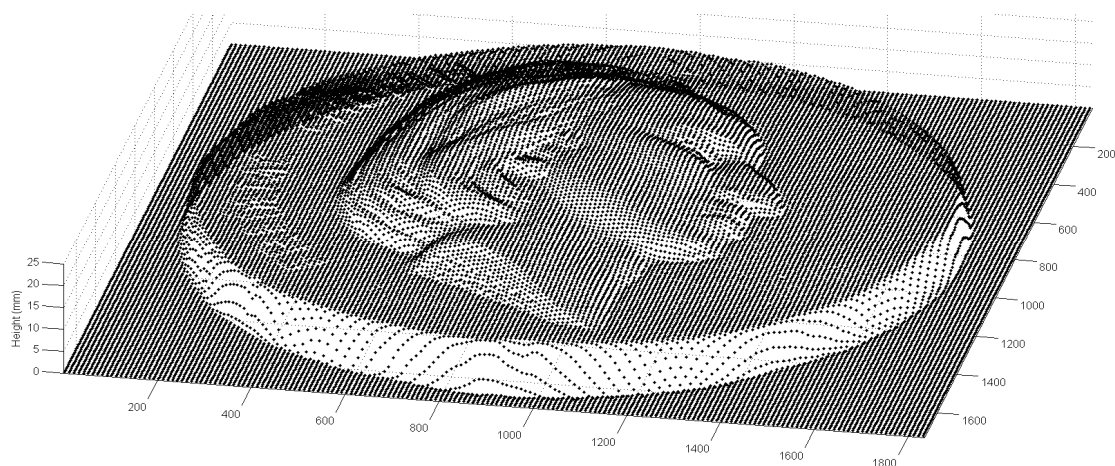


Figure 5.80: Reconstructed surface from combined low and high frequencies.

A cross-section along the horizontal midline shows the close correspondence between the Arius surface and the reconstructed surface (Fig. 5.81). The height everywhere is very close, but the reconstructed surface has visibly more detail. The local detail is evident where the line

passes (left to right) through the left edge of the basal disc, the lettering, the hair in two masses, the cavum and tragus of the ear, the cheek, the ala above the nostril, the tip of the nose, and the right edge of the basal disc. Particularly in the region across the ear (shown in the enlarged detail of Fig. 5.81), the greater depth of penetration into the intertragic incisure and resolving of the structures therein are both apparent.

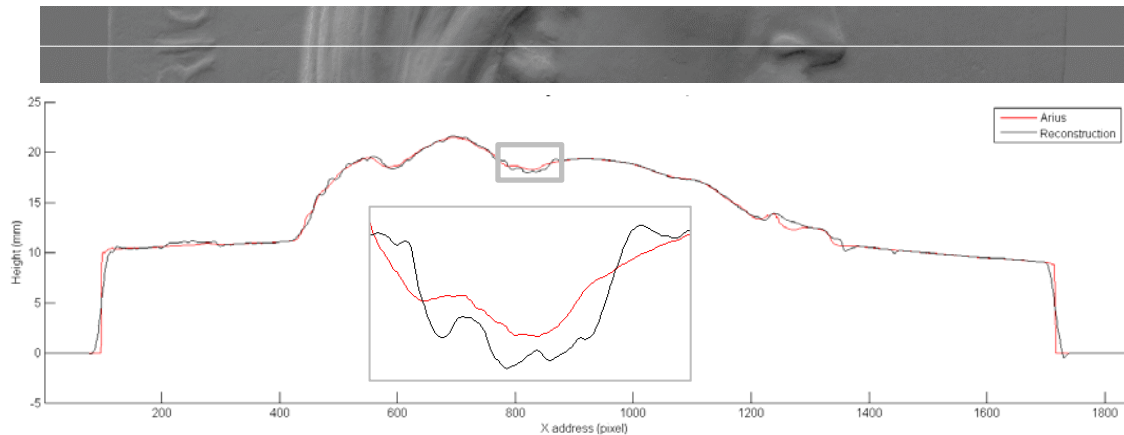


Figure 5.81: Cross-sections through Arius 3D surface (red) and reconstructed surface (black). The inset shows an enlarged section of the curves between $X=770$ and $X=870$.

Per-pixel differences between the Arius height map and the reconstruction height map are colour-coded in Fig. 5.82 left, in units of hundredths of a mm ($10\text{ }\mu\text{m}$). A difference of zero shows as green (value 127), so the range of most areas of the face between cyan and yellow indicates a difference in height of ± 30 units = ± 0.3 mm. The mean absolute difference is 0.13 mm. Of particular note is the negative difference all around the outline of the head, which indicates an enhancement of gradient contrast akin to the fringe around feature edges generated by the unsharp masking technique in two-dimensional imaging.

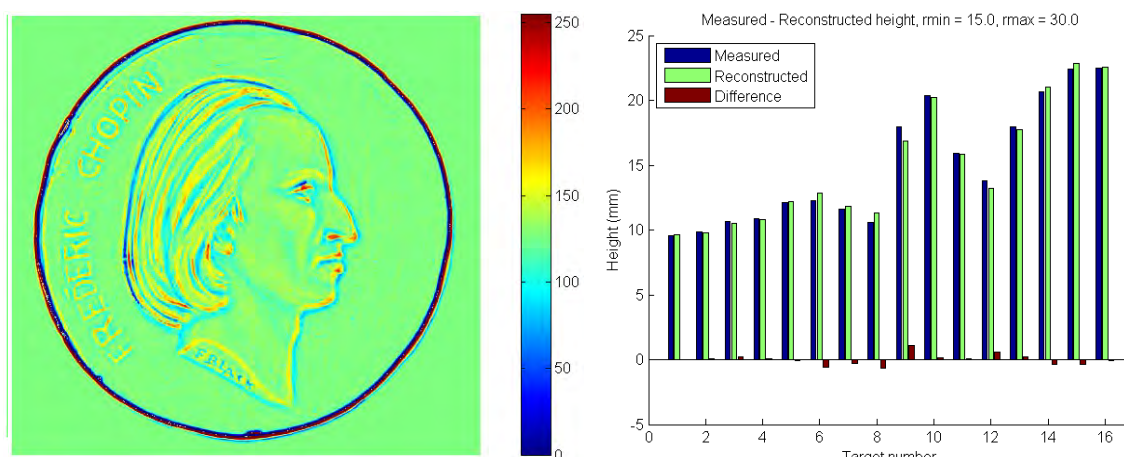


Figure 5.82: (left) Errors between Arius and reconstructed surfaces (units of 0.01 mm); (right) differences between heights measured by digital height gauge and reconstructed surface.

Differences can be calculated between the 16 points measured with the digital height gauge and the corresponding positions on the reconstructed surface. The errors are well distributed (Fig. 5.82 right), with values ranging from -0.67 to $+1.12$ mm with zero mean, mean of absolutes of 0.30 and stdev of 0.43 . These errors are similar in size and variance to those of the slanted disc reconstruction (Fig. 5.62 right) and are within the range of uncertainty of the measurements.

5.8 Process Flow for 3D Reconstruction

The whole process for reconstruction of height from a combination of the 3D point cloud and the photometric normals can be summarised by the flow diagram of Fig. 5.83. It relies on the ability to bring into registration the two views of the object: from the scanner and the camera. This is best achieved for objects that are relatively planar and flat, so that they lie securely on the baseplane of the instrument and present a similar aspect to the sensor. The combining of low and high frequencies from the two parts is analogous to texture mapping in computer graphics, i.e. placing colour and detail onto an underlying geometric framework. By adjusting the relative amplitude of the high frequency component, the degree of detail enhancement can be controlled, offering the possibility of an interactive tool for 3D surface enhancement.

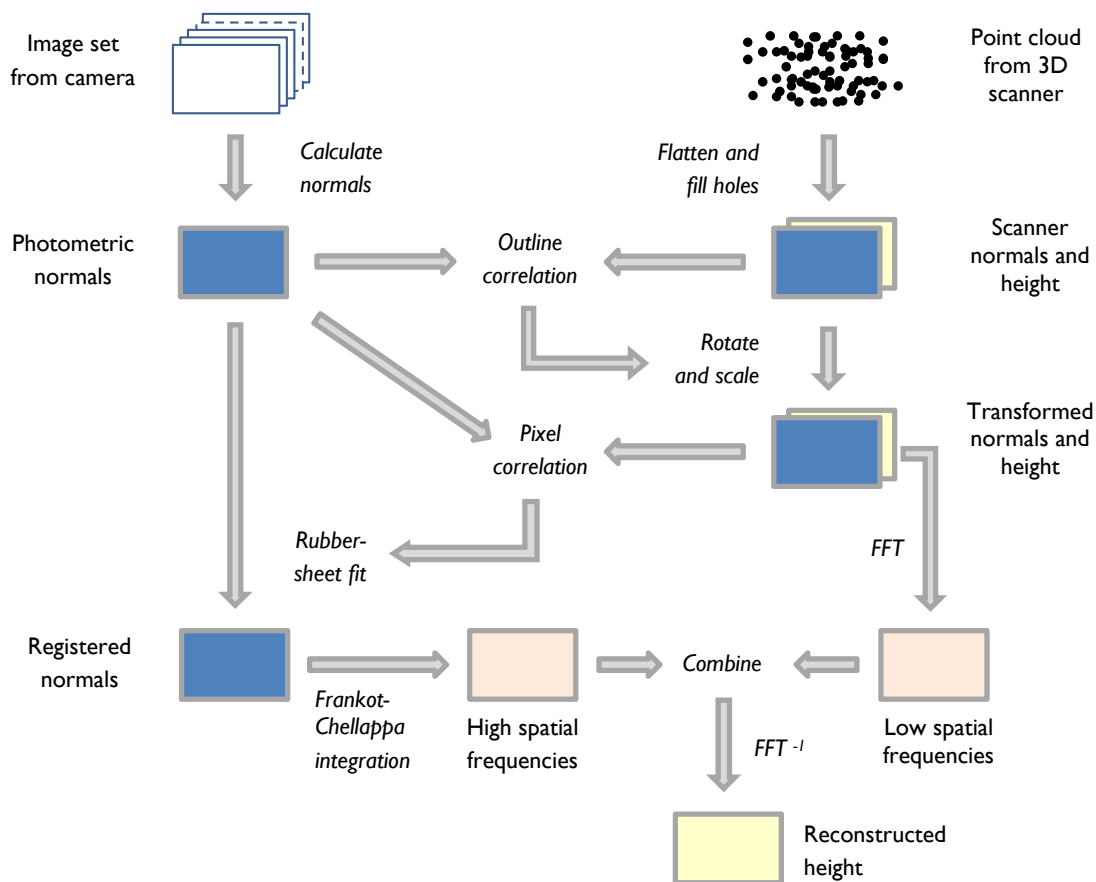


Figure 5.83: Processing flow for combining photometric and scanner normals in height reconstruction.

The result of the 3D reconstruction process is a digital terrain map, where a height is assigned to each pixel in the 2D image array. To obtain a realistic image rendering, two other essential attributes must be incorporated, namely angular reflectance distribution and colour. These are the subjects of the next two chapters.

Chapter 6 – Specularity

All that glisters is not gold;
Often have you heard that told.

– William Shakespeare (1597)
The Merchant of Venice, Act II, Scene vii

The word specular is derived from the Latin *speculum*, meaning mirror, which in turn is derived from the verb *specere*, meaning to look at. The specular component of reflection is the component reflected directly from the top layer of a shiny surface at an angle equal to that of the incident illumination, but on the opposite side of the normal vector. It is distinct from the diffuse component, which results from the light scattered by interactions of the light with colorant particles within the body of the material and is reflected in all directions. Thus it could be said that, by its etymology, the specular component of reflection is what enables us to discern the surface of an object. Specular highlights seen on a shiny or metallic surface are important visual cues to the nature of the material and its orientation and curvature.

The image sets captured in the dome, illuminated from 64 known directions over the hemisphere, contain information about the directionality of reflection from an object surface. This chapter investigates how the specularity of the surface can be modelled, first by finding the direction of the specular highlight and then by fitting a continuous function to the distribution of intensity around the specular axis. This is, in effect, half of the BRDF, i.e. the reflectance distribution function for a fixed viewpoint and variable angle of illumination. A software viewer has been developed to show the enhancement of appearance achieved by adding the specular reflection to the base diffuse reflection.

6.1 Contrast of Specular Surfaces

6.1.1 Specular Highlight for one Light Source

A white glazed bathroom tile with a glossy surface was placed in the dome and photographed with the Nikon D200 camera with 105 mm lens, giving a spatial resolution of ~ 30 pixels/mm. Fig. 6.1 shows a detail of 1000x1000 pixels ($\sim 30 \times 30$ mm on the surface) of the highlight region in four images taken with different aperture settings, all illuminated by flash lamp 64 adjacent to the lens. It can be seen that the surface is not perfectly mirror-like, but has a very fine uniform texture of dimension ~ 3 pixels (100 μm) and a macroscopic undulation with a period of ~ 100 pixels (3 mm).

As the lens aperture is closed down (i.e. higher f numbers), the area of over-exposure of the highlight region shrinks and the image of the rectangular shape of the flash lamp itself begins to become visible. Superimposed on this mirror-like reflection is a haze of light scattered from the fine surface texture, spread across a region of about three times the diameter of the light source. Away from the specular highlight the image is dark, even though the surface is white, because very much less of the incident light is reflected away from the specular angle.

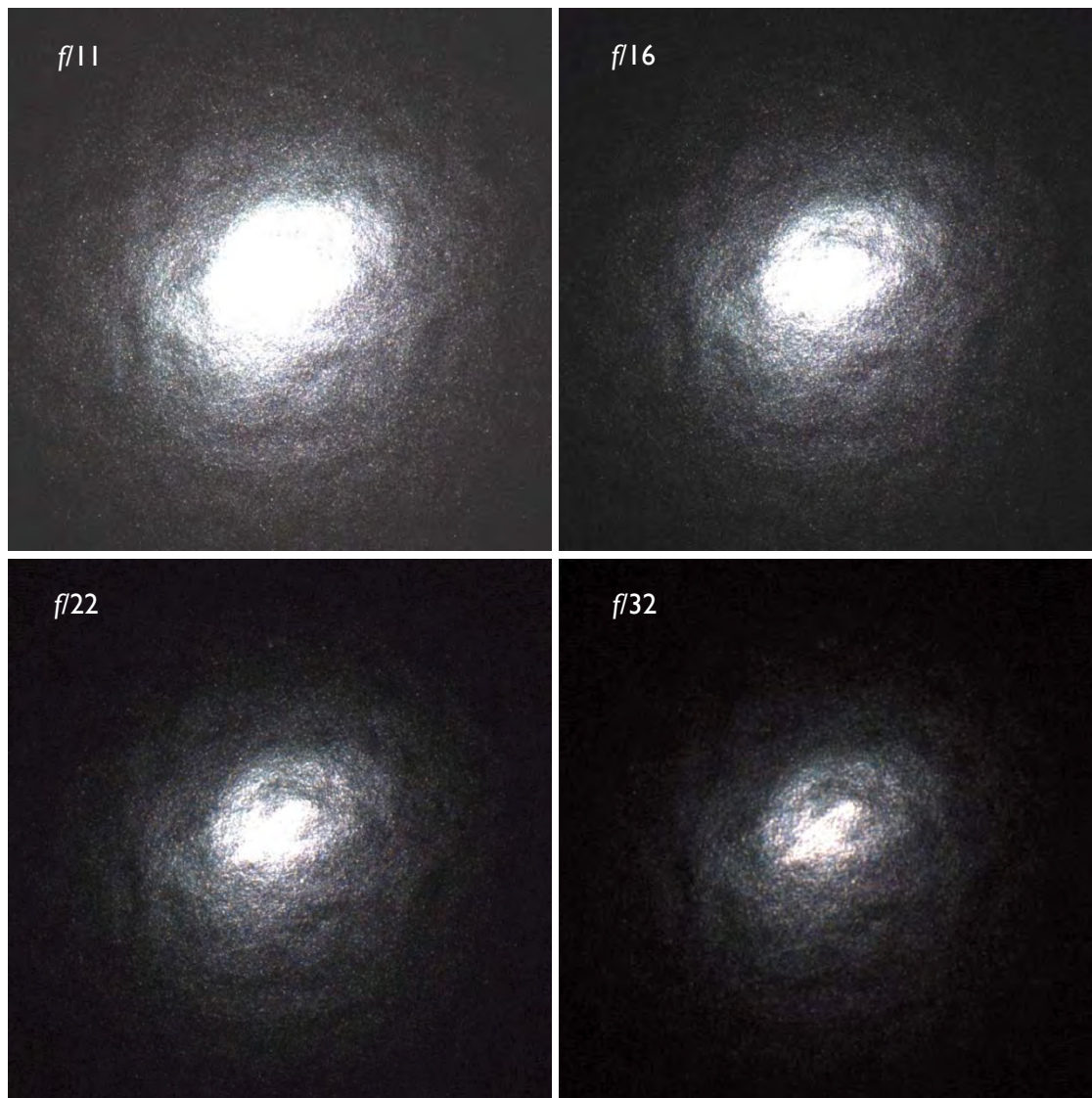


Figure 6.1: Detail of a specular highlight on a glossy white tile, taken at four aperture settings.

Plotting the intensity of the green channel along the horizontal midline of the four images (Fig. 6.2) shows that they all have a similar profile, with a central peak exceeding the exposure limit (i.e. the 16-bit maximum pixel value of 65535) and a falling away of the intensity on each side. Modulating this broad intensity distribution are large localised peaks, of much greater amplitude than sensor noise from the camera. These are the glister, or granular highlights, from the fine-grain surface texture.

An estimate of the dynamic range of the reflectance can be made by taking the ratio of the maximum value of the peak for the image at $f/32$ to the mean intensity far from the peak. Here the mean for all pixels at a distance exceeding 400 pixel widths from the centre is 230, so the dynamic range can be estimated as $65535/230 = 285$. This is a conservative estimate because even at this aperture some values are at the exposure limit, and also the intensity continues to fall with increasing distance from the peak.

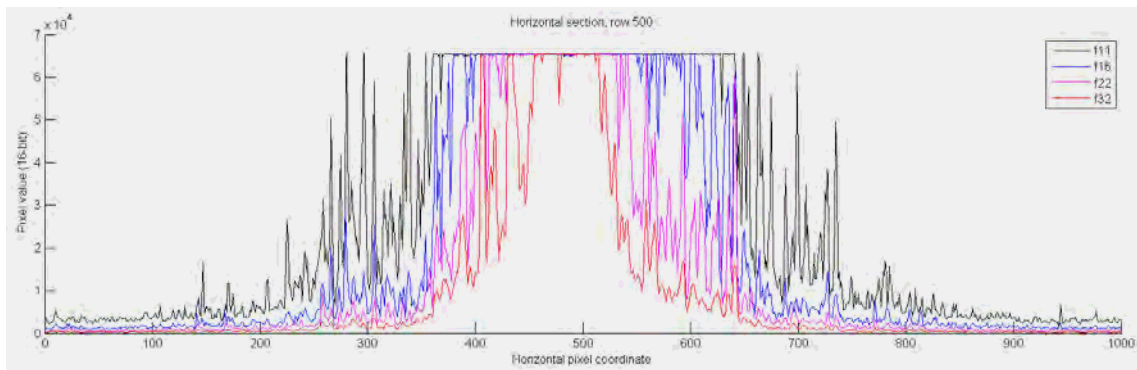


Figure 6.2: Intensity profiles of the horizontal midline through specular peak at four exposure levels.

The spatial frequencies present in the image of the tile can be seen in the power spectrum of the Fourier transform (Fig. 6.3). Taking the 1-dimensional FFT of the horizontal midline of the image taken at aperture $f/32$, windowed by a Hann function to reduce aliasing from edge effects, produces a very noisy spectrum (left), mirrored about the central frequency of 0.5 cycles/pixel (i.e. the Nyquist frequency). When the spectra of all horizontal rows in the image are averaged (right) the characteristic spectral signature of the surface emerges, up to the Nyquist limit of 15 cycles/mm. This shows a small peak at a frequency of 0.3 cycles/mm, corresponding to the gentle undulations over the surface, and an extended ridge of power over the range 2 to 7 cycles/mm, corresponding to the fine texture.

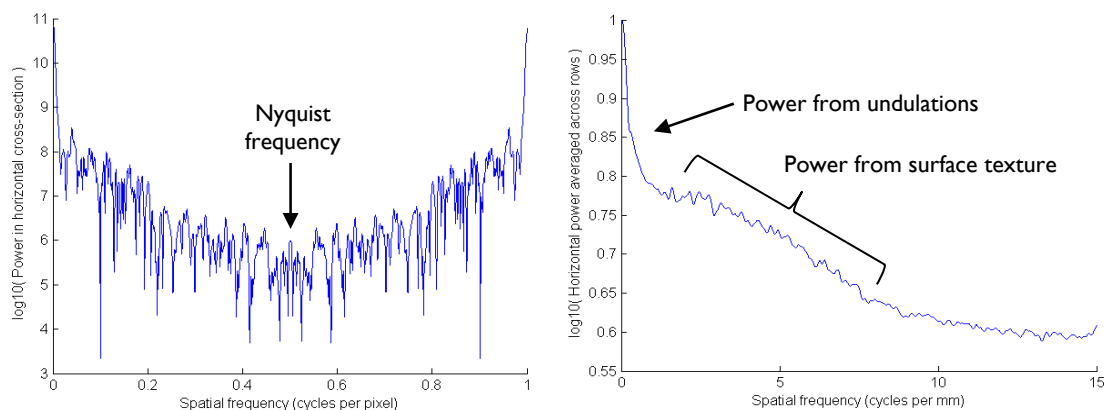


Figure 6.3: (left) Horizontal midline of FFT of image for $f/32$, in cycles/pixel; (right) Half midline after averaging over all rows in image, in cycles/mm up to Nyquist frequency.

6.1.2 Intensity Distribution in Multi-light Image Sets

The decorative test object used in this chapter is a small brass dish, 125 mm in diameter, embellished in the Damascene fashion with inlaid copper and silver arabesques (Fig. 6.4). It was chosen for its diverse metals, fine detail and highly polished shine. The brass background in the central region has been hammered with a fine point, imparting a texture to the surface. The dish was polished to a high shine with a silver cloth before being photographed in the dome with the Nikkor AF 60 mm macro lens, set to $f/11$. The cropped image is 2272 pixels square, and the diameter of the dish is 1950 pixels at a resolution of 15.56 pixels/mm.

A shiny object like this has a high dynamic range. In order to keep the specular highlights within the exposure limits of the camera the aperture was set high to $f/22$, resulting in the great majority of pixels having very low values. Analysis of the green channel of the image illuminated

by the first lamp in each tier of the dome (Table 6.1) shows that even though the maximum pixel value (on the 16-bit linear range 0-65535) in every case is 65535, the mean values are very low.



*Figure 6.4:
Damascene dish,
photographed
under four flash
lights in Tier 5
of dome.*

Table 6.1: Percentile, mean and dynamic range values for images from five tiers in dome.

Tier	Lamp	25	50	75	90	95	99	99.5	99.99	Mean	Range
1	1	0	10	32	52	70	224	658	33490	42.52	787.6
2	17	0	18	40	68	92	410	948	23189	47.57	487.5
3	33	10	34	66	104	158	626	1260	38127	78.65	484.8
4	49	36	70	118	220	394	1766	33758	65194	174.22	374.2
5	61	124	246	562	1578	3504	17490	32382	65535	1073.2	61.1

The percentile values in Table 6.1 are calculated from the pixels inside the circular outline of the dish, so exclude the area of black card on which it lay on the dome baseboard. The 50th percentile (median) is less than 246 in all cases, i.e. 0.375% of the full intensity range of the camera. The dynamic range of the image, calculated as the ratio of the 99.99th percentile/mean, decreases from 788 at the lowest angle of illumination to 61 at the highest. This is an approximate measure of the specularity of the surface of the object.

Scaling the 16-bit images from the five tiers by the same factor shows how little of the object surface is detected by the camera at the lowest angles of illumination (Fig. 6.5). In this case the 16-bit image is divided by 1000 in generating the 8-bit image for display, so that all pixels exceeding an intensity of 65 are mapped to maximum (white) in the figure. There is a significant trade-off in choosing the exposure setting when photographing the object in the dome: if too low then most of the non-specular pixels are of very low intensity (as in this case) and hence greatly affected by sensor noise; if too high then most of the specular pixels are over-exposed, producing the maximum output value and causing blooming in neighbouring pixels by spill-over of photoelectrons in the sensor.



Figure 6.5: Scaled images for five tiers in dome.

It would be better to make multiple exposures at different sensitivities and then to combine them into an HDR image, but this is not so easy with flash light because the camera exposure time setting has no effect. Instead the aperture would have to be changed, which would affect the optical characteristics of the lens and make it more difficult to register pixels from multiple images. Changing the ISO setting of the camera does little more than amplify noise. The decision was therefore taken in this study to capture only a single set of images at relatively high aperture setting, accepting that the noise in non-specular image areas would influence the results.

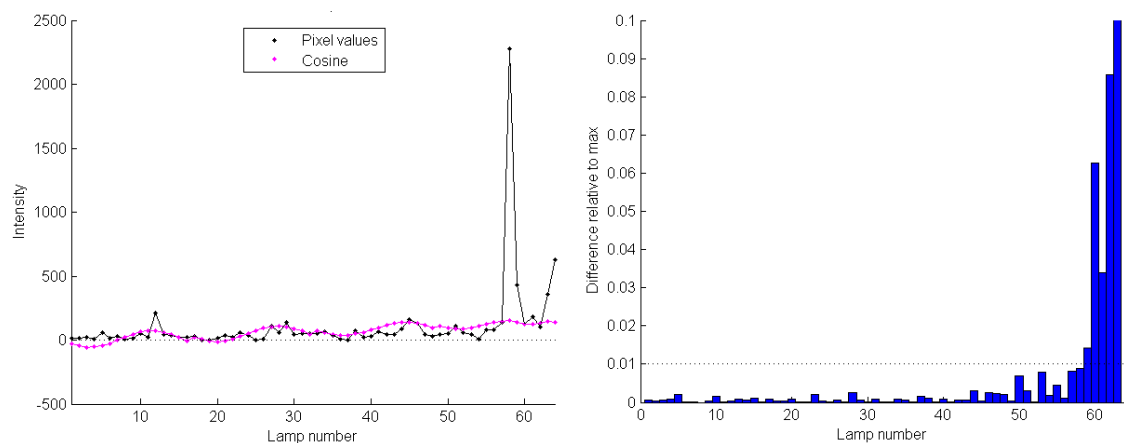


Figure 6.6: (left) distribution of 64 intensity values at a single pixel with corresponding cosine values; (right) first-order differences between normalised intensity values in sorted distribution.

The cumulative distribution of sorted non-specular intensity values is then thresholded to determine the range of lamps within which to perform the regression. The lower and upper threshold values are set to 0.15 and 0.5 respectively, giving the range between values 26 and 46 in the sorted intensity distribution (Fig. 6.7 right). Note the variation in the method from

what was applied for the Chopin terracotta in Chapter 5, where the two thresholds were applied to the full cumulative sorted distribution of 64 values, including the specular intensities (Fig. 5.8). In this case it is preferable to identify the specular intensities, and then to exclude them from the distribution for thresholding.

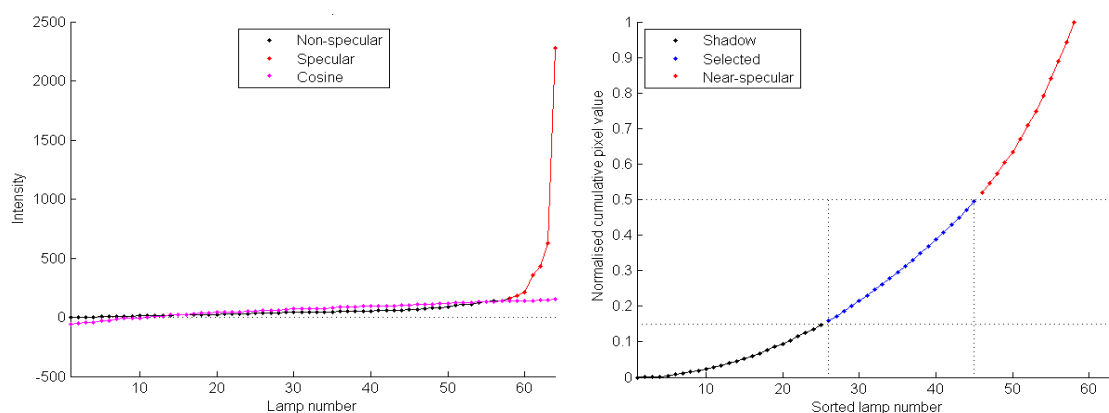


Figure 6.7: (left) Sorted distribution of 64 intensity values, identifying specular values (red); (right) thresholded cumulative distribution of non-specular values.

The subset of lamps selected by the procedure can also be visualized by plotting 3D vectors of length proportional to the pixel intensity for each lamp (Fig. 6.8). The vertical black line represents the view vector V and the green line the normal N . The black, blue and red lines correspond to the points in Fig. 6.7 for the 64 lamps. The vectors nearest the specular angle (in red) are very much longer than the vectors selected in the range for calculation of the normal (in blue). This difference in magnitude is much more pronounced here than for the corresponding diagram for the semi-gloss surface of the Chopin terracotta (Fig. 5.8).

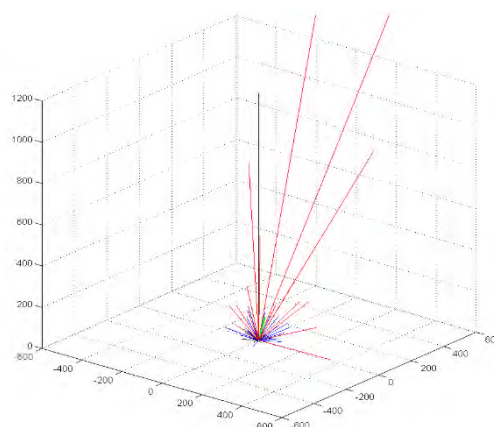


Figure 6.8: Visualisation of intensities at one pixel as 3D vectors toward lamp coordinates.

By applying this procedure, the normal vector and albedo are computed for every pixel (Fig. 6.9). The normals are quite subtle because most of the surface is horizontal and the relief of the decoration is shallow. The albedo is surprisingly dark and chromatic, representing the diffuse ‘base colour’ of the surface without any specular component.

6.2 Determining Specular Angle

The general direction of the specular vector is clear from Fig. 6.8, where the intensities from the four lamps with the nearest directions are much greater than any others. The figure also indicates the coarseness inherent in sampling angular space with only 64 points over the hemisphere, where every adjacent pair of lamps is separated by an interval of 20–25 degrees. The strategy adopted in this study is to calculate the ‘specular quotient’, i.e. the ratio between the actual intensity for each lamp direction and the intensity that would be produced by a perfect diffuser in the same direction. This would be the ratio of the black divided by the magenta values for the pixel illustrated in Fig. 6.6.

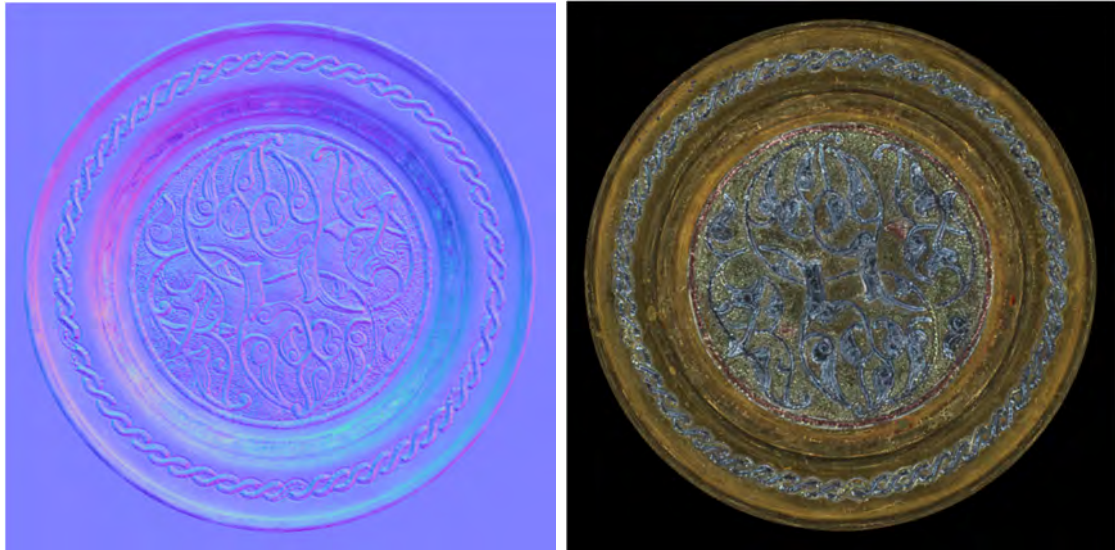


Figure 6.9: (left) Normal vectors in false colour and (right) albedo of Damascene dish.

The more shiny the surface, the greater the quotient value (Fig. 6.10 left). The specular vector is calculated as a weighted sum of the lamp vectors exceeding a threshold, multiplied by the corresponding specular quotient values. The same weighted sum gives the colour of the specular reflections (Fig. 6.10 right). For most materials these would be the same colour as the illumination, i.e. white, but for metals the specular component carries the colour of the metal. Here for the Damascene dish the colours of the brass, copper and silver are clearly defined.

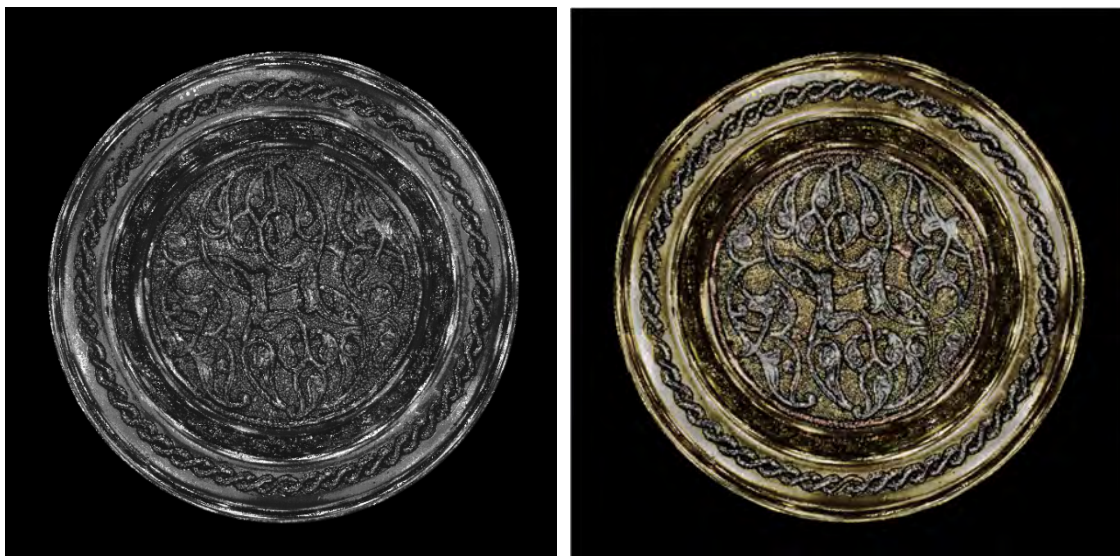


Figure 6.10: (left) Specular quotient = ratio of specular/diffuse intensities; (right) Specular colour.

The resulting specular direction vectors (Fig. 6.11 left) have the same general appearance as the normals (Fig. 6.9 left) but are more chromatic because the specular angles are greater with respect to the view vector. In conventional practice in computational photography it is almost universally assumed that the specular angle should be exactly double that of the normal, and for a perfect mirror this would of course be true. But the surfaces of real objects have a meso-structure with fine texture and granularity in the range 1 to 20 μm . One pixel as sampled by the camera may span a number of micro-facets at different angles, which reflect light differently from the incident illumination.

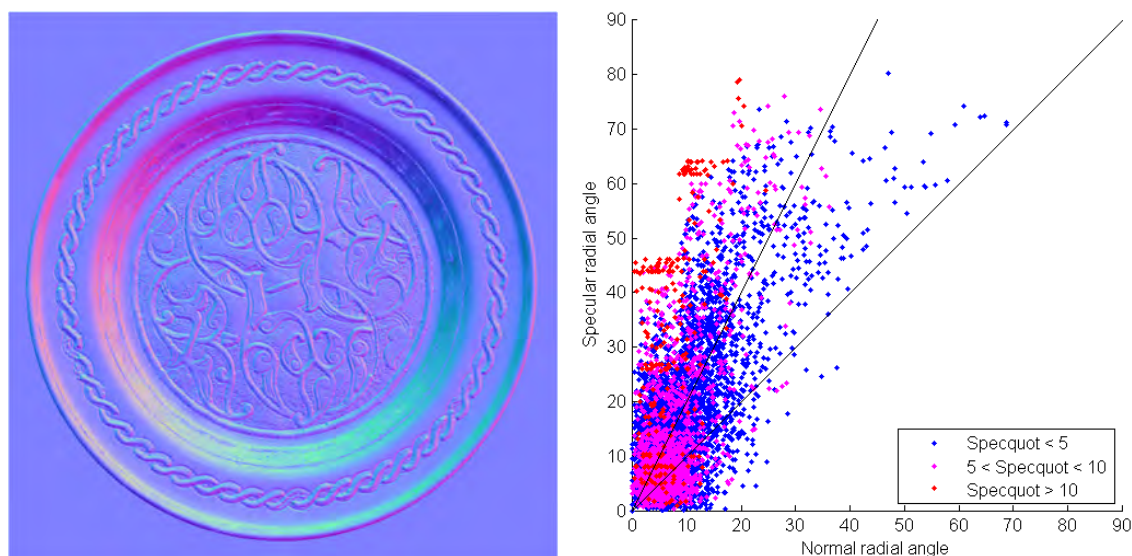


Figure 6.11: (left) Specular vectors; (right) Specular vs normal angles, classified by quotient value.

The approach taken here is to use the ideal specular (at double the angle of the normal) as a guide to where the specular angle should be. A weighted sum is taken of all lamp vectors within a cone of 45° around this direction, weighted by their quotient values. It is clear from scatterplotting the specular vs normal angles for a random selection of 10,000 pixels (Fig. 6.11 right) that there is a considerable amount of scatter around the line of slope 2 (i.e. specular angle = $2 \times$ normal angle), which is a genuine indication of the roughness of the surface. Pixels with low values of specular quotient (blue in the figure) generally have a greater scatter. Some clustering onto the five tier angles of the dome is evident in the pixels of high quotient values (red in the figure). Plotting the specular and normal angles along a horizontal section of the image (Fig. 6.12) illustrates their relationship, and also shows in this case that the maximum normal angle that can be quantified by the photometric stereo technique is $c.30^\circ$, with corresponding maximum specular angle of $c.60^\circ$.

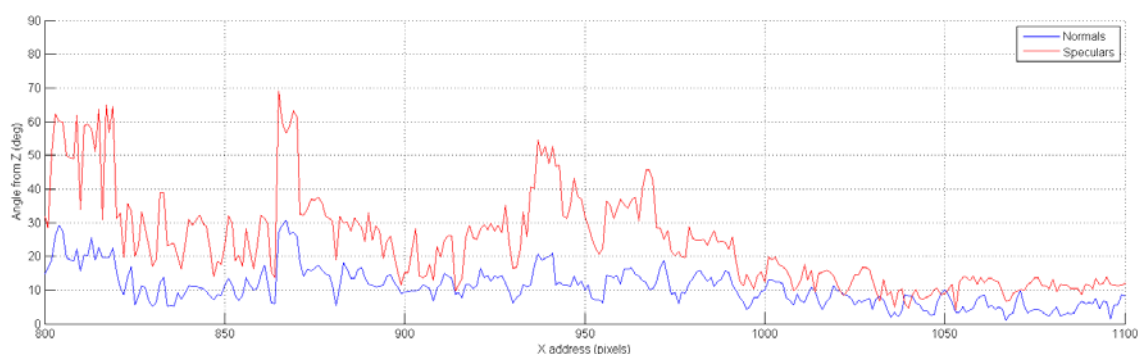


Figure 6.12: Specular and normal angles along a horizontal image section.

6.3 Modelling Angular Distribution

The aim is to model the luminance variation at each point on the object surface as a function of the angle of illumination, in such a way that the reconstructed images are indistinguishable from the original photographs. This would also enable views of the object to be 'relit' for a continuous range of illumination angles in between those of the fixed lamps in the dome.

A complete model would accommodate the bidirectional reflectance distribution function (BRDF), with four degrees of freedom, giving the reflectance of the surface at any viewpoint when illuminated from any direction. In the case of dome imaging, however, the viewpoint is fixed with the camera always at the ‘north pole’ of the hemisphere and the object lying in the equatorial plane. So the problem is simplified to finding a two-dimensional function of the reflectance factor toward the camera, given the normal and lamp vectors. A further simplification is to assume that the function of reflectance is isotropic and therefore rotationally symmetric, i.e. dependent only on the radial angle ω from the peak but not on the phase angle around the peak. (This is not the case with many real materials, but is a useful assumption for the first version of the model.) The required function needs to be positive, continuous and monotonic, with a peak at $\omega = 0$ and asymptotic to zero as $\omega \rightarrow 90^\circ$ (excluding the Fresnel component at grazing angles).

Returning to the highlight images of the white tile (Fig. 6.1), the general form of the reflectance distribution as a function of angle can be inferred by smoothing the image with a two-dimensional box filter and scatter-plotting the intensities as a function of radius from the centre of the distribution. The results in Fig. 6.13 were obtained by first applying a 21x21 averaging filter, then finding the intensity-weighted centroid coordinates in the image for $f/32$, then calculating the distance of every point in the image from the centroid, then plotting 20,000 points selected at random over the image for distances not exceeding 500 pixels. The radial distribution (right) has a very distinctive form, with a smooth descent from a peak down an inflection point at a radius of approximately 100 pixels, then a long tail extending out to the edge of the image. The width of the distribution can be explained by the elliptical shape of the intensity profile (see Fig. 6.1). This arises not because the surface is anisotropic but because the flash lamp is not a point source, but has a finite size and rectangular shape, causing the elongation of the highlight in the image.

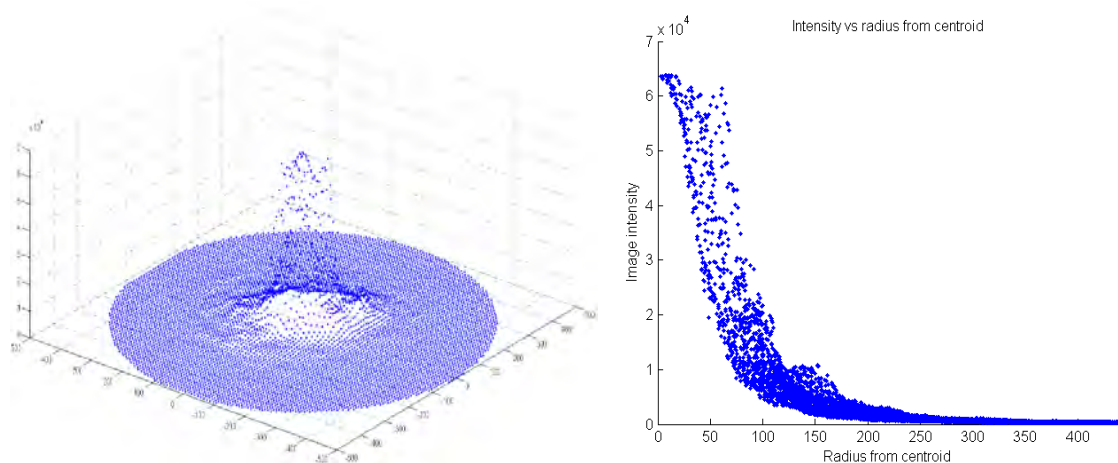


Figure 6.13: (left) 3D smoothed distribution of reflected intensity from white gloss tile around the specular peak; (right) intensity vs radial distance in pixels from centroid for 20,000 points.

The illuminated area of the flash lamps used in the dome is approximately 15x3 mm, so at a distance of 500mm (the dome radius) from the centre of the baseboard a flash light subtends an angle of 2×0.4 degrees. This is one of the factors limiting the ability of the dome system to resolve angular resolution.

6.3.1 Lorentzian Function

The model adopted in this investigation to fit the specular peak is based on the Lorentzian function, because it naturally fits the observed shape and is mathematically convenient. In particular the broad flanks enable the scattered light at perispecular angles to be modelled more effectively than the Gaussian function, which falls too quickly to zero. The comparison can be seen by fitting both functions empirically to the distribution (Fig. 6.14). An offset in the X axis has been made to accommodate the horizontal scatter, and the scale factors (divisors of X value) are different. But it is clear that the Gaussian approaches zero too rapidly and therefore underestimates the reflected intensity in the critical intermediate angles between peak and flank.

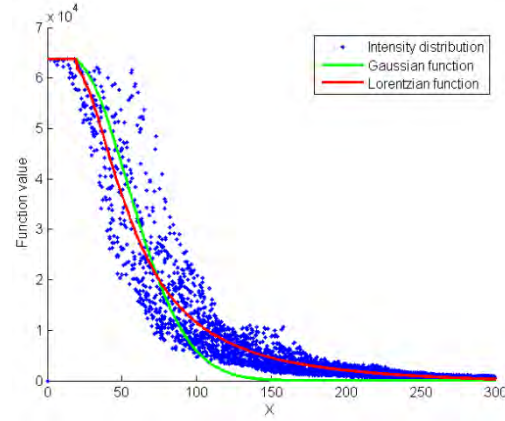


Figure 6.14: Comparison of Gaussian and Lorentzian functions against real distribution of reflected intensities.

The Lorentzian can be written as a function of three variables:

$$f(x) = \frac{a}{1 + \left(\frac{x}{s}\right)^2} + c \quad (6-1)$$

where a is the amplitude of the peak, s is the scale factor (horizontal spread), and c is a constant (uplift). Given any three points $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ in monotonic decreasing sequence such that $x_1 < x_2 < x_3$ and $y_1 > y_2 > y_3$, Eq. 6-1 can be solved for the value of c :

$$c = \frac{n_1 - n_2 - n_3 + n_4}{d_1 - d_2 - d_3 + d_4} \quad (6-2)$$

where the numerator terms n_i and denominator terms d_i are given by:

$$\begin{aligned} n_1 &= (y_1 - y_3)y_2x_2^2 & d_1 &= (y_1 - y_3)x_2^2 \\ n_2 &= (y_1 - y_3)y_2x_1^2 & d_2 &= (y_1 - y_3)x_1^2 \\ n_3 &= (y_1 - y_2)y_2x_3^2 & d_3 &= (y_1 - y_2)x_3^2 \\ n_4 &= (y_1 - y_2)y_2x_1^2 & d_4 &= (y_1 - y_2)x_1^2 \end{aligned}$$

and hence solved for the values of s and a :

$$s = \sqrt{\frac{(y_2 - c)x_2^2 - (y_1 - c)x_1^2}{y_1 - y_2}} \quad (6-3)$$

$$a = (y_1 - c) \left(1 + \left(\frac{x_1}{s} \right)^2 \right) \quad (6-4)$$

This is illustrated in Fig. 6.15 (left), where the red line is the Lorentzian function and the horizontal black dotted line is the constant c , which is the asymptote for the curve. In this case the coordinates of the three points are (1,9), (5,3.5) and (9,2). The calculated parameter values are: $a = 8.753$, $s = 3.096$ and $c = 1.074$.

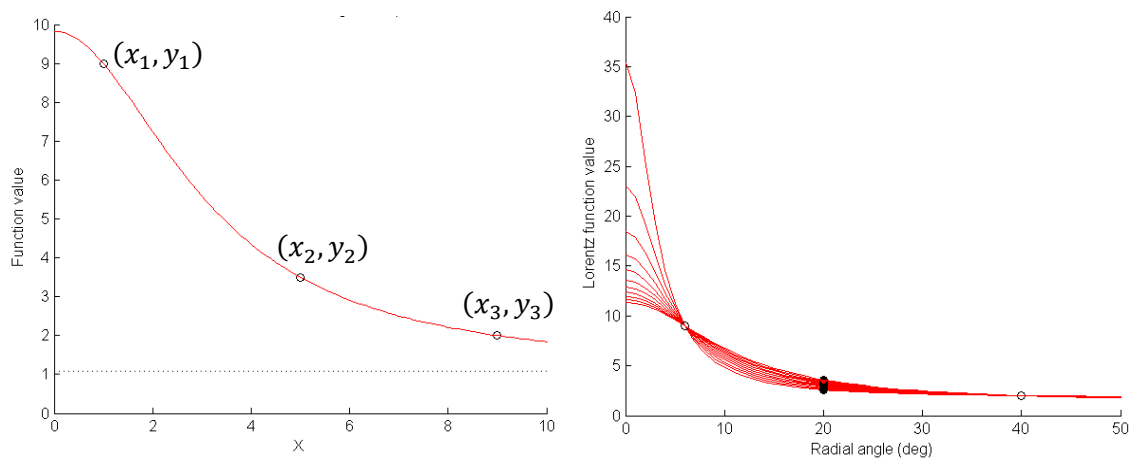


Fig. 6.15: Lorentzian function fitted through three points.

The curve fitted by the above procedure is very sensitive to the ordinate of the middle control point, as shown in Fig. 6.15 (right). Here the X axis has been extended to the range $[0, 50]$ representing a radial angle in degrees, and the abscissae of the three points are 6, 20 and 40, similar to the angles that will be used for modelling the specular distributions. The family of curves shows the solutions of Eq. 6.1 for fixed positions of the first and third control points, with the Y value of the second point y_2 changing by increments of 0.1 over the range $[2.6, 3.6]$. As y_2 decreases the amplitude of the peak increases more and more quickly; the change from $y_2 = 2.7$ to $y_2 = 2.6$ causes the peak amplitude to increase from 23.06 to 35.53, an increase of 54%. In general the value of y_2 should be at least 1.3 times y_3 for the curve to be well-behaved. This is intrinsic to the Lorentzian function. The alternative would be to add another term to the function, or to change the exponent of x in the denominator to a value less than 2, giving a shallower peak.

6.3.2 Specular Quotient Distributions

To understand better the form of distribution of the specular quotient as a function of angle for different surfaces, three test objects are considered: a grey teflon-coated sphere (a calibration object for the Arius scanner), a semi-matte white ceramic tile, and a glossy white tile (same as analysed in Sec. 6.1). All three are homogeneous across the surface, with no visible variation in gloss, apart from some blemishes on the tiles and a few scratches and nicks on the sphere. Image sets were captured in the dome in the usual way, with the Nikon D200 camera and 105 mm lens set at aperture $f/11$, and converted to linear 16-bit TIFF files.

For the tiles an area of 1080 pixels square was cropped from the centre of each image and a border of width 40 pixels imposed; for the grey sphere, the images were cropped and reduced to 1080 pixels square, within which the sphere diameter was 960 pixels. Normal vectors, albedo and specular normals were calculated for each of the three test objects (Fig. 6.16). The distribution of specular angles vs normal angles is compared in Fig. 6.17 (cf. Fig. 6.11 for the Damascene dish) and they are somewhat problematic, because it is not entirely clear how much of the scatter of the results is due to the method and how much to the nature of the surface:

1. For the grey sphere the distribution is narrow and well-defined, but the specular angles lie between the bounds of 1x and 2x normal angle. This suggests a microstructure on the

surface, such as flakes or scales of a size below the threshold of visibility to the unaided eye, i.e. less than c. 20 μm , causing greater reflected intensity toward the camera than would be expected for a perfectly uniform surface.

2. For the semi-matte tile the specular angles are highly scattered and almost all are greater than 2x the normal angle. Given that the normal angle barely exceeds 5° , this indicates the randomness of the just-visible fine surface texture that imparts the sheen to its appearance.
3. For the glossy tile there are no normal angles less than $\sim 4^\circ$, probably caused by the slight tilt of the superstructure of the dome, and therefore of the camera, relative to the baseboard. Hence there are very few specular angles less than $\sim 8^\circ$, and they are distributed around the line of 2x normal angle. But the extent of the normal angles from $4\text{--}13^\circ$ and speculars from $8\text{--}30^\circ$ is surprising for a surface that appears completely planar and uniform.

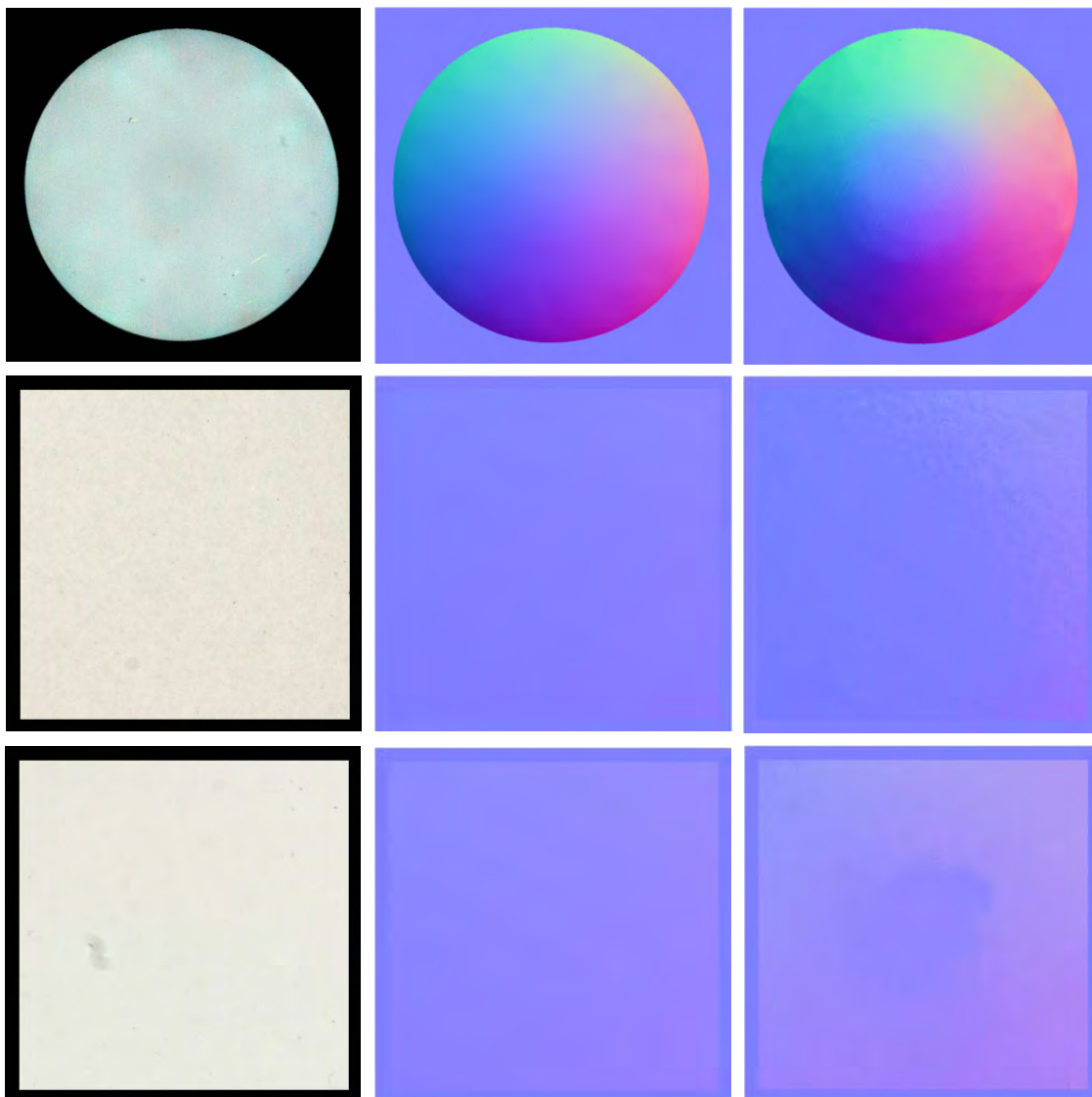


Figure 6.16: (left to right) albedo, normal, and specular normal for three test objects: (top) grey teflon sphere; (centre) semi-matte white tile; (bottom) glossy white tile.

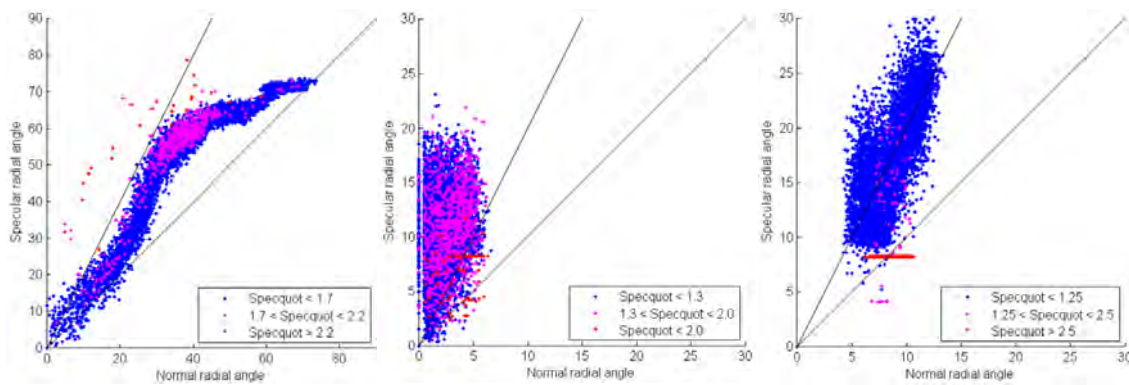


Figure 6.17: Specular vs normal angles, classified by quotient value of 20,000 random points for: (left) grey teflon sphere; (centre) semi-matte white tile; (right) glossy white tile.

The fitting procedure for calculation of the specular components is as follows. For each pixel the specular quotients are calculated as the ratio of the actual intensity for each lamp divided by the Lambertian intensity (i.e. the albedo modulated by the cosine of the normal). The pseudo-monochrome pixel intensity and albedo are used instead of the colour values in the code in Fig. 6.17, although an alternative would be to take the mean of the specular quotients for the individual red, green and blue channels. The specular angle is calculated between the specular vector for the pixel and the illumination vector of each lamp. The specular angle and specular quotient intensity values for each point can then be treated as the X and Y axes of a distribution, enabling the fitting of a continuous function.

```

P = Imageset(:,jp,ip,:);           % vector of RGB intensities
Lvec = (P(:,1)+2*P(:,2)+P(:,3))/4; % convert to pseudo-monochrome
No = Normals(jp,ip);               % normal for pixel
alb = Albm(jp,ip)+1;               % monochrome albedo for pixel
Sp = Spec(jp,ip);                  % specular for pixel
Ncos = Vlampcos*No/norm(No);        % cos angles subtended with lamps
Lam = alb*Ncos;                    % Lambertian component
Spa = Vlampcos*Sp/norm(Sp);         % cos specular angles with lamps
Spa(Spa>1) = 1; Spa(Spa<-1) = -1; % ensure within valid range
Sa = acosd(Spa);                   % radial angle lamps wrt specular
T = (Ncos>0.5);                    % use only normal angles < 60
Squot(T) = Lvec(T)./Lam(T);         % quotient on div by Lambertian
T = (Sa<90)&(Squot>0.1);            % select values in range 0-90 deg
X = Sa(T);                         % X axis is radial angle (deg)
Y = Squot(T);                      % Y axis is specular quotient

```

Figure 6.17: Fragment of Matlab code to extract specular quotients for one pixel.

In practice, there are not enough points in the 64 intensities for a single pixel to enable reliable fitting, so a region of 3x3 pixels is used, giving $9 \times 64 = 576$ points. This also has the beneficial effect of reducing noise by averaging the specular quotients of the pixel's nearest neighbours, akin to a 3x3 box filter. Plotting the specular quotients against radial angle from the specular vector (Fig. 6.18) shows the typical distribution for each test object, with large values at small angles $\omega < 10^\circ$ falling to an inflexion between 20° and 30° and then a long tail out to 90° , asymptotic to unity.

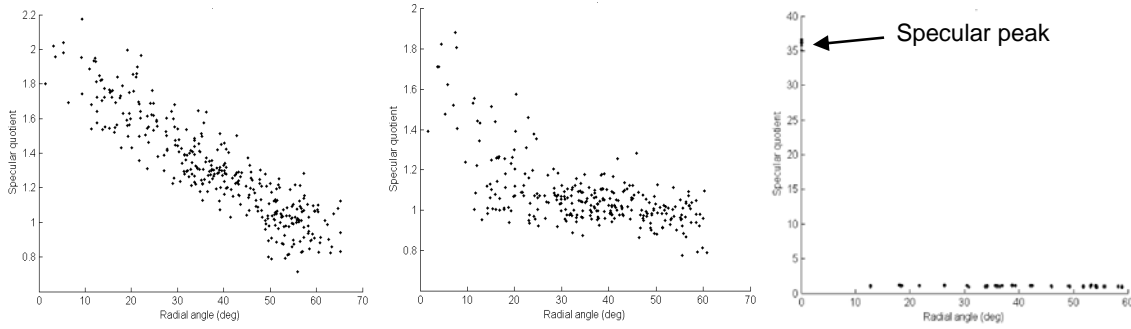


Figure 6.18: Distribution of specular quotients in a region of 3x3 pixels for: (left) grey Teflon sphere; (centre) semi-matte white tile; (right) glossy white tile.

During analysis of various materials it was found that the flank of the specular quotient distributions, for angles $\omega > 30^\circ$, varies markedly in slope according to the roughness of the surface, as is evident from Fig. 6.18. Various different functions were tried for fitting the distribution, and eventually two separate models were adopted. The first, called the ‘cluster model’, fits a one-dimensional function to the specular quotient as a function of specular radial angle, and is described in section 6.3.3. The second, called the ‘power model’, fits a two-dimensional function to the specular quotient over the polar plane of specular radial angle and hence of the rotational angle about the specular peak.

6.3.3 Cluster Lorentzian Model

The cluster model has two components, a Lorentzian peak and a linear flank, both functions of one-dimensional specular radial angle ω , each with two parameters for amplitude and scaling:

$$f(\omega) = \frac{p_a}{1+(\omega/p_s)^2} + (f_a\omega + f_s) \quad (6-5)$$

where $\omega = \text{acosd}(\mathbf{L} \cdot \mathbf{S})$ is the specular radial angle in degrees between the lamp vector and specular vector. Thus the distribution is characterised at each pixel in terms of four parameters: $[p_a, p_s, f_a, f_s]$. The first step in fitting the distribution is to fit a linear flank, using the points at radial angles $> 30^\circ$. The Matlab `polyfit` function performs a linear regression on the selected points and gives the slope f_a (always negative) and offset f_s (always positive). The offset is lowered to 0.85 of its fitted value to facilitate the subsequent fitting of the Lorentzian.

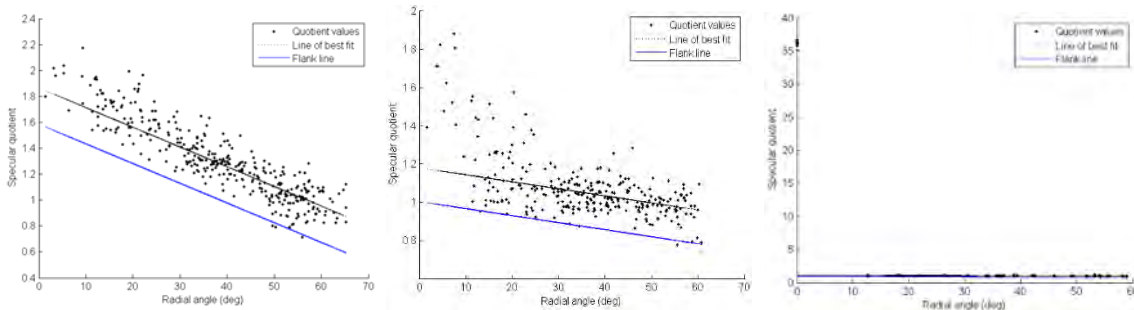


Figure 6.19: Line of best fit and linear flank functions fitted to specular quotient distribution for: (left) grey Teflon sphere; (centre) semi-matte white tile; (right) glossy white tile.

The lowered flank (blue line in Fig. 6.19) is subtracted from the specular quotients to give a distribution of residual values. Note that the zero points and scales on the Y axis are different for the three graphs. The Lorentzian is fitted to three control points at radial angles of 6° , 20°

and 40° . The Y value of the first control point is calculated as the 70th percentile of residual values in the range 0° to 12° . The Y values of the second and third control points are calculated as the 80th and 50th percentile (median) of points within the ranges 18° – 22° and 36° – 44° respectively. The use of percentiles has proved to be more stable and resistant to noise and outliers than the mean. The three parameters a, s, c are then derived from the three control points by solving Eq. 6.2 to 6.4.

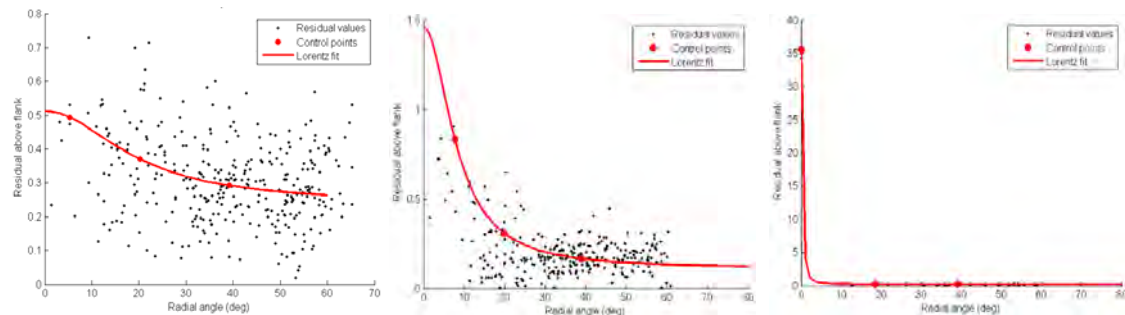


Figure 6.20: Lorentzian function fitted to the residuals of specular quotient distribution minus flank for: (left) grey teflon sphere; (centre) semi-matte white tile; (right) glossy white tile.

The combined function is the sum of the linear flank and the Lorentzian peak fitted to the residual values (Fig. 6.20). The offset value c in the Lorentzian (Eq. 6.2) compensates the lowering of the linear flank and is added into the f_s coefficient value in Eq. 6.4 for storage in the component files.

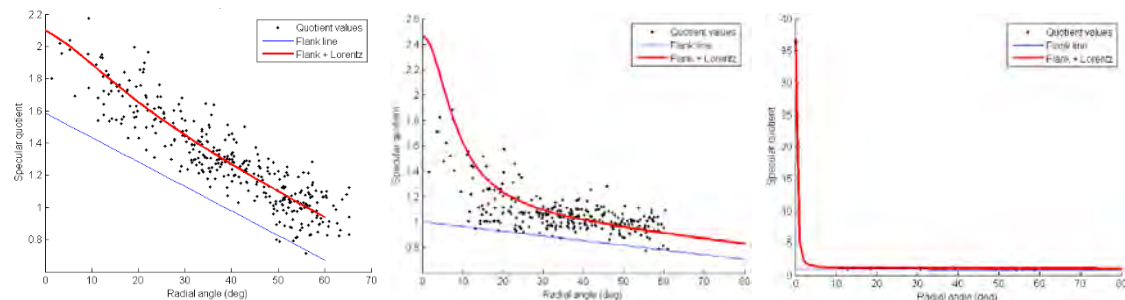


Figure 6.21: Combined function (sum of Lorentzian and linear flank) fitted to specular quotient distribution for: (left) grey teflon sphere; (centre) semi-matte white tile; (right) glossy white tile.

The difficult step in the above process is fitting the Lorentzian function to the quotient (residual) values because: (a) the values are accumulated from a 3×3 pixel region and there is a natural variation in intensity values amongst the nine pixels; (b) the pixel values themselves are subject to noise from the sensor in the camera; (c) errors in the albedo (the denominator) are magnified by the division operation when calculating the specular quotient; (d) the differencing operation of subtracting the flank line increases the relative differences; (e) the Lorentzian function is very sensitive to the ratio between the first and second control points, as described previously. In consequence, careful conditioning of the values during the fitting process is needed to maintain a consistent performance.

An indication of the variability of the parameters over the whole image is shown in Fig. 6.22 for the white semi-matte tile. This is a surface that is highly uniform and planar, so ideally all parameters should be constant at all pixels. The four cluster fitting parameters are shown as monochrome images, scaled independently to fill the tonal range between black and white. The

outer border is 40 pixels wide and is zero in every case and shows as black in three images, but for the flank slope it shows as white because the slope values are all negative. It is evident that only the flank offset is anything like constant over the image area, whereas the other three show a great deal of variation, which is an indication of the natural texture of the surface. The whole process of calculating specular quotients and components effectively differentiates and amplifies the grain, and so enlivens the final rendering through addition of sparkle and variety to the appearance of the surface.

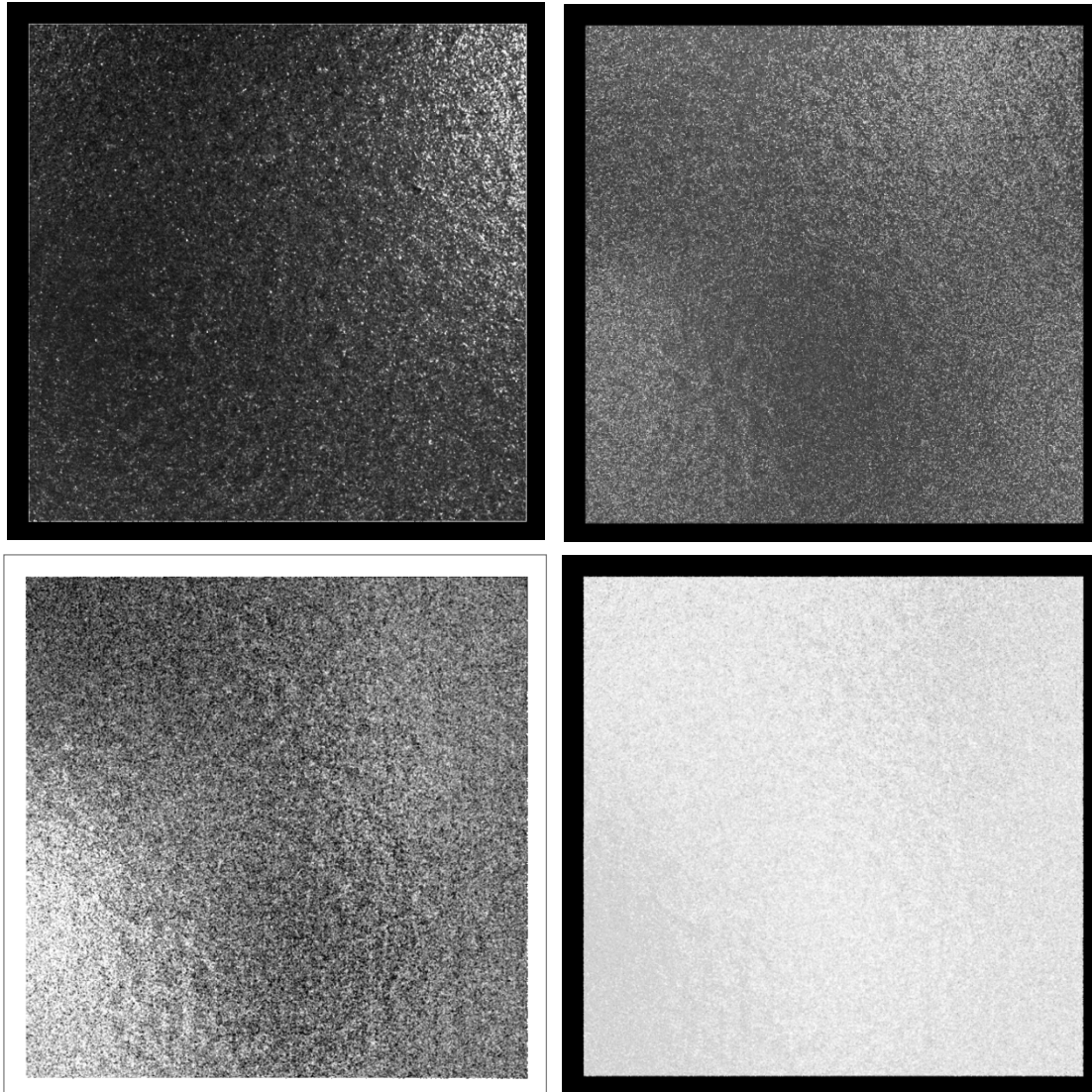


Figure 6.22: Four parameters calculated for every pixel of semi-matte tile, represented as images: (top left) peak amplitude; (top right) peak scale; (bottom left) flank slope; (bottom right) flank offset.

Fig. 6.23 shows cross-sectional plots of the horizontal midline of each parameter across the corresponding images of Fig. 6.22. The black horizontal line on each graph is the median of the values in a square region at the centre of the image. Again the wide range of variation is evident, particularly the isolated spikes in all parameters. For the gloss white tile the corresponding parameters show less variation because the granularity of the surface is less. Also the peak amplitude is greater and the scale smaller than for the semi-matte tile.

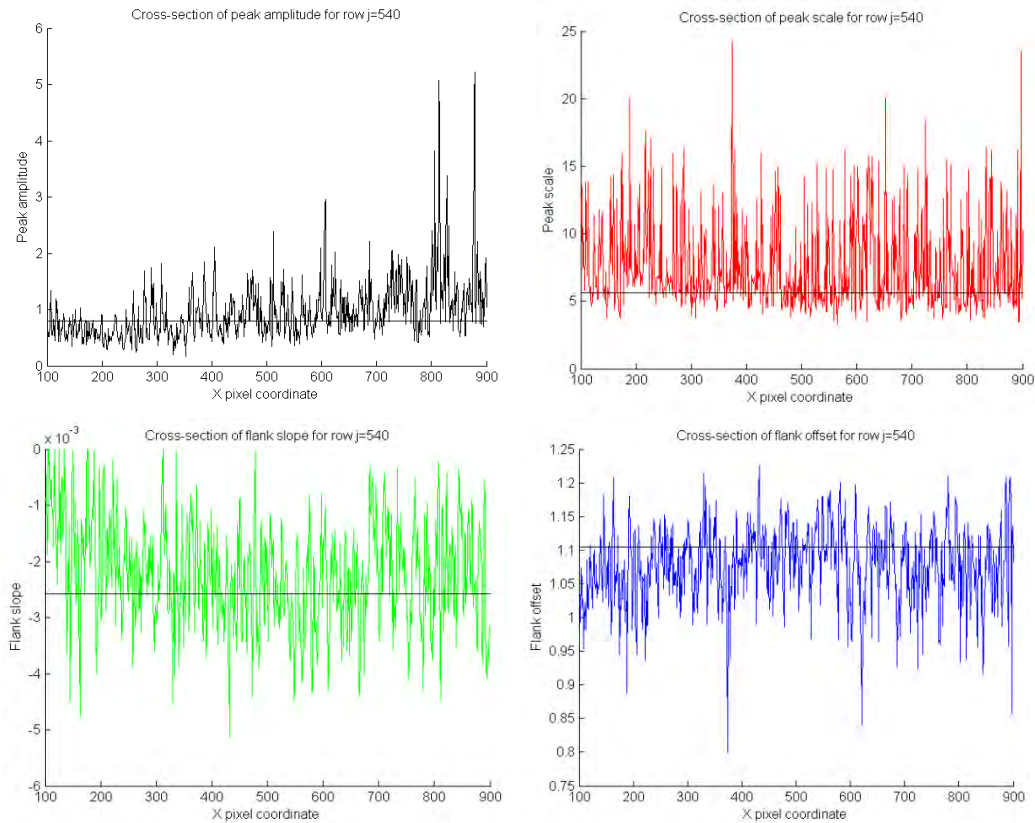


Figure 6.23: Horizontal midline sections of Lorentzian cluster parameters for semi-matte tile: (top left) peak amplitude; (top right) peak scale; (bottom left) flank slope; (bottom right) flank offset.

Histograms of each component, taken over the whole image area (Fig. 6.24) indicate the spread of values about the median (shown in each graph by a vertical white line). The curious asymmetry of the distribution of the peak scale parameter results from the conditioning of the control points for generating the Lorentzian.

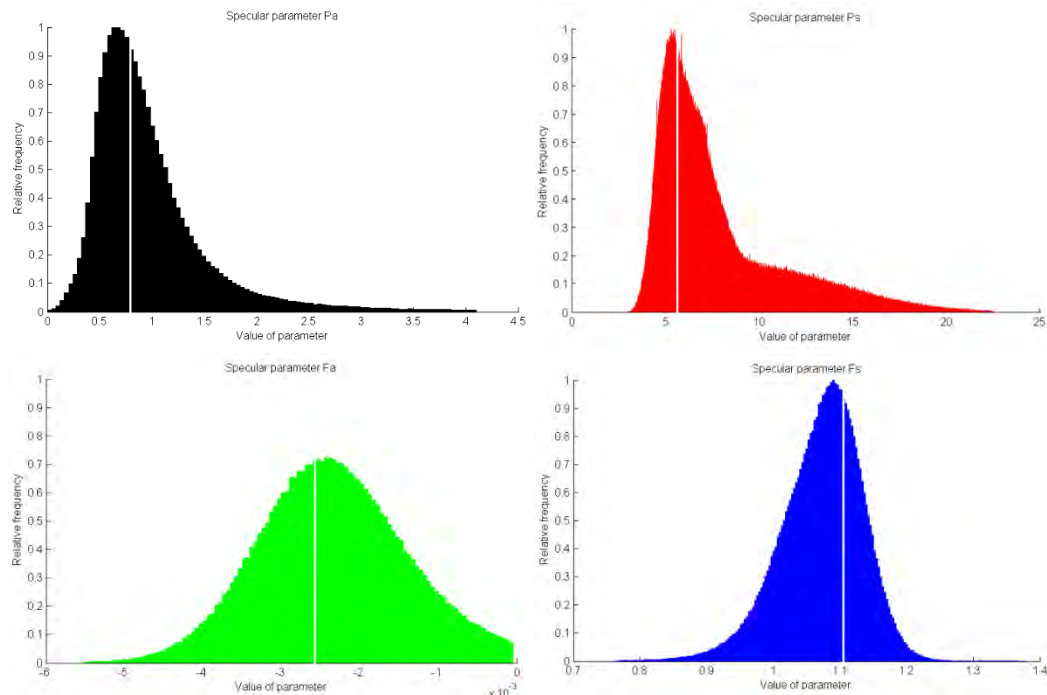


Figure 6.24: Histograms of Lorentzian cluster parameters for semi-matte tile: (top left) peak amplitude; (top right) peak scale; (bottom left) flank slope; (bottom right) flank offset.

The median values from the four distributions in Fig. 6.24, namely $p_a = 0.7947$, $p_s = 5.6417$, $f_a = -0.0026$, and $f_s = 1.1049$, comprise a parametric descriptor for the specular reflectance from the surface of the semi-matte tile. Plotting the cluster Lorentzian function of Eq. 6.5 with these parameters gives the specular function of Fig. 6.25, in which the peak reflectance is ~ 1.9 times the diffuse reflectance and falls away to ~ 1.2 times the peak at 10° from the specular angle.

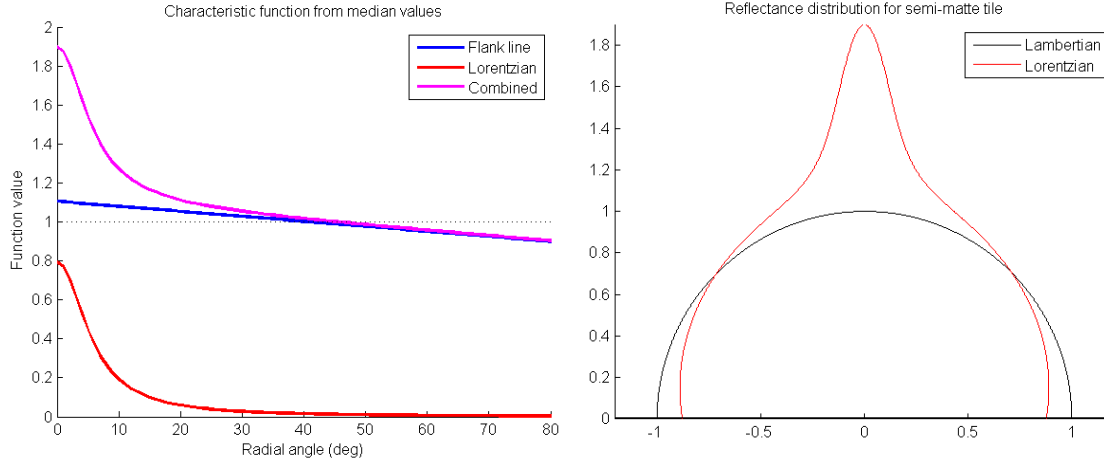


Figure 6.25: (left) Median reflectance of semi-matte tile as a function of angle from the specular peak; (right) reflectance distribution plotted as a function of hemispherical angle.

6.3.4 Power Lorentzian Model

The cluster model described in the previous section reduces the specular radial angle to a single dimension, and ignores any variation with angle of rotation around the specular peak. Thus it can be applied only to isotropic materials where there is no change in reflectance dependent on this angle. Also the method of fitting the parameters to clusters of points is somewhat unstable for noisy images and granular surfaces, resulting in large changes of coefficient values from one pixel to the next.

The power model again consists of a sum of two components, a Lorentzian-like peak and a linear flank, both functions of the one-dimensional specular radial angle ω , each with two parameters for amplitude and scaling:

$$f(\omega) = \frac{p_a}{1 + (\omega/p_s)^{p_e}} + \left(1 - \frac{\omega}{180}\right) \quad (6-6)$$

where $\omega = \text{acosd}(\mathbf{L} \cdot \mathbf{S})$ is the specular radial angle in degrees between the lamp vector and specular vector. Thus the distribution is characterised at each pixel in terms of three parameters: $[p_a, p_s, p_e]$. The flank is not fitted to the distribution of quotient values, but is assumed to be constant, a cone of value 1 at the apex and 0.5 at 90° from the specular vector. In this model the exponent p_e of the peak term is not constant $= 2$, as for a standard Lorentzian, but can vary over the range 1-10. The influence of the scale and exponent parameters on the shape of the function is shown in Fig. 6.26.

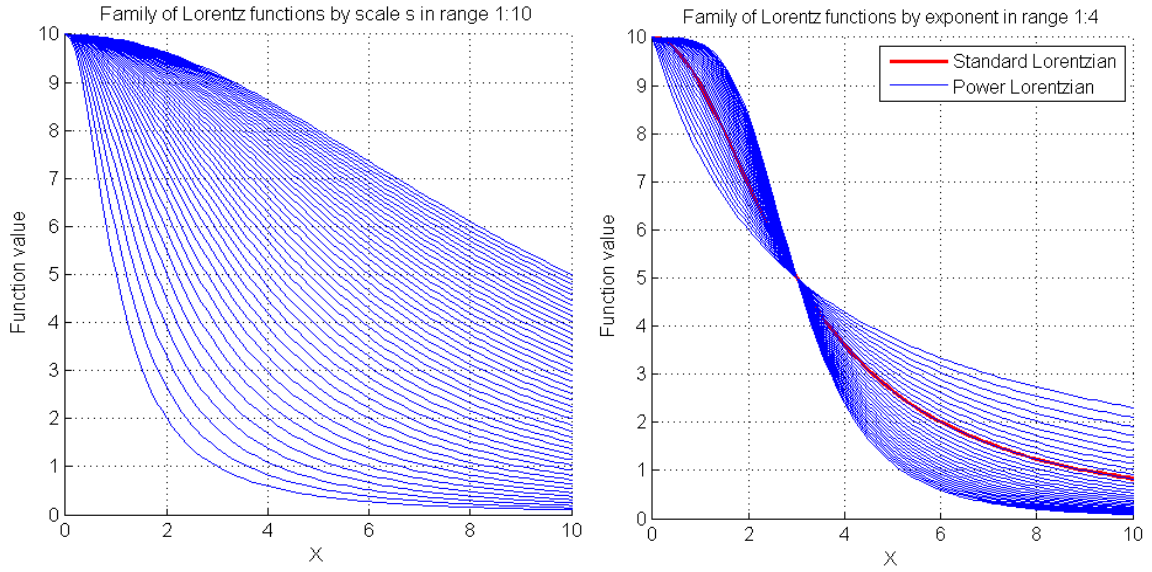


Figure 6.26: Influence of two parameters on the Lorentzian function: (left) scale p_s over the range $[1, 10]$ in steps of 0.2, with constant exponent $p_e = 2$ (standard Lorentzian); (right) exponent p_e over the range $[1, 4]$ in steps of 0.1, with constant scale $p_s = 3$. The red curve is the standard Lorentzian.

The peak function of Eq. 6-6 can be rewritten more compactly as:

$$y = \frac{a}{1+(x/s)^p} \quad (6-7)$$

and inverted to give:

$$\frac{x}{s} = \left(\frac{a}{y} - 1\right)^{1/p} \quad (6-8)$$

The principle of fitting the peak function to the distribution of specular quotient values at a pixel is to construct the curve through three control points derived from determining the x values corresponding to y values with amplitudes of a_0 , a_1 and a_2 , as described below. The parameters s and p are then calculated as follows. For the first control point $y_0 = a_0 = a$, the peak amplitude. For the second control point $y_1 = a_1$ it follows from Eq. 6-8 that:

$$s = x_1 / \left(\frac{a}{a_1} - 1\right)^{1/p} \quad (6-9)$$

Hence:

$$x_1 / \left(\frac{a}{a_1} - 1\right)^{1/p} = x_2 / \left(\frac{a}{a_2} - 1\right)^{1/p} \quad (6-10)$$

$$\therefore \log(x_1) - \frac{1}{p} \log\left(\frac{a}{a_1} - 1\right) = \log(x_2) - \frac{1}{p} \log\left(\frac{a}{a_2} - 1\right)$$

$$\therefore p = \frac{\log\left(\frac{a}{a_2} - 1\right) - \log\left(\frac{a}{a_1} - 1\right)}{\log(x_2) - \log(x_1)} \quad (6-11)$$

The relationship of the three control points to the function is illustrated in Fig. 6.27 for parameter values $a=10$, $s=3$ and $p=2.5$. Note that when $a_1 = a/2$ then $s = x_1$ (independent of p).

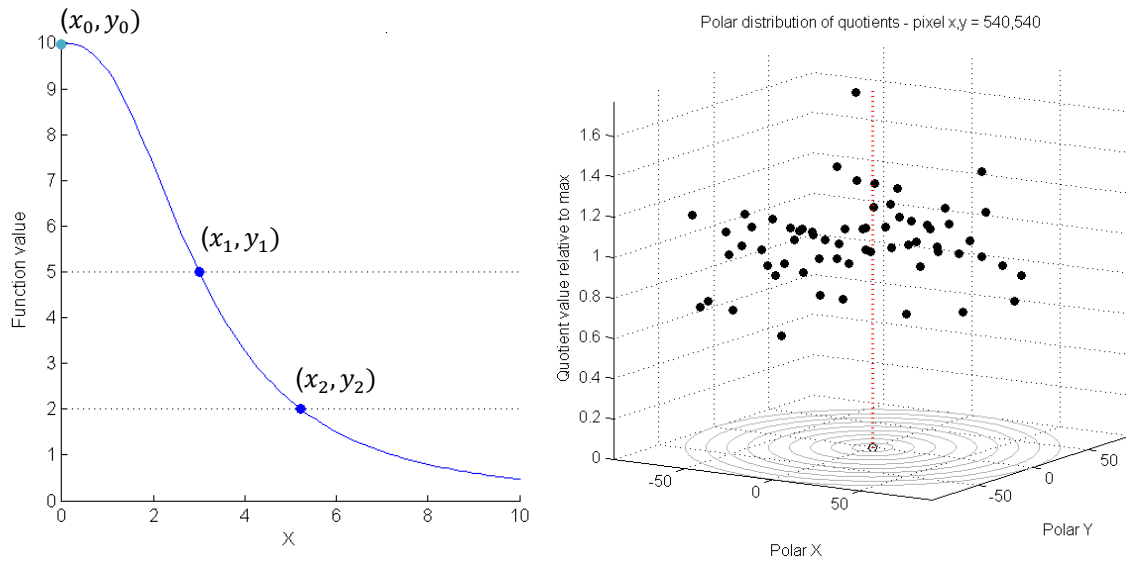


Figure 6.27: (left) Three control points for power Lorentzian function, with $a_1 = a/2$ and $a_2 = a/5$; (right) Distribution of specular quotient values for one pixel of the semi-matte tile, plotted on the polar plane with specular angle at origin.

The radii x_1 and x_2 are determined for each pixel from the distribution of specular quotient values in the polar plane. The angle of the specular vector is translated to the centre of the plane, and the angles of all incident lamp vectors are translated by the same amount. The resulting 64 points form a peaked distribution in 3D, as shown in Fig. 6.27 (right) for the semi-matte tile. Delaunay triangulation is applied to the two-dimensional distribution of points on the polar plane; when plotted as a mesh in 3D, with quotient value on the Z axis, the distribution forms a polyhedral central peak surrounded by an irregular flank (Fig. 6.28 right).

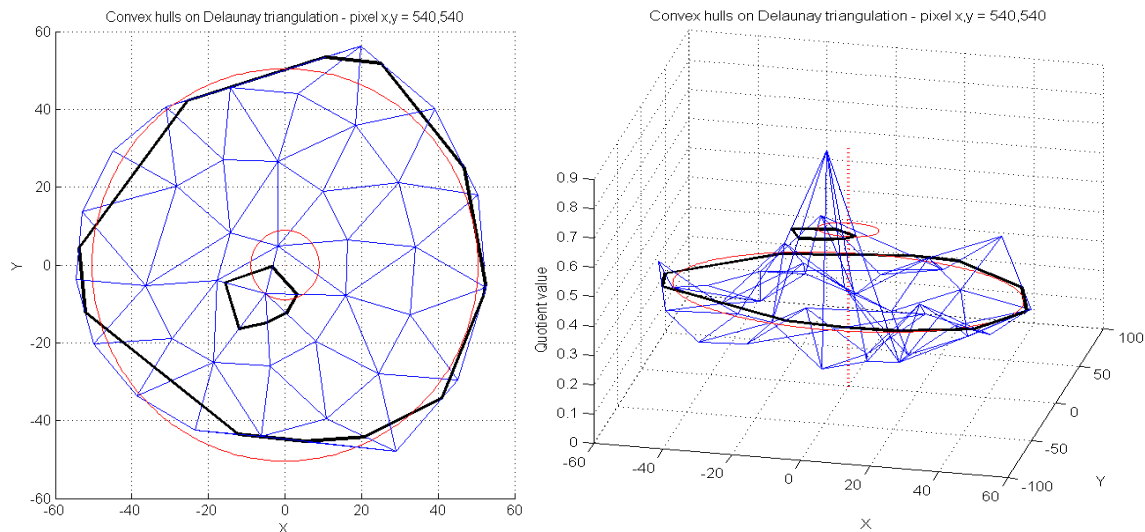


Figure 6.28: Delaunay triangulation of specular quotient values, after subtraction of the flank, for a single pixel of the semi-matte tile, showing convex hulls and circles of equal area at two heights: (left) plan view; (right) orthographic 3D view.

By intersecting the peak with a horizontal plane at a constant height and taking the convex hull, a polygon is obtained, representing the region of angular space within which the specular quotient exceeds the given value. The corresponding radius r gives a circle of area $A = \pi r^2$

equal to the area of the polygon $A = \sum A_i$ where A_i is the area of each of the constituent triangles formed by one of the sides and a vertex at the centre.

The resultant fit of the power Lorentzian model to the data for the central pixel of the semi-matte tile is shown at the centre of Fig. 6.29, together with the fitting for the grey Teflon sphere (left) and the gloss tile (right). These figures may be compared with the results of fitting the cluster Lorentzian model to the same pixel of the same three objects in Fig. 6.21. Note that the fitting in Fig. 6.29 was done on just the 64 quotient values of a single pixel, whereas the fitting in Fig. 6.21 was done on $9 \times 64 = 576$ quotient values in a 3×3 pixel region around the selected pixel.

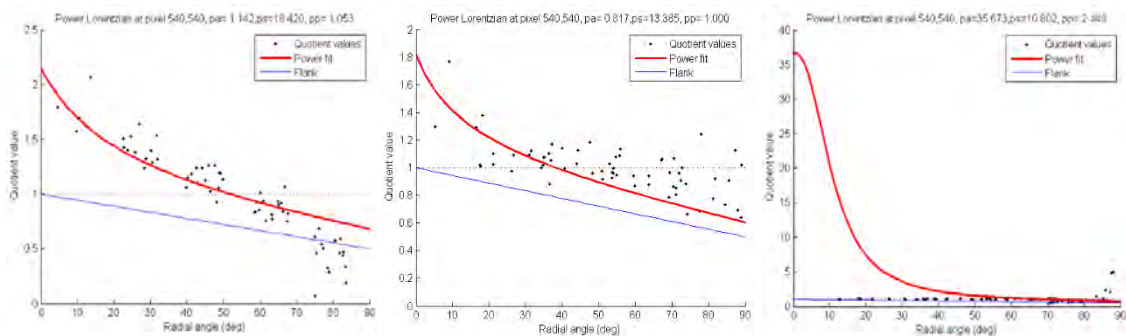


Figure 6.29: Fitting of power Lorentzian model to centre pixel of: (left) grey Teflon sphere; (centre) semi-matte white tile; (right) glossy white tile. The blue line is the flank (invariant).

Displaying the parameters of the power Lorentzian for the semi-matte tile as image components (Fig. 6.30) shows greater uniformity over the area than the parameters for the cluster Lorentzian algorithm (Fig. 6.22). This suggests that the power method of fitting over the 2D polar plane is more stable than the cluster method of fitting over a one-dimensional function of radial specular angle.

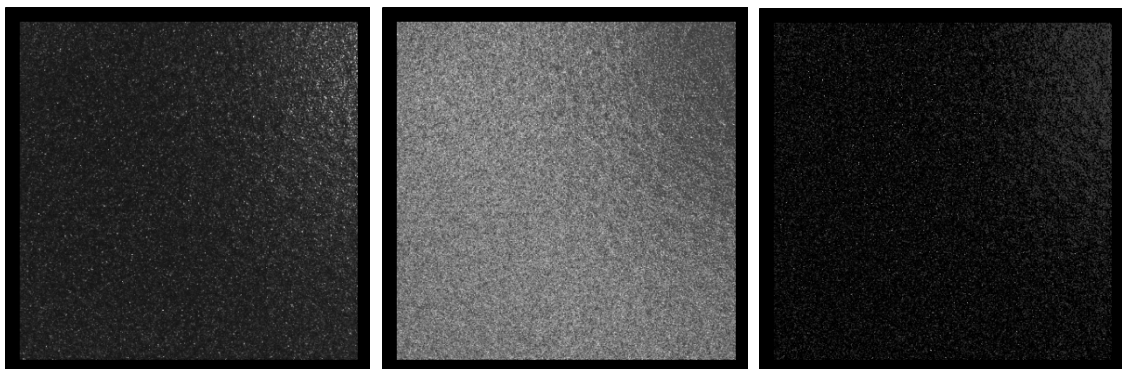


Figure 6.30: Three power Lorentzian parameters calculated for every pixel of semi-matte tile, represented as images: (left) peak amplitude; (centre) peak scale; (right) exponent.

The power model tends to produce a peak of lower amplitude and greater width than the cluster model. In particular it does poorly for highly specular objects such as the gloss tile, where there is a very high amplitude quotient value at one single angle and very low values everywhere else. The reason is clear from Fig. 6.31, where the construction of the Delaunay triangulation produces a ‘witch’s hat’ shape. The slope of the edges from the apex to the nearest points makes it inevitable that the lower convex hull will have a diameter of at least 20° and hence that a broad Lorentzian distribution will be generated.

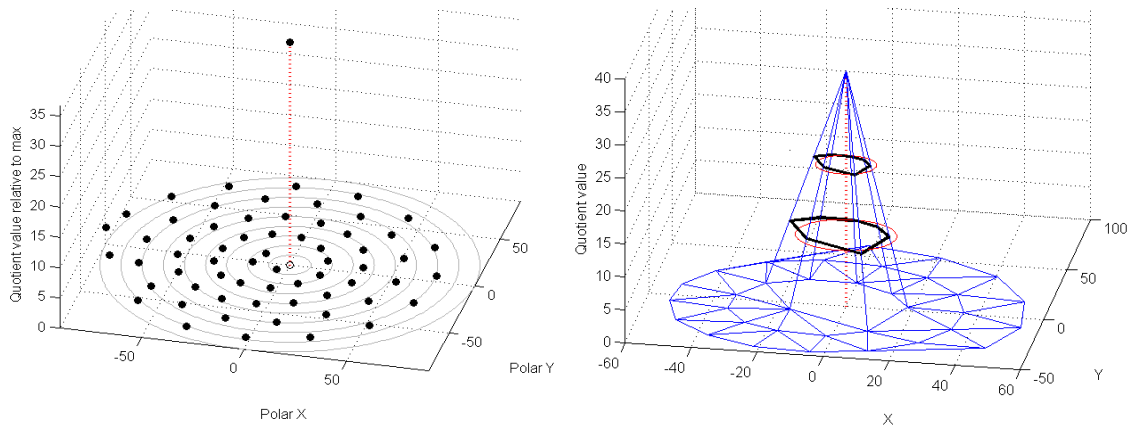


Figure 6.31: (left) Distribution of quotient values around the specular axis for a single pixel of the gloss white tile; (right) Delaunay triangulation and convex hulls.

Conversely, for the grey sphere, the power distribution does well in representing the conical mound of the quotient distribution (Fig. 6.32).

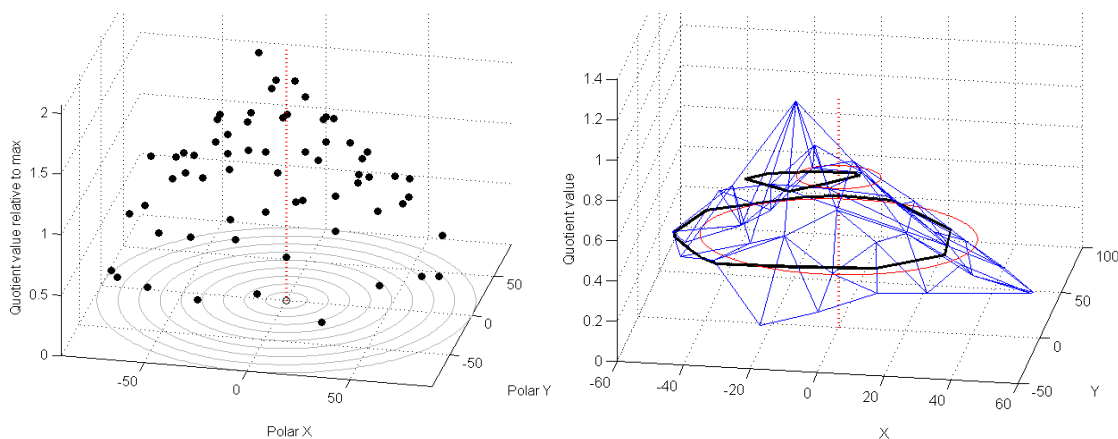


Figure 6.32: (left) Distribution of quotient values around the specular axis for a single pixel of the grey Teflon sphere; (right) Delaunay triangulation and convex hulls.

6.3.5 Comparison of Lorentzian Models

It is interesting to compare the functional forms of the two versions of the Lorentzian model described in the previous two sections by plotting the two curves on the same graph. For an object with a semi-gloss surface we return to the Chopin terracotta studied in the previous chapter (see Fig. 5.1) and choose a single pixel on the cheekbone of the subject (white spot in Fig. 6.33). For this pixel both algorithms produce similar curves that fall below 1 at an angle of 45° . Superimposing the individual curves for all the surrounding region of 15×15 pixels (Fig. 6.34 right), corresponding to a surface area of 1 mm^2 , shows that they all have the same general form, but that the amplitude of the peak varies by a factor of about 2. The invariant conical flank of the power Lorentzian slightly underestimates the quotient values in the flank at higher radial angles.



Figure 6.33: Selected point on cheekbone of Chopin terracotta.

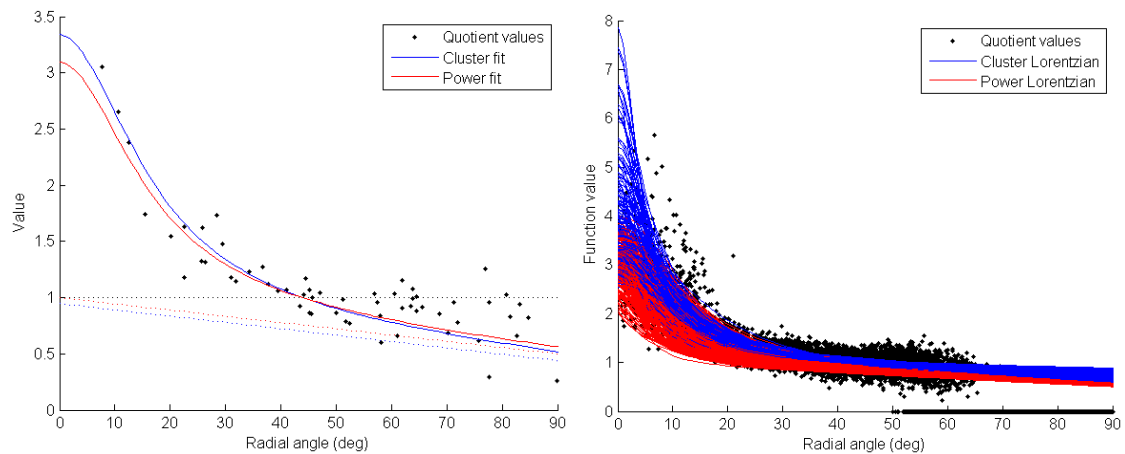


Figure 6.34: (left) Fitting of two Lorentzian models at a single pixel of the Chopin terracotta; (right) Superimposed curves for fitting of single pixels in a region of 21x21 pixels.

The characteristic curves can be generated by taking the median of each of the fitting parameters over all the pixels in the same region of 15x15 pixels. These are shown in two projections in Fig. 6.35. The cluster algorithm in general produces a taller, narrower peak and is more nearly asymptotic to 1, whereas the power algorithm produces a shallower curve that falls away at higher radial angles. At illumination angles greater than 45° from the specular, both models predict that the tile will have a lower reflectance than a Lambertian surface.

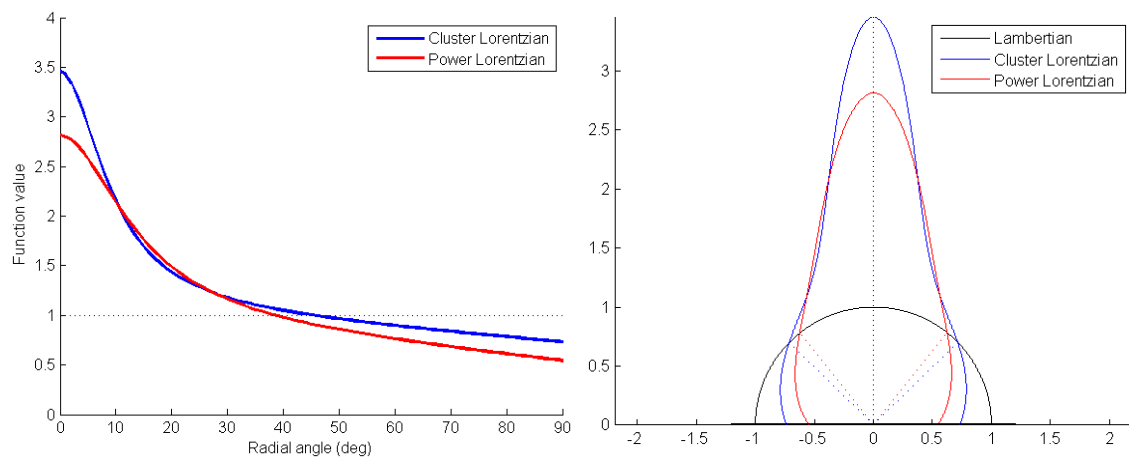


Figure 6.35: Characteristic curves for two Lorentzian models, generated from medians of parameters of a region of 15x15 pixels: (left) linear function of radial angle; (right) hemispherical function.

A detail of the specular colour image of the Damascene dish shows clearly that the specular highlights of the three metallic components carry the colour of the metal (Fig. 6.36 left). Scatter-plotting 10,000 points chosen at random by their colours and locations in RGB space shows that they lie in an oblate region around the long diagonal of the colour cube, i.e. the neutral axis. There is a surprising amount of colour variation for what appears to be a surface composed of only three materials, and the tonal variation from black to the lightest points at about 0.7 of the full range is continuous.

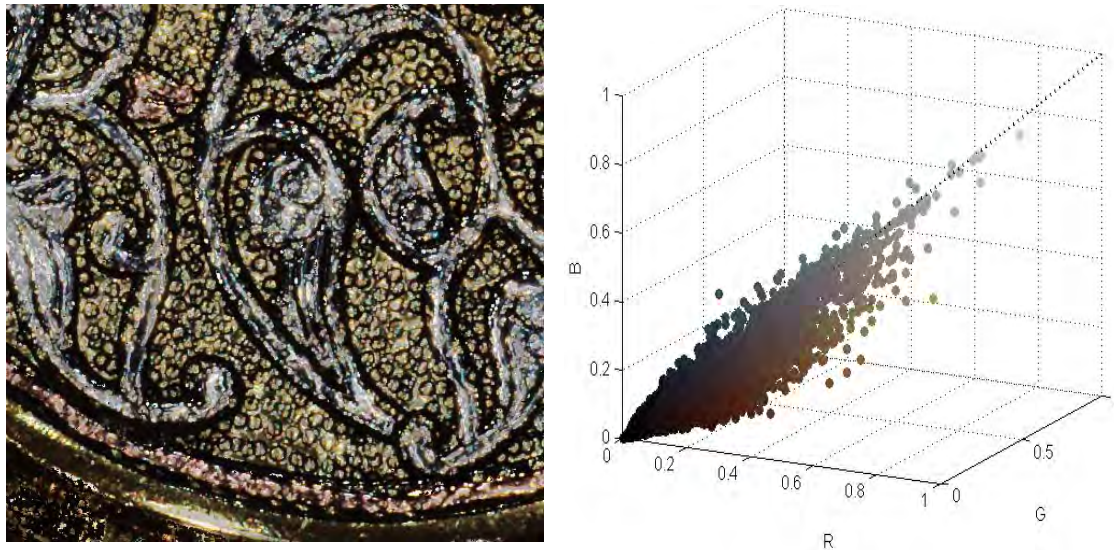


Figure 6.36: (left) Detail of specular colour image of Damascene dish, size 500x500 pixels; (right) scatter plot of 10,000 pixels in R,G,B colour space.

The pixel colours are converted from RGB via XYZ to CIELAB, assuming the sRGB colour space and the standard 2° observer with D65 white point. Plotting the same points on the a^* – b^* chromatic plane shows that the colours of the three metals, brass, copper and silver, have distinctly different hue angles (Fig. 6.36). This provides the opportunity to segment the image pixels into four categories, namely the three metals plus black. The simple way to do this is by hue angle around the centroid of the three category centres, where reddish colours with angles in the range -45° to $+70^\circ$ correspond to copper; the yellow-greenish colours with angles in the range 70° to 180° correspond to brass; and the slightly bluish colours with angles in the range 180° to -45° correspond to silver. A more effective method is to categorise each pixel by its nearest distance to one of the category colours for the three metals. The resulting ‘posterised’ image is equivalent to a K-means classification with four cluster centres (including black).

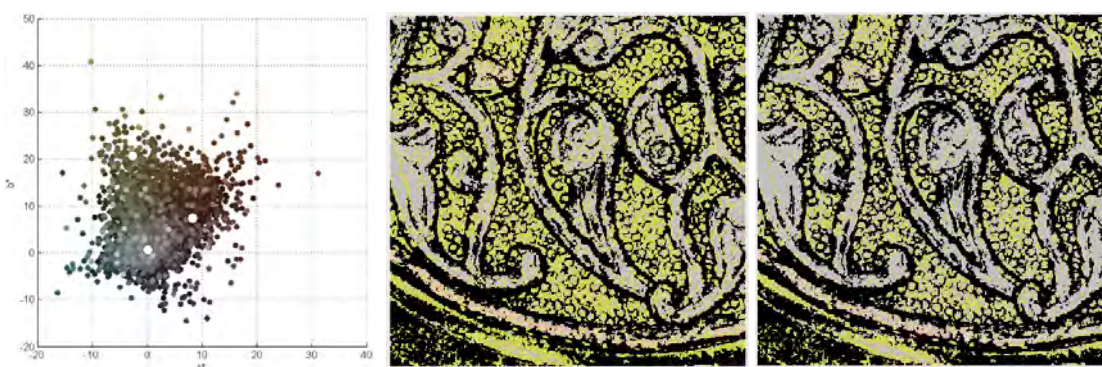


Figure 6.37: (left) 10,000 pixels plotted on the CIELAB a^* – b^* chromatic plane, with the centroids for the three metals; (centre) classification by hue angle; (right) classification by nearest colour.

The ability to classify different regions of a heterogeneous surface according to their gloss enables each region to be modelled and rendered in a different way. This is an important capability for objects that are made of multiple materials, such as inlays, and also for objects that were once homogeneous but have weathered variably across the surface.

In an attempt to differentiate the specular curves of the three metals in the Damascene dish, the map generated by image classification (Fig. 6.37 right) was used to select 500 random samples of each of the three metals. Curves were fitted by both the cluster and power Lorentzian models and plotted in superimposition (Fig. 6.38). All three sets of curves show a similar behaviour, with a few cases having peak values of very high amplitude, in excess of 500, but the majority much lower.

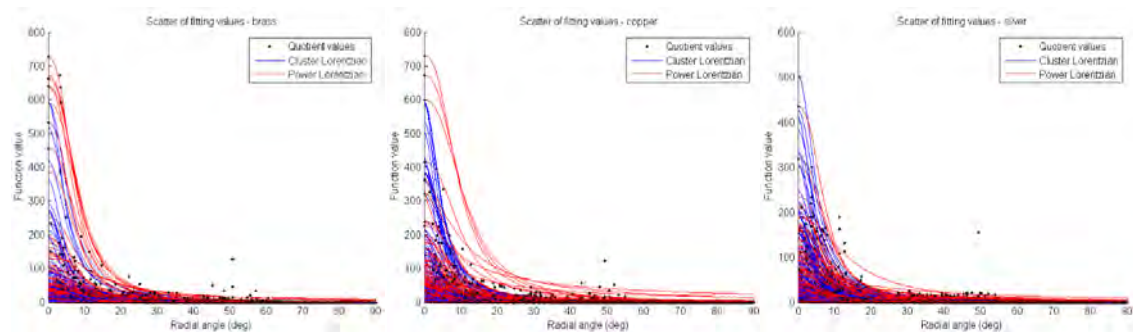


Figure 6.38: Specular curves fitted by the cluster and power Lorentzian models to 500 pixels each of the (left) brass, (centre) copper and (right) silver on the Damascene dish.

Taking the median of the parameter sets for each metal for those cases where the amplitude exceeds 15, and using the median parameter values in the Lorentzian model gives the indicative curves of Fig. 6.39. For both models the ordering of amplitude is the same: silver highest, copper second and brass lowest, but the relative differences are small and the variance of the curves for individual pixels is so great that these curves could not be used as a reliable diagnostic to determine the type of metal. The colour in the albedo and specular highlights is a much more reliable guide. All that one can say in this case is that the freshly polished silver is likely to be slightly brighter in the specular highlights than the other two metals.

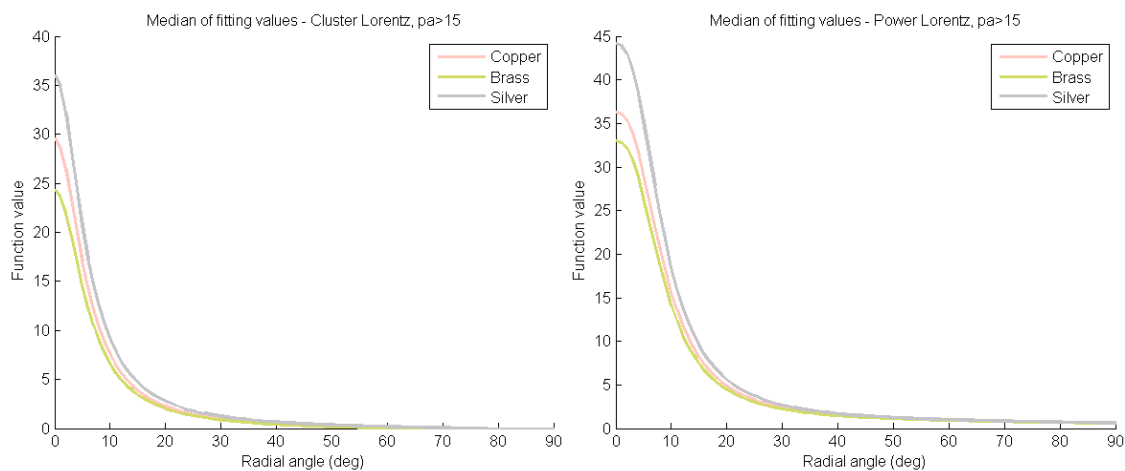


Figure 6.39: Characteristic curves for the three metals of the Damascene dish for peak amplitudes >15 generated by the two Lorentzian models: (left) cluster model; (right) power model.

6.4 Image Rendering

6.4.1 Lorentzian Reflectance Distribution

To visualise the effect of modelling the specular component by a Lorentzian function, a virtual sphere was rendered, using realistic parameters as if a real object in the dome. The sphere radius was 10 mm and the camera parameters set as for the Nikkor 105 mm macro lens. A single point source at one of the dome lamp coordinates was used to illuminate the sphere. Using the geometric transform method developed in Chapter 3, each pixel of the image was projected onto the baseboard and the intersection calculated of the ray from the front perspective centre of the lens with the top surface of the sphere. The incident intensity was modulated by the inverse square distance from the light source to the point of incidence. Initially the surface was treated as perfect Lambertian, with only diffuse and no specular component, and the reflected intensity was proportional to the cosine of the angle between the incident and normal vectors, giving the image in Fig. 6.40 (left).

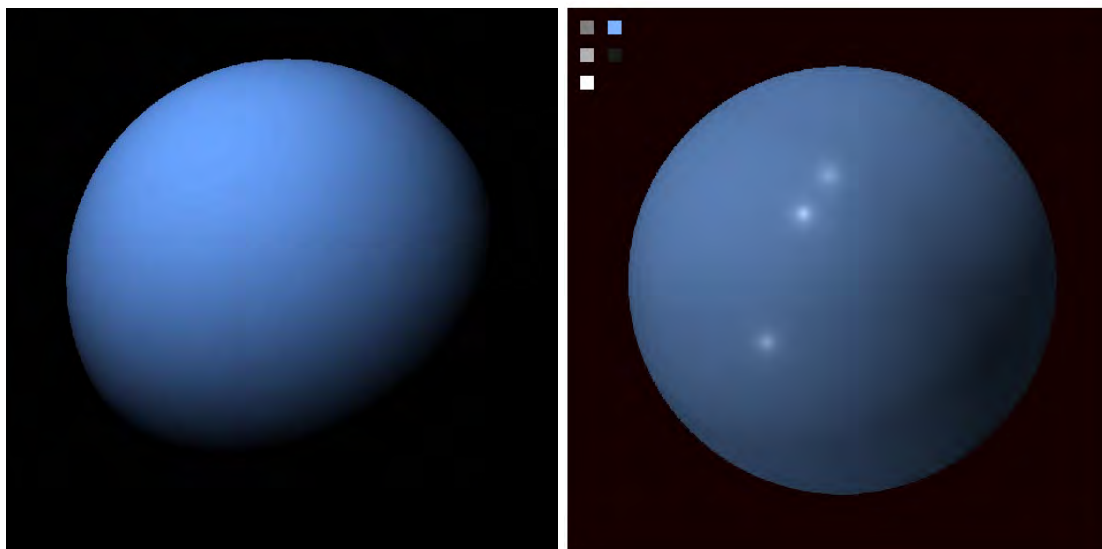


Figure 6.40: Rendered images of a blue sphere under white light: (left) diffuse component only; (right) with addition of specular highlights from three white lights of different intensities.

The specular component was then added by a Lorentzian distribution around the specular angle, assumed in the case of a perfect sphere to be double the normal angle. The radial scale factor (parameter s in Eq. 6-1) was set to 3. Fig. 6.40 (right) shows the result for the same blue surface of the sphere with summation of three lights, all achromatic but of different intensities. Each specular highlight appears as a bright spot surrounded by a hazy pool of scattered illumination (the phenomenon described by Hunter as lustre).

The rendering was repeated for a grey sphere (neutral reflectance factor = 0.5) with three light sources: a red light on the X axis, a green light on the Y axis, and a white light in Tier 1 of the dome at an elevation of $\sim 10^\circ$. The result is shown in Fig. 6.41 (left). The colours of the respective light sources appear both in the diffuse reflectance on the facing side of the sphere and in the highlight. The overall appearance is a convincing simulation of a semi-gloss ball.

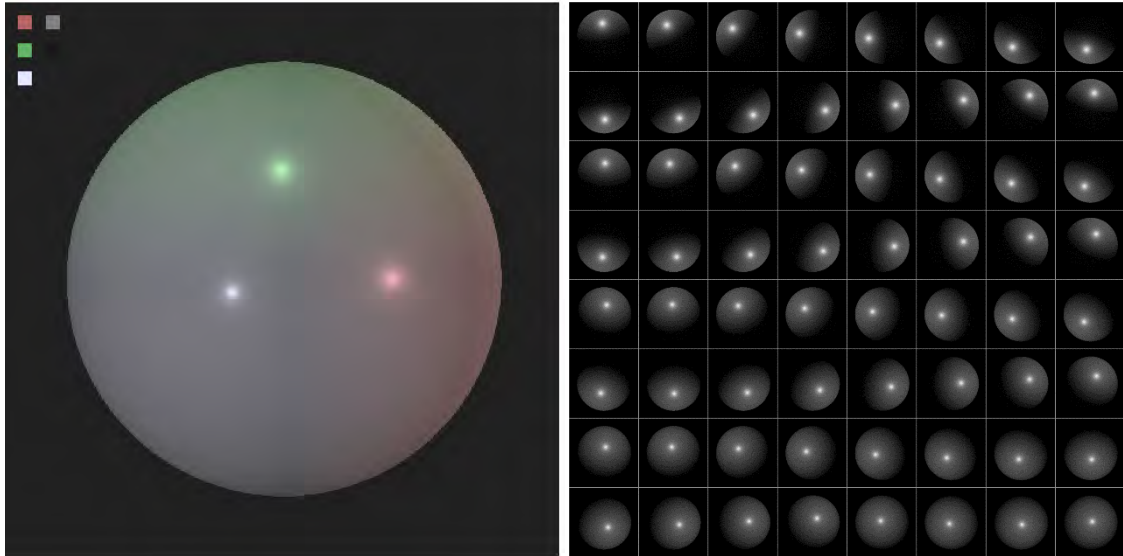


Figure 6.41: (left) Rendered image of a grey sphere under three lights (white, green and red); (right) grey sphere rendered for each of 64 lamps in dome.

The grey sphere was rendered with a broader specular highlight (scale factor = 10) for a single light source at every one of the 64 coordinates of the lamps in the dome. The results are shown as an 8x8 image array in Fig. 6.41 (right). The resulting set of images was used as a virtual test object for development and evaluation of algorithms for extracting normal and specular vectors and fitting specular distributions and for 3D reconstruction.

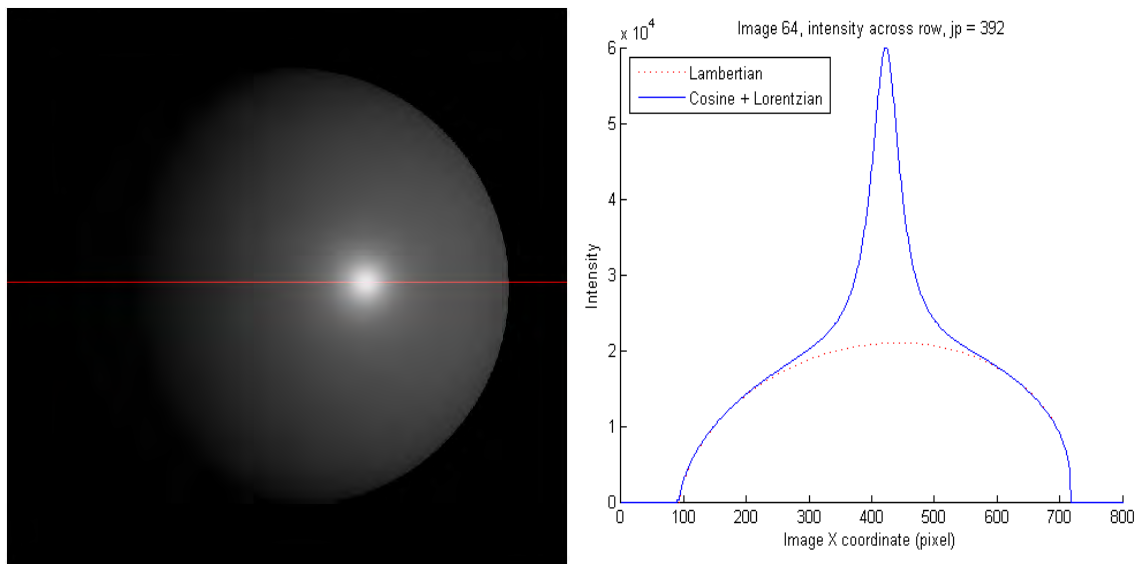


Figure 6.42: (left) Rendered image 64 from the set in 6.41; (right) intensity profile on a horizontal cross-section through the specular peak.

A cross-section through the last image in the sequence, corresponding to lamp 64, shows the combined intensity distribution vs the basic Lambertian distribution (Fig. 6.42 right). The shape is slightly asymmetrical because of the intensity correction for inverse square distance from the lamp to the surface of the sphere.

6.4.2 Image Computation

The complete model for rendering images under a single light source makes a sum of the diffuse and specular terms:

$$I = \mathbf{L} \cdot \mathbf{N} A_{rgb} + (f(\omega) - 1) S_{rgb} A_m / S_m \quad (6-12)$$

where A_{rgb} is the albedo colour, S_{rgb} is the specular colour, A_m is the monochrome albedo (weighted sum of the R,G,B channels of A_{rgb}), and S_m is the monochrome specular intensity.

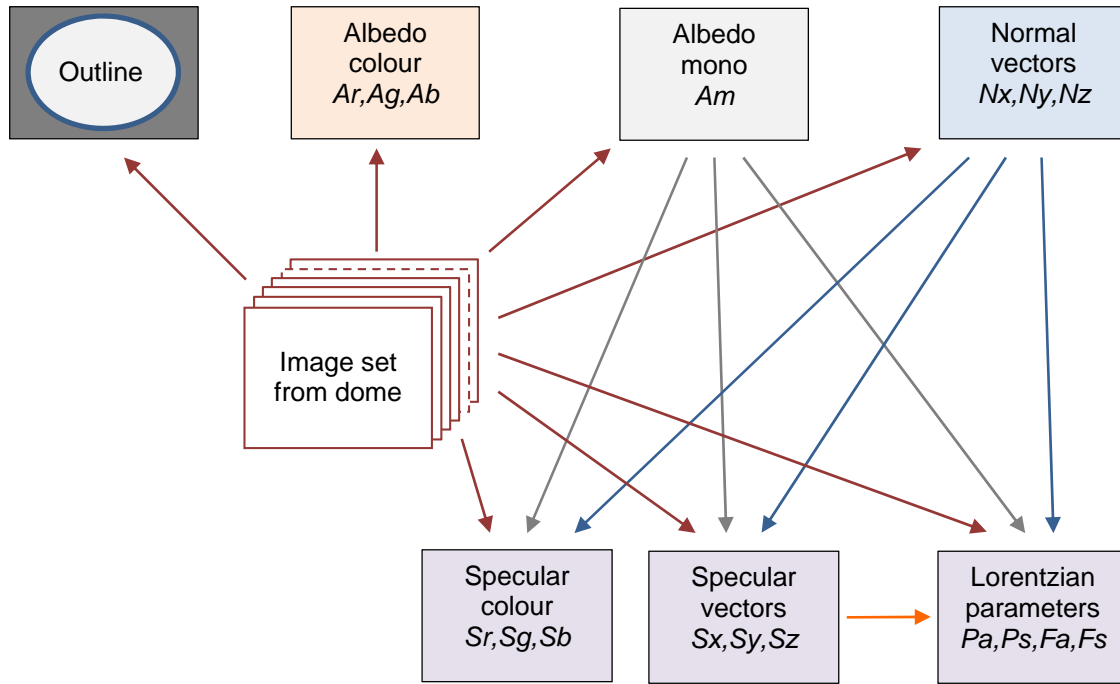


Figure 6.43: Process flow for extracting specular components from image set.

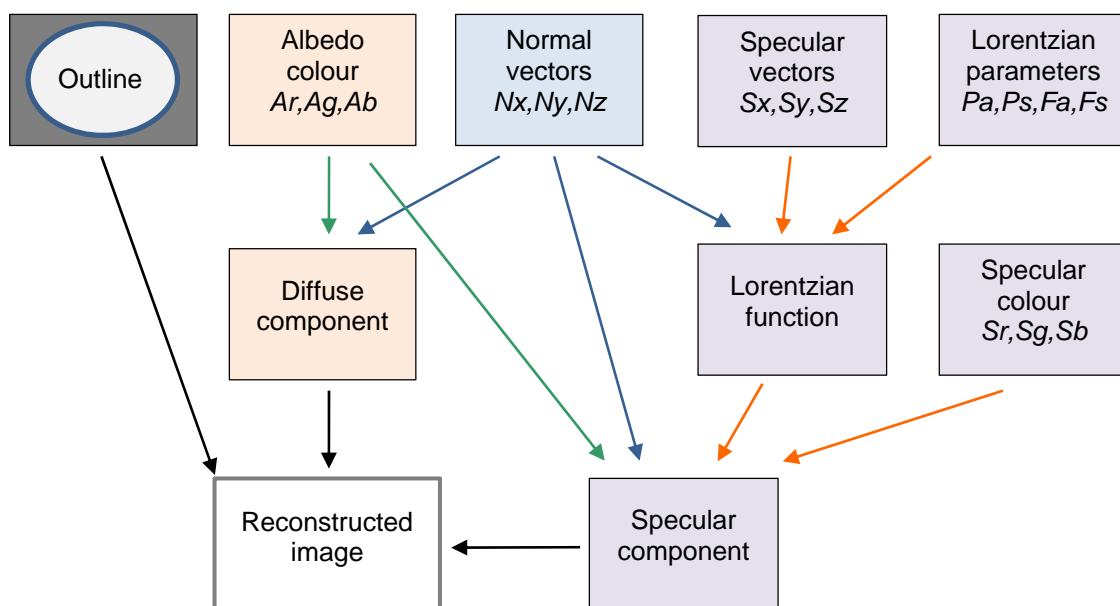


Figure 6.44: Process flow for reconstructing image from diffuse and specular components.

6.4.3 Image Viewer

In the course of this research, the author has developed a suite of software routines in Matlab as a viewer for the photographic images taken by the Nikon camera in the dome and the various derived image components. It fills a full HD display of 1920x1080 pixels, and the window layout is fixed and cannot be moved or resized (see screen grab in Fig. 6.45). The majority of the display area is given over to the image in a panel of 1570x1080 pixels. Down the right side is a panel of 350x1080 pixels containing the controls for the graphic user interface. The design is modal, so that the image is static and remains in a given state until a mouse button is clicked. Unlike the public-domain PTM and RTI viewers, there is no interactive change of the displayed image with mouse movement. The value of the viewer is as a visualisation tool for comparing different renderings of the object with the original photographic images.



Figure 6.45: Image viewer, showing the Roman medallion illuminated by lamp 48.

One of the most useful functions of the viewer is to select any one of the source images corresponding to the 64 lamps in the dome, and hence the appearance of the object when illuminated along this vector in the hemisphere. This is achieved by clicking on the circular target area in the upper right (Fig. 6.46) or by selecting the lamp number directly in the edit box or by clicking the up/down buttons.

The central area of the control panel has a set of radio buttons for display rendering mode selection. These enable the method of rendering the image from the given illumination direction to be set to one of six different modes:

- Original image, i.e. the photograph as taken by the camera;
- Lambertian image, assuming a perfect diffusely reflecting surface;
- PTM image, based on reconstruction from the six PTM coefficients;
- HSH image, based on reconstruction from the sixteen hemispherical harmonics;
- Lorentzian cluster, using the point cluster model for the specular component;
- Lorentzian power, using the 2D power model for the specular component.

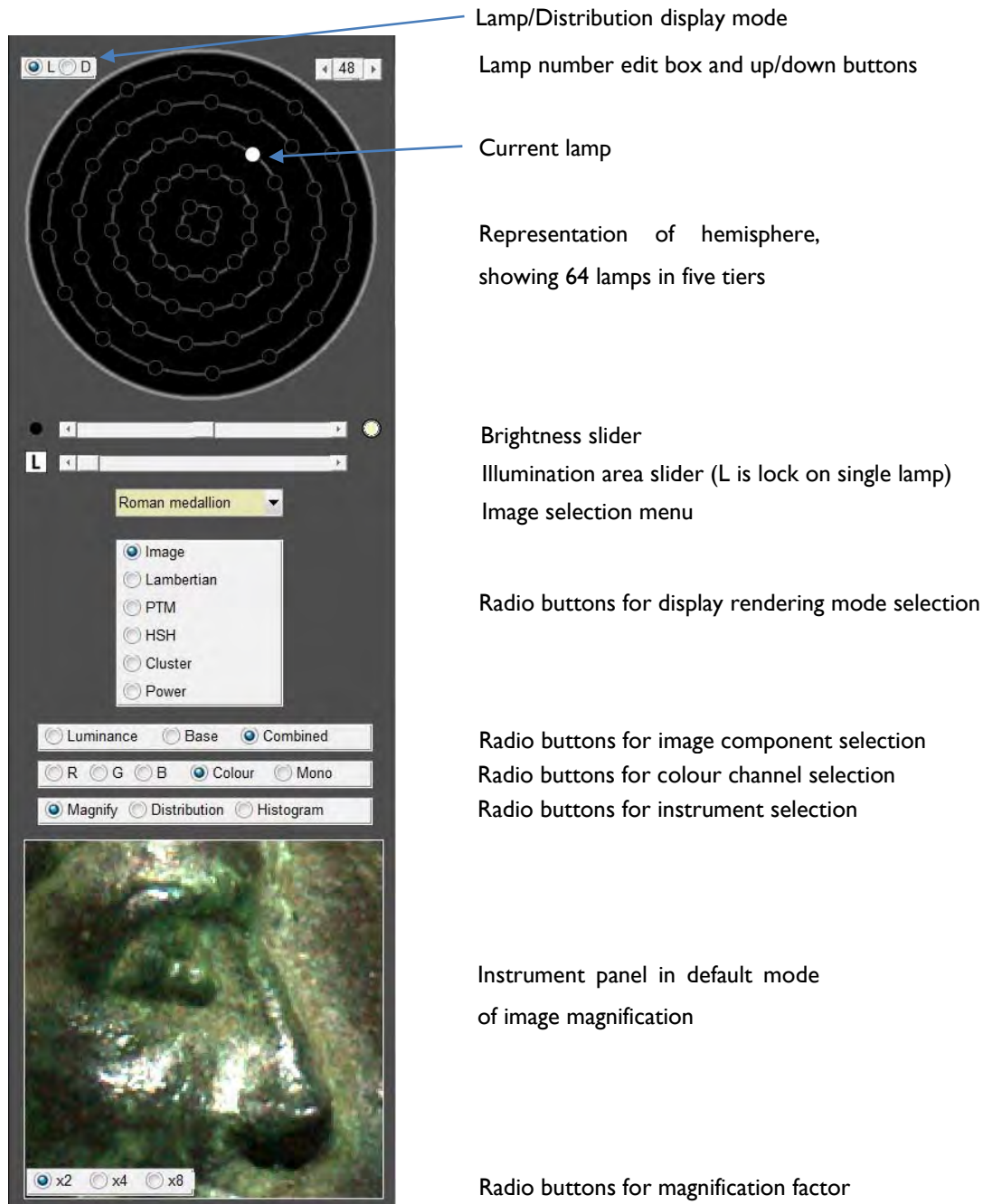


Figure 6.46: Control panel of image viewer, with explanations of functions.

All images are rendered for display by a technique borrowed from the graphic arts, consisting of gamma correction followed by histogram-based identification of black and white points (percentile values in tonal range) and mapping of these points to specified black and white values for the 8-bit display (code in Fig. 6.48). The darkest pixels, corresponding to self-shadowed regions of objects illuminated by low raking light, are excluded from the histogram. For a typical image the histogram has a long tail for the brightest values, especially for a glossy surface, and so the white point is selected to be a value well above the 99th percentile and is mapped to a much lower display value, here 70% of the display white level, to allow plenty of 'headroom' for the display of specular highlights (Fig. 6.47). The colour balance has previously been corrected for the R,G,B channels in the linear images (before non-linear gamma and tonal mapping) by applying scaling factors derived from the Nikon white point.

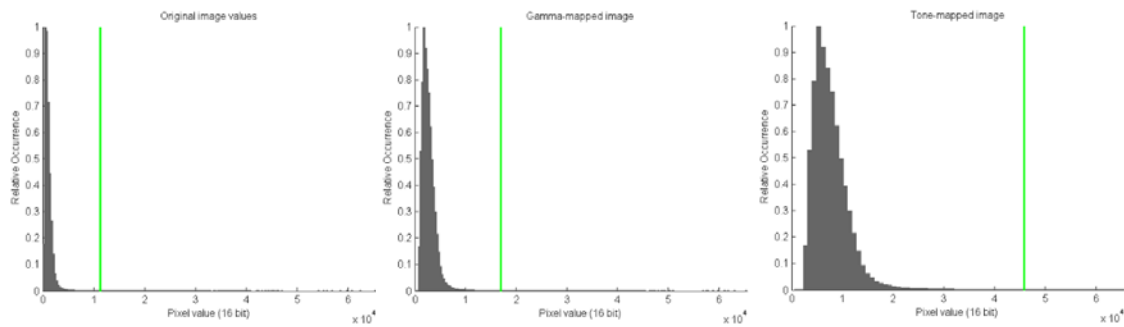


Figure 6.47: Histogram of green channel of image (16-bit range), with 99.8th percentile: (left) original image values; (centre) after gamma-mapping; (right) after tone-mapping.

The histogram for the green channel of the image of the medallion illuminated by lamp 35 is shown in Fig. 6.47. The original image (left) is very dark, with a median value of 910 and 99.8th percentile value of 11,340, indicated by the green line. But it does have higher intensity values, caused by sparkling highlights from the shiny regions of the surface, over the full range of the 16-bit pixel value scale from 0 to 65,535. The effect of gamma mapping (centre), in this case by an exponent of $1/1.3 = 0.77$, is to raise the dark tones, giving a median value of 2,441 and 99.8th percentile value of 16,999. Tone-mapping (right) further raises all the tonal values, giving a median value of 7,051 and 99.8th percentile value of 45,875, which is the set point of 70% of white. By this means, images captured at a low exposure level to capture the specular highlights and processed in a linear signal domain can be optimised for realistic display.

```

im_gamma = 1/1.3; % image gamma correction
lfac = Light_Intensity/100; % light intensity scaling
im_pcb = 3; % percentile value for black
im_pcw = 99.8; % percentile value for white
im_black = 0.05*maxval; % output level corr to black
im_white = 0.7*maxval; % output level corr to white

RGBbuf(RGBbuf<0) = 0; % clip any negative values
Ibuf = maxval*(RGBbuf/maxval).^im_gamma; % gamma-map whole image
Rbuf = Ibuf(:,:,1);
Gbuf = Ibuf(:,:,2); % separate out R,G,B components
Bbuf = Ibuf(:,:,3);
Ilum = 0.3*Rbuf+0.6*Gbuf+0.1*Bbuf; % form luminance (approx)
T = (Imap>0)&(Ilum>500); % use only pixels inside outline
len = sum(sum(T)); % number of pixels selected
ss = 5; % sub-sampling factor
RT = Rbuf(T); GT = Gbuf(T); BT = Bbuf(T); % 1-D pixel vectors
rbw = prctile(RT(1:ss:len),[im_pcb im_pcw]); % R percentiles
gbw = prctile(GT(1:ss:len),[im_pcb im_pcw]); % G percentiles
bbw = prctile(BT(1:ss:len),[im_pcb im_pcw]); % B percentiles
im_bl = min([rbw(1) gbw(1) bbw(1)]); % black point (colour)
im_wh = max([rbw(2) gbw(2) bbw(2)]); % white point (colour)
cfac = (im_white-im_black)/(im_wh-im_bl); % scaling factor
Tbuf = im_black+lfac*cfac*(Ibuf-im_bl); % tone map whole image
for k = 1:3
    Dbuf(:,:,k) = uint8(Imap.*Tbuf(:,:,k)); % zero outside outline
end

```

Figure 6.48: Fragment of Matlab code to render image for display.

All images for the viewer have been previously sized or cropped to fit into the 1570x1080 display window, so that one or other dimension fills the rectangle, depending on image aspect ratio. No resizing of the images is done within the viewer. All of the following image components are pre-computed and stored for each subject (see Fig. 6.44 and 6.45):

- 64 photographic images from camera (RGB)
- Outline map
- Albedo colour and normal vectors
- Albedo monochrome
- Specular colour and specular normal vectors
- Specular cluster model coefficients (4 values)
- Specular power model coefficients (3 values)
- PTM RGB and six luminance coefficients
- HSH RGB and sixteen luminance coefficients

These images are all of the same dimensions and in pixel register. The total number of components per subject is thus $64 \times 3 = 192$ photographic image frames plus 49 derived frames = 241 frames in total. The format for all stored images is 16-bit TIFF, in which values are assumed to be unsigned integers. For the normal and specular vectors a standard mapping is assumed from the numerical range $[-1, +1]$ of normalised vectors into encoded integer range $[0, 65535]$, implemented by the code in Fig. 6.49. The quantising error caused by rounding of the lsb is thus approximately $\pm 1.5 \times 10^{-5}$. Note that for convenience the Nz component of the normal vectors is converted by the same formula as Nx and Ny, even though its range is $[0, +1]$.

```
Nim = uint16(32767*(Normals+1)); % convert normals to image file
Normals = (double(Nim)/32767)-1; % convert image file to normals
```

Figure 6.49: Fragment of Matlab code to convert normals to and from 16-bit integer format.

Albedo values are always positive but the range may vary from one image to another. They are represented in memory as single-precision floats and in file as unsigned 16-bit integers. The range is stored in the first two 16-bit pixels (i.e. 4 bytes) in the top left corner of the image, converted by the `typecast` function, as shown in Fig. 6.50. For colour images, the process is repeated for each of the three R,G,B channels. This is a kind of metadata that is carried along with the image, but without the need for a special header or associated file.

```
% Write scaled mono albedo image to file
A_scale = max(Albedom); % range 0-max
Al6im = uint16(maxval*Albedom/A_scale); % convert to 16-bit
Al6im(1,1:2) = typecast(single(A_scale), 'uint16'); % encode single
afilename = fullfile(rootdir, adir, 'Albedo mono.tif');
imwrite(Al6im, afilename, 'tif');

% Read normals and albedo
Al6im = imread(afilename, 'tif');
A_scale = typecast(Al6im(1,1:2), 'single'); % scaling to 2 pixels
Al6im(1,1:2) = 0; % delete from image
Albedom = single(Al6im)*A_scale/maxval; % restore original
```

Figure 6.50: Fragment of Matlab code to convert monochrome albedo to and from 16-bit integer.

Coefficient values, as in the PTM and HSH models, can be both positive and negative. They are represented in memory as double-precision floats and in file as unsigned 16-bit integers. The offset and range are each stored in four 16-bit pixels (i.e. 8 bytes) in the top left corner of the image, converted by the `typecast` function, as shown in Fig. 6.51. For colour images, the process is repeated for each of the three R,G,B channels.

```
% Write scaled coefficient image to file
cmin = min(min(Coeff)); % min value = offset
cmax = max(max(Coeff)); % range
crange = cmax-cmin; % scale image to 16 bits
Cim = uint16(maxval*(Coeff-cmin)/crange); % double offset 8 bytes
Cim(1,1:4) = typecast(cmin,'uint16'); % double range 8 bytes
Cim(2,1:4) = typecast(crange,'uint16'); % write to file
imwrite(Cim,cname,'tif');

% Read scaled coefficient image and reconstruct values
Cim = imread(cname,'tif'); % read coefficients
offset = typecast(squeeze(Cim(1,1:4)), 'double'); % extract offset
range = typecast(squeeze(Cim(2,1:4)), 'double'); % extract range
Cim(1:2,1:4) = 0; % delete from image
Coeff(:, :, n) = offset+range*double(Cim)/maxval; % reconstruct values
```

Figure 6.51: Fragment of Matlab code to convert coefficient values to and from 16-bit integer.



Figure 6.52: Six renderings of the Roman medallion illuminated by lamp 48: (top row) Lambertian, PTM, HSH; (bottom row) photographic image, cluster Lorentzian, power Lorentzian.

A feature of the viewer is the ability to compare different renderings of an object illuminated from the same direction, by selecting the radio buttons for display rendering mode in the centre of the control panel. Fig. 6.52 shows six renderings of the Roman medallion, illuminated by lamp 48 in Tier 3 of the dome, at an azimuthal direction from the north-east and elevation of 51°. A detail of 150x150 pixels square of each image, showing the region of the emperor's eye and nose, is in Fig. 6.53. The Lambertian rendering (top left) is flat and devoid of any specular

reflection, giving the impression that the object is covered in a matte powder or a product of oxidation or corrosion. Its modulation of the shading is proportional to the cosine of the angle between the surface normal and the illumination vector (the latter assumed to be constant over the whole surface). The PTM rendering (top centre) is also rather flat but less powdery in appearance, and the relief on the surface is better defined. The HSH rendering (top right) shows more shine on the highlight areas and also greater image contrast with the shaded areas tending toward black. Compared with the original photographic image (lower left), the two new renderings using a Lorentzian model for the specular component are much closer. The cluster model (bottom centre) emphasises the glitter of individual pixels, giving a slightly grainy appearance. The power model (bottom right) gives a smoother rendering of the specular highlights without losing the spatial resolution of the surface features.



Figure 6.53: Detail of 150x150 pixels from the images of Fig. 6.52, showing the same six renderings.

Another feature of the viewer is the ability to display the intensity of the reflectance distribution for a single pixel at all points over the hemisphere. This is achieved by clicking the 'D' radio button at the top left corner of the control panel (Fig. 6.46), and the distribution is shown in the circular window instead of the lamp layout diagram. Fig. 6.55 shows the distributions at the single pixel (668,197), located at the edge of the front of the emperor's hair, above the eye. The first three renderings (top row) all exhibit the same hue, arising from the albedo for the Lambertian and the RGB base image for the PTM and HSH respectively, for which lightness is modulated by the angular distribution of luminance. For the original image (lower left) the starting point is the set of RGB values at the single pixel coordinates for each of the 64 lamps. To make a continuous

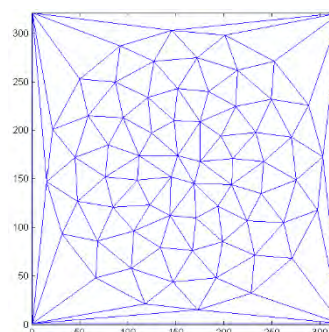


Figure 6.54: Delaunay triangulation of X,Y coordinates of dome lamps.

distribution that fills the whole hemisphere, a triangular mesh is first generated by Delaunay triangulation over the plane of the X,Y coordinates of the 64 lamps plus the four corner points of the rendering plane (Fig. 6.54). Within each triangle bilinear interpolation is then used to compute a weighted sum of the RGB values at each of the three corners. The resulting distribution shows the degree of natural variation in both intensity and colour of the object surface over the hemisphere of illumination angles. The two specular renderings (Fig. 6.55 bottom centre and right) show a 'base colour' with the same distribution as the Lambertian, with a superimposed Lorentzian highlight in the specular colour (in this case white). The peak generated by the cluster model is here more narrow and focussed than the peak generated by the power model.

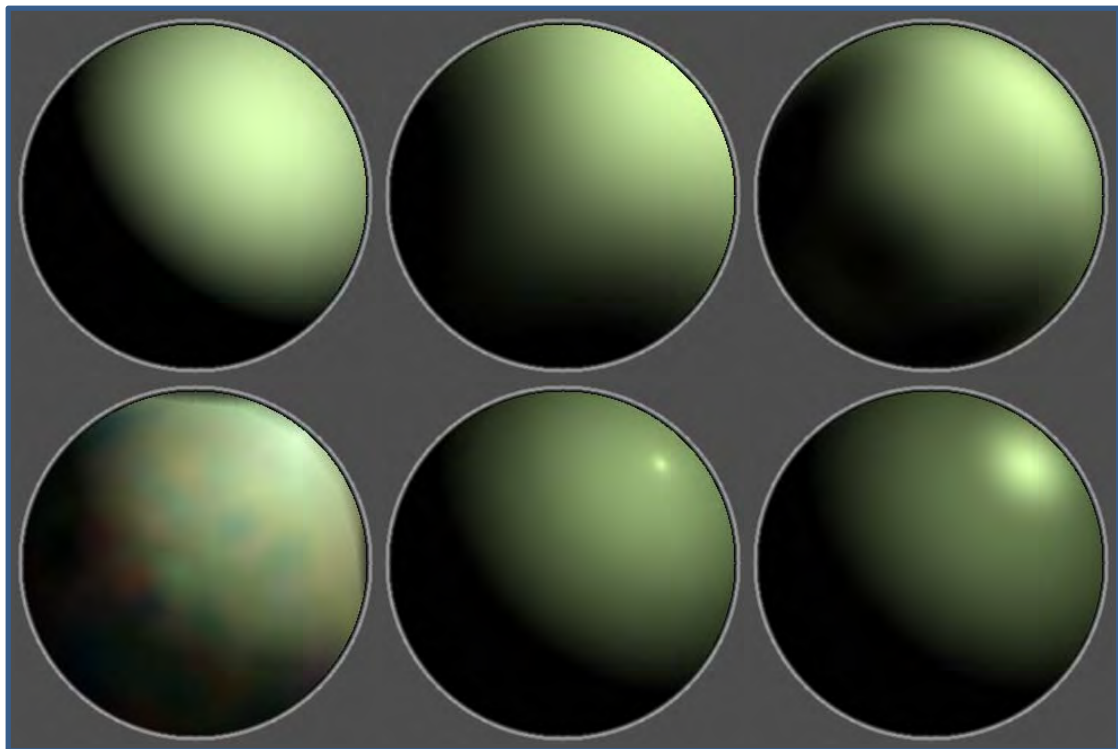


Figure 6.55: Intensity distribution at a single pixel over the hemisphere of illumination angles: (top row) Lambertian, PTM, HSH; (bottom row) photographic image, cluster Lorentzian, power Lorentzian.

Another feature of the viewer is the ability to select multiple light sources. By using the right mouse button to click on the circular lamp selection panel, any number of lights from 1 to 64 can be selected in any combination. The image is rendered individually for each light source selected and the results averaged. This simulates the positioning of lamps around the object in studio copy lighting. In the example in Fig. 6.56 four lamps have been selected: a group of three together at the north-east (numbers 47, 48 and 60 in tiers 3 and 4) to illuminate the emperor's face with a broader, beam and a single lamp at low elevation (number 3 in tier 1) at the north-west to provide a narrow beam to highlight the hair. Again the power Lorentzian model provides a good likeness of the original image (Fig. 6.57), despite the more complicated illumination geometry.

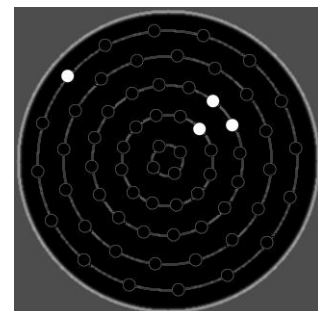


Figure 6.56: Selecting positions of four lamps.



*Figure 6.57: Two renderings of the Roman medallion illuminated by four lamps:
(left) photographic image, (right) power Lorentzian.*

Chapter 7 – Colour

Mere colour, unspoiled by meaning,
and unallied with definite form,
can speak to the soul in a thousand different ways.

- Oscar Wilde (1891)
The Critic as Artist, Part 2.

There are two realms of colour: the physical world and the mind of the observer. The first is objective, defined by measurable quantities such as time, power, wavelength and reflectance. The second is subjective, described by words and dependent on adaptation, judgement and emotion. These two realms are interdependent and both are necessary for colour to be meaningful: without the real-world stimulus the observer would have only dreams and imagination; without an observer the world would never be seen. As Newton (1704) famously said: “And if at any time I speak of Light and Rays as coloured or endued with Colours, I would be understood to speak not philosophically and properly, but grossly, and according to such Conceptions as vulgar People in seeing all these Experiments would be apt to frame. For the Rays to speak properly are not coloured. In them there is nothing else than a certain Power and Disposition to stir up a Sensation of this or that Colour...”.

So the challenge of displaying objects in realistic colour is to generate the proper stimuli so that the observer has the impression that he or she is seeing the same colour as the actual object, or at least something that conforms to the memory of how such an object should appear. This chapter, however, deals hardly at all with those aspects of colour science relating to human visual perception. Instead it takes a more metrological approach, focussing on the physical world. It commences with a series of studies on characterisation of the spectral sensitivity of the camera (Section 7.1), then the colorimetry of colour laser scanners and choice of laser wavelengths (Section 7.2). The reflectance spectra of a number of datasets are analysed by their principal components (Section 7.3). Finally a new method is presented for using very large datasets of synthetic for optimal loading of a 3D lookup table (LUT) to transform R,G,B signals from a camera or laser scanner in a way that minimises colorimetric error (Section 7.4).



Figure 7.1: Front, top and rear views of the Nikon D200 digital camera, with an 18-70 mm zoom lens.

7.1 Spectral Characterisation of Camera

The Nikon D200 (Fig. 7.1) used throughout the study is a good quality mid-range digital camera introduced in 2006 for the ‘prosumer’ market. It produces 12-bit colour images of 3900x2600 pixels in the classic photographic 3:2 aspect ratio. The CCD sensor dimensions of 24.6 by 15.8 mm are slightly less than the APS format (25.1 by 16.7 mm), enabling the more compact Nikkor DX-Series lenses to be used. Because the sensor is only 2/3 the linear dimensions of the classic 35 mm film format (36 by 24 mm), the lenses have an effective magnification factor of 1.5 relative to standard lenses of the same focal length.

Raw images from the camera are saved as NEF files, which when converted to 16-bit TIFF format by the DCRAW utility have been shown to give very good linearity with reflectance (MacDonald, 2010). The R,G,B colour channels are formed by a Bayer mask layered over the sensor and the sensitivity of each channel to wavelengths throughout the visible spectrum is a product of the transmittance of the filters and infra-red cutoff filter and the sensitivity of the silicon sensor. Because Nikon provides no data on colour sensitivity, it has been necessary to characterise the response of the three camera channels as a function of wavelength. This section describes four different methods that were employed in the laboratory, all working from basic physical principles.

7.1.1 Monochromator

The response of the three channels of the Nikon D200 camera was measured directly by exposure to a series of near-monochromatic lights, using equipment in the Colour Science research centre at the University of Leeds. The Bentham FSGM-150 instrument passes white light from a pulsed xenon source through a monochromator to produce light of 10 nm bandwidth (at half maximum amplitude), variable over a wavelength range from 380 to 780 nm. The light was transmitted through an optical fibre to one port of an integrating sphere, and was imaged through the second port (Fig. 7.2). The wavelength of the output light was controlled by the voltage from an external stabilised power supply, calibrated by a set of measurements by a Minolta CS-1000 spectroradiometer. The radiant power was also calibrated from a series of measurements by the CS-1000 throughout the spectrum (Fig. 7.3), and the emissive power determined by integrating the output distribution at each wavelength.



Figure 7.2: (left) Setup on laboratory bench for measurement of spectral response of the Nikon D200 camera; (right) the camera's view of 540 nm monochromatic light through one port of the integrating sphere.

To minimise the influence of ambient light, all images were captured in a darkened room. For each wavelength setting an image of the illumination seen through the port of the integrating sphere was taken by the camera, with shutter speed of 4 seconds, sensitivity of ISO100 and lens aperture $f/5$. The images were linearised via DCRAW and a region of 200x200 pixels cropped from each image within the area of the sphere port.

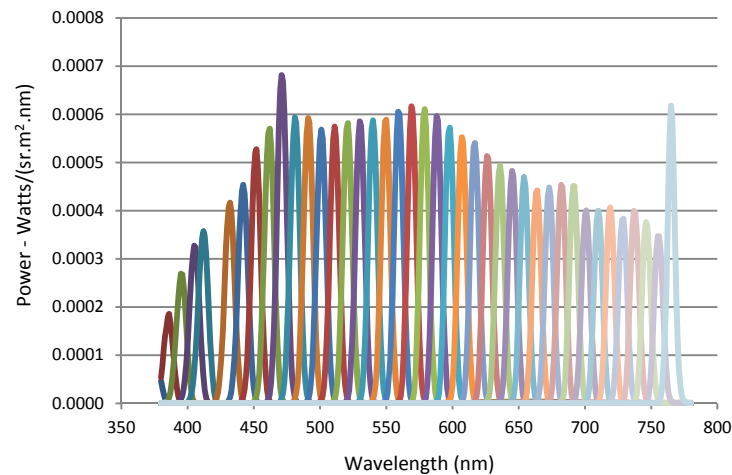


Figure 7.3: Radiant power output of monochromatic light source for each wavelength setting.

The 16-bit pixel values in the cropped area of each image were averaged to give the 'raw' camera response of each channel as a function of wavelength (Fig. 7.4 left). They were then corrected by dividing at each wavelength by the integrated power of the lamp (Fig. 7.4 right), yielding sensitivity curves for the three channels of the camera (Fig. 7.5).

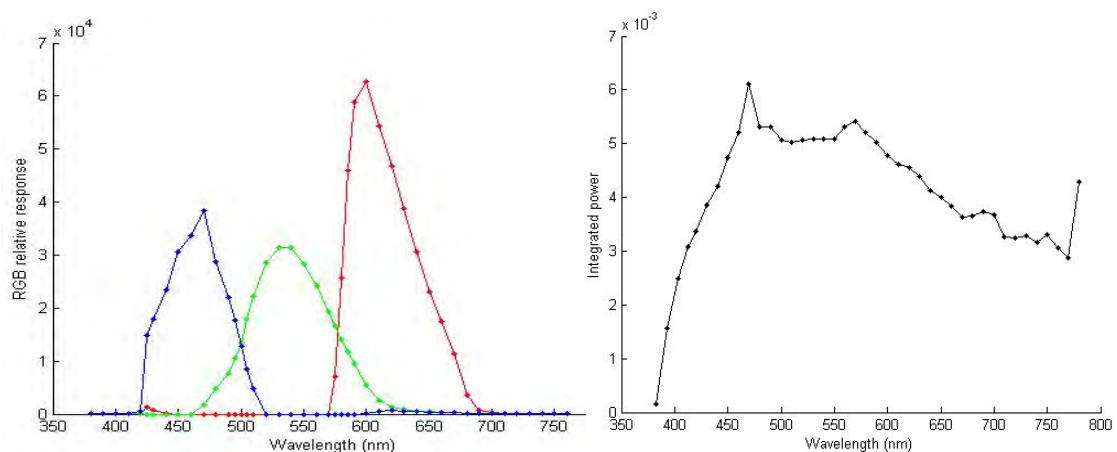


Figure 7.4: (left) Uncorrected values for sensitivity of R,G,B channels of camera, from averaged pixel values in linearised images; (right) radiometric correction factor derived from measurements of source.

The curves for the three channels are broad and overlapping, giving good coverage of all parts of the visible spectrum, with an average bandwidth at half maximum amplitude of approximately 80 nm. It is unclear why the peak sensitivity of the red channel is nearly double that of the green and blue channels, but probably it is the result of the greater sensitivity of the silicon sensor at longer wavelengths. (Because of the RAW image capture method, no white balance scaling has been performed by the camera on the signal values.)

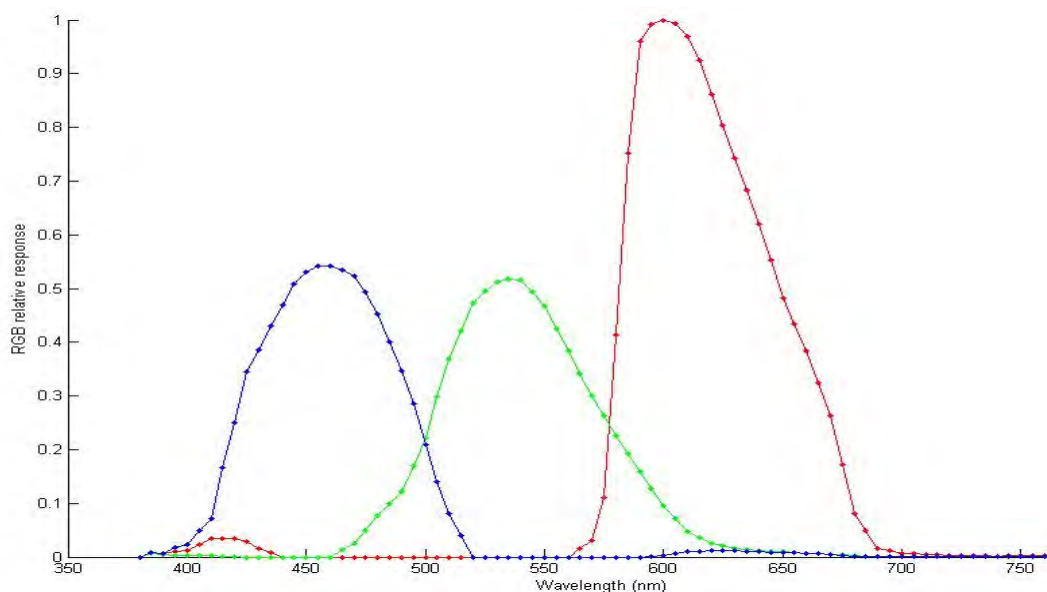


Figure 7.5: Normalised sensitivity vs wavelength at 5 nm intervals for the red, green and blue channels of the Nikon D200 camera.

7.1.2 LED Spectral Target

A prototype LED colour target (Hawkins, 2008) was tested for a quick ‘one shot’ characterisation of the camera’s spectral sensitivity (Fig. 7.6). It consists of 36 luminous circular patches in a box with an overall area of approximately 10x10 cm. Each patch has a narrow-band interference filter in front of an LED source, with an approximate bandwidth of 10 nm. The overall ensemble samples the spectrum from 380 to 730 nm at approximately 10 nm intervals. It therefore enables the spectral sensitivity of each channel of the camera to be determined in a similar manner to the use of a variable monochromatic source, but with a single exposure by the camera.

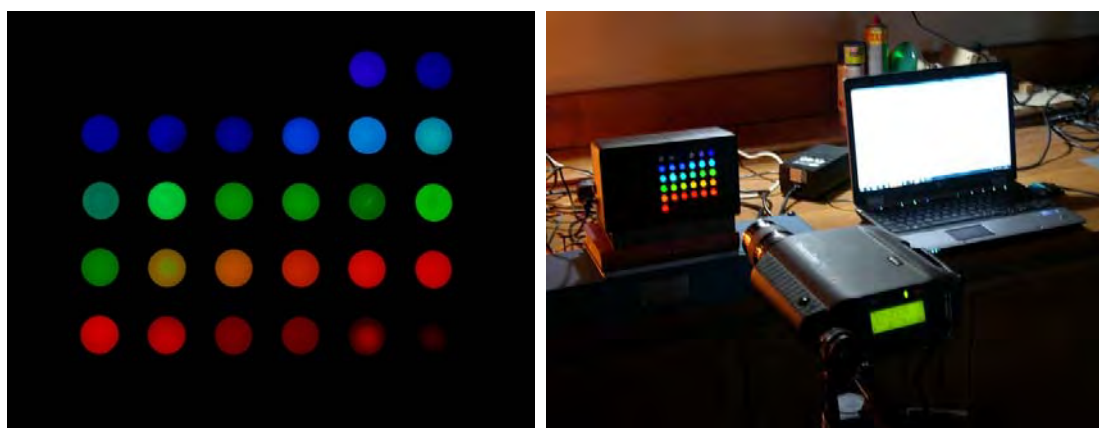


Figure 7.6: (left) LED target captured as image by Nikon D200 camera; (right) laboratory setup for measurement by PR-650 (foreground).

The power of each patch of the target was measured with a calibrated PR-650 spectroradiometer, giving the radiance at 4 nm intervals over the range 380-780 nm in a vector of 101 values per patch. Considerable variation was observed in both bandwidth and amplitude across the different patches (Fig. 7.7 left).

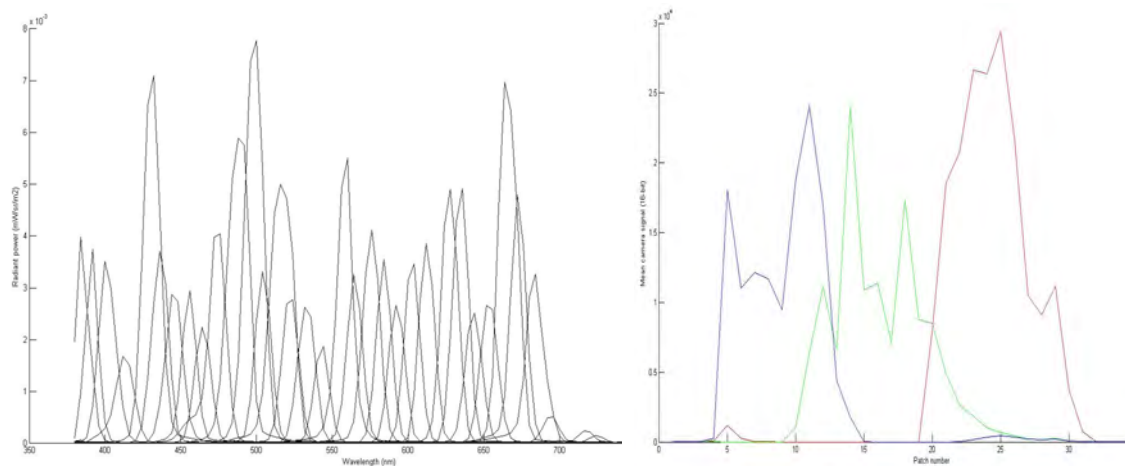


Figure 7.7: (left) Radiant power of each LED target patch, measured by PR-650; (right) Mean pixel value (16-bit) in R,G,B channels vs patch number.

The target was captured by the Nikon D200 with the 105 mm lens at $f/8$ at $1/4$ sec, at a distance of approximately 80 cm. The image was saved in raw format (NEF file) and converted via DCRAW to 16-bit linear TIFF. The mean values of the R,G,B channels were determined from an area of 101×101 pixels at the centre of each patch in the image. Fig. 7.7 (right) shows the mean value of each channel plotted against patch number. The irregularity caused by the variation in power of the patches is evident.

Correction factors could have been derived from the maximum power of each lamp, assuming that they were evenly spaced at 10 nm intervals. Improved factors were determined from the measured power for each patch in the target as follows: the 'true' wavelength was calculated as the centroid of the curve, using a weighted summation over all wavelengths, with the amplitude as a linear weighting coefficient; the total (integral) power was calculated as the area under the curve, by summation over all wavelengths (Fig. 7.8 left).

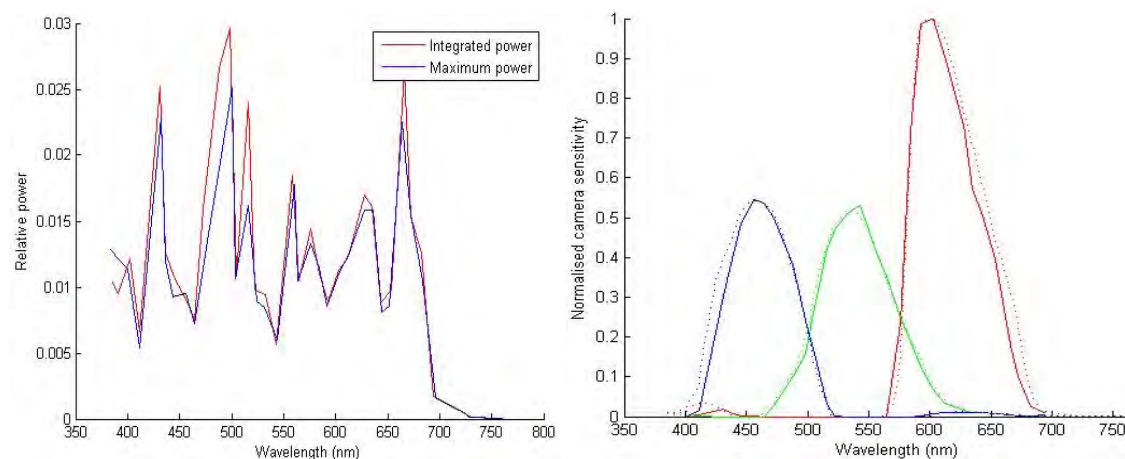


Figure 7.8: (left) Integrated power at centroidal wavelengths vs maximum power at nominal wavelengths; (right) Corrected sensitivity of R,G,B channels as functions of wavelength for LED target (solid lines) vs sensitivity calculated from variable monochromator (dotted lines).

Dividing the pixel values by the respective radiant patch power and plotting against the correct centroid wavelength gave the spectral sensitivity of each channel of the camera (Fig. 7.8 right). The three curves were normalised to set the maximum value (red channel) to 1.0.

The results show good agreement with the characterisation by the Bentham variable monochromator (Sec. 7.1), except for a sliver of difference on the short-wavelength side of blue and the long-wavelength side of red. This could perhaps be explained by a lower response from the camera than expected for the blue patches in the upper rows of the chart and the red patches in the lower rows of the chart, because of vignetting of intensity by the lens and/or angular attenuation of the light transmitted through the interference filters. Both effects could be mitigated by imaging the target from a greater distance so that it subtended a smaller angle of view to the camera.

7.1.3 Multispectral Filters on Illumination

A set of 16 transmission filters from a multispectral image capture system was used, with central wavelengths at intervals of 20 nm throughout visible spectrum from 400 to 700 nm inclusive, each with a bandwidth of approximately 20 nm. The transmittance of each filter was measured directly with an Ocean Optics HR2000+ spectrometer (Fig. 7.9 left). Light from a tungsten halogen source (left) was passed through a diffuser (centre) and two measurements of the power were made, one without and one with the filter in the light path. The transmittance factor was computed as the ratio at each wavelength of the two measurements, at intervals of ~ 0.5 nm, and a 5-point median filter applied to each curve to reduce noise (Fig. 7.9 right).

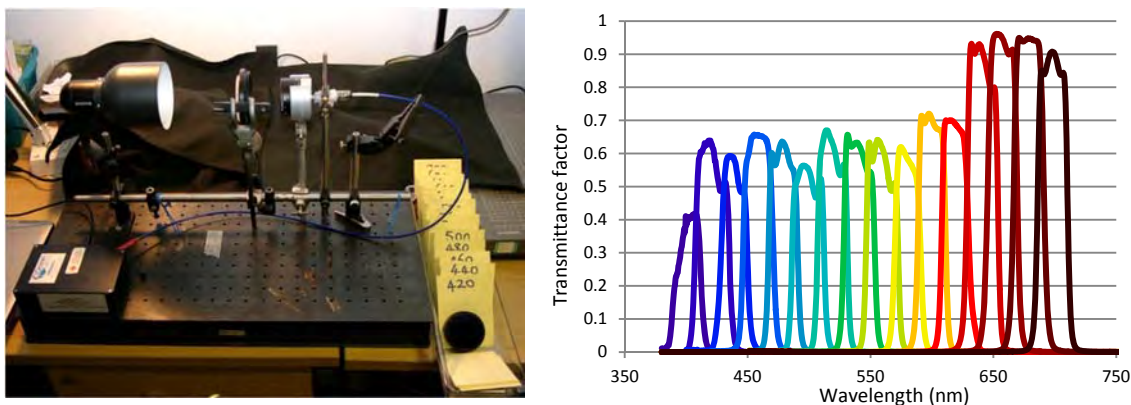


Figure 7.9: (left) Lab bench set up for measurement of filter transmittance; (right) calculated transmittance factors for 16 filters.

The Nikon D200 camera with zoom lens set to 55 mm focal length and aperture $f/2.8$ was mounted in its usual position at the top of the dome and an X-rite ColorChecker White Balance target placed on the baseboard. The illumination was generated by a Balcar xenon arc studio flash head inside a customised snoot with a transmission filter mounted at the front. The light passed through a 25 cm tube to collimate the beam, which was directed through one of the portholes at the side of the dome, at an elevation of $\sim 50^\circ$ (Fig. 7.10 left) onto the white card below. The fibre optic probe of the Ocean Optics HR2000+ spectrometer was fixed in position alongside the lens, parallel to the optical axis, pointing at the centre of the intensity distribution on the white card (Fig. 7.10 right). A sequence of 16 images was captured, one for each filter, and the spectrum of the reflected illumination was simultaneously recorded by the spectrometer. The advantage of this method is that it is not necessary to know the spectrum of either the light source or the white card, because the spectrometer measures the same radiant spectrum that is entering the lens of the camera.

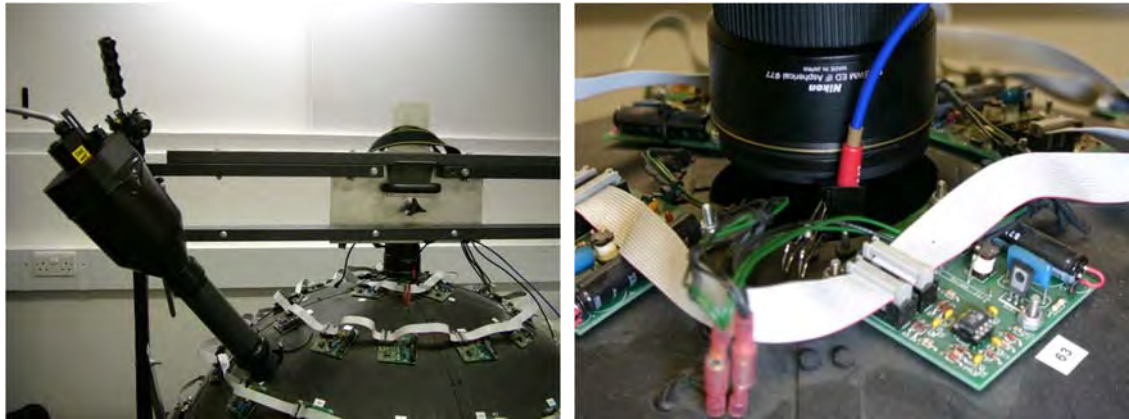


Figure 7.10: (left) Illumination setup with studio flash directed through narrow-band filter into dome; (right) Optic fibre probe for spectrometer inserted beside lens, pointing toward white target.

The camera responses were determined as the mean values in each of the R,G,B channels in a region of 200x200 pixels near the position of maximum intensity of the beam profile in the image. Repeating this for each of the 16 images corresponding to the 16 narrow-band transmission filters gave 16 points across the spectrum (Fig. 7.11).

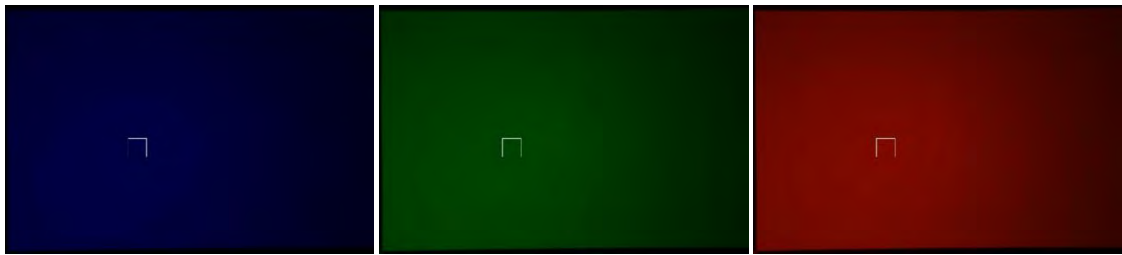


Figure 7.11: Three images captured through the filters at wavelengths of 460 nm, 540 nm and 620 nm. The white square shows the sampling region of 200x200 pixels.

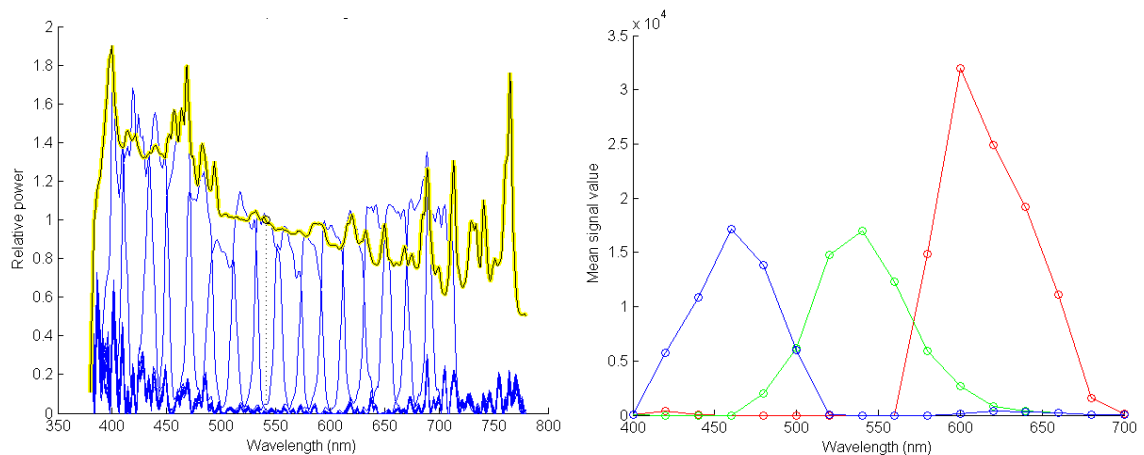


Figure 7.12: (left) Measured spectral power distribution of Balcar flash (yellow) and of radiation transmitted through each filter (blue); (right) Mean image intensity values for 16 filters.

After interpolation to 1 nm intervals and correction for the integral power transmitted through each filter (Fig. 7.12 left), the normalised sensitivity of the three channels could be compared with that measured by the monochromator (Fig. 7.15 left). The correspondence for green is excellent, but the widths of the curves for red and blue are slightly narrower by about 20 nm in the upper half of their peak. This may have been caused by the effects of measurement noise at either end of the spectrum.

7.1.4 Multispectral Filters on Camera

The camera was mounted on a photographic copystand and the X-rite ColorChecker White Balance target placed on the baseboard. The same multispectral filter set was used as for the previous method, but this time the circular glass filters were fitted into 52 mm rings for screw-mounting directly onto the front of the camera lens. This arrangement ensured that no stray light entered the lens. The target was illuminated by a Bowens studio monobloc flash lamp with a large softbox diffuser (Fig. 7.13 left). This setup with an incident angle of over a range 20° – 60° allowed the camera to capture light reflected from the surface without specular highlights. The 105 mm lens was used, with aperture set to $f/9$, and the filters were screwed one-by-one into a 62-52 mm step-down adapter ring on the front of the lens.

Images were processed in a similar way to the previous method, by cropping an area of 200×200 pixels out of the centre of each of the sixteen images (Fig. 7.13 right) and calculating the mean R,G,B values (Fig. 7.14 right). Although there was a slight luminance gradient across the image because the flash illumination was all from one side, nothing moved between exposures and so all samples were affected to the same degree. Measurement of the reflected illumination from the white card was made by the PhotoResearch PR-650 spectroradiometer, mounted on a tripod. It would also have been possible to measure the SPD of the flash from the front of the softbox, but then the reflectance spectrum of the white card would also have been needed in the computation. The advantage of measuring directly from the white card was that it gave the same spectrum as was entering the filter in front of the lens (Fig. 7.14 left). Unlike the previous method described in Section 7.1.3, where the illumination was filtered and so changed for every exposure, in this case the illumination was constant (subject to the repeatability of the flash light) and so needed to be measured only once.

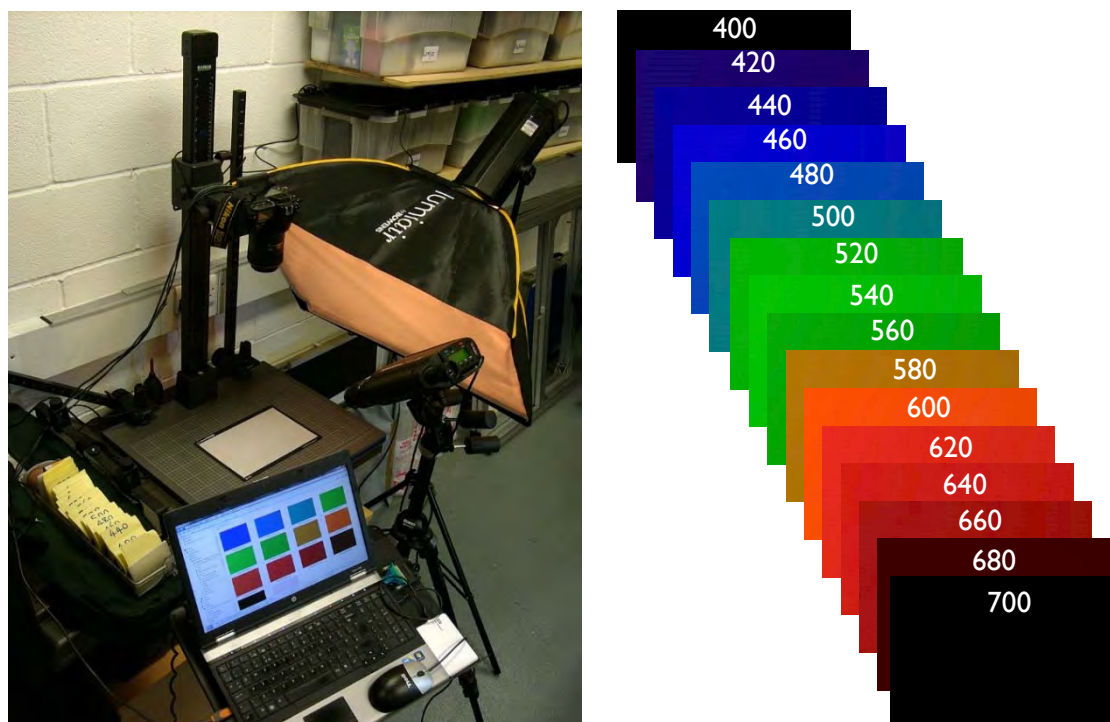


Figure 7.13: (left) Lab setup with Nikon D200 camera mounted on a copystand with filtered flash lighting from a softbox in front of a studio flash lamp (upper right). The light reflected from the white card was measured by the PR-650 spectroradiometer under control of the PC; (right) Sixteen images taken through the filters.

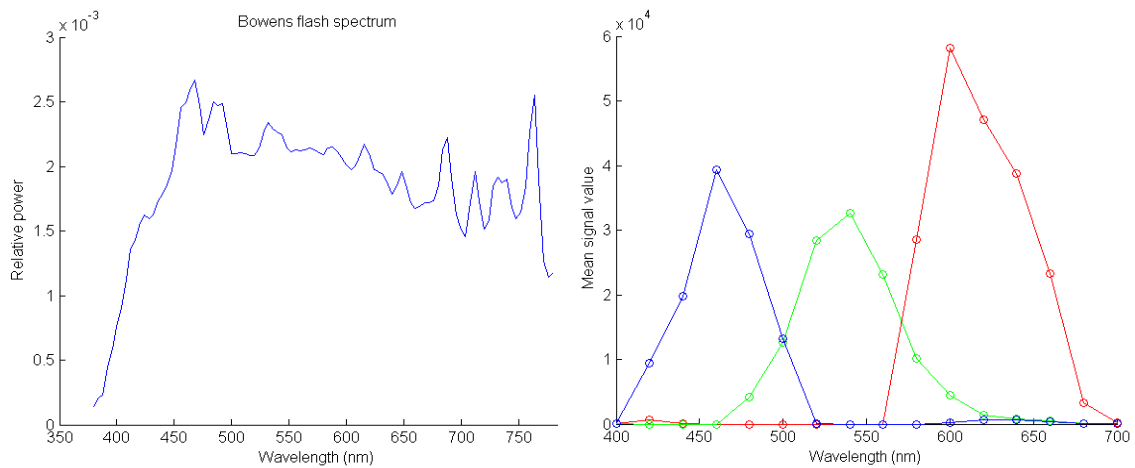


Figure 7.14: (left) Spectral power distribution measured for the Bowens flash; (right) Mean image intensity values for images through 16 filters.

After interpolation to 1 nm intervals and correction for the integral power transmitted through each filter, the normalised sensitivity of the three channels (Fig. 7.15 right) can be compared with that measured by the monochromator (Fig. 7.5) and by the filtered camera technique (Fig. 7.15 left). Although the curves are broadly similar, two differences are apparent: the base of the red curve is shifted by c.10 nm to shorter wavelengths, and the amplitude of the green and blue channels is greater.

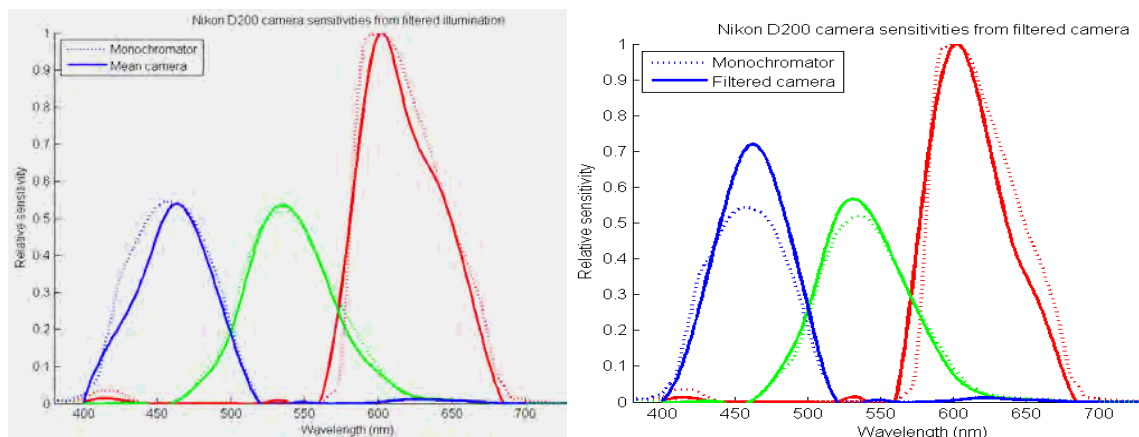


Figure 7.15: Corrected sensitivity of R,G,B channels as functions of wavelength for multispectral filter technique (solid lines) vs sensitivity calculated from variable monochromator (dotted lines). (left): Filtered illumination; (right) filtered camera.

To explore the discrepancy further, the procedure with filters on the camera was repeated with a different source of lighting and a different spectroradiometer. The Kaiser fluorescent graphic arts lights were fitted on either side of the copystand, and a new Ocean Optics USB2000+ spectrometer was coupled with an optical fibre, measuring the spectra at wavelength intervals of 0.35 nm. Its radiance scale was corrected to that of the PR-650 by using the ratio of measurements by both instruments of light from a tungsten halogen lamp introduced into a 25 cm integrating sphere by an optical fibre bundle (Fig. 7.16 left). The spectral power distribution (Fig. 7.16 right) shows several very strong emission lines, with two in particular at 436 and 546 nm, characteristic of the mercury vapour discharge used to excite the phosphors of a fluorescent lamp.

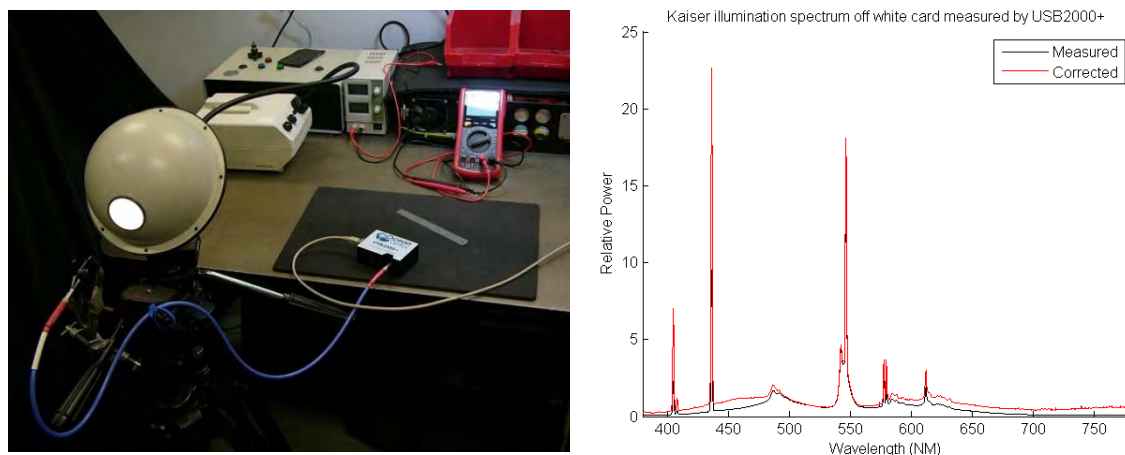


Figure 7.16: (left) Laboratory setup for calibration measurement of USB2000+ spectrometer; (right) spectral power distribution of Kaiser copystand illumination reflected from white card.

The results for the camera responsivity after the same calculation as before (Fig. 7.17 right) show the red curve similar to that for the flash illumination, i.e. shifted by 10nm to the left, but the green and blue curves considerably enlarged and distorted.

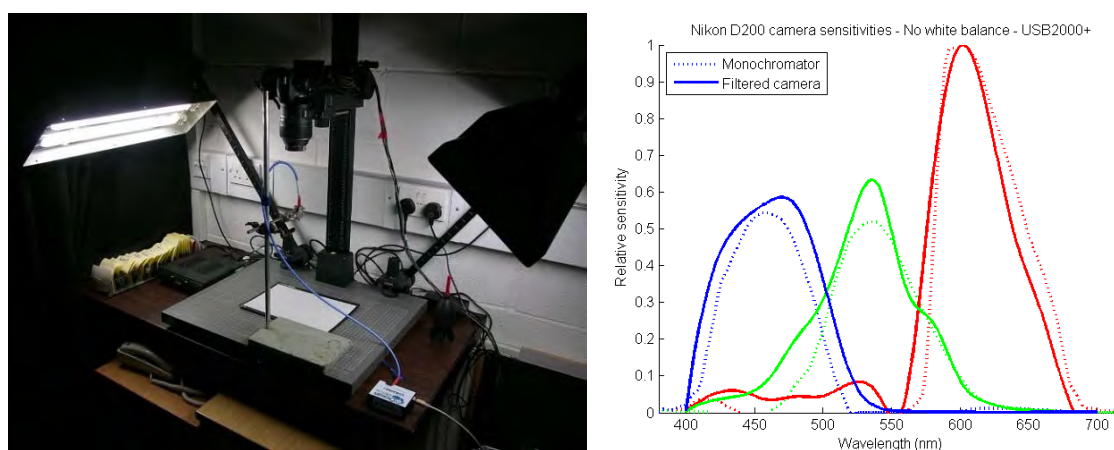


Figure 7.17: (left) Kaiser copystand lighting with camera and measurement from white card by USB2000+ spectrometer; (right) Corrected sensitivity of R,G,B vs monochromator.

7.1.5 Camera White Balance

The greater blue sensitivity found when the filters were fitted on the camera lens (Figs. 7.15 and 7.17) suggests that somehow the setting for colour balance (ratio of blue to red channels) in the camera may have been different in this later procedure than when images were taken with the monochromator. The procedure was therefore repeated with six different settings of white balance in the Nikon software to determine the effect of this parameter. The settings were named in the Nikon menu as: Incandescent, Fluorescent, Sunlight, Flash, Cloudy, Shade.

It was found that if the raw images (NEF file format) were converted via DCRAW with the white balance suppressed then all the resulting TIFF files were identical, apart from the effects of noise. If the white balance was enabled in DCRAW, however, then the resulting TIFF files were different. Analysis by sampling a patch of 400x400 pixels from the centre of each image revealed further that the green channel never varied in the white-balanced images, but that the blue and red channels each were scaled by a constant factor relative to the corresponding raw

values (Fig. 7.18 left). These were related inversely to the integrated power in the short and long wavelengths of the typical spectral power distribution for the given illumination (Table 7.1 and Fig. 7.18 right). Thus for incandescent illumination, for example, B was increased by a factor of 1.96, while R was reduced by a factor of 0.52, relative to G = 1.

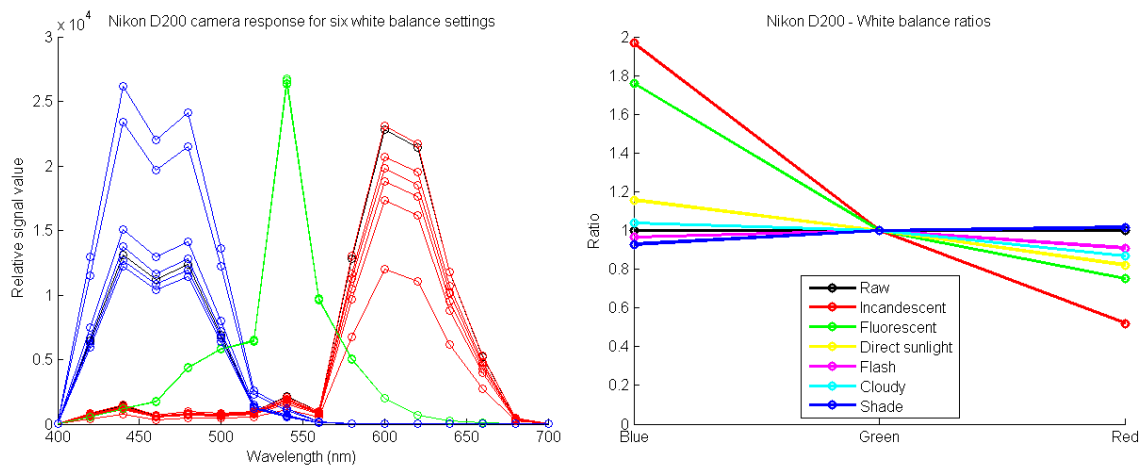


Figure 7.18: (left) Mean RGB values for raw (black) and six settings of white balance; (right) Ratios of blue and red channels relative to green for six white balance settings.

The conclusion from this analysis is that the white balance setting does not affect the image data stored in the raw image file. It is just a parameter in the metadata associated with the image file. But, unless it is explicitly suppressed, the camera itself and all file conversion programs such as the Photoshop image processor, will apply these multiplicative factors to the image data, resulting in an apparent change of spectral responsivity. This is suggested as an explanation for the different relative amplitudes of the spectral sensitivity curves for the Nikon D200 published by Farrell *et al* (2008).

Table 7.1: Nikon D200 – White balance correction ratios B:G:R

	Incandescent	Fluorescent	Sunlight	Flash	Cloudy	Shade
B	1.9687	1.7610	1.1575	0.9656	1.0372	0.9286
G	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
R	0.5201	0.7516	0.8212	0.9095	0.8679	1.0149

The white balance scaling could not account for the shift in wavelength or changes in shape of the response curves observed in Figures 7.15 and 7.17. There must have been some other influence causing the observed change in responsivity. Either the spectral power measured by the spectrometer was too low or the signal response of the camera was too high in these regions of the spectrum, causing the responsivity ratio of camera response to incident power to be too large. One explanation is that the spectroradiometric measurement of the spectral power distribution of the fluorescent light (Fig. 7.16 right) failed to capture the true amplitude and width of the spikes, resulting in a low estimate of the luminous power. This could have been caused, for example, by saturation of the detector for the high amplitude of the spikes, or the limited spectral resolution. The spectral dispersion of the grating is likely to be larger than the wavelength interval of the data delivered by the instrument. In either case the integral power of the illumination would be under-estimated.

Another explanation is that some sort of ‘spectral leakage’ was occurring through the filters. Because they are dichroic the filter transmittance spectra (Fig. 7.9 right) apply only for the normal angle of incidence and propagation of light through the body of the filter. Like all interference effects, the transmission changes with angle of incidence, shifting toward shorter wavelengths, approximated by (Quijada *et al*, 2004):

$$\lambda(\theta) = \lambda_0 \sqrt{1 - \left(\frac{\sin \theta}{n}\right)^2} \quad (7-1)$$

where: θ is the angle of incidence relative to the normal, λ_0 is the unshifted wavelength of the transmission at normal incidence, and n is the effective index of refraction inside the filter, which varies with polarisation. In the arrangement of the camera mounted on the copystand, rays reflected from the white target can pass through the filter and thence through the lens in a cone of angles of up to 20° off the optical axis, depending on the lens aperture setting. The effect of Eq. 7-1 is shown in Fig. 7.19, assuming an index of refraction of 1.6 and unshifted wavelength of 580 nm. Thus for an oblique ray of 20° incidence the filter normally transmitting in the 20 nm waveband 570–590 nm would instead be transmitting in the waveband 556–576 nm. The tail of the transmission spectrum would therefore be close to the spike in the emission spectrum at 546 nm, resulting in a larger signal value from the camera than expected, leading to the high values for both G and R responsivity curves at nominal 580 nm (Fig. 7.17 right). This effect could be minimised by masking the white card to a small area at the centre of the image field and increasing the distance between camera and target.

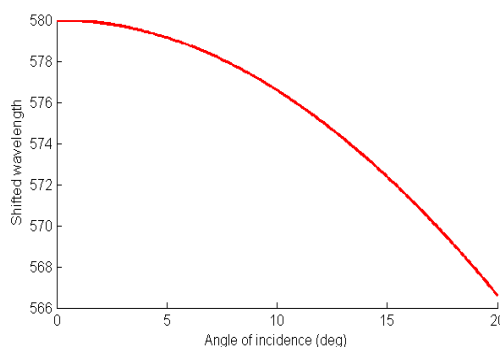


Figure 7.19: Shift in filter transmission wavelength with angle of incidence.

In conclusion, all four methods of spectral characterisation gave consistent results, though all required careful measurement and calibration. The one-shot LED target is very promising as an alternative to a monochromator. For best performance an illumination source with a smooth spectral power distribution such as tungsten or tungsten-halogen is recommended. Sources such as flash or fluorescent with line emission spectra, producing high-amplitude spikes in the spectral power distribution, should be avoided because they can lead to inaccuracies in both measurement and camera response.

7.1.6 Correction of Illumination Profile

To characterise the camera with a 2D target array, the signals from the camera must be related to the luminance of each patch. Ideally the illuminance should be uniform over the whole area of the image, so that the known reflectance factor of each patch can be converted to the corresponding luminance. In practice, illuminance is never perfectly uniform, so a correction procedure must be employed.

The flash lights in the hemispherical dome do not produce a uniform distribution of illumination across the target area. A single flash lamp approximates a point source, and the distribution of light across the baseplane has a strong gradient, with power inversely proportional to the

square of distance from the source. Fig. 7.20 shows the images of a matte white card illuminated by each of the 64 lamps. Contrast is highest for the lights in Tier 1 (top two rows), where the 16 lamps are closest to the card. Contrast is lowest in Tier 5 (last four images of last row), where the 4 lamps are close to the camera at the ‘north pole’ of the dome.

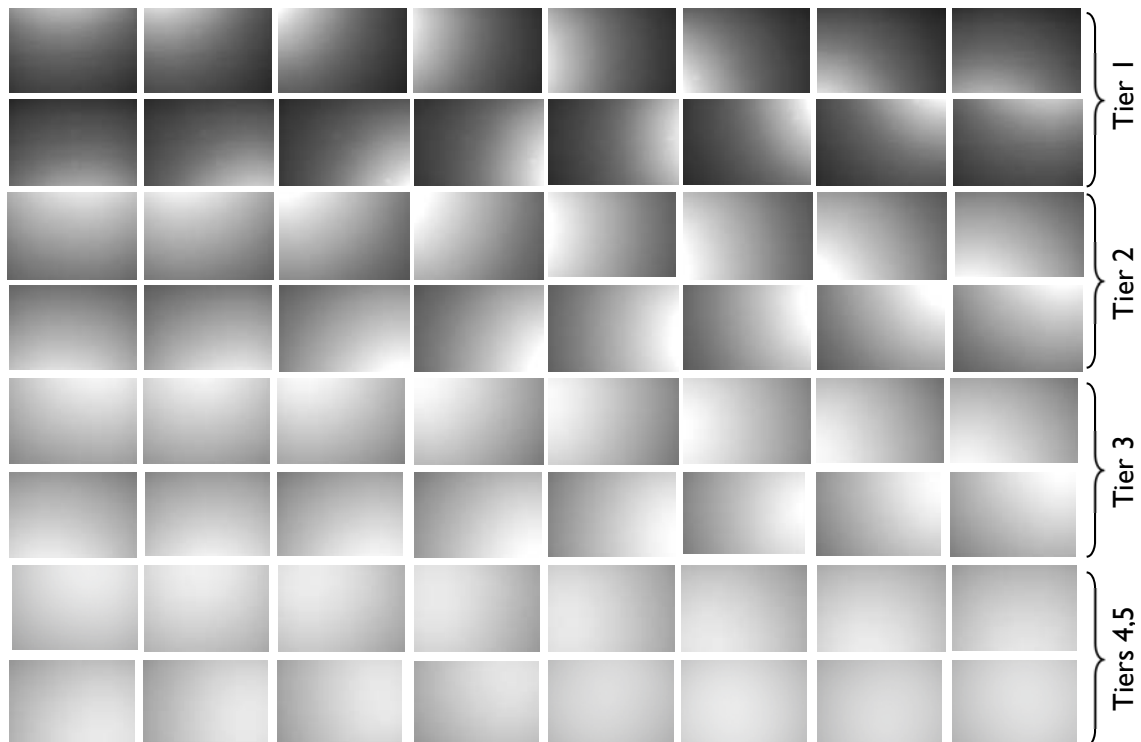


Figure 7.20: Images of a white card under all 64 single flash lights in the dome, showing luminance distributions. The images were taken with the 17-55 mm zoom lens set at 28 mm, f/8.

When the ‘ring flash’ mode with 12 or 16 lights is employed, the illuminance profile is approximately circularly symmetric, as shown in Fig. 7.21. For Tiers 1 and 2 (low angle of illumination) the intensity is higher at the periphery; for Tier 3 it is approximately uniform across the whole area, and for Tiers 4 and 5 (highest angles of illumination) it is higher in the centre and falls away toward the periphery.



Figure 7.21: Images of a white card taken by ‘ring flash’, using all lights in a tier, showing the luminance distributions. The images were taken with the 17-55 mm zoom lens set at 28 mm, f/8.

The correction principle is to divide the target image at each pixel position by the relative luminance at the corresponding pixel position. In order to achieve this both the target image and the white card calibration image were linearised by processing the NEF files from the camera through the DCRAW utility (Fig. 7.22). In fact, the image of the white card proved to be very noisy, from a combination of camera noise and the mottle or surface texture of the card (Fig. 7.23 top).

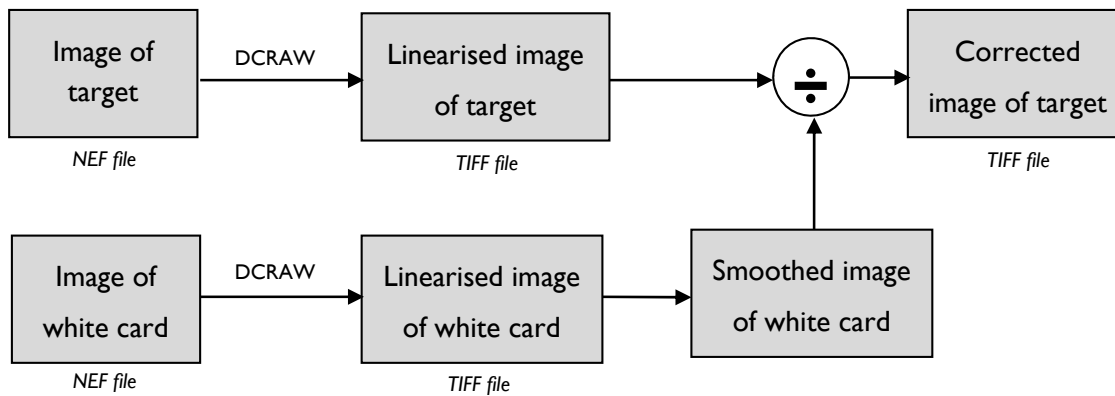


Figure 7.22: Processing sequence for correction of image for non-uniform illumination profile.

The noisy white card image was smoothed by reducing it to one quarter size, applying an averaging filter on an area of 41 pixels square (arbitrarily chosen large diameter) around each pixel, then enlarging the image back to original size with bicubic interpolation. In effect the smoothing filter applied to the high resolution image was 164 pixels square.

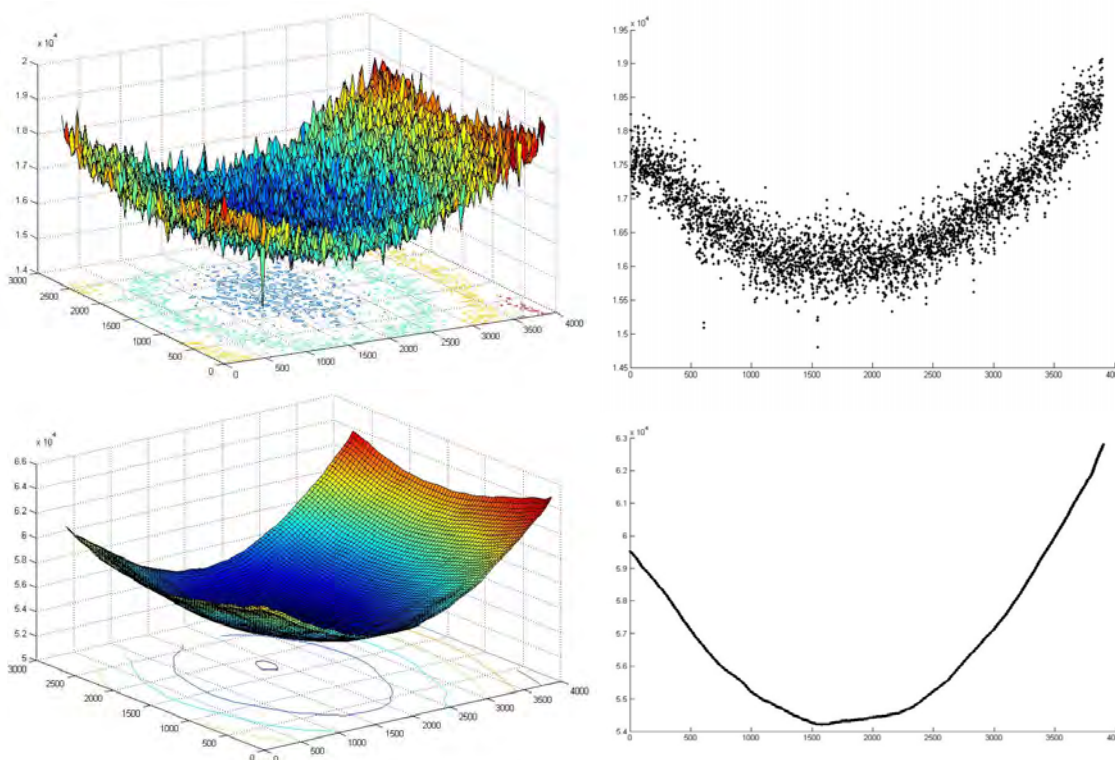


Figure 7.23: Linearised green channel of white card taken under Tier 2 ring flash using 17-55 mm lens set at 45 mm. The noisy original image (top) was heavily filtered to produce the smooth profile (bottom). (left) 3D surface showing illumination across whole surface; (right) horizontal mid-line cross-section.

The GretagMacbeth DC chart was chosen as a test object, because its matte surface is tolerant to variations in the direction of incident light. The overall target dimensions are 320x200 mm, with a rectangular array of 12 rows by 20 columns of square patches each of size ½ inch (12.7 mm). At the centre is a large white patch surrounded by a 12-step grey scale, and around the periphery is a ring of achromatic white, grey and black patches. The number of colour patches is thus $240 - 16 - 60 = 164$. The target was illuminated in the dome by the 16 lights of tier 3, at an incident angle of $\sim 45^\circ$, to minimise specular reflection. A second image was taken of a white

card with the same camera settings and processed as above. The corrected image was saved as a 16-bit RGB TIFF file, scaled so that the data filled the whole range. Because a small number of pixels were outliers with very high values, caused by specular highlights or noise, a histogram technique was used to identify the 99.9th percentile and set it to a nominal white point value at 99% of the full 16-bit range, i.e. 64881. Figure 7.24 shows the ‘before and after’ versions of the chart image. The range of the linearised image has been extended to fill the full 16-bit data scale, and the non-uniformity of the flash illumination has been corrected. In this form the image is ready for extraction of data for colour characterisation of the camera.

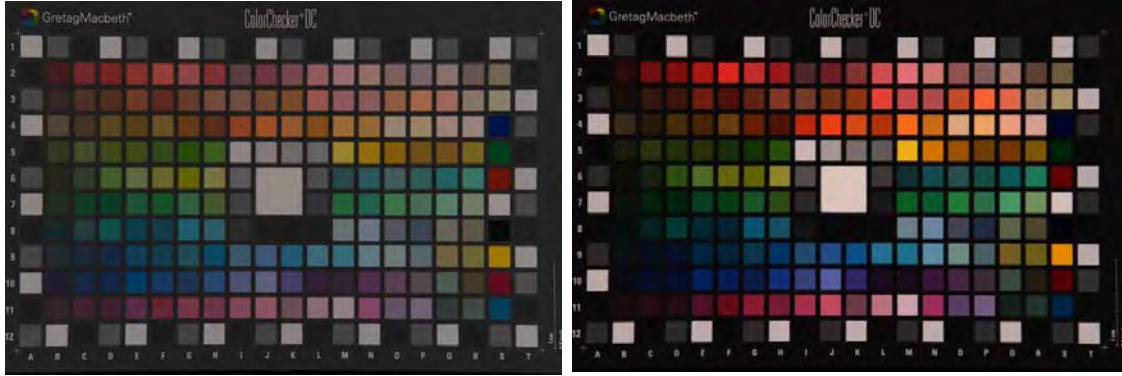


Figure 7.24: GretagMacbeth DC colour target, 45 mm zoom, f/8, under Tier 2 ring illumination. (left) Original linearised image of target; (right) after correction for non-uniform illumination profile.

7.1.7 Prediction of Camera RGB

Having established the spectral sensitivity of the Nikon D200 by the methods described in Sections 7.1.1 to 7.1.4, it is possible to calculate the R,G,B signal values that the camera would generate for a patch of any reflectance spectrum. Using the discrete form of the colorimetric equations, at each 5 nm wavelength interval, the three factors were multiplied together and summed:

$$V_{n,k} = \sum_{380}^{730} S(\lambda) \cdot R_n(\lambda) \cdot C_k(\lambda) \cdot \Delta\lambda \quad (7-2)$$

where: $V_{n,k}$ is the camera signal for channel k for patch n

$S(\lambda)$ is the power of the flash illumination

$R_n(\lambda)$ is the reflectance of the surface colour for patch n

$C_k(\lambda)$ is the sensitivity of camera channel k (R,G,B)

λ is the wavelength (nm)

The spectral power distribution of the dome flash lights was measured by a Minolta CS-1000 spectroradiometer at 5 nm intervals. The reflectance spectrum of all the patches of the GMDC colour chart were measured individually by a hand-held X-rite i1Pro spectrophotometer at 10 nm intervals, and interpolated to 5 nm intervals. The eight glossy patches in column 5 have been removed from the dataset. The image of the chart, taken under ring flash 3, was linearised and corrected for the luminance profile. Camera signal values were derived as the mean values in an area of 98x98 pixels at the centre of each colour patch in the full resolution 16-bit TIFF image (size 3900x2616 pixels). The predicted values were normalised by a factor calculated as

the ratio of the mean camera value for the large central white patch divided by the predicted value for the same patch. The results of the computation (Fig. 7.25) show a convincing linear relationship between the predicted and actual values for 232 patches in the chart, including all the neutral patches at the centre and around the periphery.

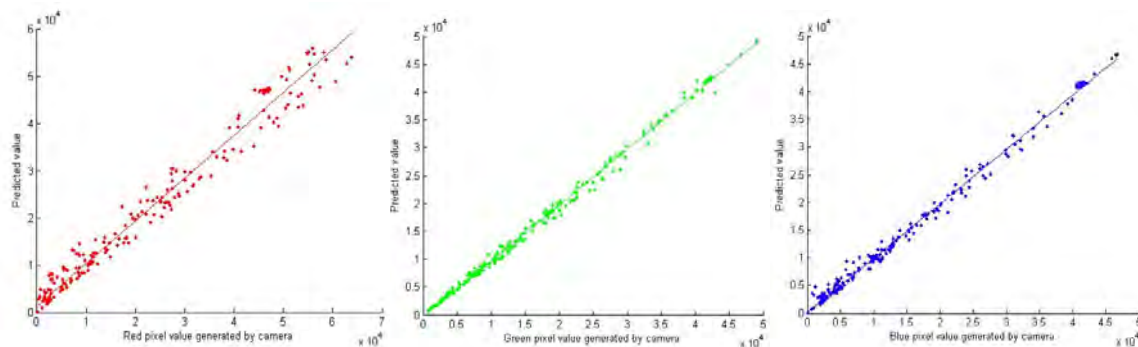


Figure 7.25: Predicted (y axis) vs actual camera (x axis) values for 232 colour patches in the GretagMacbeth DC target, for the red (left), green (centre) and blue (right) channels of the Nikon D200 camera.

The scatter of values around the line of best fit was greatest in the red channel. This distribution is clearer when the difference is divided by the mean, as shown in the histograms of Fig. 7.26.

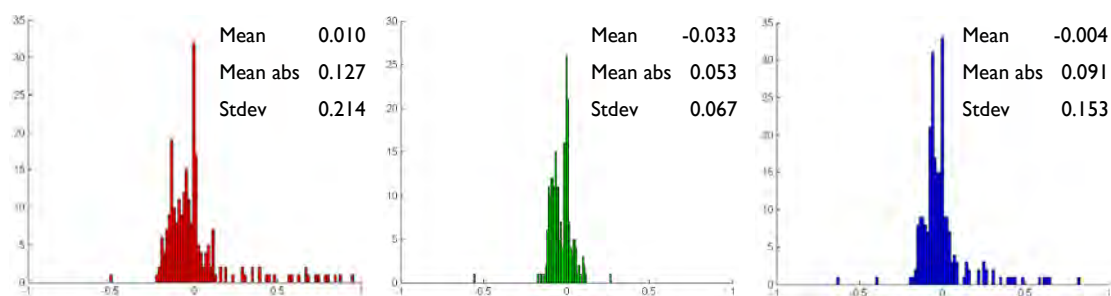


Figure 7.26: Histograms of relative differences between predicted and actual RGB values.

There are various possible reasons for the discrepancy between the predicted and actual values:

- Errors in measurement of the spectral power distribution of the flash light;
- Errors in measurement of the reflectance spectra of colour patches in chart;
- Errors in measurement of the spectral sensitivity of camera;
- Imperfect correction of non-uniform illumination on target;
- Non-Lambertian reflectance (i.e. angular variation) from colour chart and white card;
- Differences in geometry between colour chart and white card in the dome;
- Variability of spectral power distribution of the flash light;
- Optical flare in the lens;
- Ambient light and inter-reflections;
- Noise in the images from camera sensor and electronics.

All of these factors probably contributed to the error distribution seen above, but the correspondence between predicted and actual values is sufficiently close to prove the principle.

7.1.8 Image Noise

The camera was fitted with the 100 mm lens and mounted in the dome. A series of images was taken of the central area of the GretagMacbeth DC colour chart (Fig. 7.27 left) illuminated by flash ring 3. The aperture was set to $f/7.1$ to give the maximum exposure of the central white patch that could be achieved without some pixels being clipped to maximum value. Ten images were captured in NEF format and converted by DCRAW to 16-bit TIFF. From each image a region of 200x200 pixels was extracted from each of the 16 greyscale patches and the mean taken over all pixels in all images. The reflectance factor of each patch was measured with an X-rite i1Pro spectrophotometer.

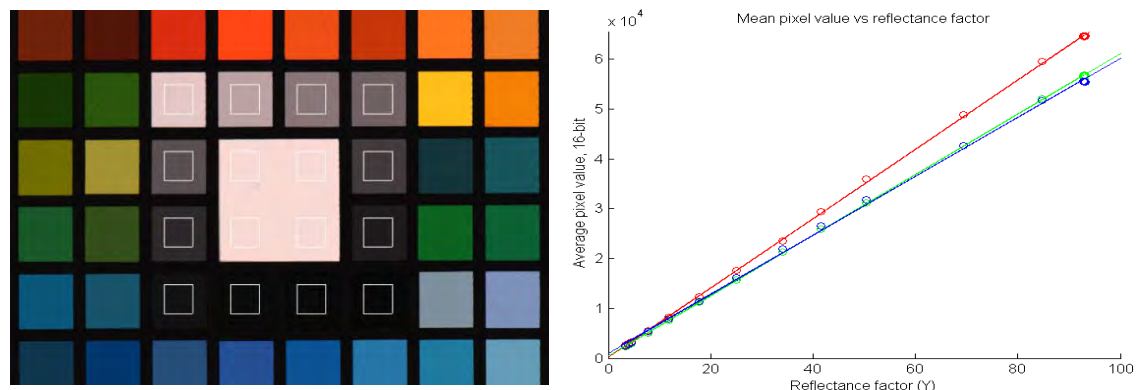


Figure 7.27: (left) image of central area of GMDC colour chart, showing sampling boxes in 16 patches of greyscale; (right) Mean pixel values vs measured reflectance factor with lines of best fit.

Plotting the mean pixel value of each patch against reflectance factor (Fig. 7.27 right) shows excellent linearity in all channels, with all points lying very close to the lines of best fit over full signal range 0-65535. The ratio of slopes of the three lines (relative slopes $R=1.000$, $G=0.873$, $B=0.851$) is a measure of the white balance of the camera, with the red predominating. This can be seen in the pinkish tinge of the image. To correct this image for correct neutral white, the multipliers for the red and blue channels would be 0.869 and 1.022 respectively.

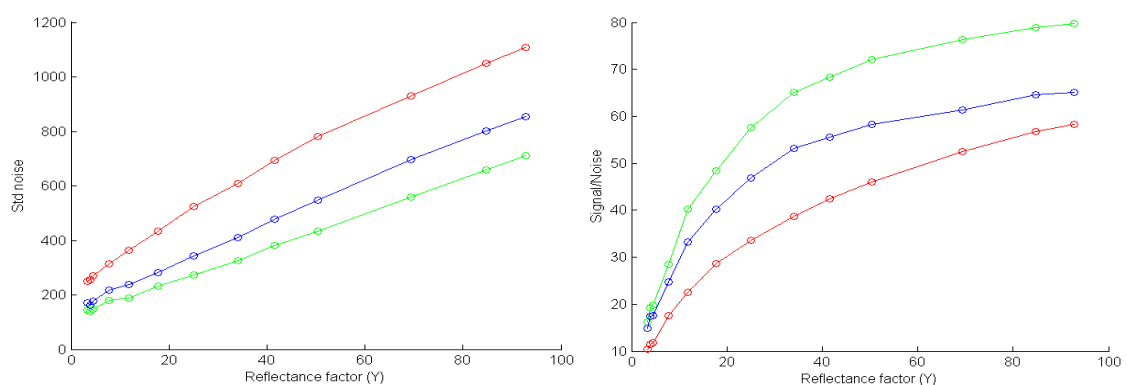


Figure 7.28: (left) Mean noise vs reflectance factor; (right) Signal-to-noise ratio vs reflectance factor.

Calculating the standard deviation of pixel values in each patch, and taking mean of stdev values over 10 images, shows that noise increases almost linearly with light intensity (Fig. 7.28 left). Calculating signal-to-noise ratio (SNR) as mean signal for each patch divided by corresponding mean noise, shows that noise rises rapidly from low values for dark tones and flattens out for light tones approaching white (Figure 7.28 right). Red has the lowest SNR and green the highest. For mid-grey (reflectance factor = 50%), the SNR values are $R=52.4$, $G=76.3$, $B=61.3$.

7.2 Optimum vs Actual Wavelengths of Colour Laser Scanner

Laser scanners pose a colorimetric challenge for the digitisation of reflective objects because, unlike the human visual system and digital cameras with broadband sensitivity, they sample the spectrum at single wavelengths. More properly, they illuminate the sample at one or more monochromatic wavelengths, and sense the intensity of the reflected light. For a colour scanner with red, green and blue lasers the RGB triplet at each pixel position depends only on the wavelength of each laser and the surface reflectance of the object at those three specific wavelengths (assuming a Lambertian surface and linearity of the detector). Although this arrangement is highly metameric, it produces acceptable results for a range of media because natural reflectance spectra are broadband and can be sampled with sufficient accuracy at a few widely-spaced wavelengths. The question then arises as to which three wavelengths would be optimum for trichromatic imaging of typical colorants?

7.2.1 Method

Tristimulus values X, Y, Z were calculated for each reflectance sample in a dataset under daylight at 6500 K (D65) and under multi-laser illumination, using the Judd-Vos (1978) modification of the CIE 2-degree standard observer. These were converted to CIELAB values using the respective spectra of D65 and equal unit power of the lasers as reference white, and the colour difference ΔE^*_{ab} was calculated. The median and mean values were calculated over all the samples in the dataset.

For an initial assessment, three independent sets of colour reflectance spectra were obtained, all measured from physical paint samples:

1. GretagMacbeth DC Chart – 170 patches (no gloss), measured at 10 nm intervals with *i/Pro*;
2. Berns-Taplin RIT pigment set – 100 patches, measured at 10 nm intervals with *i/Pro*;
3. National Gallery pigment set – 64 patches, measured at 2 nm intervals with *Monolight*.

Each dataset was interpolated to wavelength intervals of 5 nm over the range 380–730 nm (a vector of 71 reflectance factors). They were combined into a single large dataset containing 334 spectra (Fig. 7.29).

Suppose that the wavelengths of the three lasers in a laser scanner were arbitrarily chosen, spaced at uniform intervals across the visible spectrum. The results (Table 7.2) show an unpredictable relationship between the mean colorimetric errors and the number of laser wavelengths, with four lasers at the chosen wavelengths

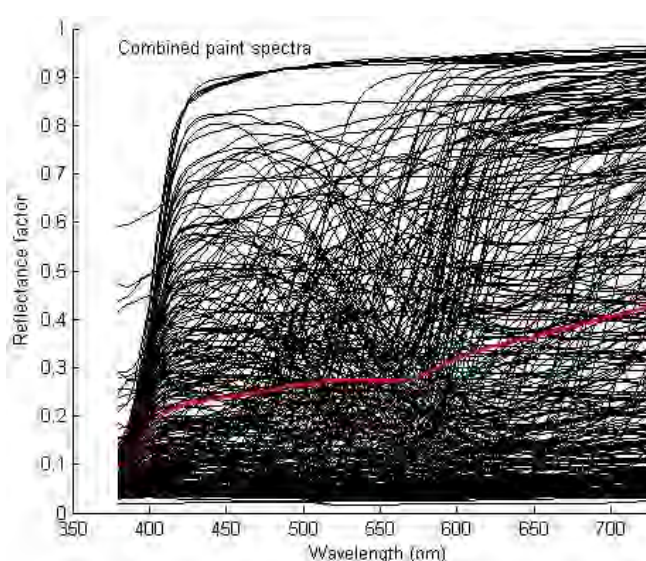


Figure 7.29: Measured reflectance spectra of 334 paint samples and their mean (red).

producing a much better result than five lasers, and seven lasers being hardly better than six. The colorimetric performance is more complex than reconstruction of the reflectance spectra, because it depends on the form of the tristimulus functions.

Table 7.2: Evenly-spaced laser wavelengths and colorimetric error values for the combined paint dataset.

Number of lasers	Wavelengths (nm)	Interval (nm)	Median	Mean	Maximum
3	450, 550, 650	100	13.73	15.23	53.06
4	435, 510, 585, 660	75	4.96	7.64	34.87
5	430, 490, 550, 610, 670	60	14.45	16.24	51.36
6	425, 475, 525, 575, 625, 675	50	1.60	1.99	9.69
7	430, 470, 510, 550, 590, 630, 670	40	1.30	1.51	9.72

The alternative approach is to choose the laser wavelengths to minimise the colorimetric errors over the full set of spectra. By exhaustive search, the combination of laser wavelengths at 5 nm intervals was found that minimised the mean colour difference between D65 and laser illumination over the pigment set. The search space was limited to all combinations of laser wavelengths, at 5 nm intervals, in the three ranges: blue 380-495 nm; green 500-570 nm; and red 575-730 nm.

The three curves in Fig. 7.30 (left) represent one-dimensional sections through the three-dimensional distribution of mean ΔE^*_{ab} values for the combined dataset. The global minimum point was well defined with respect to all three variable laser wavelengths, and both mean and median measures of error distributions gave the same result. The optimum wavelengths (to the nearest 5 nm) were 460, 535 and 600 nm.

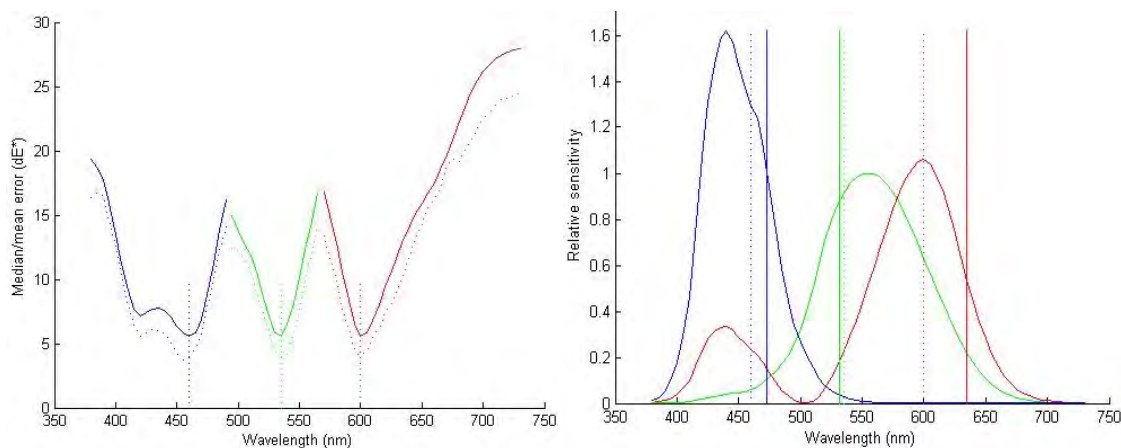


Figure 7.30: (left) Mean (solid line) and median (dotted line) values of colour difference for 334 paint spectra; (right) CIE standard observer (Judd-Vos modification) with wavelengths of Arius scanner (solid lines) and optimum (dotted lines).

The results were similar for all three datasets when evaluated individually (Table 7.3). The mean error at the optimum wavelength triplet was consistently in the range 5-6 ΔE^*_{ab} and the median error 3-4 ΔE^*_{ab} .

Table 7.3: Optimum laser wavelengths and error values for three separate datasets.

Dataset	Samples	Optimum laser wavelengths B, G, R (nm)	Median	Mean	Maximum
National Gallery chart	64	460, 535, 600	4.02	6.01	21.50
Berns-Taplin acrylic data	100	465, 540, 605	3.53	5.34	21.46
GretagMacbeth DC chart	170	460, 530, 595	3.12	5.28	46.40
Combined dataset	334	460, 535, 600	3.99	5.60	34.15

7.2.2 Arius 3D Scanner

The optimum results were compared with the colorimetric performance of the Arius 3D colour laser scanner, in which the laser wavelengths were initially 473, 532 and 635 nm (Fig. 7.30 right). The mean error of the Arius scanner for the patches in the combined dataset was 3.4 times greater than for the optimum (Table 7.4 and Fig. 7.31 left). The maximum error for the optimum wavelengths was produced by colour patch 21, a vivid red, which has maximum slope in the reflectance spectrum between 600 and 620 nm. The maximum error for the Arius scanner wavelengths was produced by colour patch 284, a strong yellow, with maximum slope in reflectance between 545 and 565 nm.

Table 7.4: Error values for Arius laser scanner for three datasets and combined dataset.

Chart	Samples	Median	Mean	Maximum
National Gallery chart	64	21.54	23.84	80.49
Berns-Taplin acrylic dataset	100	13.64	16.83	66.29
GretagMacbeth DC chart	170	15.09	18.16	63.43
Combined dataset	344	15.31	19.02	80.84

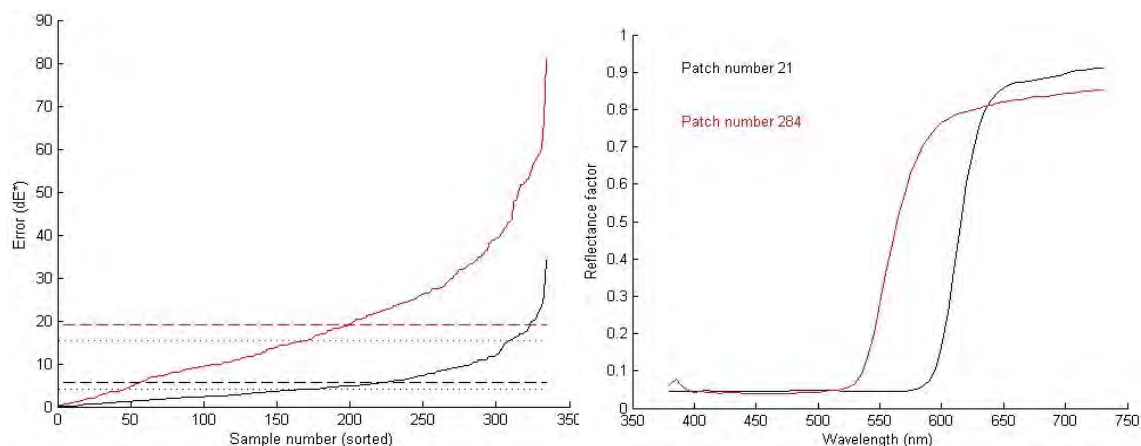


Figure 7.31: (left) Colour difference values for each of 334 spectra, sorted into ascending order, with median (dotted line) and mean (dashed line). The black line shows the errors for optimum laser wavelengths and the red line the errors for approximate Arius wavelengths; (right) reflectance spectrum of patches giving the largest error for the optimum wavelengths (black) and for the Arius wavelengths (red).

The difference in colorimetric accuracy between the Arius scanner and the optimum laser wavelengths can be seen clearly in the error vectors plotted on the CIELAB colour plane between the a^*-b^* coordinates of each sample under D65 and its corresponding coordinates under the laser illumination (Fig. 7.32). For the Arius wavelengths there is a large displacement of colours in the hue range from magenta to orange towards red ($+a^*$) and a lesser displacement of greens and blues towards green ($-a^*$). For the optimum laser wavelengths the displacements are much smaller and their directions rather random, indicating a less systematic colour distortion.

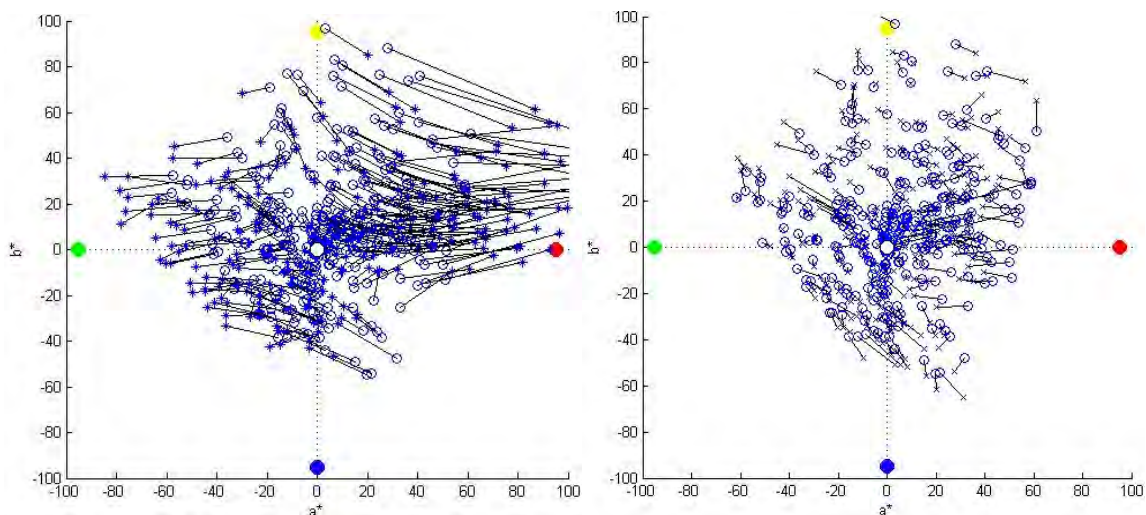


Figure 7.32: Error vectors in CIELAB a^*b^* plane between samples of combined dataset under D65 and: (left) Arius scanner laser wavelengths; (right) optimal three scanner wavelengths. Circles represent colours under D65 and crosses the corresponding colour when illuminated by the three lasers.

If the Arius scanner with its current laser wavelengths could be augmented with one additional laser, to sample the surface at four discrete wavelengths, what would be the optimum wavelength of the fourth laser? The same method as before was used to calculate mean colour difference over the set of pigments for each possible laser wavelength. The results (Table 7.5 and Fig. 7.33 left) show that the optimum wavelength would be 575 nm and mean error would be reduced to less than $8 \Delta E^*_{ab}$. Similarly, adding a fifth laser at an optimum wavelength of 415 nm would reduce the mean error to less than $3 \Delta E^*_{ab}$ (Fig. 7.33 right). In a configuration where the three RGB lasers are fixed and one is allowed to vary, the variation in mean error is less marked than where all three laser wavelengths can vary, but there are large fluctuations in the maximum error. A sixth laser would produce little improvement.

Table 7.5: Error values with combined dataset for Arius RGB laser wavelengths plus additional lasers.

Configuration	Wavelength(s) (nm)	Median	Mean	Maximum
Arius 3D scanner – three lasers	475, 530, 635	15.31	19.02	80.84
Arius with fourth laser	575	5.94	7.26	30.87
Arius with fifth laser	415	2.36	2.94	16.99
Arius with sixth laser	710	2.28	2.93	16.99

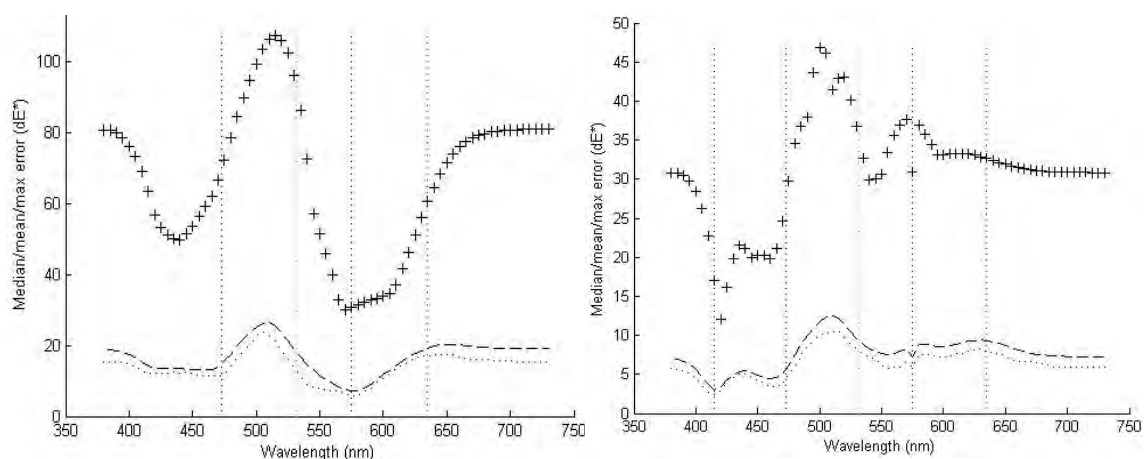


Figure 7.33: Mean (dashed line), median (dotted line) and maximum (crosses) values of colour difference for the combined dataset for the Arius scanner plus (left) a fourth laser channel and (right) a fifth laser channel.

In the course of this investigation the Arius company changed the wavelength of the blue laser from 473 nm to 445 nm, so-called “sky blue”. The previous analysis was repeated for the new Arius wavelengths of 445, 530 and 635 nm (approximating 532 by 530 nm for the green). The colour errors are plotted in Fig. 7.34 (left), the same as Fig. 7.31 (left) with the addition of the results for the new wavelength combination in blue. The mean error was reduced from 19.02 to 15.47 and the maximum error from 80.8 to 54.3. The distribution of colorimetric errors in Fig. 7.34 (right) shows that the red shift is reduced, compared to Fig. 7.32 (left), but there is a greater shift of green colours toward yellowish green.

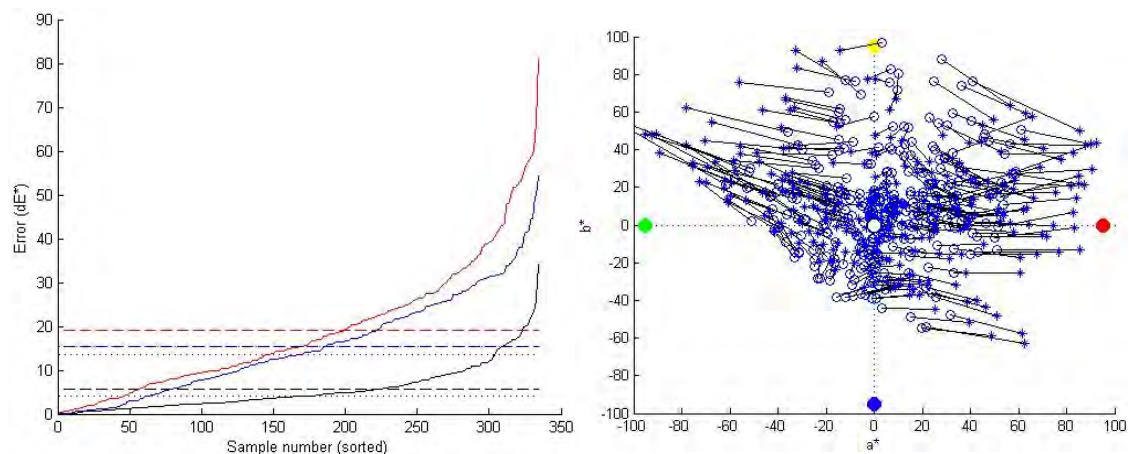


Figure 7.34: (left) Colour difference values for each of 334 spectra, sorted into ascending order, with median (dotted line) and mean (dashed line). The black line shows the errors for optimum laser wavelengths and the red line the errors for old Arius wavelengths, and the blue line for new Arius wavelengths; (right) Error vectors in CIELAB a^*b^* plane between samples of combined dataset under D65 and new Arius laser wavelengths.

A set of 494 measured spectra of natural materials, mostly green and brown leaves, was tested. These occupy a relatively small volume of colour space but are typical of real-world vegetation. The optimum sampling wavelengths were 460, 530 and 590 nm (Fig. 7.35 left), with the minimum for the green laser being particularly sharply defined. For the Arius scanner with the new blue laser the results were much worse, nearly double the colorimetric error, than for the old Arius (Fig. 7.35 right).

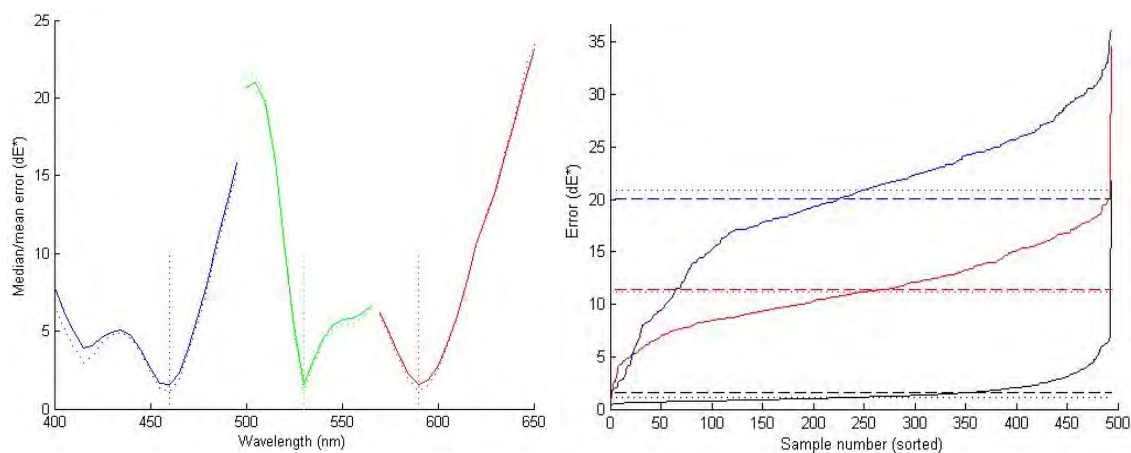


Figure 7.35: (left) Mean (solid line) and median (dotted line) values of colour difference for 494 natural spectra; (right) Colour difference values, sorted into ascending order, with median (dotted line) and mean (dashed line). The black line shows the errors for optimum laser wavelengths and the red line the errors for the old Arius wavelengths, and the blue line for the new Arius wavelengths.

A set of 899 measured spectra of polyester fabric samples with synthetic dyes in the Chinese Natural Color System (CNCS) was tested. These occupy a large volume of colour space. The optimum sampling wavelengths were 460, 535 and 605 nm (Fig. 7.36 left), with the minima for the green and red lasers being well defined. The minima for blue were spread across the range 420–460 nm. For the Arius scanner with the new blue laser the results were a little better than for the old Arius (Fig. 7.36 right), but both versions of the Arius scanner give an error of over double that of a scanner with the optimum wavelengths.

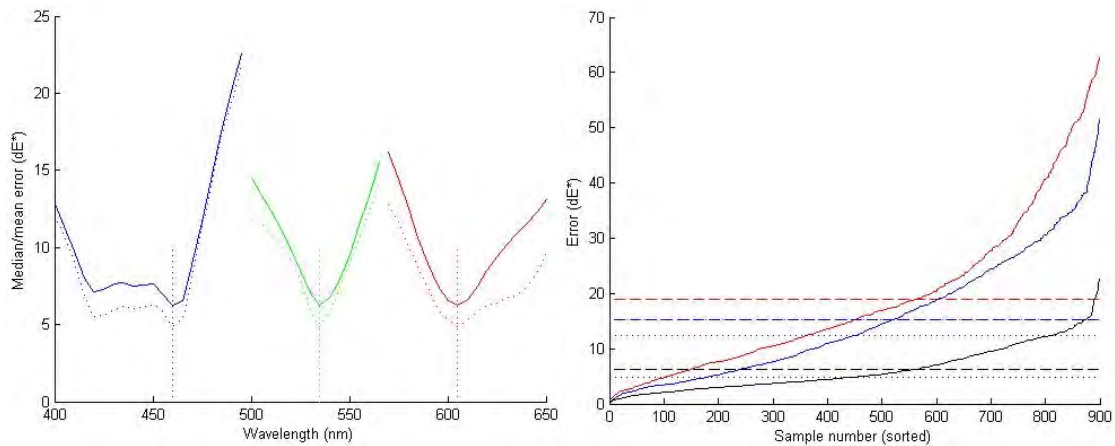


Figure 7.36: (left) Mean (solid line) and median (dotted line) values of colour difference for 899 spectra of Chinese CNCS dyed textile samples; (right) Colour difference values, sorted into ascending order, with median (dotted line) and mean (dashed line). The black line shows the errors for optimum laser wavelengths and the red line the errors for the old Arius wavelengths, and the blue line for the new Arius wavelengths.

A set of 1950 measured spectra of the Swedish Natural Colour System (NCS) atlas was tested. These occupy a large volume of colour space and sample the space quite densely, especially for pastel colours. The optimum sampling wavelengths were 460, 535 and 600 nm (Fig. 7.37 left), with the minima for the green and red lasers being well defined. The minima for blue were broader across the range 420–460 nm. For the Arius scanner with the new blue laser the mean errors were a little better than for the old Arius (Fig. 7.37 right), but in both cases the errors are over triple those of the optimum wavelengths.

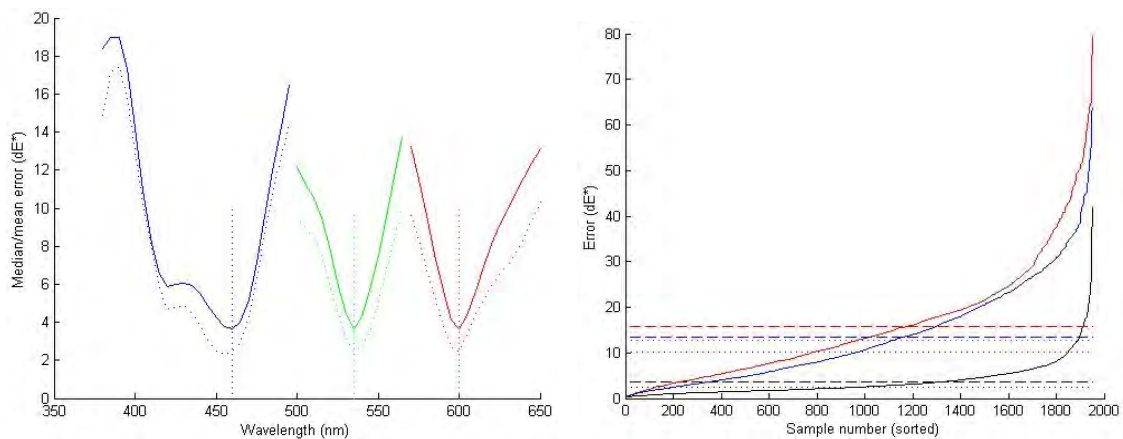


Figure 7.37: (left) Mean (solid line) and median (dotted line) values of colour difference for 1950 samples in the Swedish NCS atlas; (right) Colour difference values, sorted into ascending order, with median (dotted line) and mean (dashed line). The black line shows the errors for optimum laser wavelengths and the red line the errors for the old Arius wavelengths, and the blue line for the new Arius wavelengths.

The procedure was repeated for the Munsell gloss atlas, using the full 1 nm dataset over the range 380-780 nm. The 1 nm datasets were also used for the D56 illuminant and the CIE 1931 Standard Observer, both taken from the CVRL database. The results (Fig. 7.38 left) show a well-defined minimum value for both mean and median at the optimum wavelengths of 461, 529, 597 nm. In this case the new Arius blue laser wavelength gives only slightly larger errors than the old (Fig. 7.38 right).

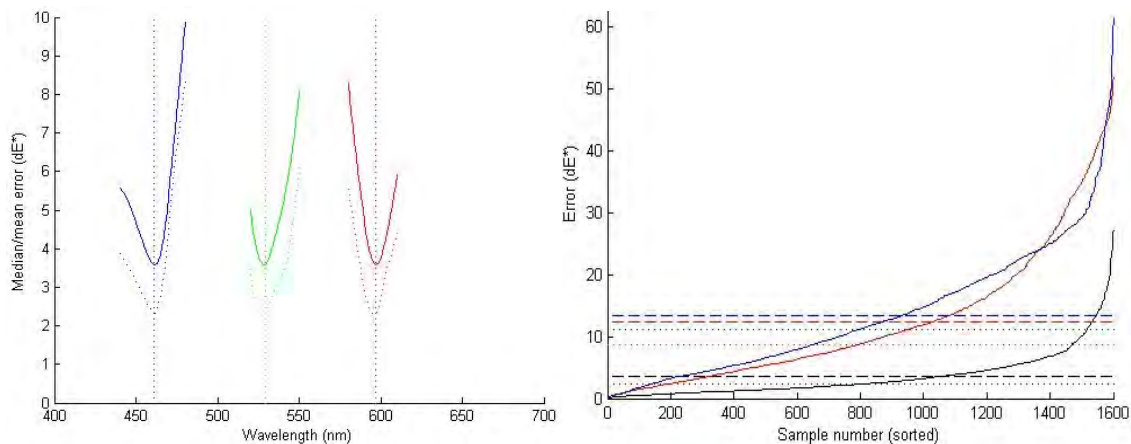


Figure 7.38: (left) Mean (solid line) and median (dotted line) values of colour difference for 1600 samples in the Munsell gloss atlas at 1 nm intervals; (right) Colour difference values, sorted into ascending order, with median (dotted line) and mean (dashed line). The black line shows the errors for optimum laser wavelengths and the red line the errors for the old Arius wavelengths, and the blue line for the new Arius wavelengths.

Finally a dataset of synthetic reflectances was generated, with 500 spectra derived from Gaussian, logistic and sine functions. These are extreme in that the waveforms span the full range of reflectance factors from 0 to 1, band-limited to a maximum slope of ± 0.15 per 5 nm interval, corresponding to a maximum spectral frequency of approximately 1.5 cycles over the 300 nm range. The optimum sampling wavelengths were 450, 535 and 595 nm (Fig. 7.39 left). For the Arius scanner with the new blue laser wavelength of 445 nm, the results were worse than for the previous 475 nm.

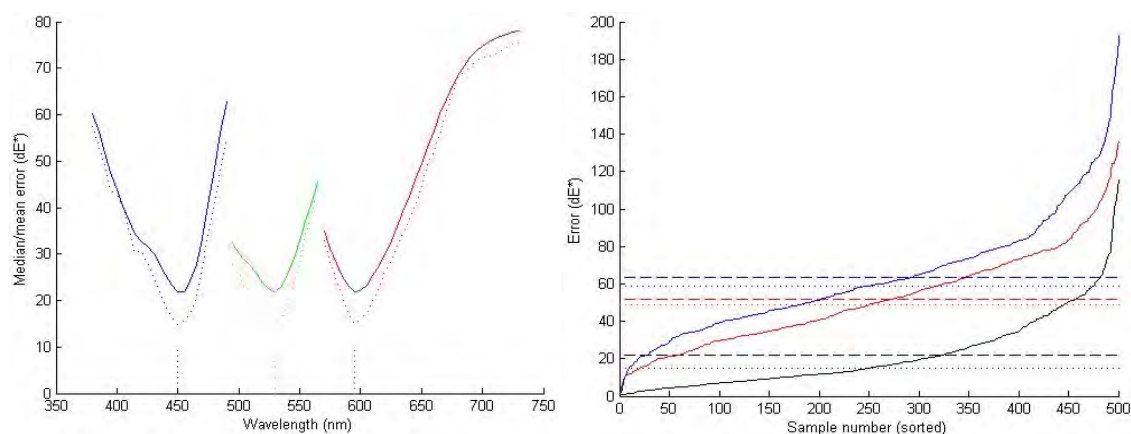


Figure 7.39: (left) Mean (solid line) and median (dotted line) values of colour difference for 500 synthetic spectra; (right) Colour difference values, sorted into ascending order, with median (dotted line) and mean (dashed line). The black line shows the errors for optimum laser wavelengths and the red line the errors for old Arius wavelengths, and the blue line for new Arius wavelengths.

The results of trials with the various datasets are compared in Table 7.6. In all cases the optimum wavelengths are consistently in the range R 590-605, G 529-535, B 450-461 nm.

Table 7.6: Mean/max error values for seven reflectance datasets for the Arius scanner with the old and new blue laser wavelengths and for the optimum laser wavelengths.

Dataset	Samples	Arius, B=475		Arius, B=445		Optimum		Wavelengths
Combined art pigments	334	19.02	80.84	15.47	54.30	5.60	34.15	460, 535, 600
Natural materials	494	11.44	34.51	20.03	36.03	1.58	21.65	460, 530, 590
Chinese CNCS textiles	899	19.03	62.73	15.20	51.55	6.24	22.58	460, 535, 605
Swedish NCS atlas	1950	12.97	55.24	11.86	65.33	3.65	41.99	460, 535, 600
Munsell atlas 1 nm	1600	12.32	51.90	13.40	61.45	3.59	27.17	461, 529, 597
Synthetic spectra 5 nm	500	51.58	135.94	63.60	192.87	21.89	115.28	450, 535, 595
Synthetic spectra 1 nm	1000	44.75	126.57	40.54	146.46	20.60	106.96	456, 533, 601

7.2.3 Influence of White Balance

All of the previous trials were made using the CIELAB colour difference metric (ΔE^*_{ab}) and with equi-energy white for the laser illumination, i.e. unit power for each of the three laser sources. For CIELAB the white balance has no effect because of the Von Kries normalisation of the tristimulus values.

The conversion from laser intensity values L_r, L_g, L_b to tristimulus values X, Y, Z is simplified from the usual summation because the laser spectra can be regarded as delta functions, with a non-zero value at a single wavelength. Each tristimulus value is then simply the sum of the tristimulus function at each of the three laser wavelengths, weighted by the laser intensity:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = K \begin{bmatrix} \bar{x}_r & \bar{x}_g & \bar{x}_b \\ \bar{y}_r & \bar{y}_g & \bar{y}_b \\ \bar{z}_r & \bar{z}_g & \bar{z}_b \end{bmatrix} \begin{bmatrix} L_r \\ L_g \\ L_b \end{bmatrix} \quad \text{i.e. } \mathbf{T} = \mathbf{KML} \quad (7-3)$$

where $K = \frac{100}{\bar{y}_r L_r + \bar{y}_g L_g + \bar{y}_b L_b}$ is a normalising factor to make $Y = 100$.

and \bar{x}_r = value of \bar{x} tristimulus function at the red laser wavelength, etc.

Old Arius laser wavelengths

For the former Arius laser wavelengths of 473, 532, 635 nm, the values of the coefficients (using the Judd-Vos modification of the tristimulus functions) are:

$$\mathbf{M} = \begin{bmatrix} 0.5340 & 0.1908 & 0.1501 \\ 0.2170 & 0.8849 & 0.1037 \\ 0.0001 & 0.0363 & 1.0125 \end{bmatrix} \quad (7-4)$$

The equations can be inverted to express the laser intensities as a function of tristimulus values:

$$\mathbf{KL} = \mathbf{M}^{-1}\mathbf{T} \quad (7-5)$$

The inverse matrix is:

$$\mathbf{M}^{-1} = \begin{bmatrix} 2.0478 & -0.4310 & -0.2594 \\ -0.5043 & 1.2410 & -0.0524 \\ 0.0179 & -0.0444 & 0.9896 \end{bmatrix} \quad (7-6)$$

The power of the three lasers can be varied to change the white balance. We seek to set the laser intensity values so that the tristimulus values of white are equal to those of the D65 illuminant, i.e. $X = 94.33$, $Y = 100$, $Z = 104.155$. Note that because we are using the Judd-Vos tristimulus functions these are different from the customary values of $X = 95.04$, $Y = 100$, $Z = 108.89$ associated with the standard CIE 1931 2-degree observer. Then, setting $K=1$, we have:

$$\mathbf{L}_{\text{Wold}} = \mathbf{M}^{-1}\mathbf{T}_w = \begin{bmatrix} 2.0478 & -0.4310 & -0.2594 \\ -0.5043 & 1.2410 & -0.0524 \\ 0.0179 & -0.0444 & 0.9896 \end{bmatrix} \begin{bmatrix} 94.33 \\ 100 \\ 104.155 \end{bmatrix} = \begin{bmatrix} 123.05 \\ 71.08 \\ 100.32 \end{bmatrix} = 71.08 \begin{bmatrix} 1.73 \\ 1.00 \\ 1.41 \end{bmatrix} \quad (7-7)$$

New Arius laser wavelengths

Following the change of the Arius blue laser wavelength to 445 nm (so-called 'sky blue'), different values of the coefficients are required:

$$\mathbf{M}_{\text{new}} = \begin{bmatrix} 0.5340 & 0.1908 & 0.3170 \\ 0.2170 & 0.8849 & 0.0424 \\ 0.0001 & 0.0363 & 1.5700 \end{bmatrix} \quad (7-8)$$

The inverse matrix is:

$$\mathbf{M}_{\text{new}}^{-1} = \begin{bmatrix} 2.0453 & -0.4246 & -0.4015 \\ -0.5021 & 1.2356 & 0.0680 \\ 0.0115 & -0.0285 & 0.6354 \end{bmatrix} \quad (7-9)$$

Thus

$$\mathbf{L}_{\text{Wnew}} = \begin{bmatrix} 108.65 \\ 83.28 \\ 64.41 \end{bmatrix} = 83.28 \begin{bmatrix} 1.30 \\ 1.00 \\ 0.77 \end{bmatrix} \quad (7-10)$$

Optimum laser wavelengths

From analysis at 5 nm intervals (Table 7.2), the optimum laser wavelengths were found to be 460, 535, 600 nm, giving values of the coefficients:

$$\mathbf{M}_{\text{opt}} = \begin{bmatrix} 1.0600 & 0.2280 & 0.2330 \\ 0.6310 & 0.9150 & 0.0600 \\ 0.0009 & 0.0294 & 1.2900 \end{bmatrix} \quad (7-10)$$

The inverse matrix is:

$$\mathbf{M}_{\text{opt}}^{-1} = \begin{bmatrix} 1.1037 & -0.2690 & -0.1868 \\ -0.7622 & 1.2803 & 0.0781 \\ 0.0166 & -0.0290 & 0.7735 \end{bmatrix} \quad (7-11)$$

Thus

$$\mathbf{L}_{\text{Wopt}} = \begin{bmatrix} 57.75 \\ 64.27 \\ 79.24 \end{bmatrix} = 64.27 \begin{bmatrix} 0.90 \\ 1.00 \\ 1.23 \end{bmatrix} \quad (7-12)$$

The variation of laser power around the optimum is shown in the three graphs of Fig. 7.40, with values normalised to 1.0 for the green laser at 535 nm. The red laser power is minimised for wavelengths in the range 600-605 nm, and for green laser wavelengths in the range 535-540 nm. The minimum for blue is at 440 nm, but the curve for the blue laser is anomalous because of the kink in the \bar{z} Judd-Vos tristimulus function between 440 and 460 nm.

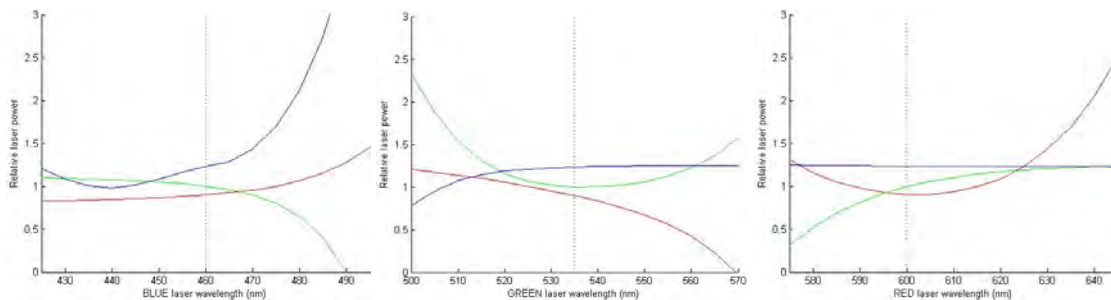


Figure 7.40: Variation in power of red, green and blue lasers required to match tristimulus values of D65 white for variation in wavelength of: (left) blue laser; (centre) green laser; and (right) red laser.

The behaviour of the error minimisation procedure was tested for the Swedish NCS dataset with difference metrics in other colour spaces, giving the results shown in Fig. 7.41 and Table 7.7. All the uniform colour spaces give broadly similar results for the optimum wavelengths, although the behaviour of the minima for the blue laser wavelength varies. The XYZ metric gives smaller minimum error values because it is linear in light intensity, and so de-emphasises dark colours relative to the lightness as a cube root of intensity in the UCS metrics. The performance of the Arius scanner with old and new wavelengths for blue is similar in all cases.

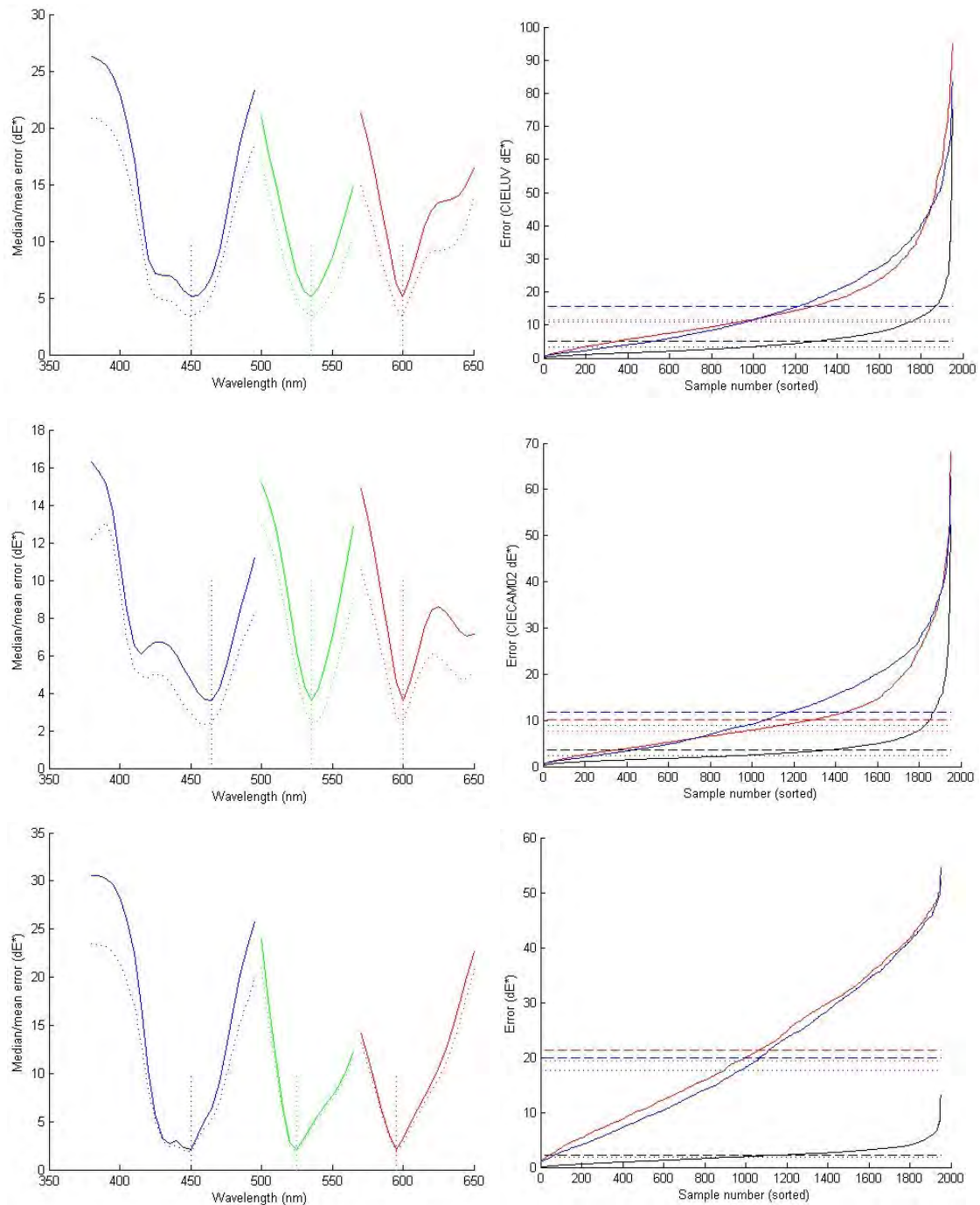


Figure 7.41: (left) Mean (solid line) and median (dotted line) values of colour difference for 1950 samples in the Swedish NCS colour atlas at 5nm intervals; (right) Colour difference values, sorted into ascending order. (top) CIELUV, (middle) CIECAM02; (bottom) XYZ.

Table 7.7: Mean/max error values using the Swedish NCS reflectance dataset for difference metrics in various colour spaces and white balance.

Model	White balance	Arius, B=475		Arius, B=445		Optimum		Ratio	Wavelengths
		mean	max	mean	max	mean	max	mean/opt	(nm)
CIELAB	any	15.81	79.33	13.38	63.70	3.65	41.99	4.0	460, 535, 600
CIELUV	equi-energy	15.69	95.01	15.70	83.40	5.11	75.06	3.1	450, 535, 600
CIELUV	D65	21.49	113.88	17.82	98.50	4.65	74.88	4.2	465, 535, 600
CIECAM02	equi-energy	10.22	67.99	11.68	62.83	3.61	52.21	3.0	465, 535, 600
CIECAM02	D65	15.50	79.95	13.05	69.42	3.57	52.16	3.1	465, 535, 600
XYZ	equi-energy	21.28	52.35	19.97	54.53	2.11	13.02	9.8	450, 525, 595
XYZ	D65	5.10	33.99	3.77	27.34	1.63	10.61	2.7	455, 535, 590

7.2.4 Effect on Colour Images

To demonstrate the visual effect of the choice of laser wavelengths, synthetic images were generated for the Macbeth Color Checker chart (Fig. 7.42) using the reflectance spectrum of each patch, previously measured at 10 nm intervals with an X-rite iIPro spectrophotometer. Tristimulus values were calculated under the corresponding light source (D65 or triple laser), then converted to $L^*a^*b^*$ using the light's tristimulus values as the reference white. The Judd-Vos observer was employed, using published 5 nm datasets for the observer and illuminant, and 5 nm reflectance data obtained by cubic interpolation of the 10 nm measurement data. Each colour was checked to ensure that it lay within the colour gamut of sRGB, and mapped if necessary onto the gamut boundary along a chord toward the $L^*=50$ point on the neutral axis. The colour was then converted from $L^*a^*b^*$ back to XYZ using D65 reference white, thence to 8-bit RGB signals for an sRGB display (Stokes et al, 1996).



Figure 7.42: Synthetic images of the Macbeth ColorChecker Chart for: (top left) D65 illumination; (bottom left) optimum laser scanner; (top right) old Arius scanner; (bottom right) new Arius scanner. The red spots indicate colours out of the sRGB gamut.

The colours of the chart under the wavelengths of the Arius scanner are grossly over-saturated, especially magenta, red, pink, orange and lime. Even the ‘light skin’ (second patch in the top row) is far too pink. The new Arius blue laser (445 nm) renders the blue patch too purple and the cyan too green. The colours of the chart under the optimum laser wavelengths give a generally acceptable visual match to the colours under D65, except notably for the red patch which is greatly desaturated because its reflectance rises sharply between 600 and 620 nm, and the laser at 600 nm falls just too short (Fig. 7.43 left). Changing the red laser wavelength to 605 nm would produce a better rendering of red without significant distortion of the other colours (see next section 7.2.5).

The relationship between the laser wavelengths and the reflectance spectra, and the consequent colour shifts in the CIELAB a^*b^* plane, can be seen in Fig. 7.43 right. The Arius red laser at 635 nm generates too large a signal for the red and magenta patches, causing a great increase in the a^* value relative to the D65 reference value. The Arius blue lasers at 473 (old) and 445 (new) lie on either side of the optimum 460 nm, and hence cause opposite directions of change for the blue patch.

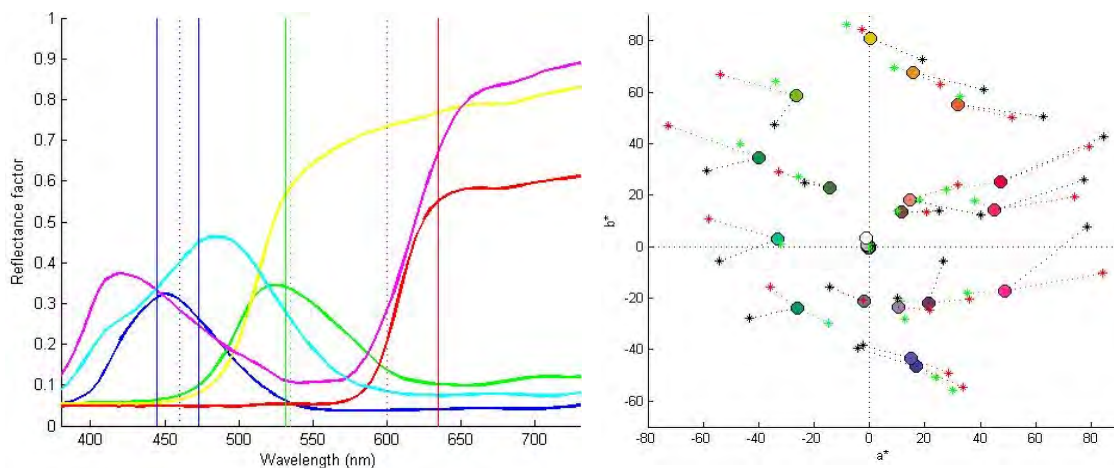


Figure 7.43: (left) Reflectance spectra of the six primary and secondary colours in the third row of the chart, with actual (solid lines) and optimal (dotted lines) laser wavelengths; (right) colour shift in the CIELAB a^*b^* plane for each of the 24 patches in the Macbeth ColorChecker Chart between D65 (circle) and: old Arius scanner (black); new Arius scanner (red); and optimum laser scanner (green).

To visualise the influence of the laser wavelengths on a more naturalistic scene, a test image was selected from the hyperspectral image set produced by Nascimento *et al* (2002). The image was originally acquired with a monochrome digital camera with a tunable birefringent filter mounted in front of the lens. The image size was 820 × 820 pixels, representing the normalised scene radiance at each point in 31 spectral bands from 400 to 700 nm at 10 nm intervals.

The image spectral data was interpolated to 5 nm intervals, giving a vector of 61 values for each pixel then rendered with the Judd-Vos observer and the four illumination sources (D65 and three laser wavelength combinations) to give the sRGB images in Fig. 7.40. The laser wavelengths of the scanner were taken to be the nearest multiple of 5 nm, i.e. 475-530-655 and 445-530-635 for the old and new scanners respectively. In each case the power of the three laser sources was adjusted to give the same XYZ tristimulus values as D65 white. The image from the optimum laser wavelengths has a reasonable overall colour balance, apart from a loss of saturation in the reds. Neither of the laser images in Fig. 7.44 is a good match to the

D65 image, with a red cast in yellow and orange and in the brown tiles in the background and in the basketball in the foreground. With either of the blue laser wavelengths the scanner grossly distorts the red areas and over-saturates the green areas.

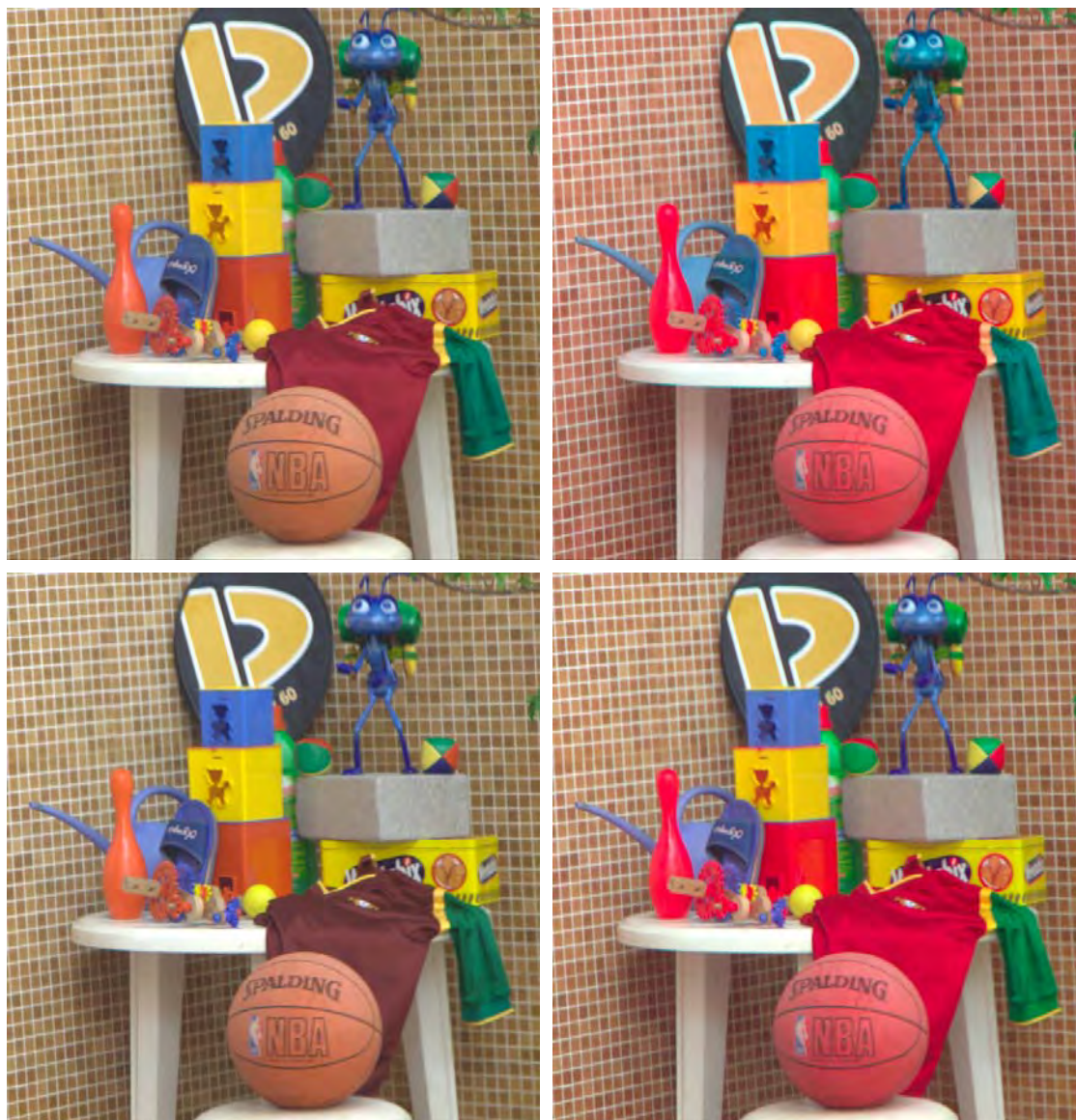
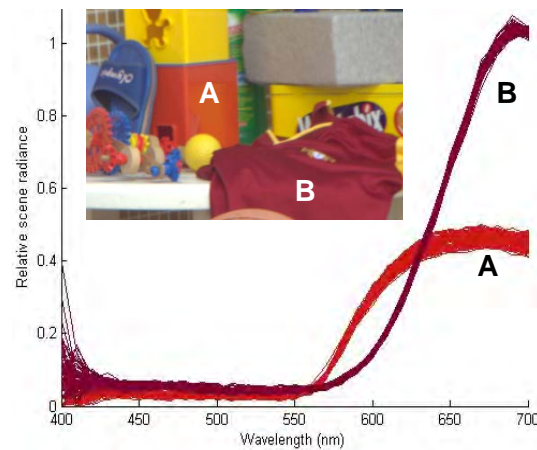


Figure 7.44: Hyperspectral image rendered with a D65 white point for: (top left) D65 illumination; (bottom left) optimum laser scanner; (top right) old Arius scanner; (bottom right) new Arius scanner.

The reason for the colour distortion of some reds with the different red laser wavelengths can be seen by comparing two areas in the image (Fig. 7.45). The spectra of one hundred points in a 10x10 pixel region are superimposed for the red plastic block at position A and the dyed cotton T-shirt at position B. Under D65 illumination the block appears to be a medium orange-red, and the T-shirt a medium-dark bluish red. At the Arius laser wavelength of 635 nm, the mean scene radiance of the two areas is nearly equal and they are rendered as similar in lightness. For the 'optimal' laser at 600 nm the reflectance of the T-shirt is much lower, and it is rendered darker.

Figure 7.45: Scene radiance spectra for two 10x10 pixel regions at A and B in the hyperspectral test image.



7.2.5 Revision of Optimal Red Laser Wavelength

Plotting the chromaticity coordinates of the 1950 samples in the Swedish NCS atlas (Fig. 7.46) shows that they all fit within the colour gamut triangle formed by the three optimal laser wavelengths except for a small cluster of red colours. Moving the red primary from 600 to 610 nm would enclose all but one of these colours. In general colours outside the gamut boundary cannot be accurately represented by positive values of the three signals, and for any out-of-gamut red colours a negative green signal would be required to match the chromaticity.

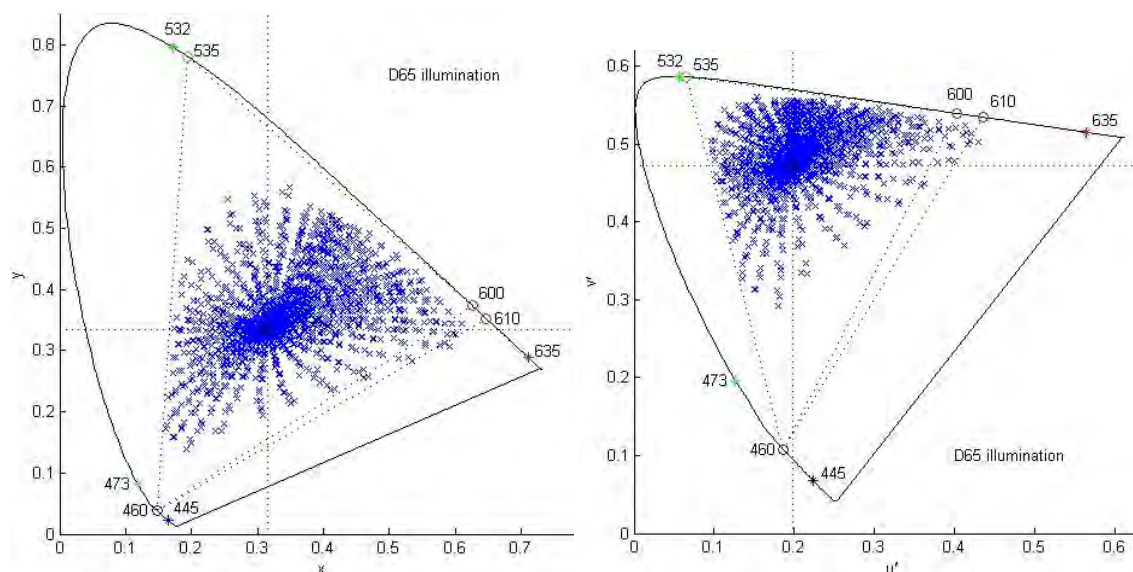


Figure 7.46: Chromaticity coordinates of 1950 samples from the Swedish NCS atlas, plotted in CIE xy (left) and $u'v'$ (right) diagrams. The outer line is the spectral locus, with wavelengths of the individual Arius lasers (*) and optimal lasers (o). The triangle shows the gamut of colours that can be represented by positive values of R,G,B.

For a quantitative assessment of colour accuracy, the images were converted from tristimulus values X,Y,Z into the CIECAM02 colour appearance space (Moroney et al, 2002). This gives better perceptual uniformity than CIELAB in all colour regions. The pixel by pixel difference between each image rendered under the laser illumination and the image rendered under D65 was averaged over all 820x820 pixels in the image and the means plotted against the red laser wavelength (Fig. 7.47). There is not much change in the mean lightness (J) or blue-yellow (b) values, but the mean red-green (a) value changes strongly, crossing the axis at approximately 603 nm. The RMS difference value (ΔE_{CAM}) accordingly reaches its minimum value at 605 nm, beyond which it rises slowly.

The hyperspectral test image was rendered for the fixed optimal blue (460 nm) and green (535 nm) laser wavelengths and with a series of red laser wavelengths from 585 to 620 nm, adjusting the power of the three lasers in each case to match D65 white. The results (Fig. 7.48) show clearly the increasing redness of the red objects in the scene. Visual comparison of the images with the D65 rendering (centre), shows that the basketball is best matched at 595 nm, the red plastic block and the background wall tiles at 605 nm, and the T-shirt at 615 nm. The background and the greens grow steadily lighter throughout the series, with the best overall match of lightness at 605 nm.

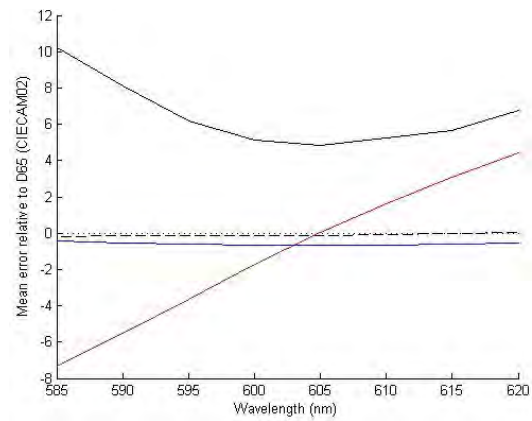


Figure 7.47: Mean CIECAM02 errors vs wave-length of the red laser: ΔE_{CAM} (solid black line), Δa (red line), Δb (blue line), and ΔJ (dotted line).

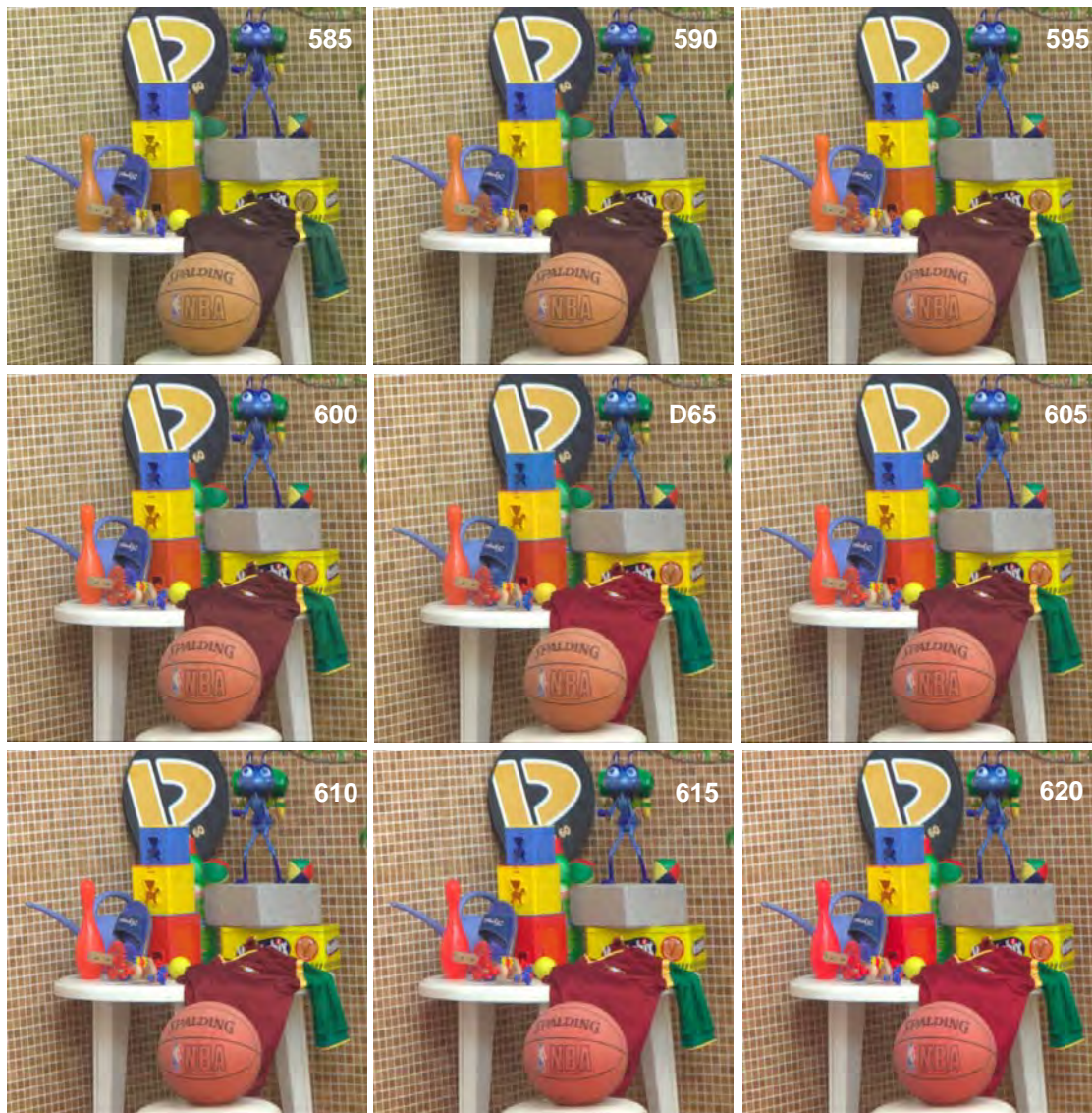


Figure 7.48: Rendering of the hyperspectral test image for: (centre) D65 illuminant; (outside) triple laser wavelengths of 460 (blue), 535 (green) and variable 585-620 (red).

The change in coordinates in the CIECAM02 a - b plane can be seen clearly in Fig. 7.49 (right), where the positions are plotted of the average colour in a 10x10 pixel square at 26 selected positions in the image. The greatest changes in the a coordinate occur for the four red objects (red skittle, red block, T-shirt and basketball), as expected from the changing visual appearance of the images in Fig. 7.48.

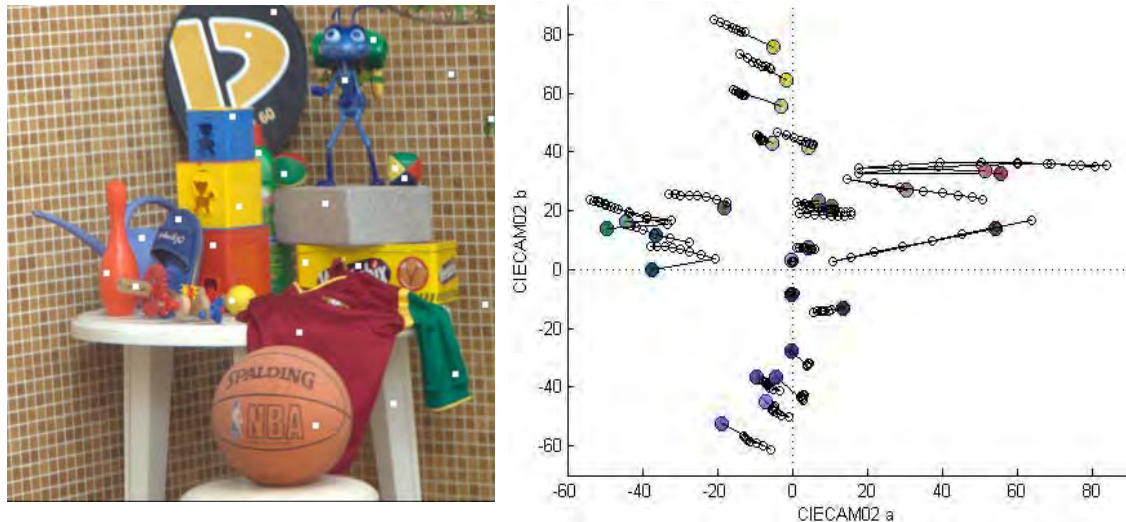


Figure 7.49: (left) Positions of 26 sample areas indicated by white squares; (right) CIECAM02 coordinates of each area under D65 (large coloured circle) and trajectory of colour under series of red laser wavelengths.

The behaviour of the coordinates with changing red laser wavelength can be better understood by examining the histograms of the errors between successive images and the D65 image (Figs. 7.50 and 7.51). For laser wavelength of 585 nm the Δa difference values (Fig. 7.50) are rather broadly spread over the range -20 to +20, but become more concentrated in the range -10 to +10 for the wavelengths 600–610 nm, before spreading again for higher wavelengths. Also the side lobe, representing the pixels of the T-shirt, progresses from left to right through the distribution as the wavelength increases, reaching the centre position at 610 nm. The lightness ΔJ difference values (Fig. 7.51) exhibit a similar behaviour, although the spread is narrower in the range from -2 to +2 approximately. Again the lobe representing the T-shirt, which is much darker at 585 nm, progresses through the distribution and is balanced at 615 nm.

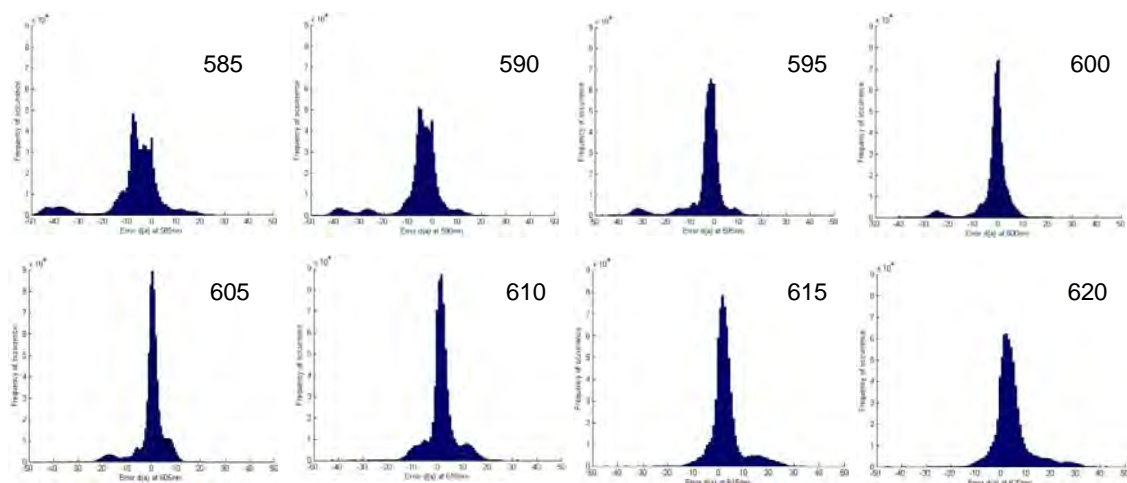


Figure 7.50: Histograms of CIECAM02 Δa differences between image illuminated by triple laser and D65.

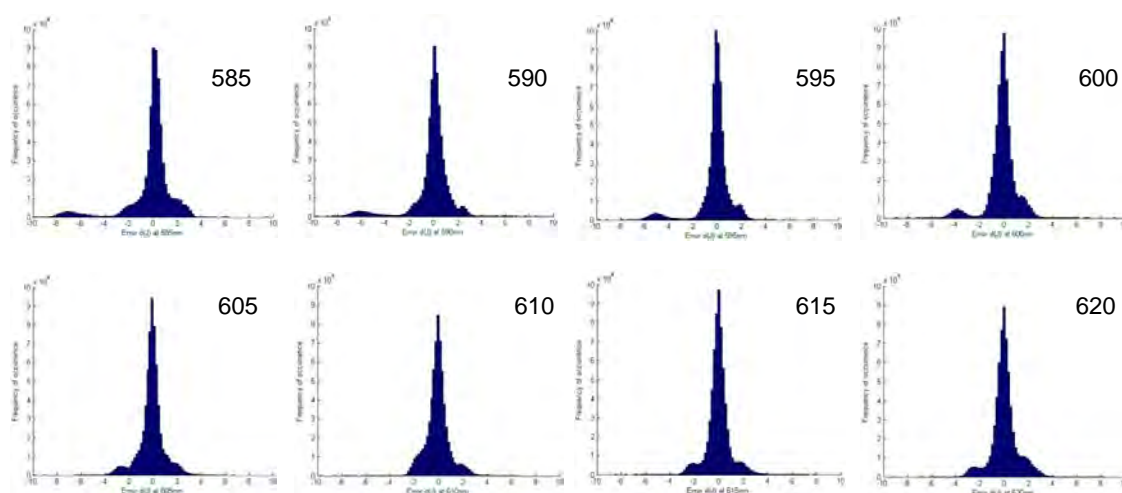


Figure 7.51: Histograms of CIECAM02 ΔJ differences between image illuminated by triple laser and D65.

Taking all of the above into consideration, it would seem that the optimal wavelength for the red laser should be in the range 605–610 nm, i.e. a little longer than indicated by the analysis of sets of colorants.

7.2.6 Relationship to Prime Wavelengths of Vision

The optimum three laser wavelengths identified in this study of 450–460, 530–535 and 595–605 nm are very close to the ‘prime colors’ of approximately 450, 530 and 610 nm identified by Thornton as the wavelengths of peak visual sensitivity¹. Thornton (1971) showed that for ‘white’ light made up of three monochromatic wavelengths, the colour rendering index (CRI) could be optimised by choosing the three prime wavelengths (Fig. 7.52 left). He demonstrated this empirically by holding constant two of the wavelengths 450, 540, 610 nm and varying the third and calculating the CRI for each combination. Repeating the calculations for various power ratios of the three sources, to achieve various correlated colour temperatures, always gave the same result. He also showed that certain other wavelengths, which he termed ‘anti-prime’, gave disastrous results, namely wavelengths near 495, 575 and beyond 640 nm.

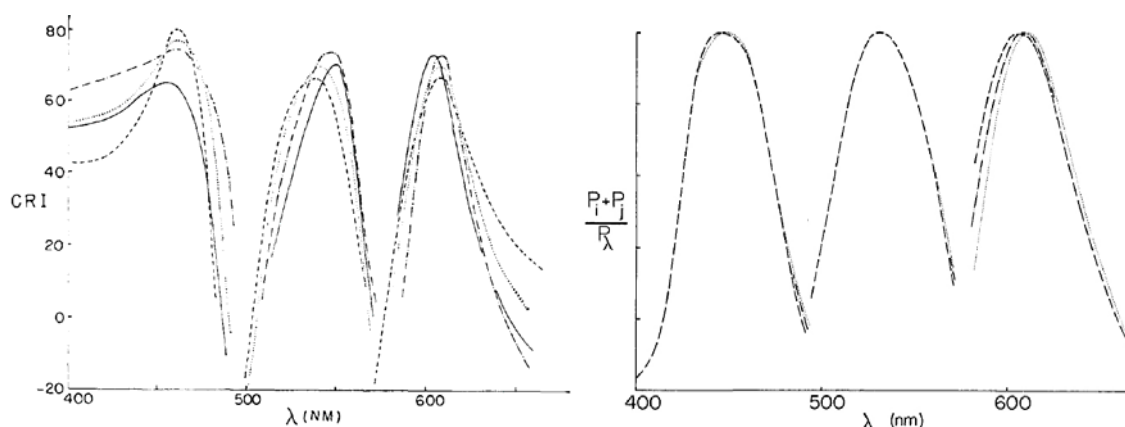


Figure 7.52: (left) CRI of white light composed of three spectral colours, from Thornton (1971); (right) Normalised power ratio in two spectral colours, from Thornton (1972).

¹ Thornton (1999) later refined the specifications for the prime wavelengths of the average human, CIE 2°, and CIE 10° observers respectively to: Blue 450, 447, 446 nm; Green 533, 541, 538 nm; Red 611, 604, 600 nm.

He concluded that “... the [optimum] three-line SPD is superior to that of the commercial fluorescent lamp of the same color in both luminosity and CRI, notwithstanding the common opinion that discontinuous distributions cannot provide good color rendering”, and further “it is apparently a more efficient use of power to sample the reflectance spectrum of an object at these three unique wavelengths than to waste power by reflection from the object in the insensitive regions near 500 nm, 580 nm and in the deep violet and deep red.”

In another study, Thornton (1972) showed that the prime wavelengths are particularly effective in altering the perceived chromaticity of a light source or illuminated object. Plotting the normalised ratio $(P_i + P_j)/P_\lambda$ of power in two monochromatic colours i, j of fixed wavelengths relative to a variable third wavelength λ from 450, 540 and 610 nm (Fig. 7.52 right) shows clearly how the efficacy falls away from the optimum as any of the lights moves away from the corresponding prime wavelength. In effect the curves may be considered as the visual response (or luminosity) per watt as a function of wavelength, i.e. the inverse of visual sensitivity. Thornton summarised that “The chromaticity of an element in any visual scene is established with minimum power input to the eye when light from that element is composed of spectral colors near the prime wavelengths.” This is confirmed by the analysis of the laser power required to balance D65 white (Fig. 7.40).

Thornton (2000) later showed that the three prime wavelengths are those at which unit-power monochromatic lights induce the largest tristimulus gamut (in terms of the volume of the parallelepiped spanned by the tristimulus vectors of these lights) and are therefore optimal as the dominant wavelengths of the primaries for displays. They are also the wavelengths at which natural metameric reflectance spectra tend to cross each other (Brill *et al*, 1998).

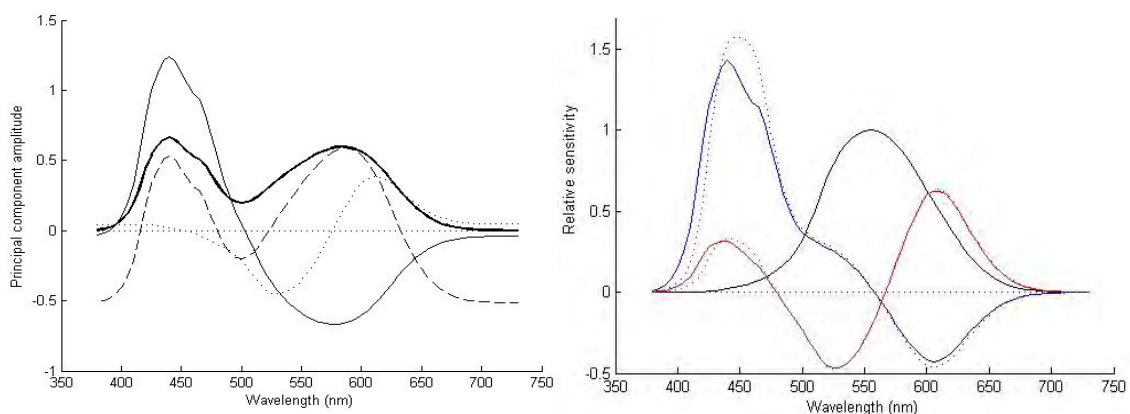


Figure 7.53: (left) Mean and principal components of the three Judd-Vos tristimulus functions; (right) Y tristimulus function and two orthonormal vectors for Judd-Vos observer (dotted lines) and CIE Standard 2-deg observer (solid lines).

Applying PCA to the Judd-Vos tristimulus functions gives the three principal components shown in Fig. 7.53 (left). The mean (bold line) has a broad saddle form and the variances of the three components are 60.3, 30.5 and 9.2 percent respectively. The first component (solid line) corresponds to the blue-yellow axis, the second component (dashed line) follows the mean, and the third component (dotted line) corresponds to the red-green axis. On this basis the system has much greater sensitivity to yellow-blue hue changes than to red-green variations (the subtlety of the latter arising from the substantial spectral overlap of the L and M retinal photoreceptor sensitivities).

Another approach to separation of the tristimulus functions is Gram-Schmidt orthonormalisation, starting with the luminance function \bar{y} as the first basis vector and then projecting the \bar{x} and \bar{z} functions onto it:

$$U_1 = \bar{y}, \quad U_2 = \bar{x} - \frac{\bar{x} \cdot \bar{y}}{\bar{y} \cdot \bar{y}} \bar{y}, \quad U_3 = \bar{z} - \frac{\bar{z} \cdot \bar{y}}{\bar{y} \cdot \bar{y}} \bar{y} - \frac{\bar{z} \cdot \bar{x}}{\bar{x} \cdot \bar{x}} \bar{x} \quad (7-13)$$

The resulting orthogonal functions (Fig. 7.53 right), which correspond closely to the decorrelated luminance and opponent chrominance signals in the neural visual pathway, have an obvious red-green and blue-yellow formation. The relationship between the orthonormal functions U_1, U_2, U_3 and the tristimulus functions $\bar{x}, \bar{y}, \bar{z}$ is:

$$\begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -a_{21} & 1 & 0 \\ -a_{31} & -a_{32} & 1 \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -0.7337 & 1 & 0 \\ -0.1114 & -0.5770 & 1 \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x} \\ \bar{z} \end{bmatrix} \quad (7-14)$$

For the CIE standard observer the maxima of the second and third orthonormal functions are at 610 and 450 nm, and the minima are at 525 and 605 nm. For the Judd-Vos observer the maxima are shifted to 605 and 440 nm. From this it is clear that the significance of the prime wavelengths is that they produce maximal excitation of the opponent channels, and hence the greatest chromatic response from the human visual system. At the anti-prime wavelength of 565 nm the two orthonormal functions cross one another close to zero, and so this wavelength has the least ability to produce a chromatic response.

Applying Gram-Schmidt orthonormalisation directly to the cone fundamentals gives the curves shown in Fig. 7.54 (left). In this case, the Stockman-Sharpe linear data at 1nm intervals are employed and the mean $(L+M)/2$ is taken as the first basis vector U_1 . Onto this S is projected to form the second basis vector U_2 and onto this is projected L to form the third basis vector U_3 . The maxima of U_3 are at 607 nm (positive) and 524 nm (negative), with the zero crossing of U_3 at 566 nm. A simpler analysis shows that the ratio M/L peaks at 461 nm (Fig. 7.54 right). Mollon *et al* (1993) note the peak in wavelength discrimination at approximately 460 nm, and attribute this to the maximum in the ratio M:L of middle to long wavelength cone sensitivity. Also the difference function M-L has minimum value at 520 nm and maximum value at 605 nm.

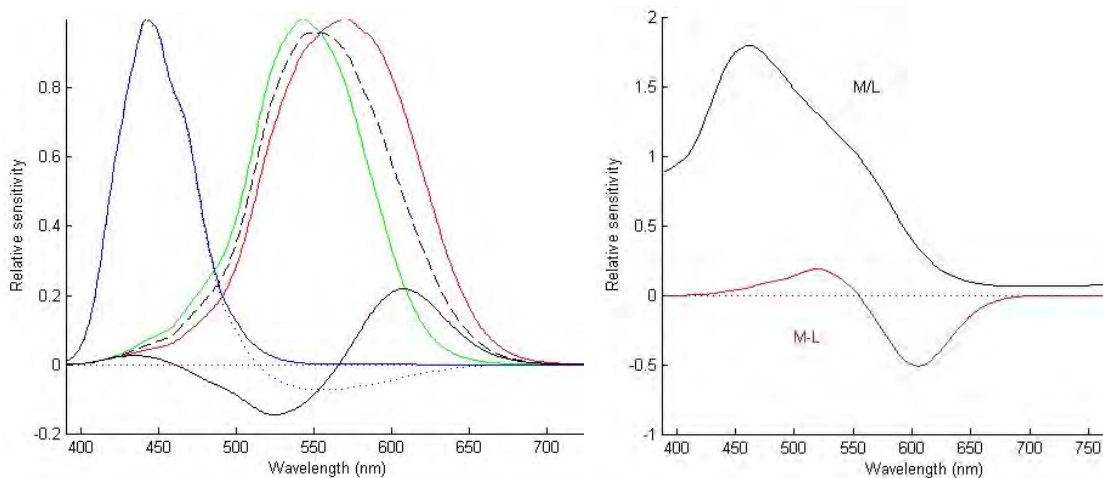


Figure 7.54: (left) Cone fundamentals shown in red, green and blue for L,M,S cones respectively (CIE, 2007) and their orthonormal basis vectors: U_1 = mean of L and M (dashed line), U_2 = opposing S (dotted line) and U_3 = opposing L (solid black line); (right) simple ratio (black) and difference (red) of M and L cone fundamentals..

Buchsbaum and Gottschalk (1983) applied principal component analysis to the Vos-Walraven cone sensitivity functions, to derive an achromatic and two opponent chromatic channels (Fig. 7.55). They showed that such channels provide an efficient transmission of information in the visual pathway, and thus could reconcile the Young-Helmholtz trichromatic receptor theories of vision with the Hering and Hurvich-Jameson opponency theories. The maxima and minima of the two opponent channels lie at 440, 530 and 610 nm, exactly the prime wavelengths identified by Thornton.

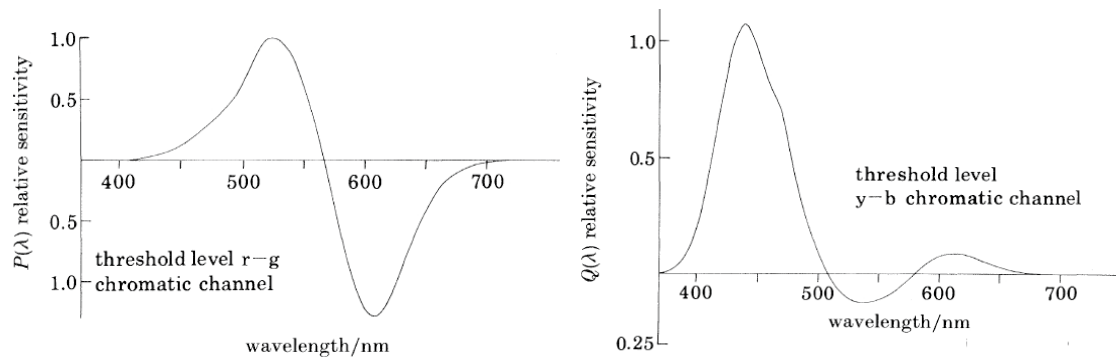


Figure 7.55: Second and third principal components of the Vos-Walraven photoreceptor sensitivity functions, from Buchsbaum and Gottschalk (1983).

The improved visual efficiency of the optimal laser wavelengths can be seen in the CIE general colour rendering index (CRI), which increases from 48.6 for the approximate Arius wavelengths of 475, 530, 635 nm to 75.4 for the optimal laser wavelengths of 460, 535, 600 nm. This is largely due to the reduced shift of colorimetric values toward red (Fig. 7.56 right). The change in the blue laser wavelength from 473 to 445 nm (Fig. 7.56 left), however, makes the colour rendering substantially worse, with the correlated colour temperature raised to over 13K and the CRI reduced to 21.4 because not only the red but also the green colours are distorted away from their true values.

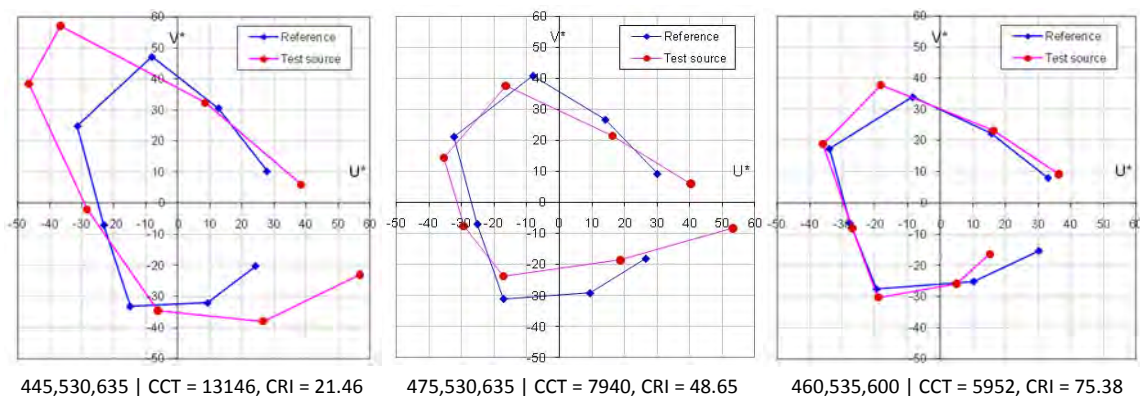


Figure 7.56: Colour rendering of the eight CIE test colours: (left) under the new Arius laser wavelengths; (centre) under the old Arius laser wavelengths; (right) under the optimal laser wavelengths;

This study confirmed the significance of the prime wavelengths of vision as the optimal wavelengths for three monochromatic lasers in a scanner. By choosing these optimal wavelengths the colorimetric performance of the scanner would be almost as good as if two additional lasers of intermediate wavelengths were added to the existing laser of the Arius scanner.

7.3 Principal Components Analysis of Reflectance Spectra

7.3.1 Introduction

Principal components analysis (Sec. 2.5.7) was applied to ten different sets of spectral reflectance data, to determine how many components are needed to represent the spectra within a specified colorimetric accuracy:

1. GretagMacbeth DC Chart – acrylic paints – 180 patches
2. Berns-Taplin pigment set – acrylic paints – 100 samples
3. National Gallery pigment set – oil paints – 64 samples
4. Cheung-Westland natural dataset – leaves and vegetation – 494 samples
5. Foster-Nascimento hyperspectral images – natural scenes – 500 samples
6. Synthetic spectra – generated from waveforms – 500 spectra
7. Chinese CNCS spectra – acrylic paints – 899 samples
8. Swedish NCS spectra – acrylic paints – 1950 samples
9. Munsell matte spectra – acrylic paints – 1269 samples
10. Munsell glossy spectra – acrylic paints – 1600 samples

The calculations were performed on spectral reflectance data in the range 380-730 nm at 5 nm intervals (71-element vector), with the corresponding data for CIE standard illuminant D65 and the Judd-Vos modification of the CIE standard 2-degree observer. The procedure is described here only for the dataset formed from spectrophotometric measurements of the GMDC chart. The details for the other datasets may be found in Appendix 3.

7.3.2 GretagMacbeth DC Chart

The GretagMacbeth DC Chart (Fig. 2.68) is photographic test target containing 12x20=240 patches overall, and 10x18=180 colour patches within the white/grey/black patches around the periphery. The chart was measured with the X-Rite eye-one spectrophotometer, and the spectra recorded at 10 nm intervals from 380 to 760 nm (39 values). The data were interpolated (by cubic spline) to 5 nm intervals and the last four values were discarded, giving a vector of 71 reflectance values 380-730 nm.

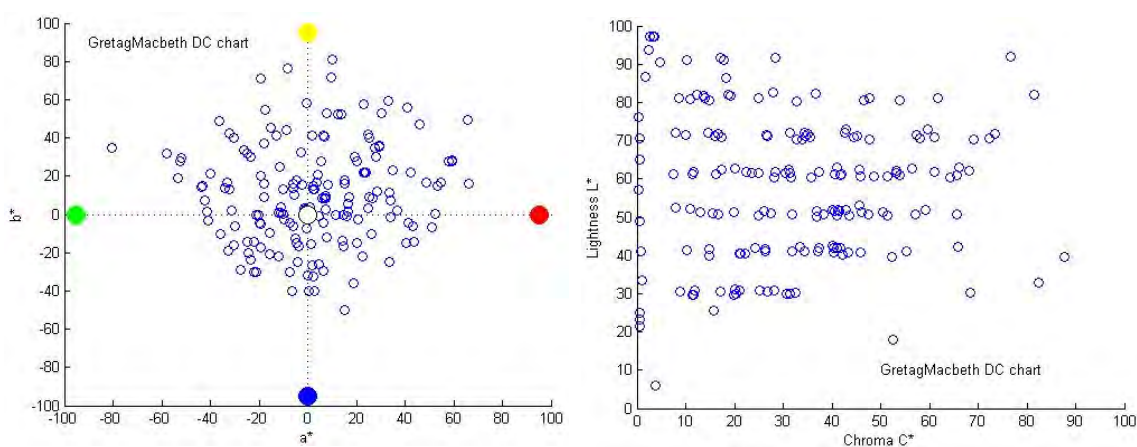


Figure 7.57: Colours of 180 patches of GretagMacbeth DC chart, plotted in CIELAB space for D65: (left) projection in a^*-b^* plane; (right) projection onto L^*-C^* plane.

The colours corresponding to the 180 patches are shown in CIELAB coordinates in Fig. 7.57. It is clear that many of the colours are set on a grid with discrete values of lightness, hue and

chroma. The spectra of all patches are shown in Fig. 7.58 (left), together with the mean (in red). The first five principal components are shown in Fig. 7.58 (right). Successive components cross the zero axis the same number of times as the component number, and their amplitude is reduced monotonically.

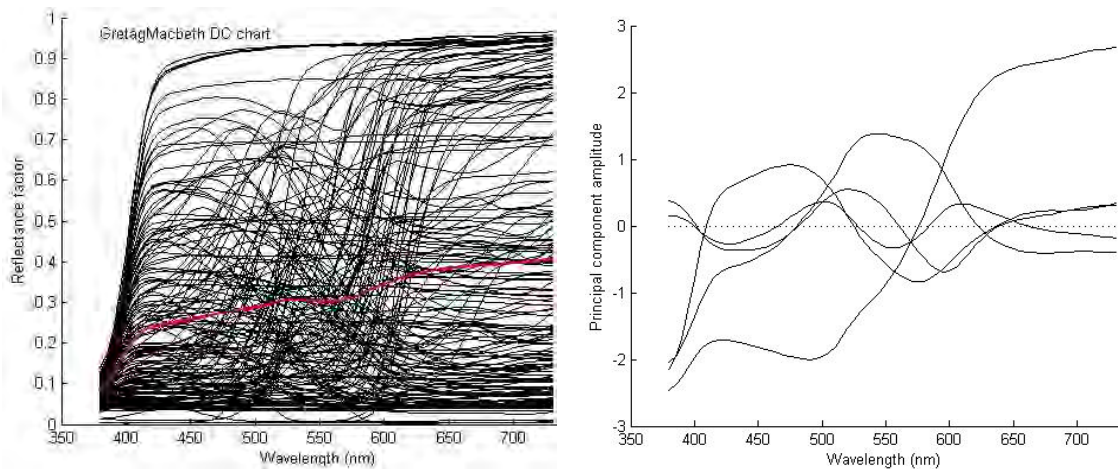


Figure 7.58: (left) reflectance spectra of 180 patches of GretagMacbeth DC chart, with mean (red); (right) spectra of first five principal components.

The variance and colour errors associated with each component and the cumulative variance are illustrated in Fig. 7.59. The first five components account for 99.3% of the variance, and over the whole set of samples the spectra reconstructed with five components have a mean colour error (ΔE^*_{ab}) of 0.77, with a maximum error of 8.32.

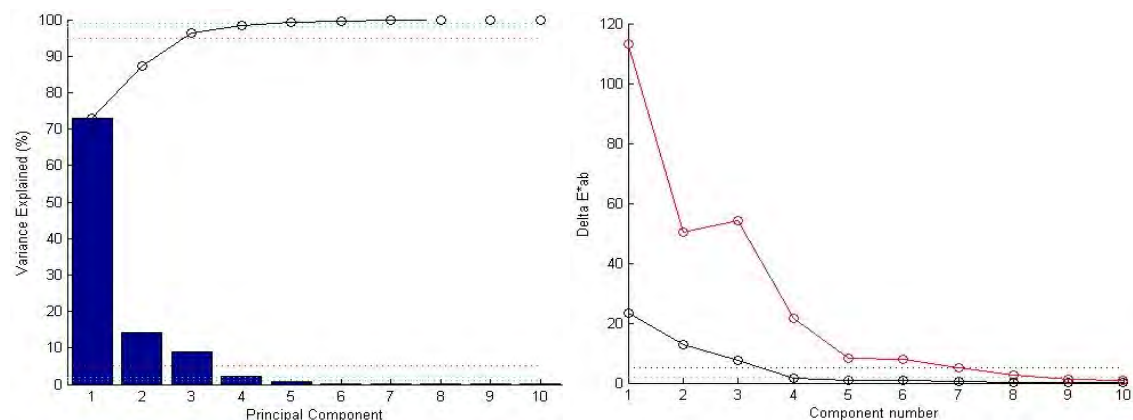


Figure 7.59: (left) variance explained by first ten principal components, with cumulative distribution; (right) maximum (red) and mean (blue) colour error (ΔE^*_{ab}) corresponding to the difference between the true spectrum of each sample and that reconstructed from increasing numbers of principal components.

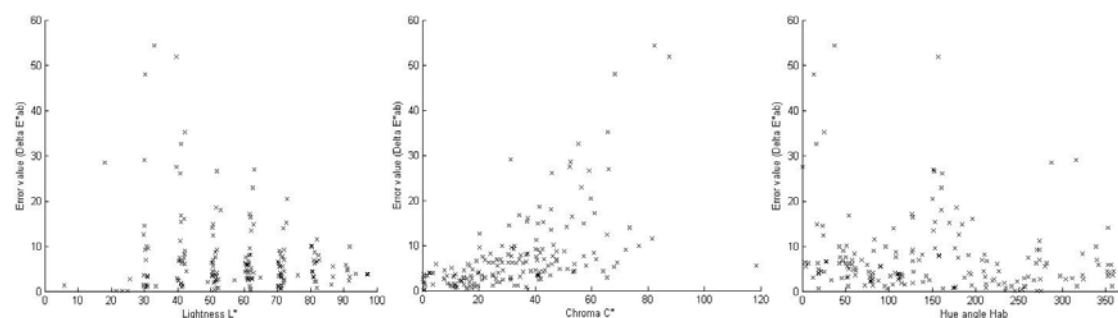


Figure 7.60: Distribution of errors between true and reconstructed spectra, plotted against: (left) lightness, (centre) chroma, and (right) hue angle in CIELAB coordinates.

The ΔE^*_{ab} values for reconstruction with three components are plotted in Fig. 7.60. Errors tend to be greater for lower levels of lightness L^* , and higher levels of chroma C^* . Three samples are outliers, with much larger errors than the others; these are the green, red and magenta gloss samples in column S of the chart. The accuracy of reconstruction of the reflectance spectrum by addition of successive components is illustrated in Fig. 7.61 (left) for the cyan sample at patch S11 of the chart, where the approximation becomes closer and closer to the true spectral waveform. The error, i.e. the difference between true spectrum and approximation, shown in Fig. 7.61 (right), tends to zero as more components are added.

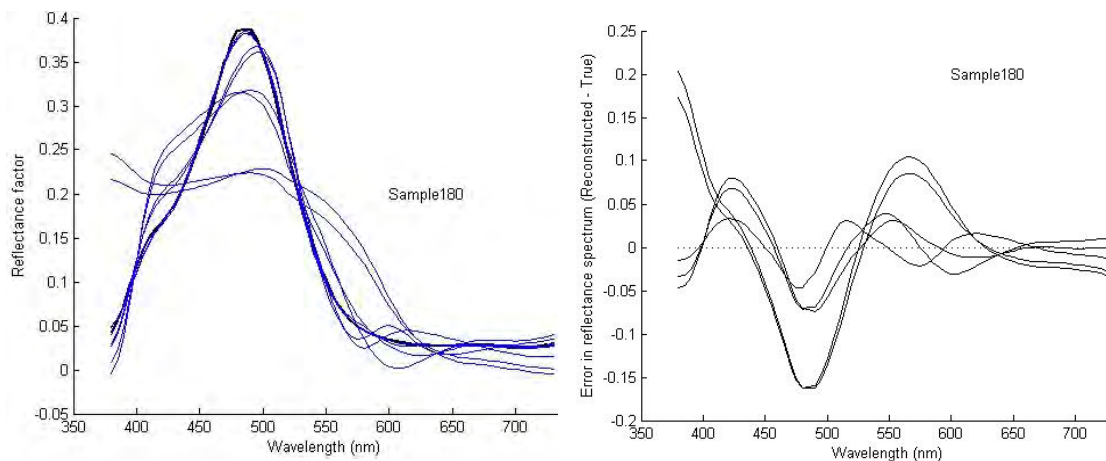


Figure 7.61: (left) reconstruction of reflectance spectrum of one sample from first ten principal components; (right) error between the true spectrum and the successive reconstructions from first five principal components.

Summarising the above results together with those of the eight datasets in Appendix 3, the number of principal components required to reconstruct the spectra in each dataset of measured surface colours is given in Table 7.8, for three criteria. For the first two criteria five or six components are sufficient, whereas for the third eight to ten components are needed.

Table 7.8: Number of components required for reconstruction to given accuracy.

Dataset	Samples	Cumulative variance > 99%	Mean ΔE^*_{ab} < 1.0	Max ΔE^*_{ab} < 3.0
Gretag-Macbeth DC chart	180	5	5	8
Berns-Taplin acrylic paints	100	6	6	8
National Gallery pigments	64	6	6	6
Cheung natural materials	494	5	4	6
Chinese CNCS dyed textiles	899	5	6	7
Swedish NCS colour atlas	1950	5	5	7
Munsell matte colour atlas	1269	6	6	9
Munsell glossy colour atlas	1600	6	6	10

7.3.3 Synthetic Reflectance Spectra

Pursuant to the analysis of spectral sampling in Sec. 2.5.6, the frequency limit of any continuous spectral waveform may be estimated by the maximum slope, obtained from first-order differences. As an example, Fig. 7.62 (left) shows the reflectance spectrum of Sample 21 in the combined pigment set, a strong red, and its first derivative which has a maximum value of 0.12, meaning the maximum change in reflectance factor over an interval of 5 nm, i.e. 0.024/nm.

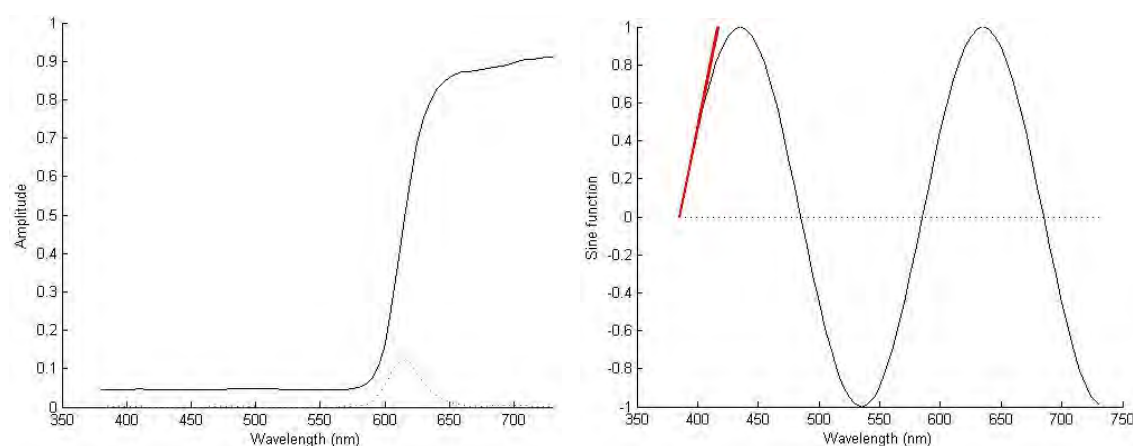


Figure 7.62: (left) Reflectance spectrum and its first derivative; (right) sine wave with a wavelength of 200 nm (= frequency of 0.005 cycles/nm). The red line shows the maximum slope.

The corresponding spectral frequency may be estimated by considering a sine wave (Fig. 7.62 right), which has the form and derivative of:

$$f(\lambda) = \sin\left(2\pi\left(\frac{\lambda-\lambda_0}{\Lambda}\right)\right), \quad f'(\lambda) = \frac{2\pi}{\Lambda} \cos\left(2\pi\left(\frac{\lambda-\lambda_0}{\Lambda}\right)\right) \quad (7-15)$$

where Λ is the period of the wave. The maximum slope of the sine function is the slope $2\pi/\Lambda$ of the tangent to the curve as it crosses the axis. In Fig. 7.62 (right) the period $\Lambda = 200$ nm corresponds to a frequency of 0.5 cycles per 100 nm, and the maximum slope is 0.0314 nm^{-1} (Table 7.9).

Table 7.9: Relationships between spectral frequency and sampling rates.

Period (nm)	Frequency (cycles per 100 nm)	Frequency (cycles/nm)	Max slope nm^{-1}	Max slope per 5 nm	Nyquist sampling period (nm)	Min samples over 400-700 nm range
50	2	0.02	0.126	0.630	25	12
100	1	0.01	0.063	0.315	50	6
200	0.5	0.005	0.031	0.155	100	3
300	0.33	0.0033	0.021	0.105	150	2

The first derivatives of all 334 reflectance samples in the combined paint dataset were calculated and the maximum and minimum at each wavelength determined (Fig. 7.63). The peak in maximum slope at 405 nm is probably an artefact of the spectrophotometer, caused by low sensitivity in the ultraviolet region (see for example the family of reflectance spectra in Fig. 7.58). The positive slopes have maxima at 520, 585, 615 and 690 nm, and the negative slopes at 510 nm. The maximum positive slope per 5 nm is $0.1354 = 0.027 \text{ nm}^{-1}$, and maximum negative slope is $-0.0523 = -0.01 \text{ nm}^{-1}$. The maximum slope corresponds to a sine wave of period 232 nm and frequency 0.0043 cycles/nm, within the Nyquist limit of three samples over the width of the spectrum.

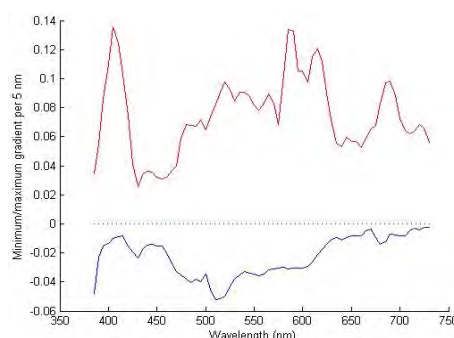


Figure 7.63: Maximum and minimum slopes of reflectance spectra.

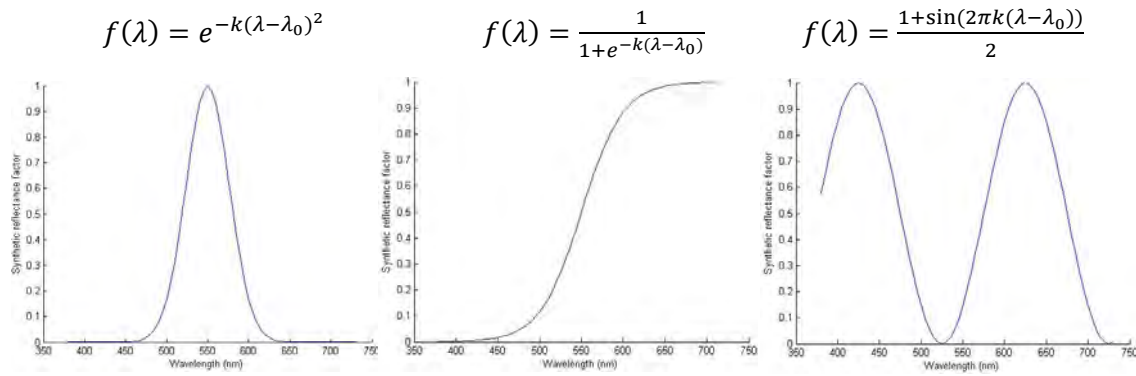


Figure 7.64: Generating functions for synthetic spectra: (left) Gaussian; (centre) logistic; (right) sine.

As an alternative to measured spectra, a set of synthetic idealised reflectance spectra was generated, using Gaussian and logistic and sine functions to generate single-valued waveforms as a function of wavelength λ (Fig. 7.64). The central wavelength λ_0 and width parameter k were randomised to produce a family of 100 curves for each function. A further 100 curves were generated for the inverse form $1 - f(\lambda)$ of the Gaussian and logistic, giving 500 synthetic spectra in total. The central wavelengths ranged from 400 to 650 nm, and the full range of reflectance factors was utilised, from 0 to 1. Limits were set on the width parameter to restrict the maximum slope of curves within range ± 0.15 per 5 nm interval (Fig. 7.65), i.e. $\pm 0.03 \text{ nm}^{-1}$.

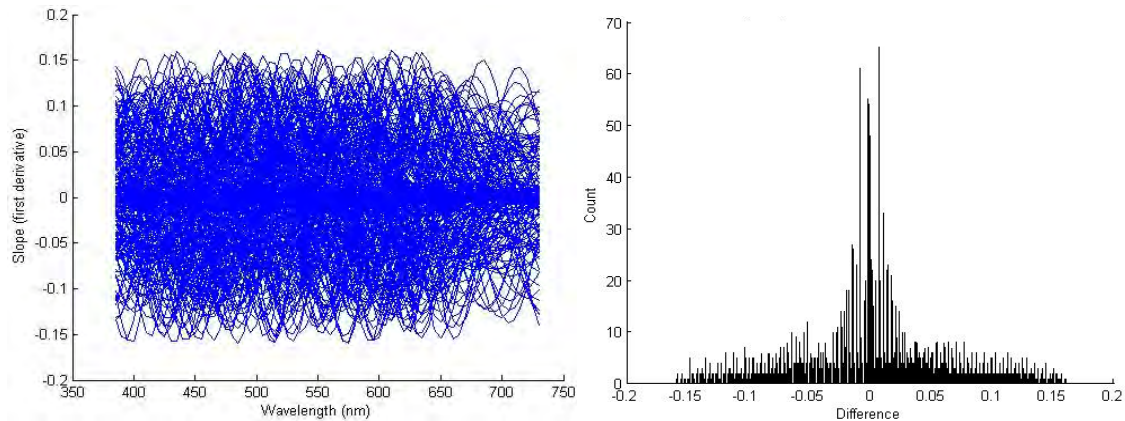


Figure 7.65: (left) First derivatives of all synthetic spectra; (right) histogram of maximum slopes.

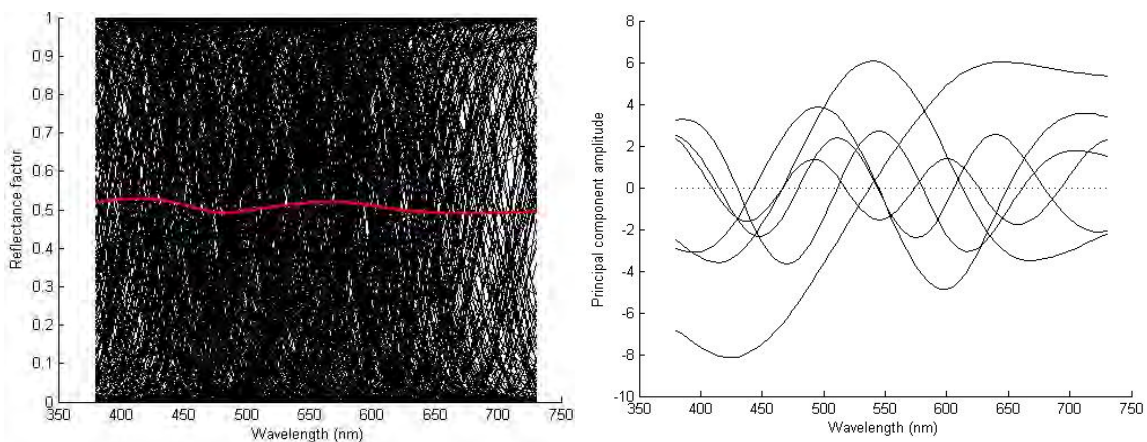


Figure 7.66: (left) 500 synthetic reflectance spectra, with mean (red); (right) first five principal components.

The combined set of spectra is shown in Fig. 7.66 (left), together with the mean (in red) which is very uniform and close to 0.5. The first five principal components are shown in Fig. 7.66 (right). Successive components are very nearly sinusoidal, rather like a Fourier series, and cross the zero axis the same number of times as the component number with diminishing amplitude.

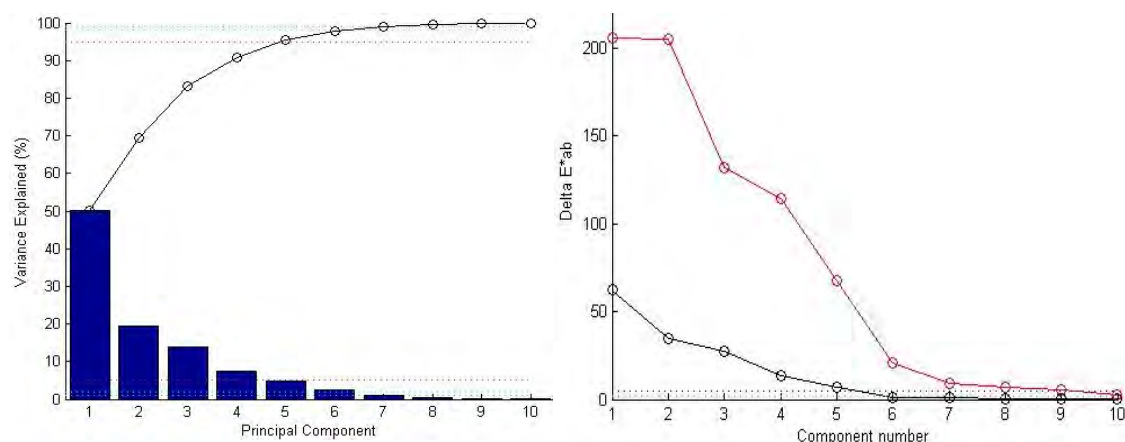


Figure 7.67: (left) variance explained by first ten principal components, with cumulative distribution; (right) maximum (red) and mean (blue) colour error (ΔE^*_{ab}) corresponding to the difference between the true spectrum of each sample and that reconstructed from increasing numbers of principal components.

The variance and colorimetric errors associated with each component and the cumulative variance are listed in Table 7.10 and illustrated in Fig. 7.67. Seven components are needed to account for 99% of the variance, and over the whole set of samples the spectra reconstructed with seven components have a mean colour error (ΔE^*_{ab}) of 0.86, with a maximum error of 9.49. The greater variance of the synthetic spectra is evident in these figures, showing the need for a seventh component in reconstruction to bring the mean ΔE^*_{ab} below 1.0, compared with five components for the measured GretagMacbeth DC chart and six components for the measured acrylic and oil paint charts (Table 7.8).

Table 7.10: Variance and colour errors associated with successive principal components.

Component	Variance	Cumulative Variance	Mean ΔE^*_{ab}	Max ΔE^*_{ab}
1	50.10	50.10	62.32	205.71
2	19.39	69.49	34.54	204.84
3	13.84	83.33	27.27	131.61
4	7.51	90.84	13.29	114.33
5	4.71	95.55	7.05	67.90
6	2.37	97.92	1.41	20.56
7	1.04	98.96	0.86	9.49
8	0.51	99.47	0.69	6.76
9	0.27	99.74	0.44	5.59
10	0.13	99.87	0.17	2.40

The $L^*a^*b^*$ coordinates of the 500 synthetic spectra were calculated for the D65 illuminant (Fig. 7.68). The locus of values away from the white point arises from the logistic function moving along the wavelength axis, toward yellow for the logistic function with positive slope (reduction of blue reflectance), and toward cyan for the logistic function with negative slope (reduction of red reflectance). The ΔE^*_{ab} values for reconstruction with seven components are plotted in Fig. 7.69. Errors tend to be greater for higher levels of chroma C^* and for hue angles of 90° (yellow) and 315° (violet).

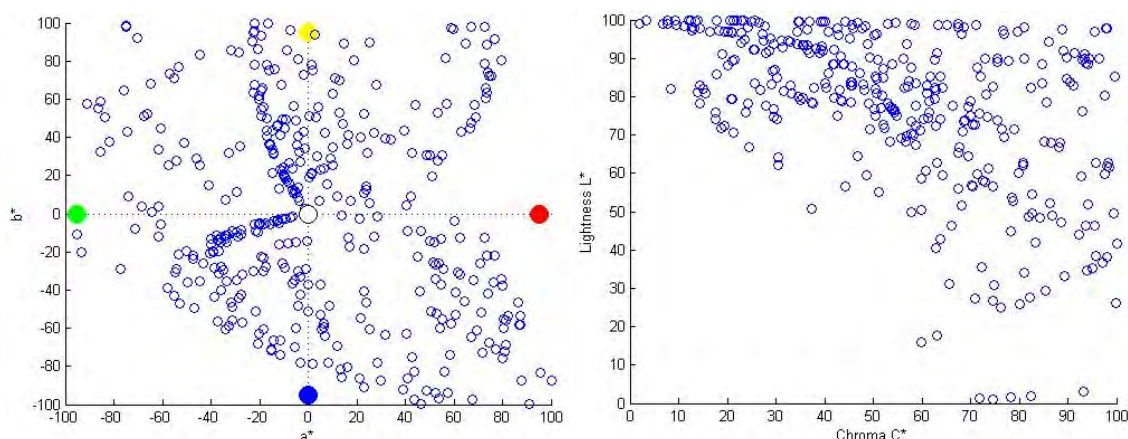


Figure 7.68: Colours of synthetic spectra in CIELAB space for D65: (left) in a^*-b^* plane; (right) in L^*-C^* plane.

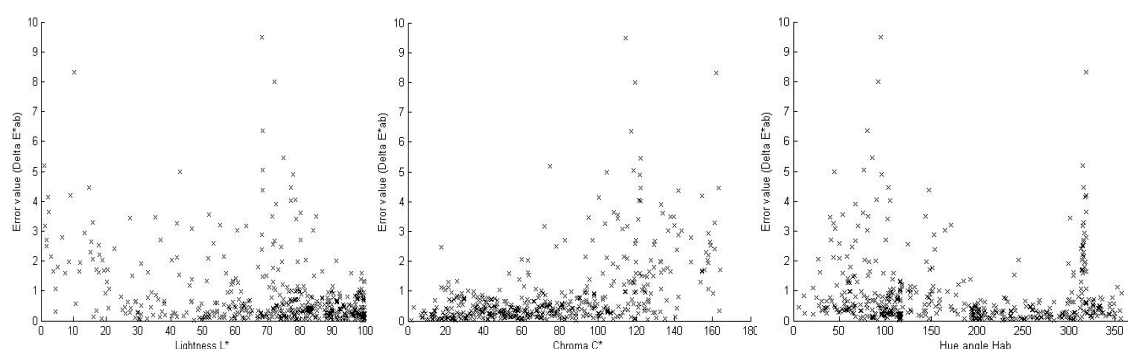


Figure 7.69: Distribution of colorimetric errors between true and reconstructed spectra for seven components, vs: (left) lightness, (centre) chroma, and (right) hue angle in CIELAB coordinates.

The accuracy of reconstruction of the reflectance spectrum by addition of successive components is illustrated in Fig. 7.70 (left) for sample 150, an inverse logistic function, where the approximation becomes closer and closer to the true spectral waveform. The error, i.e. the difference between true spectrum and approximation, shown in Fig. 7.70 (right), tends to zero as more components are added, oscillating about zero.

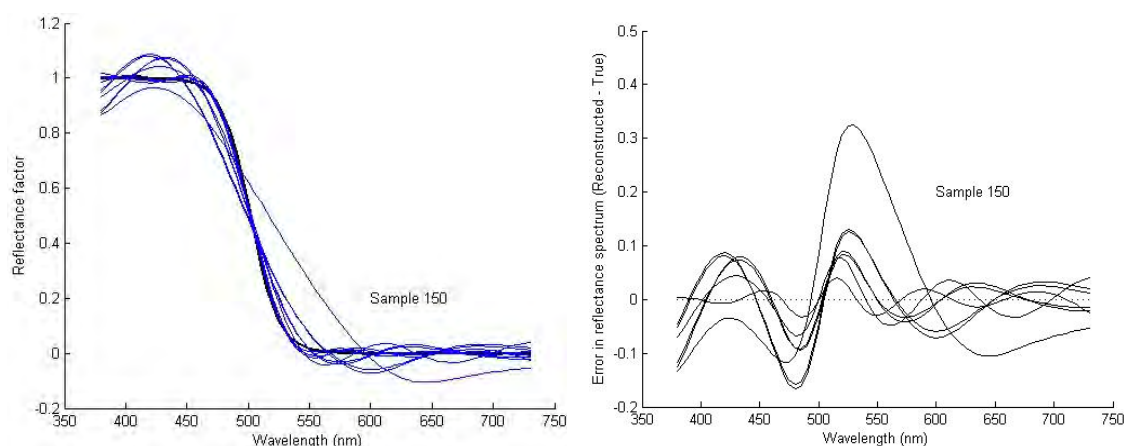


Figure 7.70: (left) reconstruction of reflectance spectrum of one sample from first ten principal components; (right) error between the true spectrum and the successive reconstructions from the first seven components.

In the next section very large sets of synthetic spectra will be used as the basis for determining the optimal loading of a 3D colour lookup table for colour transformation.

7.4 Constructing Colour Transformation by Synthetic Spectra

7.4.1 Device Characterisation

The signals produced by digital image capture devices, usually encoded at each pixel position as red, green and blue (R,G,B), are device dependent. The relationship between the object's reflectance spectrum and each signal depends on the spectral sensitivity of the corresponding channel of the device. In a digital scanner, for example, this includes the integrated product of the power at each wavelength of the inbuilt light source, the reflectance of the surface, the transmittance of the colour filter, and the sensitivity of the sensor.

Input device characterisation is the process of establishing the relationship between the colour of the original object and signals generated by the device (Fig. 7.71). The former is expressed by device-independent values related to human vision, such as CIE $L^*a^*b^*$, and the latter by device-dependent R,G,B values. Simplistically, if the signal can be represented as a function of the colour: $\text{Signal} = f(\text{Colour})$ and if the function is invertible (i.e. continuous and single-valued) then the colour may be recovered as the inverse function: $\text{Colour} = f^{-1}(\text{Signal})$.



Figure 7.71: 'Black box' model of input device behaviour

Because there is generally no analytical function available, an approximation is frequently used, by fitting the data with a low-order polynomial through a regression procedure. Hong *et al* (2001) tested six different orders of polynomial and found the best overall performance (i.e. lowest average error) with an 11-term polynomial, including a cubic term for RGB, over a set of 168 test samples. They also investigated the number of training samples needed, and concluded that 60 samples were sufficient.

The problem is that the transfer function f should be optimised over all possible reflectance spectra that the device could encounter. If the usage will always be limited to a specific population of materials or objects, for example photographic prints or paints or textiles made with known dyes, it may be sufficient to take a selection of test samples from that population. But for a digital camera or general purpose scanner that could be used for any material it is not sufficient to train it on a small number of samples with a limited range of reflectance spectra, e.g. by using a standard test target such as the GretagMacbeth DC chart, and then expect the transfer function to predict accurately the device performance on all possible spectra that might be encountered in practice.

This study aimed to characterise the Arius 3D colour laser scanner (Fig. 5.63), used at UCL for 3D digitisation of heritage objects from Museum collections, such as Egyptian tablets. Because it employs three lasers to sample the object surface at three precise wavelengths (473, 532 and 635 nm), the colour response is highly metameric (see Section 7.2.2). Its R,G,B signals are therefore much more sensitive to the form of the reflectance spectrum than a normal scanner or digital camera or the human eye, for which the channel spectral sensitivities are quite broad.

The objective was to use sets of reflectance spectra to determine the inverse transfer function f^{-1} , to transform the R, G, B signals from the scanner sensor directly into CIELAB colorimetric values. The reference tristimulus values X, Y, Z were calculated by multiplying the spectral reflectance distribution of each sample by the CIE 2° Standard Observer and the spectral power distribution of the D65 illuminant (Fig. 7.72). Knowing the spectral sensitivity of the scanner (in this case three delta functions) enabled the scanner R, G, B responses to be predicted, by sampling the reflectance spectrum at each of the three specific wavelengths. The coefficients of a 3×3 matrix were calculated by a regression procedure, enabling the corresponding X', Y', Z' to be predicted for each R, G, B . From the two sets of X, Y, Z the CIE $L^*a^*b^*$ values were calculated using the tristimulus values of D65 as the reference white.

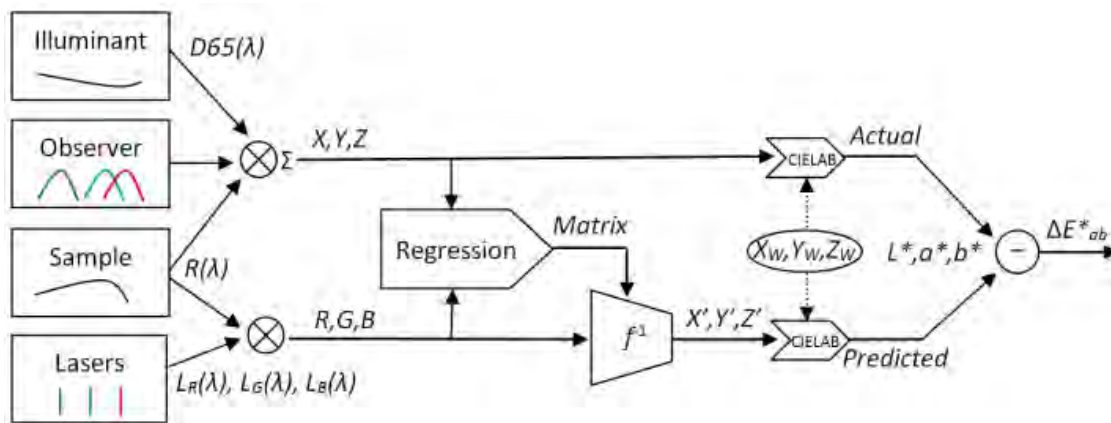


Figure 7.72: Procedure for calculating matrix coefficients and testing the accuracy of prediction.

A large dataset of 8,714 reflectance spectra was collated from readily available spectral reflectance measurement sets (Table 7.11) over the wavelength range 380 to 780 nm, and interpolated to 1 nm intervals by the Matlab function `interp1` with the cubic spline option. Where the original wavelength range was smaller, for example with no data provided below 400 nm or above 700 nm, the reflectance was set to zero.

Table 7.11: Collection of 8,714 reflectance spectra assembled from various sources.

Spectral reflectance dataset	Source	Surface type	Instrument	Geometry	Wavelength interval (nm)	Samples
NCS atlas	NCS	Matte	Gretag Spectrolino	45-0	10	1950
NCS atlas	NCS	Matte	Macbeth ColorEye 7000	d-8	10	1950
Munsell atlas	Joensuu	Gloss	Perkin-Elmer lambda 18	45-0	1	1600
Munsell atlas	Joensuu	Matte	Perkin-Elmer lambda 9	d-8	1	1269
Chinese NCS	Shanghai	Textiles	Datacolor 600	d-8	10	899
Natural materials	Derby	Leaves	Macbeth ColorEye 7000A	d-8	10	494
Natural materials	Joensuu	Plants	Acousto-Optic Tunable Filter	45-0	5	218
Gretag Macbeth DC chart	LCC	Matte	X-Rite i1Pro	45-0	10	180
Berns-Taplin dataset	RIT	Acrylic paints	X-Rite i1Pro	45-0	10	100
Art pigments	Natl. Gallery	Traditional paints	Monolight	45-0	2	64

As a baseline for the colour transform, a regression procedure was employed to find coefficients c_{ij} that would minimise the error in tristimulus X,Y,Z domain over the full dataset:

$$\begin{aligned} X_i &= c_{11}R_i + c_{12}G_i + c_{13}B_i \\ Y_i &= c_{21}R_i + c_{22}G_i + c_{23}B_i \\ Z_i &= c_{31}R_i + c_{32}G_i + c_{33}B_i \end{aligned} \quad (7-16)$$

where R_i, G_i, B_i are normalised signal values from the detectors in laser scanner for $i = 1..n$ and $n = 8174$. The system of equations can be written more compactly in matrix form as:

$$[X \ Y \ Z] = [R \ G \ B] \mathbf{M} \quad (7-17)$$

where the $n \times 1$ response vector $X = [X_1 \dots X_n]^T$ (and similarly for Y and Z), the $n \times 1$ signal vector $R = [R_1 \dots R_n]$ (and similarly for G and B) and the 3×3 matrix $\mathbf{M} = [c_{ij}]$ contains the coefficients. For the collated dataset of real samples the resulting matrix was:

$$\mathbf{M}_{\text{real}} = \begin{bmatrix} 48.61 & 28.84 & 4.93 \\ 34.96 & 68.68 & -6.69 \\ 10.18 & 1.81 & 107.02 \end{bmatrix} \quad (7-18)$$

The performance of this simple linear fit of laser R,G,B to tristimulus X,Y,Z was evaluated by converting both the reference colour and the estimated colour of each sample to $L^*a^*b^*$ and calculating the colour difference as ΔE^*_{ab} . The error distribution is shown in Fig. 7.73, with mean error of 4.98.

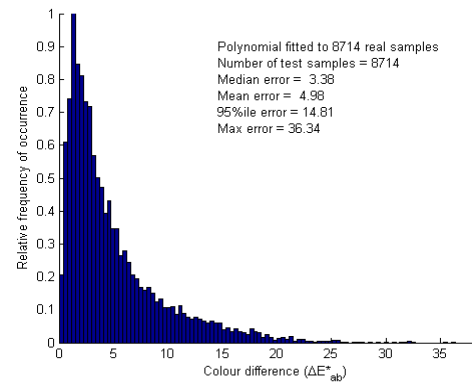


Figure 7.73: Distribution of errors from matrix fitting of Arius R,G,B laser signals to 8174 reflectance spectra.

7.4.2 Using Synthetic Spectra

To simulate the widest possible range of reflectance spectra that might be encountered by the scanner, a set of synthetic spectra was generated. These were constrained to be continuous, single-valued functions of wavelength at 1 nm intervals over the range 380 to 780 nm, with a maximum slope first derivative) not exceeding $\pm 0.024 \text{ nm}^{-1}$. The latter limit was derived from analysis of the collated spectra (Fig. 7.74), which showed that the first derivatives fall largely within the range -0.01 to $+0.02 \text{ nm}^{-1}$ (see Figs. 7.63 and 7.65). These values can be related to colour gamut (Fig. 2.59).

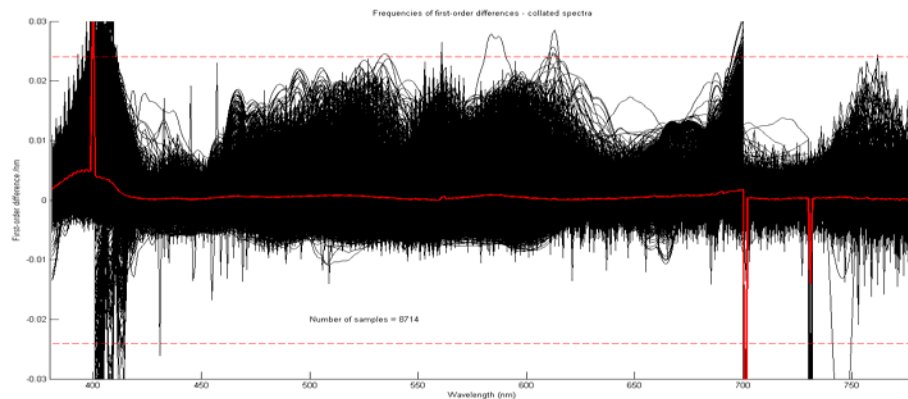


Figure 7.74: Slopes (first derivatives) of 8,714 reflectance spectra of real materials, with mean (red). The spikes at 400, 700 and 730 nm are artifacts arising from discontinuities in some datasets.

The synthetic reflectance spectra were based on eight generating functions: Gaussian, inverted Gaussian, logistic, inverted logistic, sine, upward ramp, downward ramp, sum of three Gaussians, and piecewise linear (Figs. 7.64 and 7.75). The central wavelength λ_0 and width parameter k were randomised to produce a family of 100,000 curves for each function, giving 800,000 synthetic spectra in total. The central wavelengths ranged from 400 to 700 nm, and the full range of reflectance factors was utilised, from 0 to 1. Limits were set on the slope and width parameters to restrict the maximum slope of the curves in all cases within the range ± 0.024 per 1 nm interval.

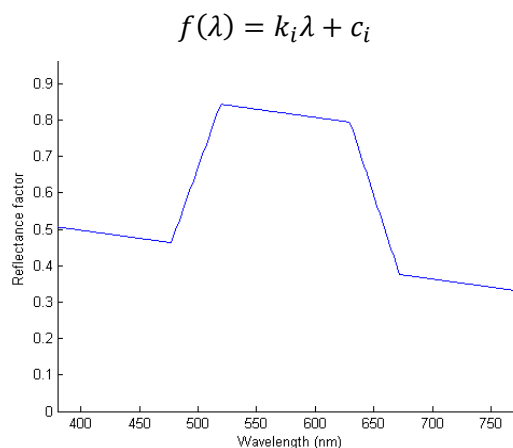


Figure 7.75: Piece-wise linear ramp generating function for synthetic spectra.

$L^*a^*b^*$ coordinates of all real and synthetic spectra were calculated for the D65 illuminant and plotted in 3D (Fig. 7.76). The real samples fit within cuboidal bounds of approximately $10 < L^* < 95$ and $-60 < a^* < +70$ and $-55 < b^* < +105$. The synthetic spectra, however, fill a much larger volume of $0 < L^* < 100$ and $-130 < a^* < +135$ and $-125 < b^* < +140$. They are therefore much more demanding as a test of the robustness of the device characterisation.

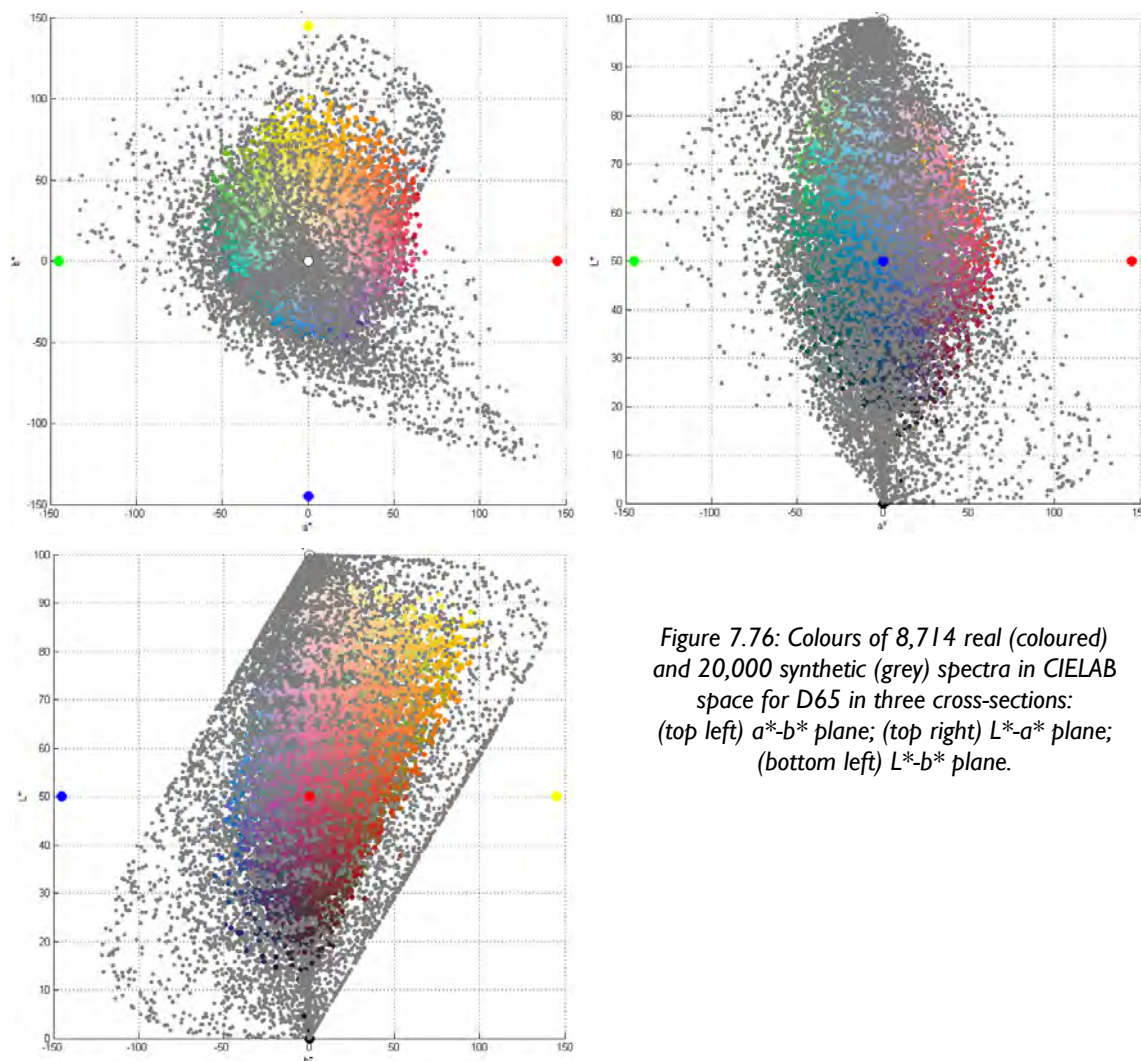


Figure 7.76: Colours of 8,714 real (coloured) and 20,000 synthetic (grey) spectra in CIELAB space for D65 in three cross-sections: (top left) a^*-b^* plane; (top right) L^*-a^* plane; (bottom left) L^*-b^* plane.

7.4.3 Colour Transformation by Lookup Table

The transformation from object colour to device RGB is conveniently implemented by a lookup table (LUT), with less computation and faster speed for real-time applications. The contents of the LUT can be pre-computed and also edited to include media-dependent corrections or operator preferences. The first use of LUTs was in digital drum scanners for the digitising of transparency photographs for graphic arts and colour printing. The logarithm of intensity of light in red, green and blue (R, G, B) channels transmitted through each point of the transparency was digitised to give density values D_R, D_G, D_B . These were used to address the correct cell of the LUT in which were stored corresponding values of print density or ink %dot in cyan, magenta, yellow and black (C, M, Y, K).

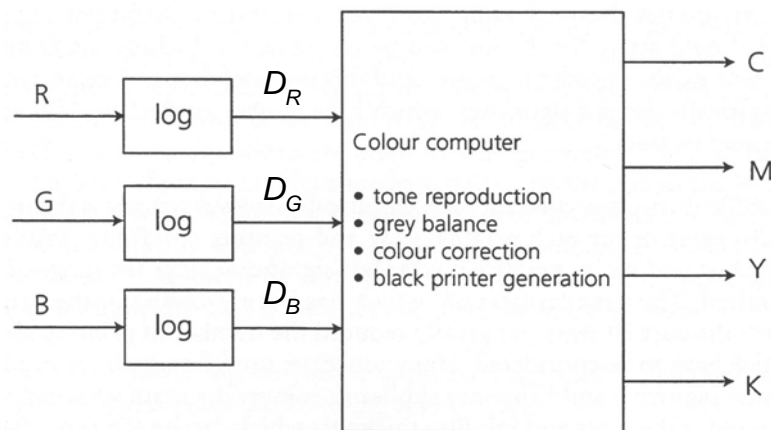


Figure 7.77: Signal processing by LUT in the Crosfield scanner, c.1975.

For good quality picture reproduction, at least eight bits per channel are required, preferably 10 bits in a logarithmic density domain or 12+ bits in a linear domain. To implement this as a LUT with a separate cell for every possible combination of input values would require a prohibitively large amount of memory. For example the arrangement in Fig. 7.77 would need $2^{24} = 16$ million cells with four output bytes per cell (64 Mbyte) for 8-bit input, and $2^{30} = 1$ billion cells (4 Gbyte) for 10-bit input. The solution has been to separate the input signals into two parts: the most significant bits (msb) address a smaller LUT, and the least significant bits (lsb) are used to interpolate values from all surrounding points in the lattice. With a 4:4 allocation of 8 input bits, the LUT requires only $2^{12} = 4096$ cells (16 Kbyte). The simplest interpolation scheme is trilinear amongst all eight corner points of the cubelet within which the input value falls (Fig. 7.78), but other schemes are widely used.

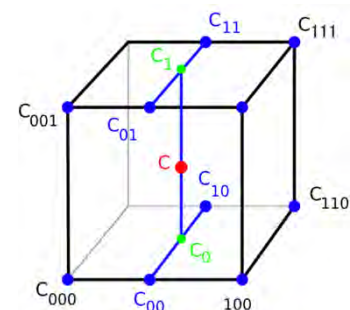


Figure 7.78: Trilinear interpolation.

To implement the colour transformation for the colour laser scanner, a lookup table with $33 \times 33 \times 33 = 35,937$ cells was constructed, addressed by the 5 msb of each of the R, G, B laser signals. The 7 lsb of the 12-bit linear input signals were used to interpolate amongst the eight corner points of the cubelet. The cells of the cube were initially loaded with the $L^*a^*b^*$ values predicted by the matrix in Eq. 7-18 followed by the standard conversion from X, Y, Z . This is similar to the arrangement used by Hung (1993) for the colorimetric calibration of imaging devices.

Plotting all samples within the normalised cube (Fig. 7.79) reveals that the real samples fill only a limited volume around the grey axis up the long diagonal from the black corner to the white corner. The coordinates corresponding to the synthetic samples, while still clustered around the grey axis, are spread much more widely and fill all of the cube volume although they are sparse in some regions. This suggests a novel way of improving the accuracy of the lookup table, by loading it with the actual $L^*a^*b^*$ values corresponding to the spectra, instead of the approximate values predicted by the regression formula. The 808,714 spectra available, if they were uniformly spread throughout the RGB cube address space, should yield an average of over 22 values in each of the 35,937 cells. In fact the distribution turns out to be very uneven where some cells, especially dark colours near neutral, have hundreds of entries whereas others have none. For conservation of memory, the maximum number of entries in each cell was limited to 101, commencing with the real samples. Fig. 7.80 (left) shows the scatter in such a distribution of $L^*a^*b^*$ values.

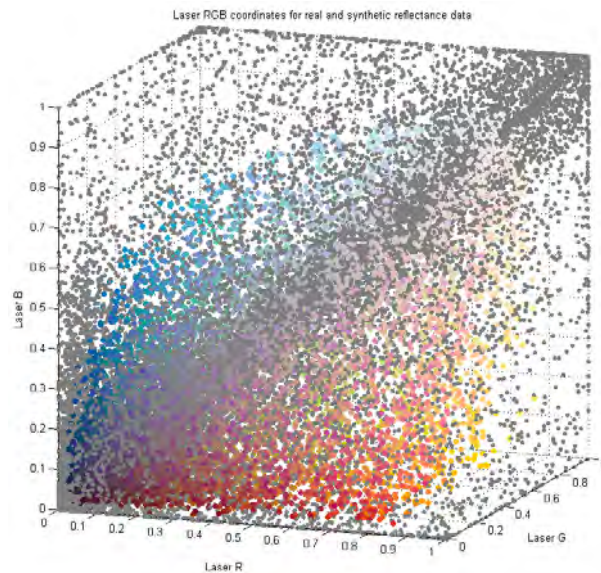


Figure 7.79: Distribution of 8,714 real (coloured) and 20,000 synthetic (grey) spectra in the RGB cube.

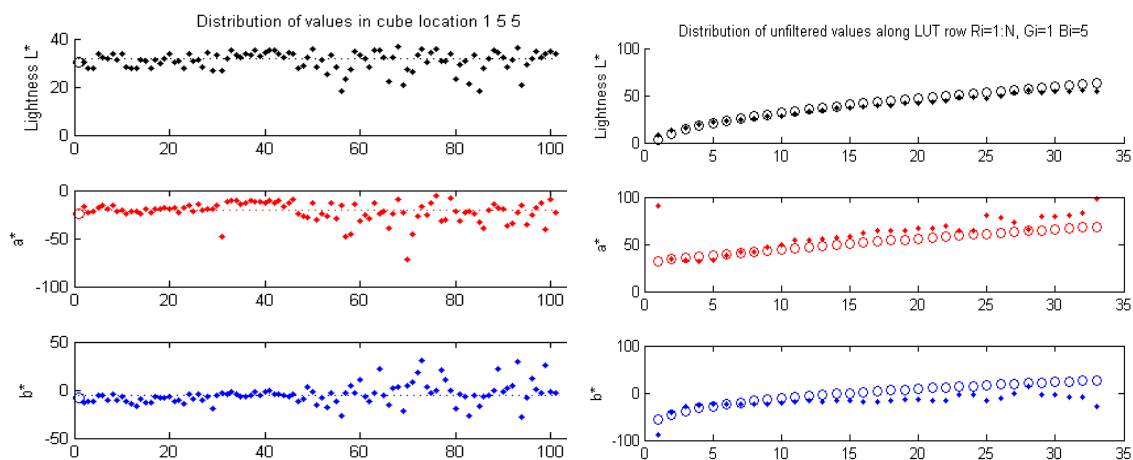


Figure 7.80: (left) Distribution of $L^*a^*b^*$ values in one cube cell, all with the same R,G,B index values. The first entry (open circle) is the predicted value; (right) Values along one row of cube after taking the median in each cell (dots). The open circles are the values predicted by the matrix.

The median was taken of all values accumulated in each cell. Because the first entry in each cell is the predicted value (from the regression matrix), there is always guaranteed to be at least one entry, even in sparse regions with no corresponding values from spectra. Fig. 7.80 (right) shows the resulting set of 33 values of L^* , a^* , b^* along one row of changing R and constant G and B . In this case the predicted value matches the actual value closely for L^* , but there are significant departures for a^* and b^* . In order to achieve better homogeneity of the function along all three dimensions of the lookup table, the data was filtered at every internal point by taking the median of the $5 \times 5 \times 5$ surrounding cells.

This method proved to be very effective as a way of removing isolated values and reducing noise (Fig. 7.81). The resulting transfer function can be visualised effectively in 3D for planes through the LUT where one of the input values is held constant (Fig. 7.82). The local deformations in the 3D function loaded into the LUT enable the colour transformation to adapt to the most probable signal produced for each spectrum, with far more local sensitivity than any polynomial fitted over the whole dataset.

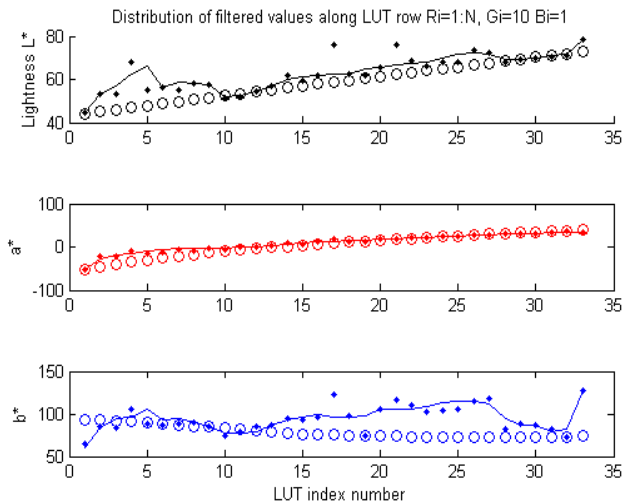


Figure 7.81: Values along one row of cube after applying the 5x5x5 median filter. The open circles are values predicted by matrix, dots are unfiltered values and lines are filtered values.

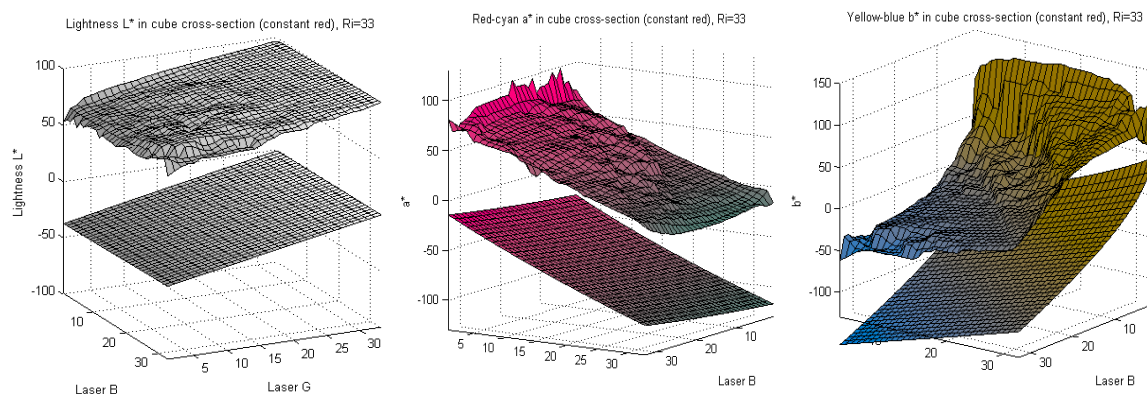


Figure 7.82: Values of L^* , a^* and b^* values for planes of constant $R=33$. The lower surface in each case shows the values predicted by the baseline matrix transform (displaced downwards for visualisation) and the upper surface shows the local deformations resulting from adaptation to the spectra.

7.4.4 Linearisation of Laser Signals

To process point cloud files from the Arius colour laser scanner, the .a3d file is read into Pointstream, then saved as an ascii file. This has one line of text for each point, consisting of 23 numerical fields. These include the raw RGB laser values for object reflectance and raw RGB laser values for the white cube. The latter are constant for each raster line, i.e. one sweep of the laser beam (y axis). The raw laser values are represented as 16-bit unsigned integers, and have a linear relationship with the object reflectance. The white cube values enable correction for possible fluctuations in laser power during the scan, although these are minor for the more stable solid-state lasers in the current scanner than for the former gas-powered lasers.

In a stable linear system, the laser detector signal L would be related to the surface reflectance factor R by the linear equation $L = mR + c$, where m is the slope and c is the signal offset from zero. In practice, the laser power can vary and an additional signal W is provided, representing the detector response for the reflected laser intensity from the white spectralon cube. This signal does not represent a perfect white, i.e. the perfect diffuser with a reflectance factor of 1.0, as might be expected but is 'whiter than white'. This may be the result of internal scattering and re-emission of the light, in a luminescence phenomenon, or may be because the

cube is closer to the laser and detector than the object surface. The black offset b applies to the detector signal in each channel and is assumed to be constant. It represents the signal produced for a perfect black with zero reflectance, as one would get from a camera with the lens cap on. It has to be subtracted from both the laser signal L and the white cube signal W . The relationship is:

$$mR = \frac{L-b}{W-b} \quad (7-19)$$

Given values of R for a measured target, and corresponding values of L and W from the point cloud produced by the laser scanner, we seek to determine m and b .

$$b = \frac{L-mRW}{1-mR} \quad (7-20)$$

This can be solved by equating the values from two points:

$$\frac{L_1-mR_1W_1}{1-mR_1} = \frac{L_2-mR_2W_2}{1-mR_2} \quad (7-21)$$

This gives a quadratic in m :

$$R_1R_2(W_1 - W_2)m^2 + (R_1L_2 - R_2L_1 + R_2W_2 - R_1W_1)m + (L_1 - L_2) = 0 \quad (7-22)$$

The equation can be solved for all corresponding points in pairs of patches of the neutral scale in the fourth row of the MiniMacbeth target, extracting the data from the laser scan point cloud data generated by Pointstream and saved as an ASC file (Fig. 7.85). For 15 pairwise combinations of the six patches in the scale, and approximately 1200 points sampled from each patch, over 12,000 values of m and b can be generated for each of the R,G,B channels, and the median of each distribution taken as the result. Care needs to be exercised for the black patch on the test chart, which in some cases is below the detection level of the laser scanner and hence contains invalid data in the file. In this case the procedure is limited to the 10 pairwise combinations of the other five patches of the grey scale. Histograms of the distributions of the slope and offset for the Arius laser scanner (with the new blue laser at 443 nm) are shown in Fig. 7.83. The actual vs predicted values of laser R,G,B for each of the 24 patches in the chart are plotted in Fig. 7.84 (right), and show a generally good correspondence except at higher values of red.

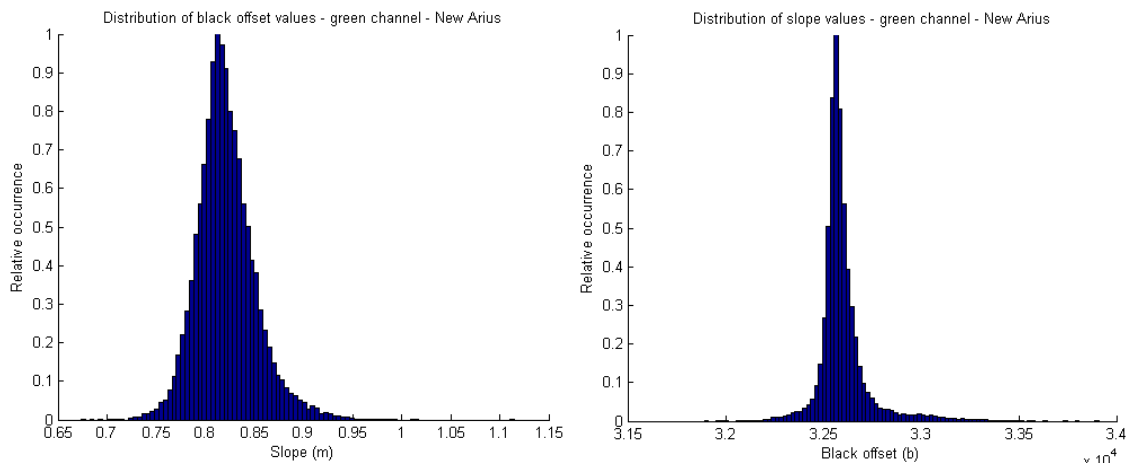


Figure 7.83: Relative distribution of computed values of slope (left) and black offset (right).

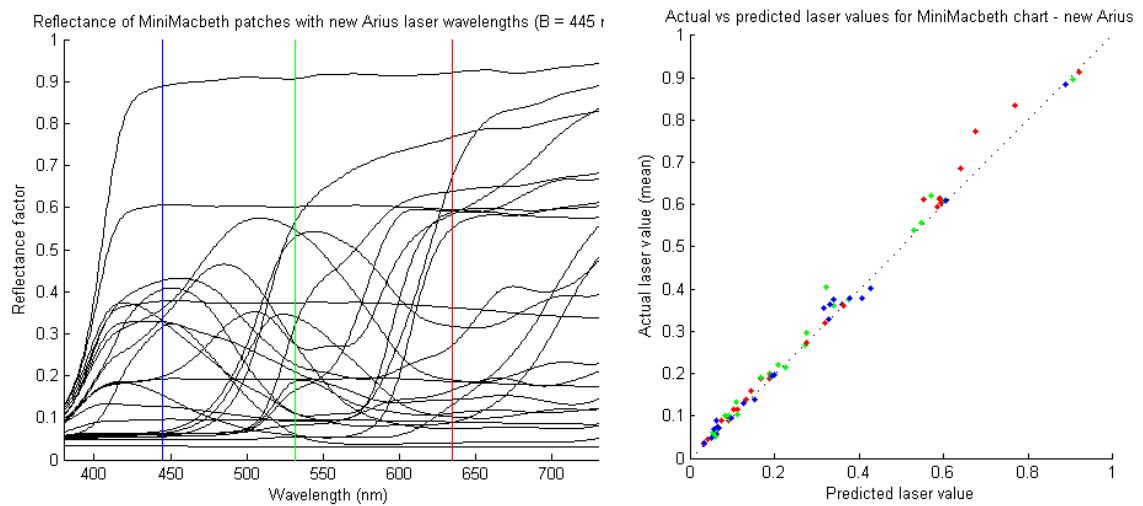


Figure 7.84: (left) Reflectance factors of 24 patches in MiniMacbeth chart, with the wavelengths of the red, green and blue lasers; (right) Value computed from mean laser signals (Y axis) vs reflectance factor of patch at the specific laser wavelengths (X axis) for each patch in MiniMacbeth chart.

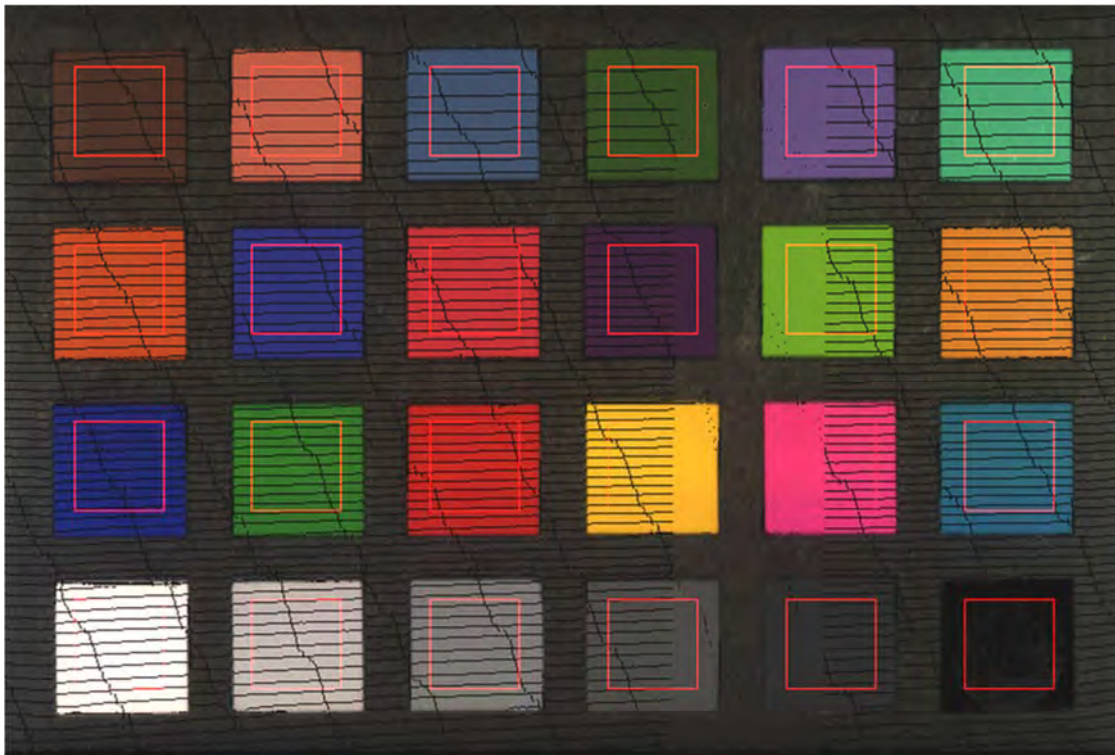


Figure 7.85: Sampling boxes (red outlines) over patches of MiniMacbeth chart. The black lines are interference patterns resulting from the projection of the XYZ geometric coordinates onto the nearest discrete pixel positions on the image plane. Pixels in the black lines are ignored in the calculations.

7.4.5 LUT Transformation

To convert the colour data in a point cloud from the scanner, the processing workflow is as shown in Fig. 7.86. The colour lookup table (CLUT) is assumed here to be loaded with 8-bit sRGB data, but other colour spaces could be offered. The R,G,B fields in the ascii line for each point are replaced by the new values.

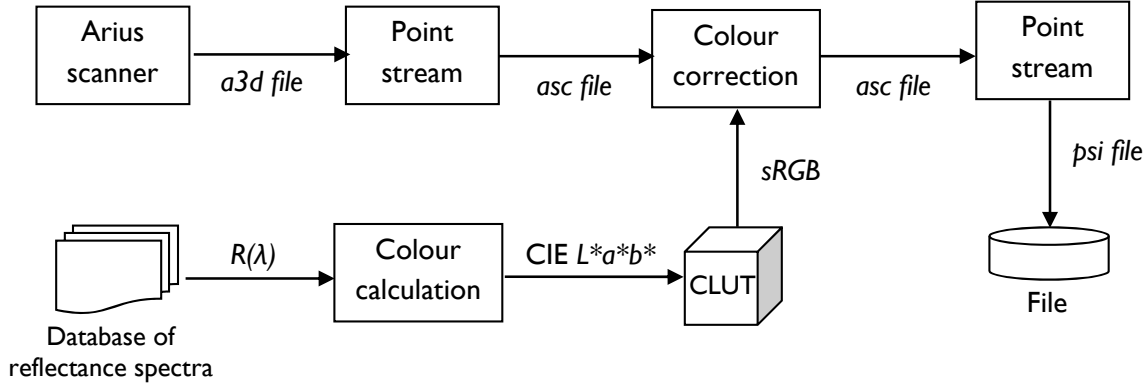


Figure 7.86: Process flow for conversion of the colour fields in Arius point cloud data.

As the baseline for the transform, regression is used to find the best fitting linear matrix transformation from the laser RGB signals to the tristimulus XYZ values, over the set of 800,000 synthetic reflectance spectra. For the old Arius scanner (i.e. with the blue laser wavelength of 473 nm), the calculated values of \mathbf{M} were:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{M} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 49.44 & 33.01 & 12.15 \\ 30.27 & 66.48 & 3.09 \\ 6.52 & -5.70 & 106.79 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (7-23)$$

where R,G,B are the normalised linear reflectance values for the sample, in the range [0,1].

For each sample the tristimulus values were calculated for the standard observer under D65 and the illumination of the three lasers. The fitting errors for the matrix were calculated for all 800,000 synthetic spectra as the ΔE^*_{ab} between the true colour and the colour predicted by the model, under D65.

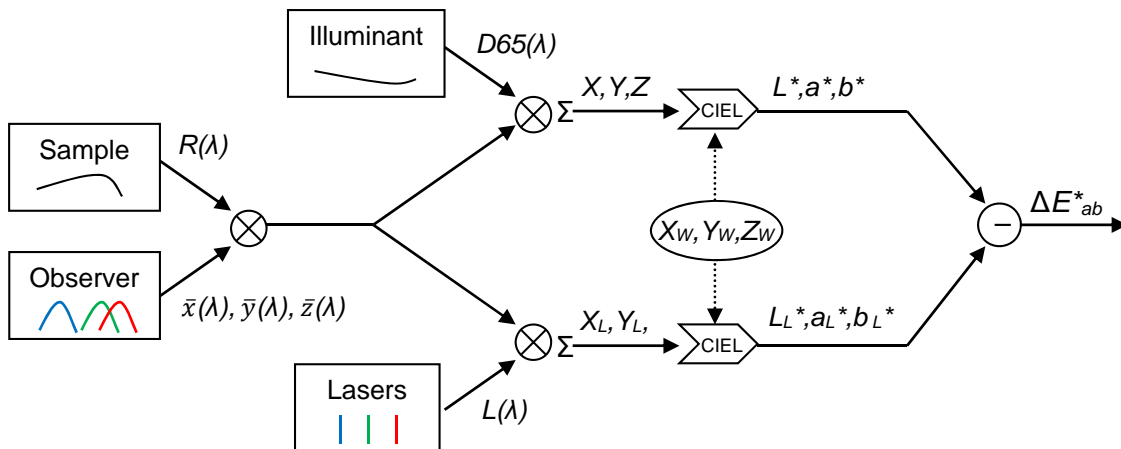


Figure 7.87: Procedure for calculating difference between colour of a sample under D65 and the corresponding colour under the three lasers. It is assumed that the relative laser power has been adjusted to match the D65 white reference.

For each synthetic spectrum in the fitting set, the lookup table is loaded as shown in Fig. 7.88. The three lasers sample the spectrum at single wavelengths and provide the three input addresses to the LUT. It was found to give improved performance in the rendering of dark colours to apply a nonlinear power function with exponent 0.7 to the R,G,B values before addressing the LUT, equivalent to a gamma correction of 1.42.

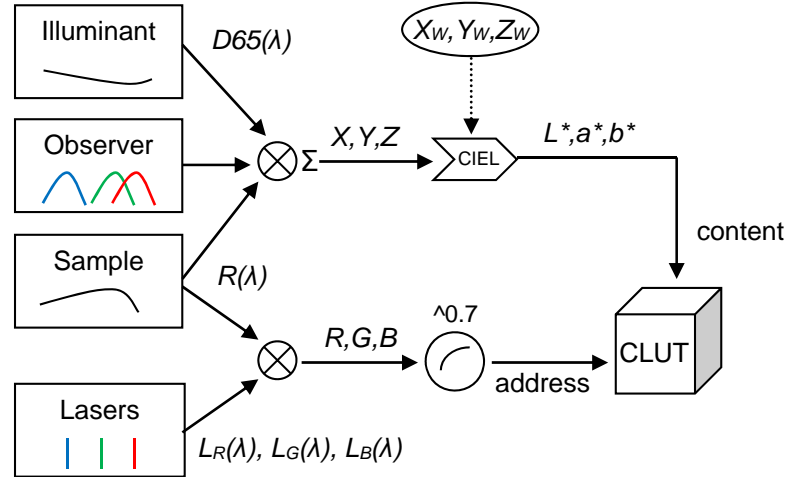


Figure 7.88: Procedure for loading the CLUT with $L^*a^*b^*$ values corresponding to the R,G,B lasers.

The colour errors produced by the system with laser sampling of the spectrum and CLUT transformation can be calculated as shown in Fig. 7.89, analogous to the error calculation in Fig. 7.87. The lasers are assumed to be balanced to match a D65 white, and the output values are converted to $L^*a^*b^*$ in order to calculate ΔE^*_{ab} .

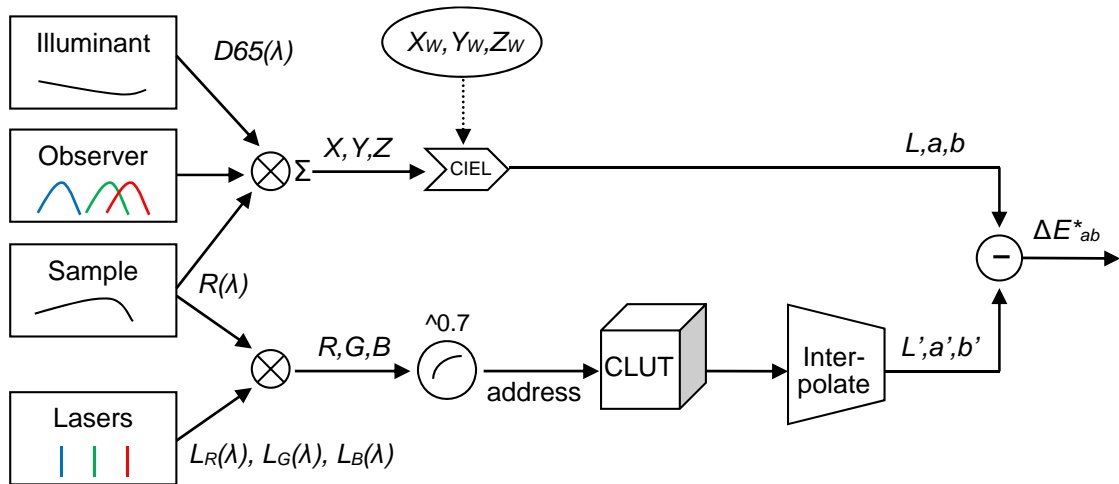


Figure 7.89: Procedure for verifying the accuracy of the $L^*a^*b^*$ values in the CLUT corresponding to R,G,B lasers.

The distributions of errors are shown in Fig. 7.90 for a set of synthetic spectra containing 700,000 samples and the old Arius laser primaries. The mean error of 9.05 with the 3x3 linear matrix transformation is reduced to a mean error of 7.25 with the trained CLUT.

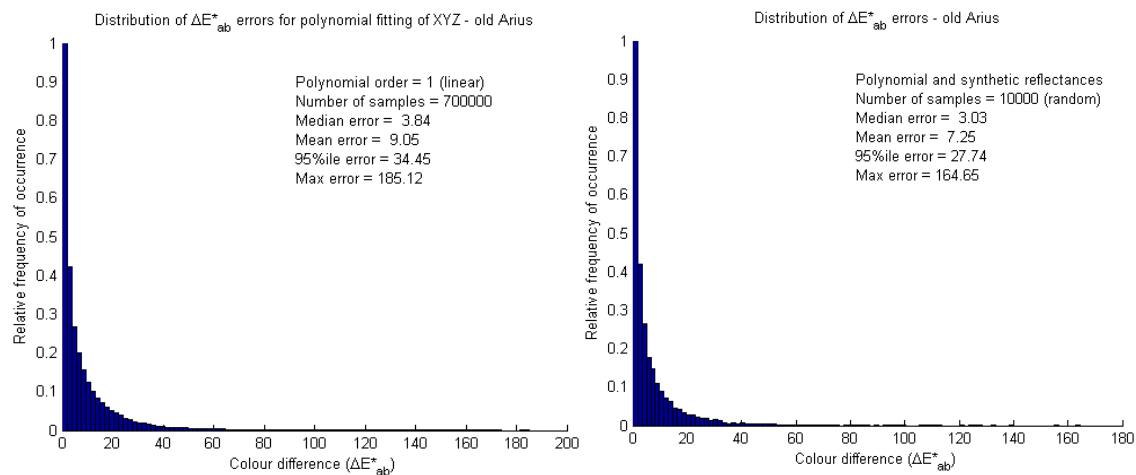


Figure 7.90: Error distributions arising from a set of 700,000 synthetic spectra: (left) using a linear 3x3 matrix on all samples; (right) using the CLUT on a random selection of 10,000 samples.

The accuracy of the resulting LUT was verified by using the NCS spectral dataset (1950 samples). Fig. 7.91 compares the performance on these samples of a 3x3 matrix transform based on a regression over 800,000 synthetic spectra with the optimised LUT technique. The mean error is reduced from 5.02 to 3.98.

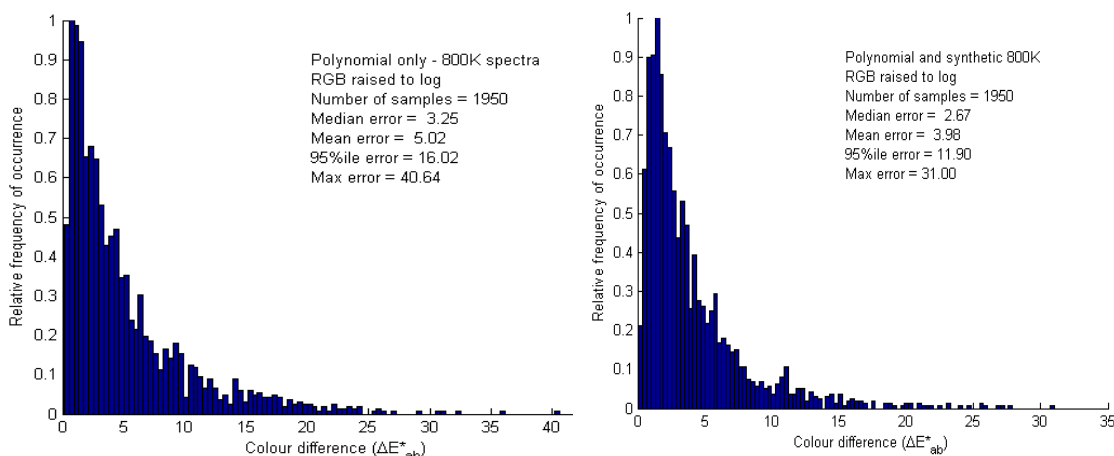


Figure 7.91: Error distributions arising from the NCS set of 1950 synthetic spectra: (left) using a linear 3x3 matrix trained on all samples; (right) using the CLUT method trained on all samples.

7.4.6 Evaluation

Testing the performance of the simple 3x3 matrix \mathbf{M} in Eq. 7-18 with the set of 8,714 real spectra gave the distribution shown in magenta in Fig. 7.92, with a median error $\Delta E^*_{ab} = 3.38$ (Table 7.12). The same matrix applied to the full set of 808,714 real and synthetic spectra resulted in a median error of 5.00 but a much longer tail in the distribution, with the 95th percentile increasing from 14.81 to 33.15. The maximum error rose to over 197, indicating that for a few spectra the monochromatic sampling by the laser scanner produces huge colorimetric errors.

The fitting procedure of Fig. 7.72 and Eq. 7-17 was repeated for the full dataset of 808,714 real and synthetic samples. The resulting matrix was:

$$\mathbf{M}_{\text{full}} = \begin{bmatrix} 49.96 & 30.56 & 7.53 \\ 32.40 & 65.87 & -5.59 \\ 12.17 & 3.39 & 105.43 \end{bmatrix} \quad (7-24)$$

Comparison of this matrix with that fitted to the real samples only (Eq. 7-18) indicates close agreement in all coefficients. Testing the performance of this matrix on the real samples shows that it has a slightly worse performance (median error of 3.66 instead of 3.38), but an improved performance for the full dataset of combined real and synthetic samples (median error of 4.89 instead of 5.00), as expected. Small further improvements to the performance of the regression could be made by introducing additional terms (for example by second- or third-order polynomials), but these are not justified by the nature of the linear system. The errors arise not from non-linearities but from the monochromatic sampling of the spectra which cannot satisfy the Luther-Ives condition.

Table 7.12: Performance of matrix and LUT on spectral datasets.

Method	#samples fitted to	#samples tested on	Med	Mean	95 th %ile	Max
Matrix	8174	8174	3.38	4.98	14.81	36.34
Matrix	8714	808714	5.00	9.87	33.15	197.22
Matrix	808714	8174	3.66	5.22	15.40	40.65
Matrix	808714	808174	4.89	9.67	33.55	180.69
LUT	808174	8174	3.38	4.74	13.22	26.79
LUT	808174	100000	4.61	8.34	28.80	175.18

Testing the performance of the lookup table method on the 8,174 real samples gave the error distribution shown in red in Fig. 7.92. Overall the LUT performed better on the set of spectra of real samples than a 3x3 matrix fitted to those samples, with the same median error of 3.38 but a reduced 95th percentile value of 13.22. A similar improvement was achieved when testing the LUT on a large set of 100,000 spectra randomly sampled from the full set. Table 7.12 shows that the errors for the LUT are lower by every measure for both the real samples and the full dataset. The LUT method reduces errors by maximising the probability that output values correspond to the spectra most likely to be encountered in practice.

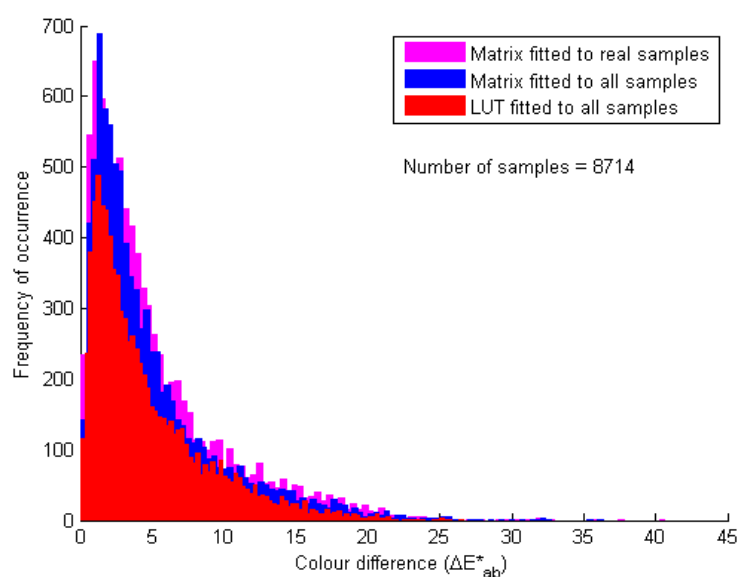


Figure 7.92: Distribution of colorimetric errors for 8,714 spectra of real samples for: (magenta) 3x3 matrix fitted to real samples; (blue) 3x3 matrix fitted to all samples; (red) LUT fitted to all samples.

Plotting the spectra of the 100 samples that produced the largest colorimetric errors (Fig. 7.93) shows the common characteristic of large changes in reflectance between the laser wavelengths. In the majority of spectra the reflectance factors, and hence the responses in the R, G, B channels of the scanner, would be well correlated. But for these spectra, the transition from a very small G response to a large R response, and vice versa, is relatively rare in the whole population of spectra. This low probability means that such spectra have little influence on the construction of the transfer function in the LUT.

In general spectra for which one of the R, G, B signals is zero are less well behaved than spectra corresponding to points near the centre of the LUT, corresponding to near-neutral colours. Plotting the differences in L^*, a^*, b^* values against R, G, B laser responses (Fig. 7.94) shows that the errors are well distributed around zero and tend to be larger for smaller signal values. The scatter is greater for the R channel than G and B .

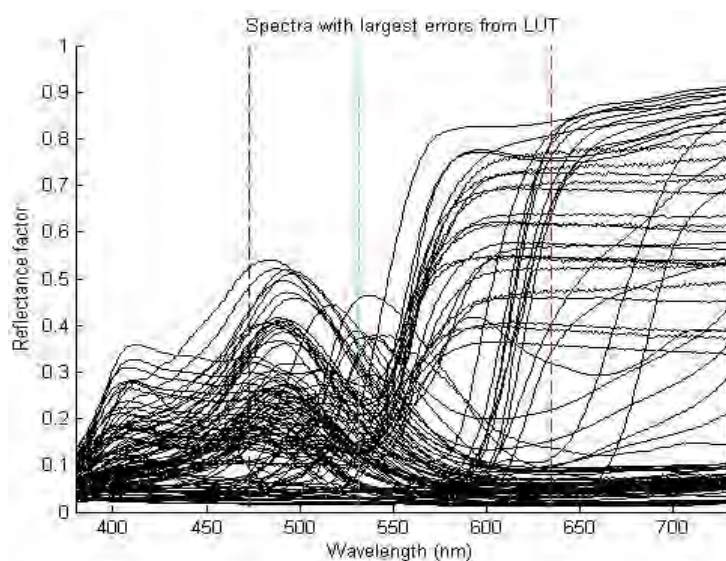


Figure 7.93: One hundred spectra of real samples producing the largest errors with the LUT technique. The dotted vertical lines are wavelengths of the three lasers at 473 nm (blue), 532 nm (green) and 635 nm (red).

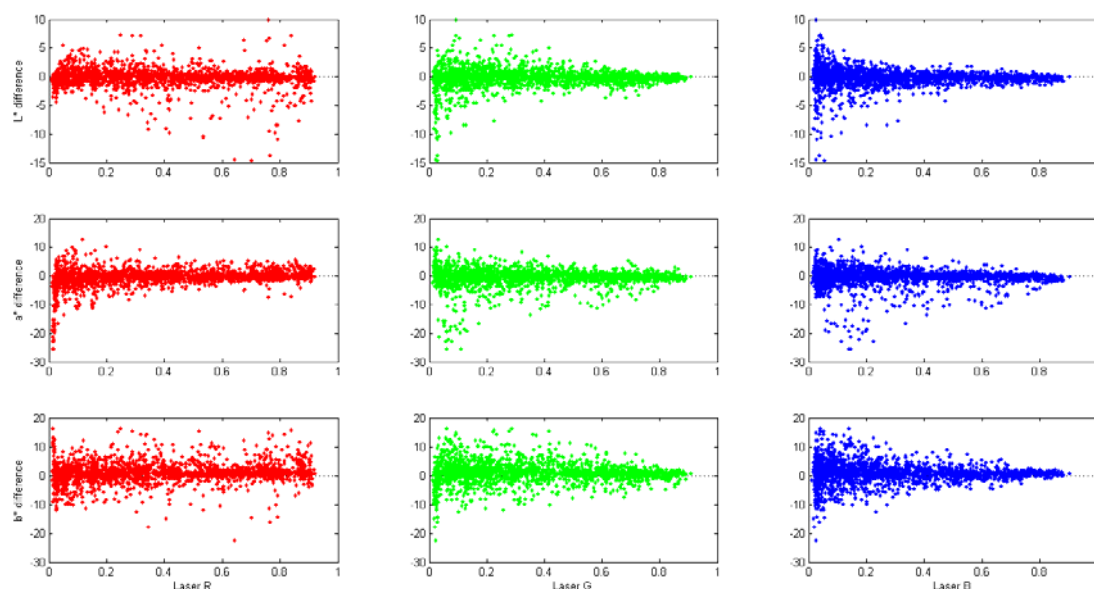


Figure 7.94: Colorimetric errors for L^* (top row), a^* (middle row) and b^* (bottom row) for normalised values of R, G, B laser response.

Plotting the cross-sectional planes of constant R , G and B in the LUT (Fig. 7.95) shows how the colour in each cell is subtly modified by the statistical influence of the spectra. For the purposes of visualisation the colours in this figure have been converted from CIE $L^*a^*b^*$ to sRGB, and out-of-gamut colours mapped onto the nearest point on the sRGB gamut boundary. The effect is most pronounced in the three planes for which one of the signals is zero, corresponding to the three faces of the cube intersecting at the black corner (Fig. 7.96).

This study has demonstrated a means of improving the colorimetric accuracy for characterisation of colour laser scanners over what can be achieved by the conventional fitting of a 3×3 matrix. The proposed method is novel because it uses the statistics of a very large population of reflectance spectra (both real spectra measured from samples and synthetic spectra generated from functions) to apply local adaptation (i.e. a localised deformation of the transfer function in the 3D LUT) to the prediction of a surface in colour space fitted by regression to the dataset. This gives optimal performance for device colour characterisation. The method is applicable to any type of input device, including scanners and cameras, for which the spectral sensitivity of each channel is known *a priori*. It could also be applied to multispectral imaging systems.

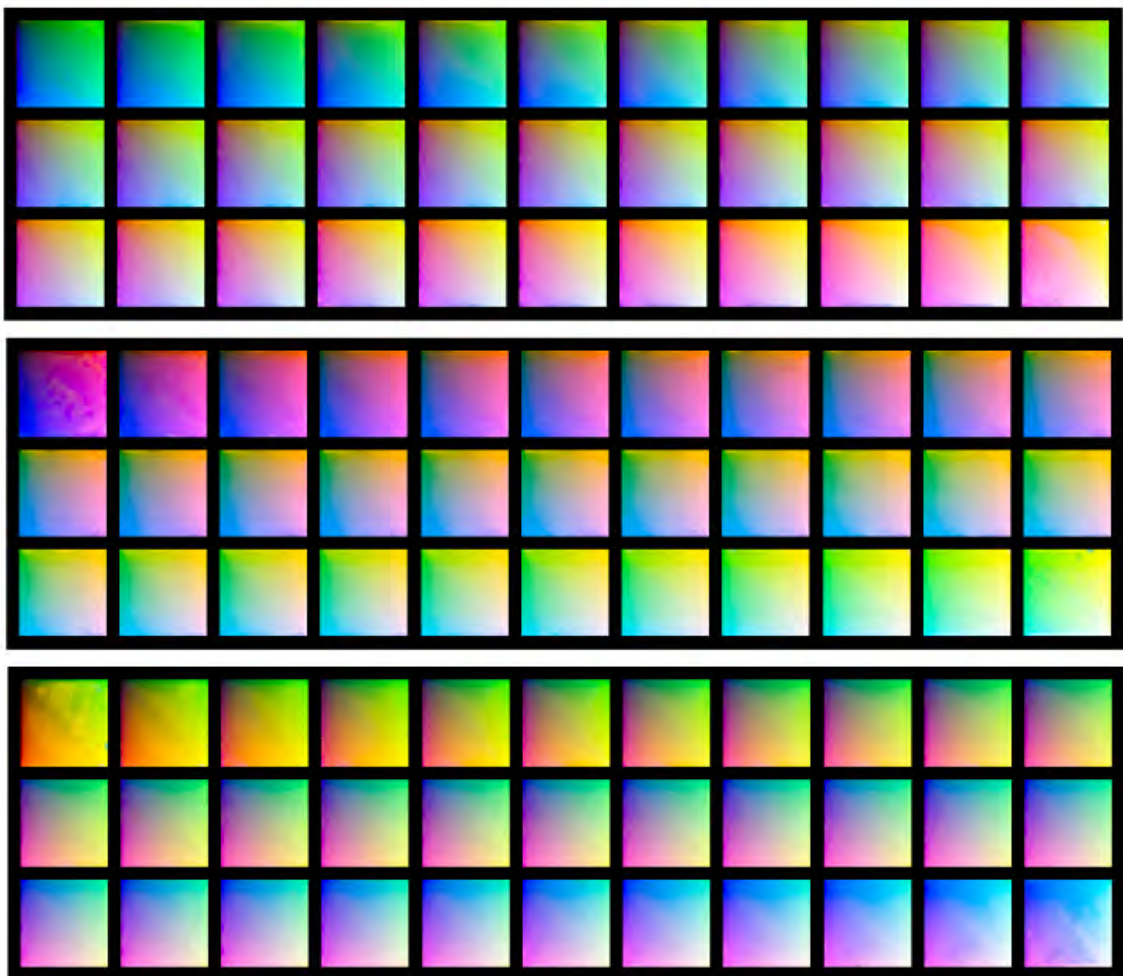


Figure 7.95: Cross-sections of the $33 \times 33 \times 33$ cube showing planes of constant red (top), green (middle) and blue (bottom) in sRGB colour space.

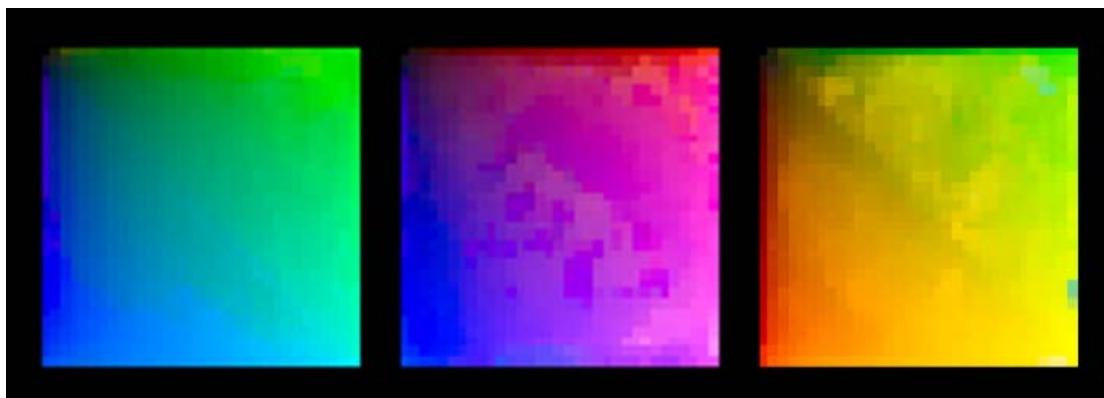


Figure 7.96: Cross-sections of the three faces of the cube for $R=0$, $G=0$ and $B=0$ showing the local effect of the spectra.

Acknowledgements

Thanks to various colleagues in the international colour imaging community for providing the spectral reflectance datasets used in this chapter and advice on how to process them:

- Prof Roy Berns and Dr Lawrence Taplin at the Munsell Color Science Laboratory at the Rochester Institute of Technology (RIT) for the reflectance dataset of 100 acrylic paint samples;
- Prof Jon Hardeberg of Gjøvik College, Norway, for the reflectance dataset for 64 oil paint samples, originally devised by Dr David Saunders and Dr John Cupitt at The National Gallery, London, in 1991. Also to John Cupitt for notes on its origins;
- Dr Vien Cheung for the measured reflectance spectra of 494 natural materials in Derby;
- Prof David Foster for the hyperspectral images, downloaded from his website at http://personalpages.manchester.ac.uk/staff/david.foster/Hyperspectral_images_of_natural_scenes_02.html and helpful discussions about their interpretation;
- Prof Jussi Pakkinen of Joensuu University, Finland, for providing the Munsell datasets;
- Berit Bergström and Anders Nilsson of the Natural Color System, Sweden, for the measured reflectance dataset of the NCS atlas;
- Zhao Shuling in Beijing for the measured reflectance dataset of the Chinese CNCS textile samples;
- Prof Janos Schanda of the University of Pannonia, Hungary, for discussions about the colour rendering of LEDs;
- Spectral datasets for the Judd-Vos and CIE 2007 observers were downloaded from the Colour & Vision Research Laboratory (CVRL) at <http://www.cvrl.org/>, based on the work of Prof Andrew Stockman at the UCL Institute of Ophthalmology.

Chapter 8 – Case Studies and Future Work

All our knowledge is the offspring of our perceptions.

– Leonardo da Vinci
Thoughts on Art and Life, v.35, 1508

This research study commenced with the hemispherical illumination dome, a digital camera, and a method developed by Malzbender. The programme of research has led to exploration of the dome geometry (Chapter 3), ways of visualising the relief of a surface (Chapter 4), 3D reconstruction from photometric normals (Chapter 5), modelling angular reflectance distribution and specularity (Chapter 6), and analysis of colour accuracy from a 3D colour laser scanner (Chapter 7). At every stage the criterion for effectiveness has been the appearance of the result: does it look realistic?

The field of digital humanities is rich with genres for application of the techniques of science and engineering to ancient objects. During the period of the research there have been opportunities to employ the newly developed techniques, and to collaborate with archaeologists, historians, classicists and conservators on objects ranging in size from a centimetre to a decametre. These interactions have helped greatly to improve the methods and to validate the significance of the results obtained. This final chapter describes five projects undertaken during the course of the research, in which the methods developed have been implemented. Indeed the projects have been key to proving that the methods described in the previous chapters are viable. The main contributions to knowledge achieved are reviewed. The chapter concludes with some ideas about directions for further research.



Figure 8.1: Fragment of Hunter's Palette, British Museum, EA 20790, 145x305 mm. Photograph taken by Kathryn Piquette in the Southampton dome at the British Museum.

8.1 Studies of Objects

8.1.1 Hunters Palette

A fragment of the Hunters Palette, also known as the Lion Hunt Palette, is held in the British Museum (EA 20790). It is an early Egyptian (c. 3100 BCE) slab of dark greyish brown mudstone, 14.5 cm in height and 30.5 cm in width (Fig. 8.1). The surface is elaborated through low relief carving on one side only, including human figures, animals and other objects which are interpreted as representing a hunting scene. Each hunter wears a short skirt with an animal tail hanging from the waist, and the majority are equipped with one or more implements in each hand, including a bow, arrow(s), a club, mace, spear(s), axe, etc. The theme of the design, in common with others such as the Two Dog Palette in the Ashmolean Museum, Oxford, seems to be subjugation – the power of hunters over animals. This suggests that the palettes might have been created to procure or commemorate a successful foray.

The object was studied for evidence of recarving, i.e. revision of the design during manufacture, particularly the figure in the upper right bending down to grasp a rope. This investigation was undertaken in collaboration with Dr Kathryn E. Piquette (Freie Universität, Berlin), who photographed the Palette at the British Museum in December 2010 as part of the collaborative AHRC-funded project between the University of Oxford and University of Southampton, 'Reflectance Transformation Imaging Systems for Ancient Documentary Artefacts'. There a new hemispherical dome of 1 metre diameter was fitted with 76 light emitting diodes (LEDs), and photography was undertaken with a Nikon D3X camera with Nikkor 50 mm lens set at $f/10$ aperture, producing images of size 6048x4032 pixels. The spatial resolution on the surface of the Palette was 18 points/mm (56 μ m pixel size).



Figure 8.2: Detail of lassoer in the plaster cast of the Hunter's Palette from the UCL Petrie Museum, illuminated from a north-north-easterly direction. Photographed by Lindsay MacDonald in UCL dome.

During the project a plaster cast of the Hunters Palette was also discovered in the dark recesses of the UCL Petrie Museum (Fig. 8.2). There is no documentation relating to this object, but it

is speculated that it may have been cast from the BM original in the early 1900s by Flinders Petrie himself, to be used as a teaching aid. The plaster is in good condition although it has clearly been handled extensively. It was therefore of interest whether analysis might reveal some additional evidence of the putative recarving. A set of 64 images of the plaster cast was taken in the UCL dome, by the Nikon D200 digital camera with Nikkor 17–55 mm zoom lens set to 40 mm focal length and $f/4$ aperture. The image size of 3900x2616 pixels gave a spatial resolution on the surface of the plaster of 11 points/mm (91 μm pixel size).

Comparison of the images of the plaster cast with those of the original Palette reveals the limitations of the casting technique for reproducing fine detail and texture (Fig. 8.3). The cast shows the same topographic features as the original, but with reduced definition. The depth of relief is shallower, incisions are less distinct, edges of features are more rounded, and the fine texture is diminished. It is an issue with plaster casts in general that they lack the material qualities of the original objects, such as colour, weight, density and 'feel', which is perhaps why they have fallen from favour in museum collections (Bilbey and Trusted, 2010).



Figure 8.3: Detail of head of the lassoer in original object (left) and plaster cast (right).

PTM visualisation was used extensively for examination of the surface relief of both the original object and the plaster cast. An innovation of the study was the development of a technique for examining two different PTM representations simultaneously on the same display screen (Fig. 8.4). Even though the two image sets were taken of different objects in different places by different people with different cameras in different domes with different lighting geometries, it was possible to combine them into a single visualisation under common interactive control of the virtual light direction. This was achieved by extracting the luminance and colour components from each individual PTM file, abutting the colour (RGB) components together, and then recomputing the biquadratic luminance components for the combined image. This facilitated the navigation and comparison of features on the two surfaces under common viewing conditions. This innovation is also valuable for the possibilities it presents for combining other RTI data. For example, the ability to view side-by-side the same or similar surface at different points in time may be useful for monitoring changes over a period for conservation and curatorial purposes.

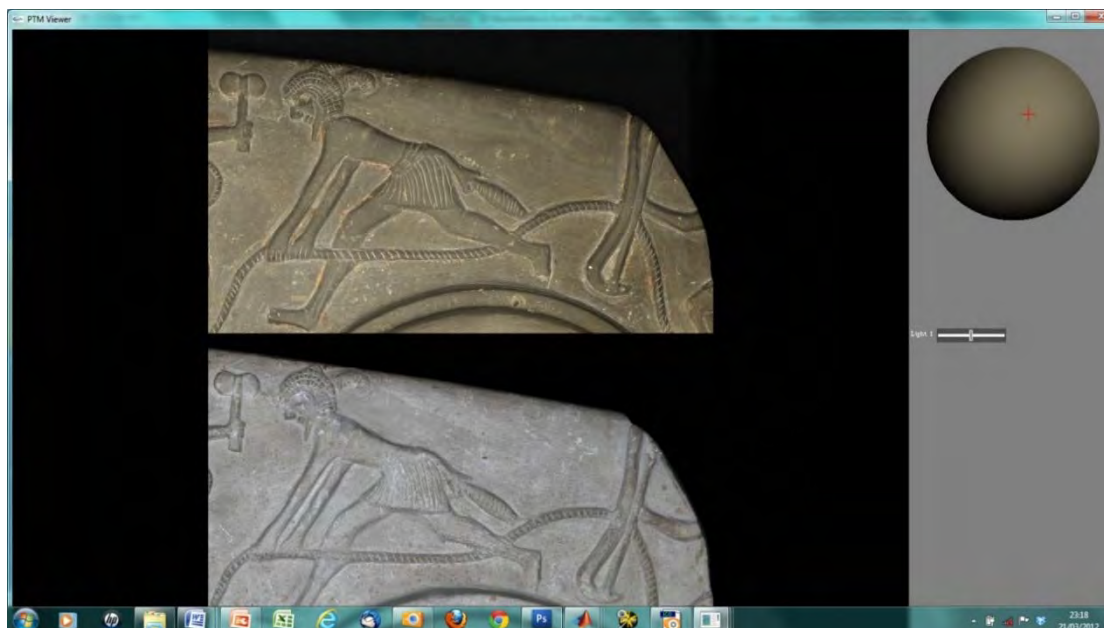


Figure 8.4: Screen shot of PTM viewer with hybrid presentation of original (above) and plaster cast (below) under combined interactive lighting control, with illumination from the same point in hemisphere.

Because the original in the BM was not available for scanning at UCL, the real value of the plaster cast in this study arose from its digitisation at UCL with the Arius 3D colour laser scanner. The resulting point cloud at a resolution of 10 points per mm, with an absolute accuracy of 25 μm , provided the precise geometric framework on which to map the higher-resolution photometric surface detail derived from the images of the original object. Surface normals were computed from the image sets, and the 3D surface was reconstructed by the method of Chapter 5. The surface normals were registered with the underlying geometry obtained from the laser scan of the plaster cast by first determining the outlines of the palette in the photographic image and in a projection of the point cloud onto the Z plane. The image was rotated and scaled (an affine transformation) to achieve the best fit between the two outlines (Fig. 8.5 left). In this case the rotation angle was 4.8° counter-clockwise and the enlargement factor was 1.76.

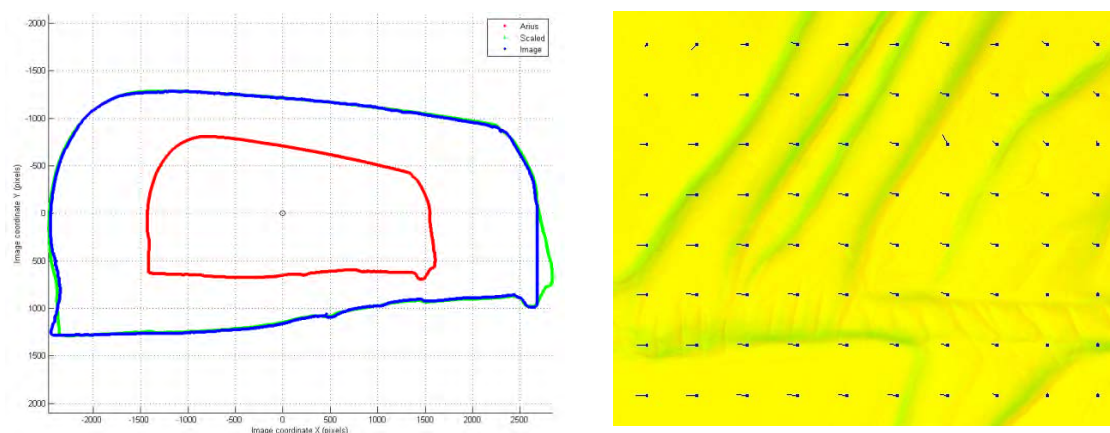


Figure 8.5: (left) Outlines of the palette in the scanner point cloud (red), the image (blue) and the rotated and scaled point cloud (green); (right) False-colour composite of the Z normals of the image (green) and point cloud (red), showing the misalignment after rotation and scaling. The blue vectors at 50 pixel intervals indicate the direction and magnitude of the displacement.

The outline is an effective feature for this purpose, akin to pattern recognition, because the normalised function of radius from the centroid provides an invariant signature of the object, enabling the rotational angle to be determined by correlation. This method gave a good result, even though a sliver of the palette had inadvertently been cropped from the right edge in the course of photography. The Z normals of the point cloud data were transformed by the same angle and scaling factor as for the outline, bringing them into approximate correspondence with the image photometric Z normals. Cross-correlation was then applied over the image on a grid at 50-pixel intervals to find the local offsets between the two images (Fig. 8.5 right).

A ‘rubber-sheet’ geometric distortion of the image was then applied to bring all areas of the image into accurate registration with the orthographic projection from the laser scanner. By this means the photometric normals were registered to sub-pixel accuracy with the height map derived from the scanner. Finally an enhanced height map was obtained by integrating the photometric normals derived from the real object into a continuous surface and blending their high frequencies with the low frequencies of the underlying geometric structure of the plaster cast (Fig. 8.6).

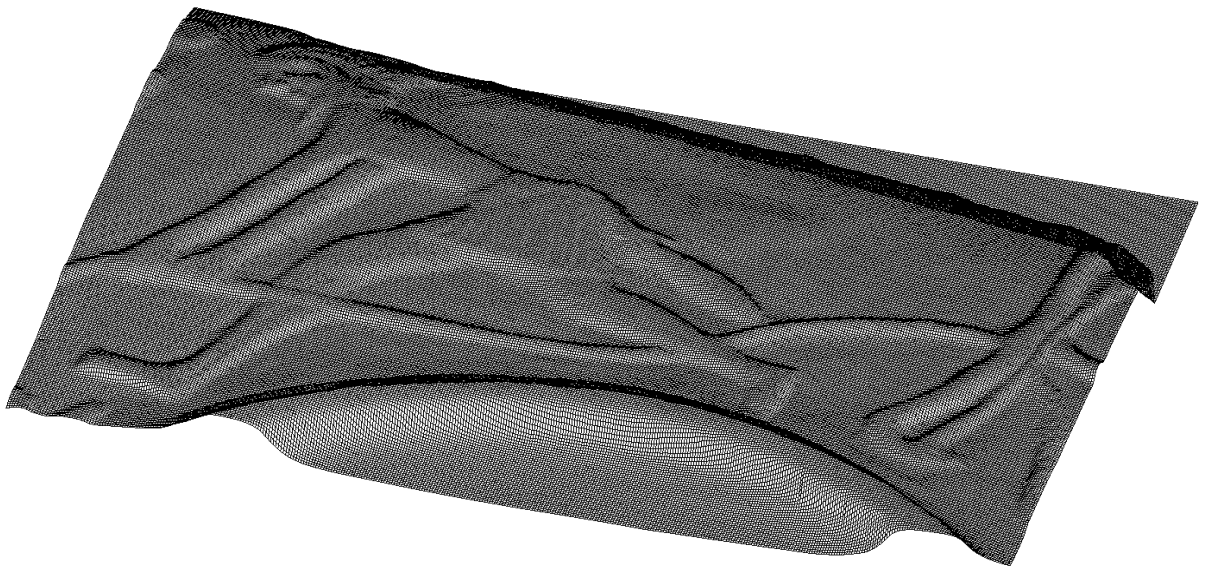


Figure 8.6: Digital terrain model plotted on a 5-pixel mesh (intervals of 0.28 mm).

To quantify the depth of the various surface features, cross-sectional elevation profiles were plotted through the digital terrain map. Fig. 8.7 shows a depth map image, in which the Z-coordinate is represented by an ordinal grey scale from black for the lowest to white for the highest. The red lines indicate the locations of the horizontal and vertical cross-sections. These show that the height of the rim around the circular grinding area of the palette is 5 mm, and elsewhere the height of the figural carving has a maximum peak-to-trough amplitude of approximately 1.5 mm, with fine details of 0.5 mm in depth, as in the pleats of the kilt. Around the major structures of the design, such as the lassoer's arms on the left and the bow on the right, are incisions (grooves) of approximately 0.3 mm in depth, suggesting that these may have been initially scribed onto the surface of the mudstone to delineate the figures.

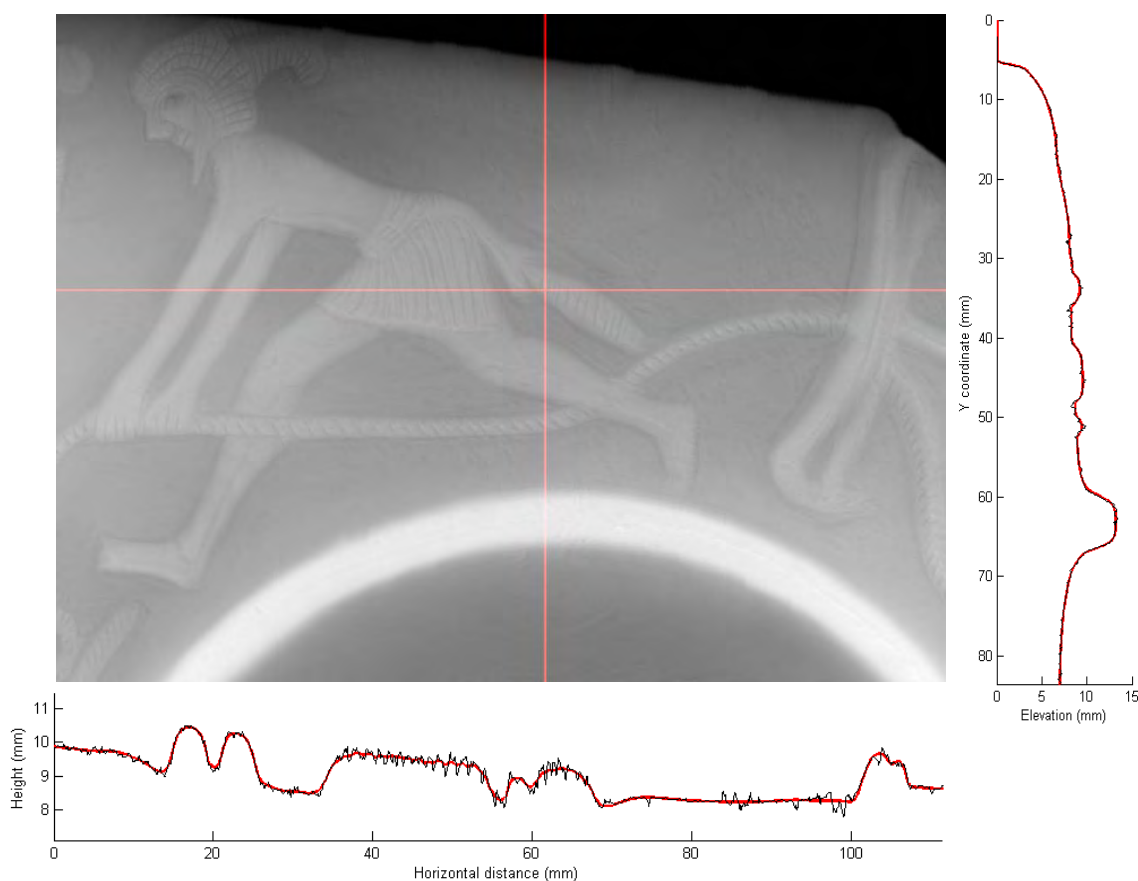


Figure 8.7: Height map with red lines showing location of elevations for a horizontal section (below) and a vertical section (right). In the elevation graphs the red lines represent the Z coordinates from the Arius 3D scanner and the black lines show the photometric enhancement of detail.

In the region of interest above the hunter's back, the furrow is seen to be approximately 5 mm in width for the majority of its length, punctuated by two features: in the centre there is a deeper bowl-shaped depression; further along is a narrow vertical ridge where the furrow briefly disappears. The dimensions of the depth and width of the furrow are consistent with the erasure of incised lines of depth 0.1-0.2 mm on either side of an earlier position of the rope. These may have been initial drafting lines before the design was finalised and the deeper outlines of 0.3-0.4 mm cut.

In this project were developed the techniques for combining two PTM objects into a single display, the 3D reconstruction of a surface, and the registration of imagery with 3D scanner data, all of which were important contributions to the overall research programme.

8.1.2 Egyptian Scarab

The CNRS PICS project 5995 entitled 'Study of Bronze Age Egyptian Gold Jewellery' involved exchanges between several French and UK institutions. The focus was on the origin of gold and composition of the alloys used for forming, soldering, and decorating, in order to provide new evidence on the technologies used in Egyptian workshops during the Bronze Age, particularly in the 2nd Intermediate Period (1650-1550 BC). Several objects from the collections of the museum partners were studied by science-based techniques, including optical microscopy, X-radiography, scanning electron microscopy, X-ray fluorescence (XRF), and ion

beam analysis (PIGE). The project team investigated the contribution of 3D imaging acquisition and reconstruction to represent the original colours and surface detail of the materials used in the production of Egyptian gold jewellery. Principal investigators were Prof Maria Filomena Guerra (CNRS), Ruven Pillay (C2RMF, Louvre) and Prof Stephen Quirke (UCL Inst. of Archaeology). Several members of the 3DImpact team were involved, including Prof Stuart Robson, Mona Hess and Ali Hosseininaveh Ahmadabadian.

An object was selected that is very representative of the period, a small scarab (Fig. 8.8) of engraved steatite with a gold band, from the UCL Petrie Museum of Egyptian Archaeology (reference UC11365). Dated to the Late Middle Kingdom (1850-1750 BC) it is inscribed in hieroglyphs on the underside with the personal name and title '*Estate overseer of the granary lufseneb*' within the scroll border. Its dimensions are 26.7(w) x 18.6(h) x 11.4(d) mm. A hole of ~2.5mm diameter is drilled through the centre, suggesting that it may have been worn on a braided thread or leather thong around the neck (Gwinnett and Gorelick, 1993).

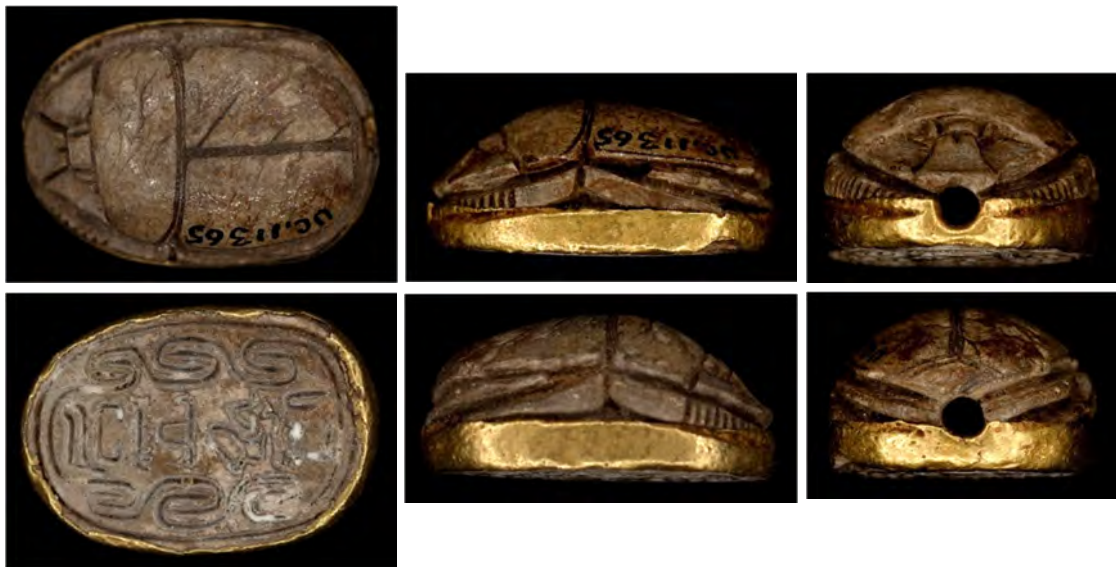


Figure 8.8: Six views of the Egyptian scarab from the UCL Petrie Museum (UC11365).

The scarab poses particular challenges for imaging due to its characteristics: a small object with a finely engraved inscription, the back is curved and polished to a semi-gloss finish, and the encircling chased gold band has a high specularity. Specific objectives were to produce a 3D representation with a resolution high enough to visualize tool marks, typically about 100 μm in dimension, and to display detailed surface geometry and correct colour. The inscribed gold band was analysed by handheld XRF together with three gold alloy standards, and showed that on average the alloy is composed of 0.3% Cu, 5.7% Ag and 94.3% Au, which is consistent with expectations for the period.

Six sets of 64 images were taken in the UCL dome imaging system, with the scarab in six different orientations (Fig. 8.8), using the Nikon D200 camera with a Nikkor 200 mm macro lens. Each image set was used to construct a representation of the scarab by variants of the polynomial texture map (PTM) technique, fitting a biquadratic function (6 coefficients) and hemispherical harmonics (16 coefficients) to the intensity distribution at each point. In addition the new cluster Lorentzian model was fitted to the specular component (Fig. 8.9).



Figure 8.9: Screenshots from the custom software viewer showing the specularity of gold in four representations of the left side of the scarab: (top left) original image; (bottom left) PTM model; (bottom right) hemispherical harmonics (HSH) model; (top right) cluster Lorentzian model.

The same six sets of images from the UCL dome were used to reconstruct the 3D surface by photometric stereo, by determining the surface normal at each pixel and then integrating them into a consistent surface, following the method of Chapter 5. The optimal subset of intensities was determined from all 64 candidates at each pixel, by sorting them into order and selecting a range between the shadow and specular regions, followed by regression over this subset for an accurate estimate of the normal. The height field was reconstructed as a digital terrain map by projecting the surface gradients onto Fourier basis functions to ensure integrability, then exported as a point cloud for 3D visualisation and comparison with the point cloud from the Arius 3D colour laser scanner.



Figure 8.10: (left) 3D reconstruction from the point cloud generated by the Arius laser scanner; showing the overall redness caused by the non-optimal laser wavelengths; (right) Detail of back.

Although the 3D laser scanner gave excellent results in terms of geometric accuracy, it was poor in colour accuracy and in the rendition of surface detail (Fig. 8.10). Even though the scanner was expected to show sufficient details with a sampling pitch of $100\text{ }\mu\text{m}$ (10 points/mm), it did not meet the expectation, and revealed less detail than shown by magnification with a

hand lens. The finest surface feature that can be wrought on an artefact by a craftsman is approximately $40\text{ }\mu\text{m}$, to capture which the scanner should have a spatial resolution of at least $25\text{ }\mu\text{m}$ (i.e. 50 pixels/mm).



Figure 8.11: The 3D reconstruction from photometric stereo produced a very dense point cloud (left and centre) showing fine detail of surface features (right).

Moreover the rendering of surface colour by the three lasers was not so accurate as with the camera, because the monochromatic sampling of the reflectance spectrum led to metamerism, as described in Chapter 7, causing the steatite to look redder than it actually was (Fig. 8.10 left). The opportunity to display diffuse colour and specular reflectance separately in the viewer made it easier to identify tool marks and material striations. The original capture resolution in the images was 65 pixels/mm ($15\text{ }\mu\text{m}$) but this was reduced to 53 pixels/mm ($19\text{ }\mu\text{m}$) to fit the full outline of the scarab into the HD display window. The rendition of the photometric stereo in 3D did not provide a metric image in the round (Fig. 8.11) but was successful, especially when reading the hieroglyphs on the underside of the scarab.

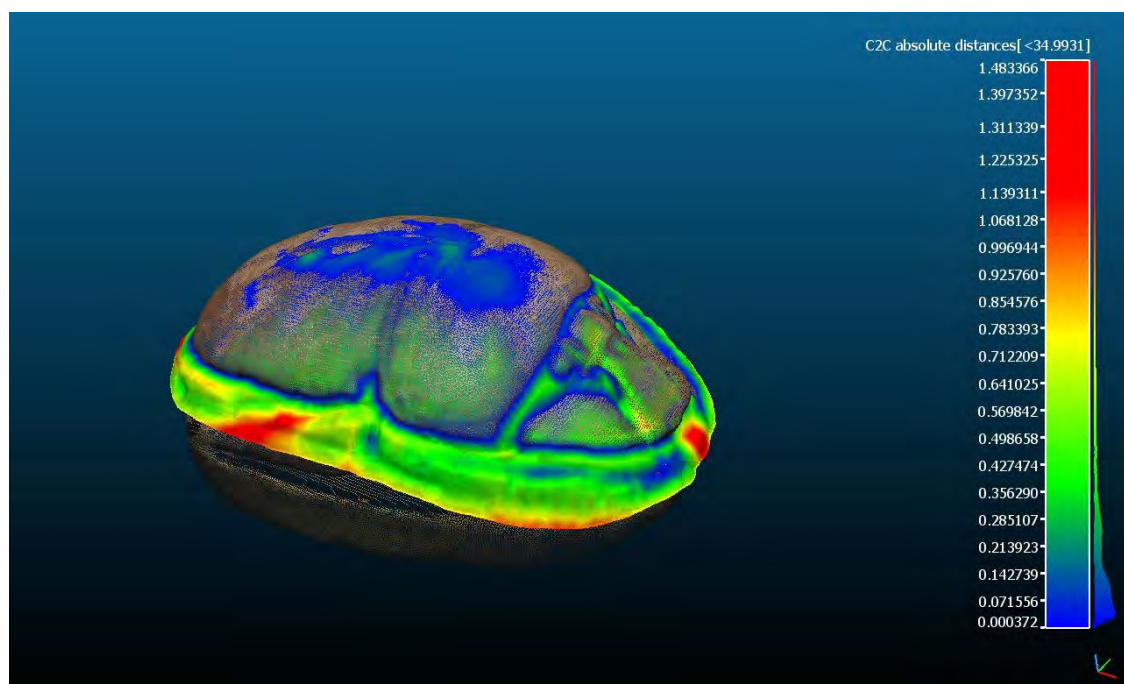


Figure 8.12: Screenshot from CloudCompare, showing the errors between the 3D reconstruction from photometric normals (test) and the point cloud from the Arius laser scanner (reference).

Comparing the point cloud generated from the 3D reconstruction with the point cloud from the Arius scanner (Fig. 8.12) showed that the majority of points over the curved top surface of the scarab fitted with an error of less than 0.3 mm.

8.1.3 Courtauld Bag

In the collection of The Courtauld Gallery, London, is a unique object made of brass inlaid with silver and gold, which is acknowledged as one of the finest and most intriguing pieces of Islamic metalwork in the world. Its sides are adorned with images of hunters, musicians and drinkers, and its lid with a scene of a couple enthroned together on a low dais. The man holds out a drink to his companion, and the pair are accompanied on either side by their retinues of servants. It was probably made in the period 1300-1320 for a patron in Il-Khanid (Mongol) Iran, and makes references to metalworking traditions in Mosul in Iraq (Ward, 2014). It portrays scenes of hunting and leisure that were common across the Near East, includes depictions of furniture and fashions unique to China, and presents an impression of authority and power. Its function was almost certainly as a handbag carried by the attendant of a high-status woman.



Figure 8.13: (left) The Islamic 'handbag' at The Courtauld Gallery; (right) placing the handbag at the centre of the baseplate of the UCL dome. For photography the hinged upper part of the dome was lowered onto the frame, so that the object was completely enclosed.

The Courtauld Gallery's Curator of Sculpture and Decorative Arts, Dr Alexandra Gerstein, approached UCL with a request to make a 3D model of the bag for an exhibition in Spring 2014, entitled 'Court and Craft: A Masterpiece from Northern Iraq'. The project team included Dr Rachel Ward, formerly Curator (Middle East), Department of Oriental Antiquities, British Museum, and John Hindmarch, a colleague at 3DImpact, UCL.



Figure 8.14: (left) Roundel on front of handbag depicting a seated drinker, illuminated by lamp 48, at a 'one o'clock' position and an elevation of $\sim 45^\circ$; (centre) albedo; (right) normal vectors.

Sets of 64 images of several views of the Courtauld bag were taken in the UCL dome by the Nikon D200 camera (Fig. 8.13). The 18-55mm zoom lens was used for the top view, set to

focal length 55 mm and aperture $f/8$. For details of three roundels, the 105mm lens was used, set to aperture $f/11$. Using the method described in Chapter 6, the images were processed and the various components were computed (Fig. 8.14). Because the object had not yet been cleaned when the photography was done, the silver was somewhat tarnished and so the specular components (Fig. 8.15) do not exhibit such a high dynamic range as, for example, the Damascene dish used as a test object in Chapter 6.



Figure 8.15: (left) specular quotient values; (centre) specular normal vectors; (right) specular colour.

A ‘flyover’ video at HD resolution was constructed, with each frame of 1920x1080 pixels computed from the model, while the light source was moved to successive positions in a scripted trajectory. Increments in illumination angle of 0.1° at 25 frames per second were sufficient to give the illusion of continuous motion. Panning was accommodated by translating the crop window by one or two pixels in each frame, and zoom by rendering a smaller crop area then enlarging to the HD video frame size. The frames were composited using Windows *MovieMaker*. The results (Fig. 8.16) were indistinguishable from real video footage and were integrated with the video sequence of a separately generated 3D model. The video was shown as part of the public exhibition at the Courtauld Gallery in summer 2014. By modelling surface normals and specular highlights from silver and gold a realistic impression was achieved.



Figure 8.16: One frame of the HD video flyover, showing the simulated appearance of the roundel with a virtual light source from a low angle in the east.

8.1.4 Painted Ceiling

The Queen's Staircase at Hampton Court Palace is part of a grand architectural scheme commissioned in the late 1720s by King George II and Queen Caroline. The room is rectangular of dimensions 12x21 m and has an imposing staircase rising to a landing of 12x5.5 m. The ceiling height is 12.5 m above the ground floor at the centre and 8 m above the principal floor landing. Daylight is admitted through three large windows and one smaller high window facing south. A large chandelier containing a cluster of tungsten lamps hangs low in the centre of the space.



Figure 8.17: Rectified view from below of the painted ceiling over the Queen's Staircase.

The painted ceiling, constructed in lath and plaster, is part of the decorative scheme completed by William Kent in 1734. The painting is executed in oil bound pigment and depicts a marble dome supported by a cornice and figures, featuring a central garter star. Extensive use is made of *trompe l'oeil* techniques with coffered panels, mouldings and elevated height to enhance the grandeur (Fig. 8.17). The design is centred on a Garter Star and supported by grisaille figures and elaborate ornament. In 2001 flaking was noted on the painted ceiling. It was unlike that observed on other wall paintings at the palace, as areas of instability were spread over the entire surface of the ceiling and resulted in the detachment of large brittle flakes of diameter up to 20 mm. Treatment with a conventional wall painting conservation material failed to address the problem (Davies, 2006). Since 2004 the conservation approach of Historic Royal Palaces (calling on the expertise of a multidisciplinary team of specialists) has been one of cautious intervention, coupled with continued research and recording to understand the causes of deterioration and to devise a suitable long-term stabilisation strategy.

The project was a collaboration between UCL and the Conservation Collection Care team at Historic Royal Palaces, working with Dr Ian Gibb and Dr Constantina Vlachou-Mogire. The aim was to develop a method of imaging the painted surface, in order to detect and characterise the process of flaking as it progresses and also be deployed at long range, preferably from floor level with minimal effect on public access and palace operations. By repeating the imaging procedure at regular intervals, it was hoped that comparison of images would provide valuable information on rates of change, such as crack growth, surface deformation and flake formation.



Figure 8.18: (left) Nikon D200 camera with 300mm telephoto lens and remote flash trigger; (right) damage to paintwork on coffered panel, with two details of areas where flakes have detached.

Photographic trials were undertaken for the feasibility of capturing suitable images by digital photography from the landing level. A Nikon D200 camera with a 55–300 mm zoom lens was set up on a tripod approximately 2 m above the landing floor level, at a distance of 6.5 m below the ceiling (Fig. 8.18 left). The ambient illumination of the ceiling was a combination of the bluish daylight coming through the windows and the yellowish light from the tungsten bulbs in the chandelier. Additional illumination was provided by a hand-held Nikon *Speedlight* flash, synchronised by a radio link. At maximum magnification of the lens (zoom set to 300 mm focal length) the field of view was $5^{\circ}20'$, so the full image of 3:2 aspect ratio covered an area of 56x38 cm. For the image size of 3872x2592 pixels this represented a surface resolution of 6.9 pixels/mm (pixel size $\sim 150\ \mu\text{m}$).

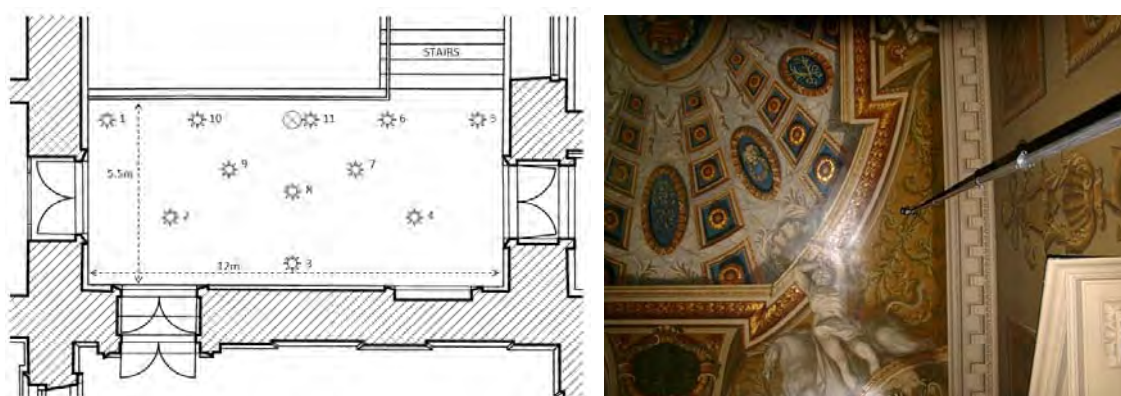


Figure 8.19: (left) Floor plan of landing of Queen's Staircase, showing locations of camera and flash; (right) Mast with flash light fully raised in location 1. The target area is the oval garland on the left.

Further investigation was undertaken into the effect of raking light on the surface of the ceiling, and whether quantitative information could be derived from images captured by this means. The camera was again mounted on a tripod, at a height of 2.0 m above the landing floor. The flash light was mounted on an 8 m telescopic mast, enabling it to be raised manually to $\sim 1\ \text{m}$

below the ceiling (Fig. 8.19 right). Images were taken from the fixed camera position with the zoom lens set at 300 mm, while the flash light was moved successively to eleven locations around the landing (Fig. 8.19 left), with five locations in each of two concentric semicircles plus one at the centre (close to camera axis). The raking light from the flash, at an incident angle of $\sim 10^\circ$, modulated the surface relief, and the array of images could be treated using the methods of dome imaging for PTM visualisation and 3D reconstruction of the surface.

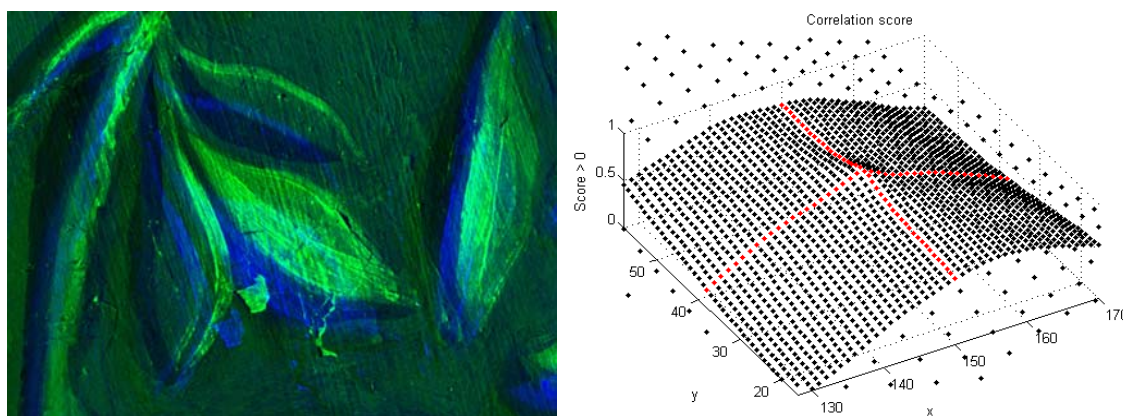


Figure 8.20: (left) The effect of the misalignment of images is revealed in a false-coloured composite of a detail of image 1 (green) and image 5 (blue); (right) Surface of correlation scores for offsets of ± 21 pixels from the maximum. The red lines indicate the row and column of maximum correlation.

The high magnification through the 300 mm lens made the process very sensitive to small perturbations of the camera and its tripod, and resulted in significant misalignment of the images (Fig. 8.20 left). The images were registered using 2D correlation on the green channel, taking a small square of 101×101 pixels of reference image, and sliding it around an enclosing region of 201×201 pixels from the test image to find maximum of correlation coefficient r (Fig. 8.20 right).

Surface normals were calculated for each of 37 three-lamp combinations for every pixel position, and the median taken. The normals were corrected to remove the macro geometry (orientation and curvature of the ceiling relative to the camera) while retaining the micro geometry (local cracks, bulges and flakes). The 3D surface was reconstructed using the Frankot and Chellappa method, in which the surface gradients are projected onto a set of integrable basis functions. This enabled cross-sections of the features to be obtained, at a surface resolution of 7.7 pixels/mm (Fig. 8.21 right). The depth (Z axis) resolution achieved by the photometric stereo technique was finer than 0.2 mm (i.e. 200 μm), which is an order of magnitude better than most 3D laser scanners, operating at a range of 6 metres.

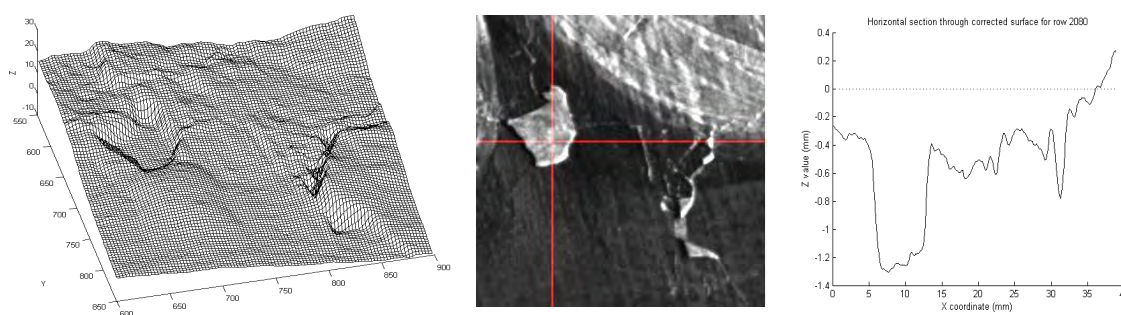


Figure 8.21: (left) 300x300 pixel detail of 3D reconstruction of surface of ceiling, showing a flake and a crack; (centre) Monochrome image, with locations of horizontal and vertical cross-sections shown by red lines; (right) depth profile along the horizontal section.

8.1.5 Degraded Parchment

Large collections of parchment documents exist in public and private libraries, archives, and museums in varying degrees of preservation. Digitisation projects generally concentrate on documents in their existing state. Typically the focus of image processing is to visualise the supposed original state of the document, before damage occurred, i.e. virtual reconstruction.

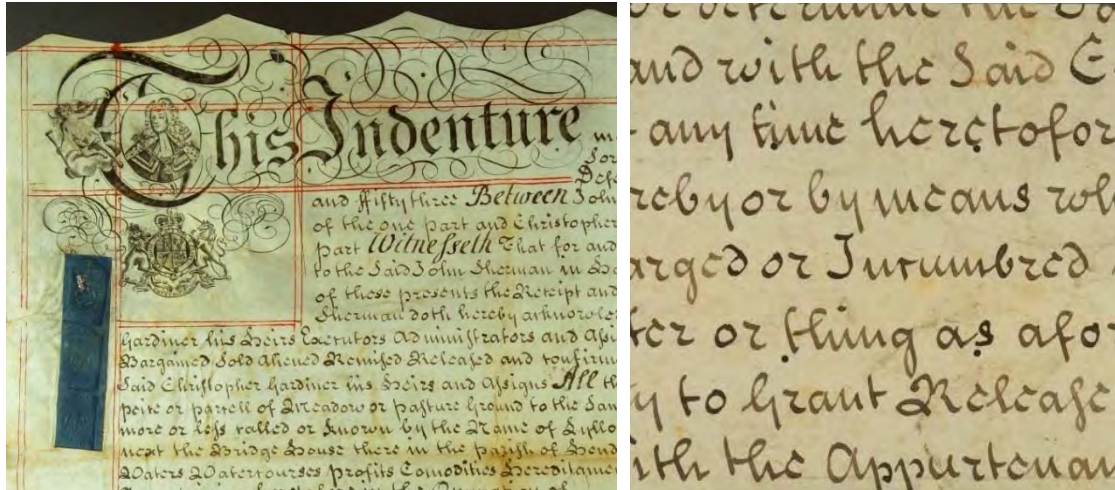


Figure 8.22: (left) Upper left corner of first page of parchment, showing engraving and the arms of King George II; (right) one section of 8x8 cm cut from parchment sheet.

This project was funded by EPSRC as a collaboration between the UCL Dept. of Medical Physics and the London Metropolitan Archives (LMA), working with Dr Adam Gibson and Dr Alejandro Giacometti (UCL), Dr Caroline De Stefani (LMA) and Dr Alberto Campagnolo (now at the Vatican Library, Rome). We investigated the processes of controlled document degradation using multispectral imaging, on an 18th-century manuscript de-accessioned from the LMA.



Figure 8.23: Copystand with four tungsten-halogen lamps for reflective imaging of samples with the Nikon D200 camera. The transmission filter is screw-mounted into the front of the lens.

Although it was in good physical condition, this manuscript was a tedious legal contract from 1753, deemed to hold no historical value, and it was donated for experimentation (Giacometti *et al*, 2012). It comprised two large sheets of parchment, c.70×70 cm, written in black iron gall ink and highlighted in red ink (Fig. 8.22). Twenty-three square sections of 8×8 cm were cut from the original manuscript, with each containing written text. Each sample was exposed to an external deteriorating agent to produce a controlled degradation, including mechanical damage, heat, humidity, abrasion and a variety of substances with different chemical properties, such as acid, alkaline, bleach, tea and human blood. These affected the appearance and condition of samples in different ways, typical of the actual damage suffered by parchment in real archives.

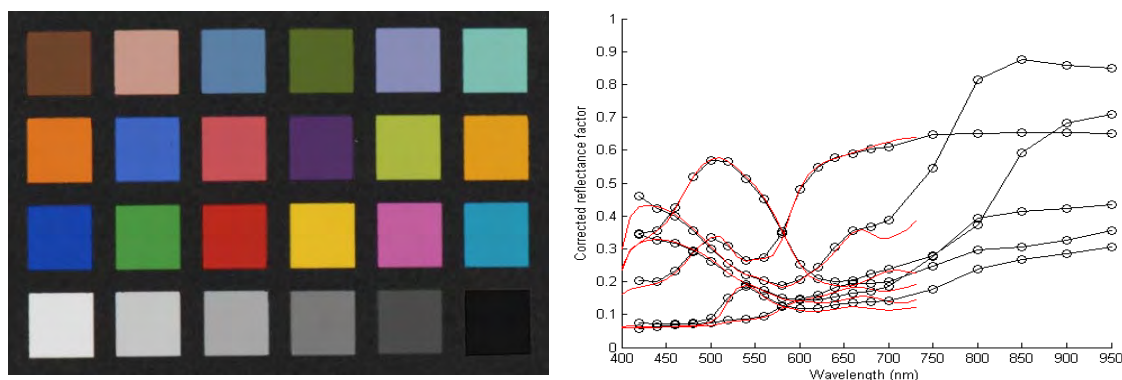


Figure 8.24: (left) MiniMacbeth target; (right) Reflectance factor vs wavelength for camera signals (black lines with circles) and spectrophotometer measurements (red lines) for top row of target.

Each of the parchment samples was imaged before and after the treatment, using two different cameras, in combination with two lighting systems for both reflective and transmissive modes, through a series of bandpass filters (Fig. 8.23). The cameras were: (1) Nikon D200 with 105 mm f/2.8 lens; (2) Kodak Megaplug 1.6i scientific monochrome camera with Nikkor 50 mm f/2 lens. For the Nikon camera, a set of 16 bandpass optical interference filters was used, spaced at 20 nm intervals across the visible spectrum. For the Kodak camera 5 additional filters at 50 nm were used to cover the near-infrared range.

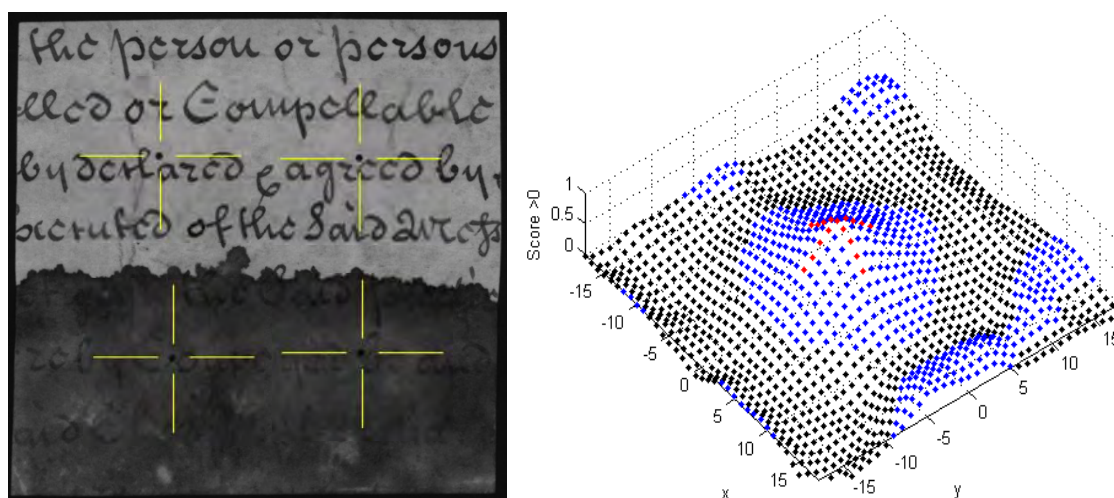


Figure 8.25: (left) Image taken by Kodak camera through 600 nm filter of parchment sample, after the treatment by blood. The yellow lines indicate the positions of the four registration holes; (right) Correlation scores for the lower left hole, visualised as a surface on a pixel grid, showing the peak score at the hole centre. Scores <0 are represented by black dots, and scores >0 by blue dots.

The tonal linearity and colour accuracy were analysed with the aid of a MiniMacbeth Color Checker target (Fig. 8.24 left). Twenty-one images were taken through the successive filters, using the same camera settings as for the parchment samples. The luminance profile was measured, and mean image intensities corrected for white and black. The reflectance spectrum of each target patches was measured with an X-Rite iIPro spectrophotometer, and demonstrated a good match (Fig. 8.24 right).

Because of differences in refractive index of the camera lens for different wavelengths of light, the geometry of the successive images in the multispectral sequence varied from one image to the next, so that no two were in perfect register. Also in some cases the treatment physically altered the parchment, causing geometric distortion. The images were registered by using the four 1 mm holes drilled in each parchment sample (Fig. 8.25 left) as the anchor points for a projective geometric transform. The image of the sample before treatment taken through the 600 nm filter was chosen as the reference and all other images, both before and after treatment, were registered to it. A template matching procedure was used to locate the centroid of each hole in each sample, by correlation with a circular template (Fig. 8.25 right).

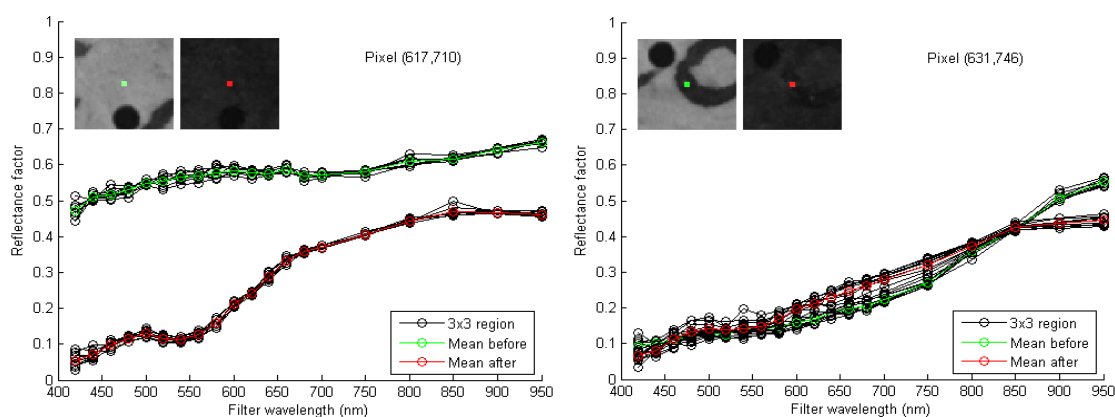


Figure 8.26: Reconstructed reflectance spectra in a 3x3 pixel region before and after treatment by blood for: (left) parchment and (right) ink.

When all of the image channels for the sample before and after treatment had been accurately registered, it was possible to compare the effect of the treatment on the reflectance spectrum at any pixel position. Fig. 8.26 shows the effect of the treatment by blood on the parchment and the iron gall ink in locations close to the bottom left hole. The spectra are plotted for all nine pixels in a 3x3 region around the selected coordinates, together with the mean in green (before) and red (after). The blood clearly affected the colour of the parchment, reducing its reflectance from about 0.5 to less than 0.1 in the short wavelengths, rising to about 0.4 at long and NIR wavelengths, giving a dark red colour, whereas the ink spectrum was apparently not much changed. Thus the study was able to demonstrate that multispectral imaging could be used to characterise the changes in the reflectance spectrum of parchment caused by treatment with a variety of chemical and physical agents.

For all five of the projects described above in this section the common theme was interdisciplinary collaboration in a team to learn more about an object of cultural heritage value. The ability to use digital image processing to analyse object surfaces in tone, space, spectrum and time is a valuable tool for digital humanities.

8.2 Contributions to Knowledge

8.2.1 Geometry of Dome

There are believed to be approximately 100 illumination domes around the world at the present time, and no two of them are the same. They differ in diameter, materials, construction technique, type of illumination, control equipment, camera mounting, method of introducing objects, etc. Crucially they also differ in the number and position of the light sources over the hemisphere. Generally the geometrical coordinates of the lamps are not known *a priori*, and so have to be determined empirically by the highlight method (see Section 2.2.4). The normalised coordinates (X,Y,Z dimensions) are stored in the LP file used to generate the PTM or RTI output file. This coordinate data is often ephemeral and not kept in the project archive and certainly not stored as metadata in the PTM file header. (Actually this is an issue with dome photography in general, that information about the geometry and spectrum of the illumination is not retained with the imagery.)

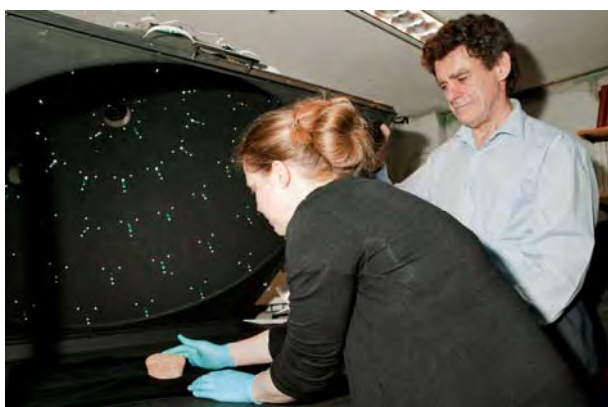


Figure 8.27: In the lab with Mona Hess, positioning an Egyptian clay funerary cone from the UCL Petrie Museum, UC37585. The retroreflective targets on the inside of the dome, four placed around each flash light, are clearly visible.

The innovation in this project was to demonstrate that the methods of photogrammetry could be brought to bear on quantifying the coordinates of the lamp centroids (Fig. 8.27). This was demonstrated in Chapter 3, both with retroreflective targets on the structure of the dome, and also by using the lamps themselves as targets. The advantage of these methods is that the interior and exterior orientation parameters of the camera could also be determined. An alternative method was also developed, using the predetermined coordinates of a pattern of targets on the baseplate of the dome to establish a direct transformation from physical positions on the baseboard to pixels in the image without recourse to the collinearity condition.

8.2.2 Photometric Normals

In conventional practice in photometric stereo, normals are derived from only three points, with the assumption that the surface is Lambertian, i.e. that the reflected light is perfectly diffuse in every direction. The consequence is that the derived normals are very noisy and for real materials, which are rarely Lambertian, they are not accurate. The effect of surface sheen or gloss is to distort the direction of the normal vector toward a greater angle from the view vector. In this study a new adaptive method has been developed on the assumption that the surface reflection can be represented as the sum of a true diffuse component, which is present at all angles, and a specular



Figure 8.28: Photometric normals calculated for an Egyptian clay funerary cone from the UCL Petrie Museum, UC37585.

component, which is negligible at angles far from the specular peak. Normals are then derived from regression over a number of intensity values (minimum 10) that lie above the shadow threshold but below the specular range (Fig. 8.28). The method was shown to work well over a wide range of materials with surfaces ranging from matte to glossy to metallic (Sec. 5.2).

8.2.3 Specular Angle

In conventional practice in computational photography it is almost universally assumed that the specular angle should be exactly double that of the normal, and for a perfect mirror this would of course be true. But real-world surfaces have a meso-structure with fine texture and granularity. The patina of old objects includes toolmarks, scratches, dents and all the accumulated effects of age and wear. One pixel as sampled by the camera may span a number of micro-facets at different angles, which reflect light differently from the incident illumination. Hence the assumption of the strict geometric relationship between normal and specular vectors is not true at the pixel level.

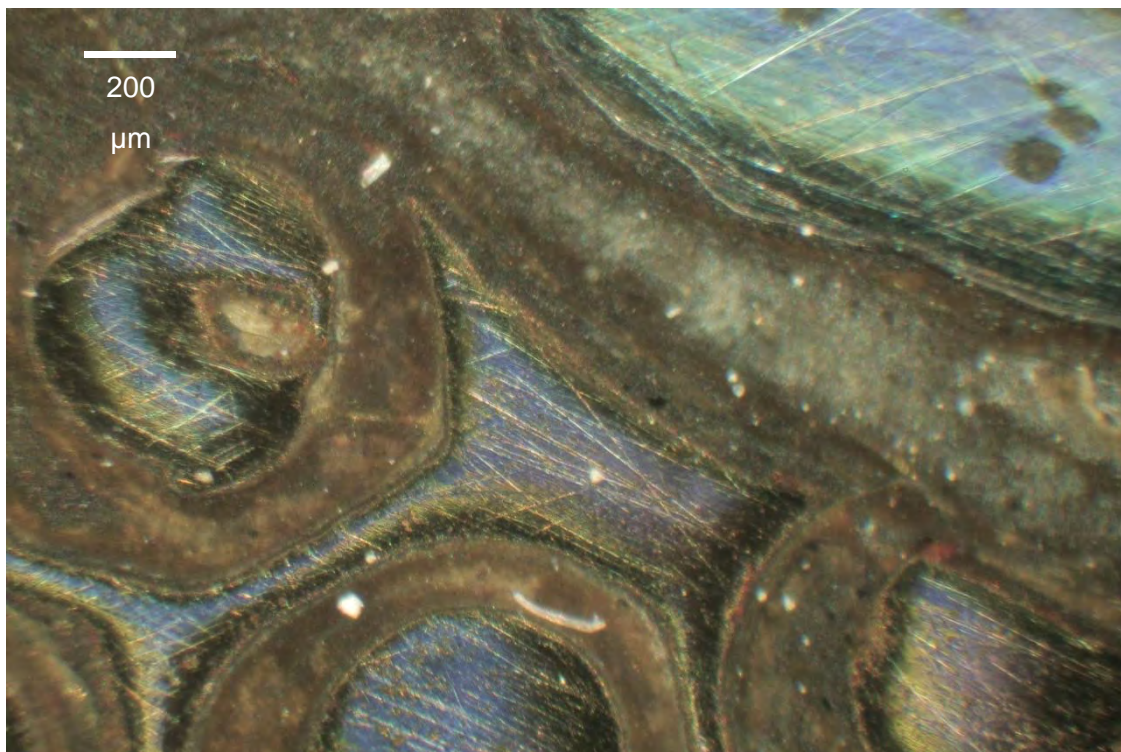


Figure 8.29: Microscope image of brass and silver surfaces on the Damascene dish.

As an example, a detail of the Damascene dish seen through a high-power microscope (Fig. 8.29) reveals a myriad of fine scratches, corrosion spots, pits and peaks on what to the naked eye appears to be a flawless highly polished surface. The image was taken with incident illumination from a tungsten halogen lamp through the x5 objective lens, and with crossed polarisers to minimise the specular reflection. The resolution of this image is ~ 1940 pixels/mm, so one pixel represents $0.51\ \mu\text{m}$ (approximately 30x the magnification achieved by the 200 mm lens on the Nikon D200 in the dome). Two further examples are shown in Fig. 8.30, taken through the x20 objective on the microscope, giving four times the magnification of the image in Fig. 8.29 and hence a resolution of 7760 pixels/mm, so one pixel represents $0.13\ \mu\text{m}$.

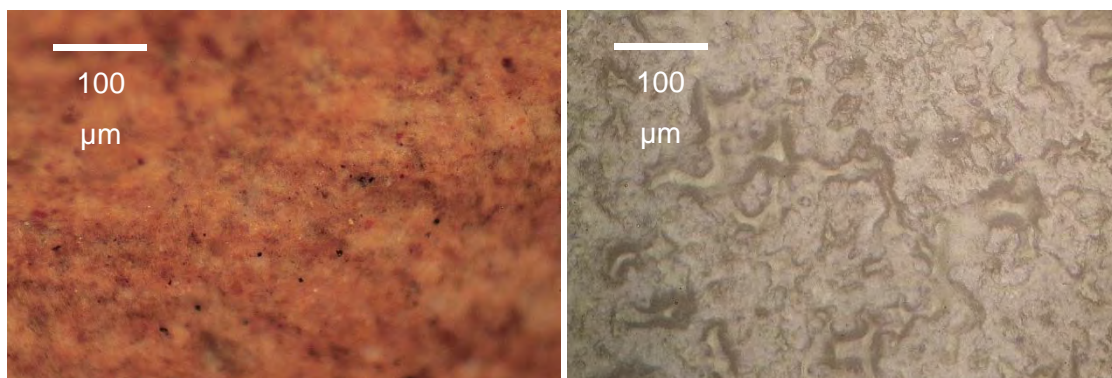


Figure 8.30: Microscope images of: (left) Chopin terracotta; (right) semi-matte white tile.

Even though it appears to be sealed with a satin varnish to a smooth uniform gloss level, the Chopin terracotta surface (left) reveals fine striations from the craftsman's spatula, and particles of quartz, iron oxide, mica and other constituents of the clay. The semi-matte white tile, despite its completely uniform appearance to the naked eye, has irregular fine blobs across the surface with dimension of the order of $20\ \mu\text{m}$. These are possibly caused by melting during the glazing process of a fine glass powder coating on the ceramic substrate.

The approach taken in this study for calculation of the specular angle has been to use the ideal specular vector (at double the angle of the normal) as a guide to where the specular angle should be. A weighted sum is taken of all lamp vectors within a cone of radius 45° around this direction, using the specular quotient value to determine the weighting of each lamp. This yields a 'specular normal' vector for each pixel. Plotting the resulting relationship between the specular and normal angles (Fig. 8.31) shows that although the average angle of the specular is indeed double that of the normal, there is considerable scatter over the full range of angles from 0° to 40° .

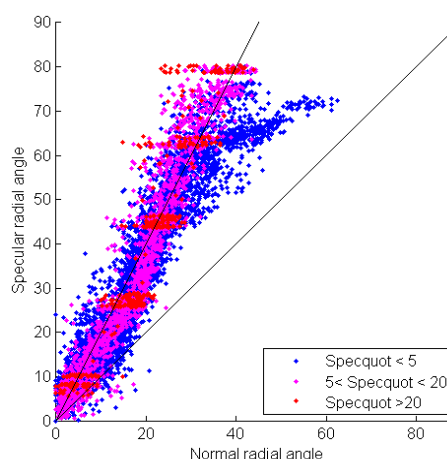


Figure 8.31: Specular vs normal angles computed for the Roman medallion.

8.2.4 Modelling Specularity

For each of the lights in the dome analysis of the ratio of reflected intensity between the actual reflection from the surface and an idealised Lambertian surface showed that the values form a distinctive pattern around the specular vector. This could be modelled effectively by a modified Lorentzian function of radial angle. Two different models were investigated, one with a standard Lorentzian plus a linear flank, and the other with a variable exponent. These specular models provide the basis for reconstructing realistic renderings of gloss and satin surfaces, bringing sparkle and vitality to imagery that would otherwise be dull and lifeless. The effect is particularly obvious for high angles of incident illumination (Fig. 8.32 left) but the broad flanks of the specular component make a visible difference (a lustre over the surface) at angles of up to 60° from the specular angle (Fig. 8.32 right).



Figure 8.32: Renderings with Lorentzian model of obverse of two silver Roman denarius coins, diameter 17 mm: (left) Republican period c. 120 BC from Narbonne in Gaul, showing the deity Roma; (right) Deification of Empress Faustina c. 140 AD. Image sets were taken in the UCL dome with a Nikkor 200 mm macro lens.

8.2.5 3D Surface Reconstruction

The normals derived from photometric stereo provide fine detail at pixel resolution, but cannot easily be integrated to reconstruct the surface for two reasons: (1) in the dome they give reliable gradients only up to $\sim 40^\circ$ because of the hemispherical lighting geometry; and (2) quantising errors limit their precision and accumulate across the width of the object surface. In consequence, surfaces derived from photometric normals alone may suffer from incorrect Z scale, with the amplitude being too great or too small, or may be distorted in the low frequencies, leading to curling at the edges or bowing in the centre of the whole surface.



Figure 8.33: Anaglyph stereo rendering in cyan (right eye) and red (left eye) of an inscribed Roman tablet found at Vindolanda in Northumberland. The 3D reconstruction was made from a set of 76 images captured in the Southampton dome by Kathryn Piquette at the British Museum in December 2010.

It was shown in Chapter 5 how the reconstructed surface can be calibrated by a small number of height measurements at selected points, or by reference to the geometrically accurate point cloud from a laser scanner. By selecting the low frequencies from the reference form and the high frequencies from the photometric normals and combining them in the Fourier frequency domain, a 3D structure can be obtained that has both the geometric accuracy of the reference scanner and the fine spatial detail of the imagery from the camera. Once the digital terrain

model has been produced it can be used in many ways, for example exported as a point cloud or rendered for stereo-pair viewing by polarising or anaglyph displays (Fig. 8.33).

8.2.6 Colour Laser Scanner Wavelengths

Because a laser is monochromatic, it samples the reflectance spectrum of a surface at a single wavelength. A colour laser scanner with three lasers can produce signals for red, green and blue, but the representation of the colour is highly metameric. Three lasers at single wavelengths can never produce the same triplets as the human visual system for all possible surface spectra because they do not have the broadband sensitivity of the retinal photoreceptors. The range of errors to be expected can be quantified by colorimetric analysis of a large number of spectra of real materials.



Figure 8.34: Trace of the three R,G,B lasers of the Arius scanner, simultaneously traversing across the surface of an Egyptian funerary cone from the UCL Petrie Museum, UC37585.

A new result arose from the question of which three wavelengths for a colour laser scanner would be optimal, in the sense of minimising the colorimetric error over a set of materials. The ideal wavelengths of 460 nm (blue), 535 nm (green) and 605 nm (red) turn out to be the same as the 'prime wavelengths of vision' in the theory developed over the past 40 years by Thornton. This points to a deeper reality in the construction of the human visual system, as well as suggesting how colour laser scanners might be optimised in future.

8.2.7 Synthetic Spectra and Colour Transformation

A convenient method for transforming colour data from one colour space to another is by means of a 3D lookup table. But the question is how best to load it in order to achieve the desired colour reproduction objectives. Typically this is done by regression over a number of samples in a colorimetric space, such as CIE XYZ or $L^*a^*b^*$. A new spectral-based method was introduced in Chapter 7, by generating large numbers (up to a million) synthetic spectra at 1 nm intervals. These spectra are parameterised to have the characteristics of the reflectances of real materials, in terms of continuity and spectral frequency, and randomised to span a wide range of variations. These spectra are then used to load all the cells of the lookup table, using a statistical approach. The result is that the transform manifold can adapt locally in a much more flexible way than a polynomial regression and can achieve optimum performance by producing the most probable output colour at every point in the 3D input colour space. A simulation has been implemented to demonstrate the improvement of colour accuracy against viewing of the original object under a D65 daylight simulator (Fig. 8.35), and discussions are underway with Arius on its efficient implementation in the hardware of the next generation of scanners.

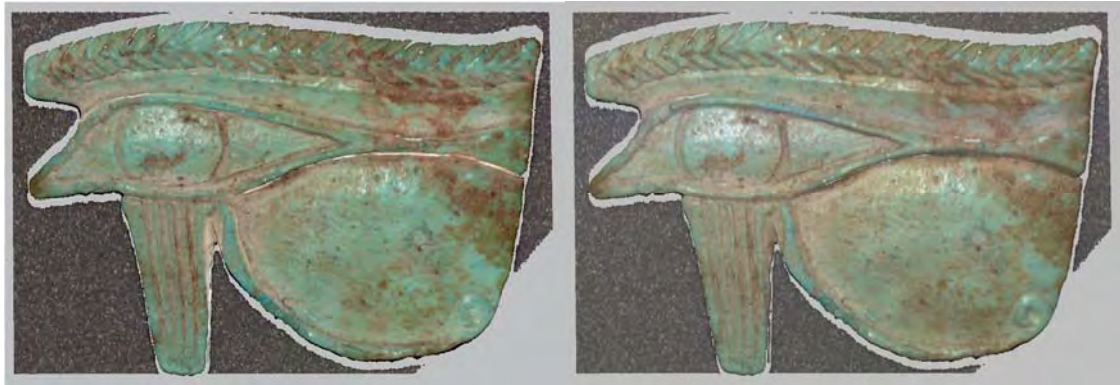


Figure 8.35: Green faience udjat eye amulet UC8504 from the Egyptian Late Period, scanned at the Petrie Museum on the Arius Rover colour laser scanner: (left) uncorrected RGB image; (right) after colour correction.

8.3 Directions for Further Research

During the course of the research, many ideas have arisen for further investigation. Very often these have been spurred by the problems associated with individual objects. Each object has its own characteristics of size, curvature, tone, colour, specularities, etc. and these affect the performance of the various algorithms at each stage of the image processing chain.

8.3.1 Dome Design

The dome used throughout this investigation has 64 flash lights positioned somewhat arbitrarily, but giving a reasonably uniform sampling of angles over the hemisphere, with a mean angle of $\sim 20^\circ$. Although the apparatus has been very reliable and enduring, its limitations have become clear:

- The flash lights are not point sources, but rectangular apertures;
- The luminance of the flash lights is barely sufficient;
- The selection of the lights is manual, not automated;
- The number and spacing of the light sources limits the angular resolution.

A new dome with a larger number of bright LEDs under computer control would improve the ability of the system to resolve fine variations in angular reflectance distribution, especially for dark surfaces. Discussions are ongoing with Dr Kirk Martinez at the University of Southampton to implement some or all of these modifications in a new dome to be used at the Centre for the Study of Ancient Documents (CSAD) in Oxford.

8.3.2 Dome Size and Automation

The existing dome is just over 1 m in diameter, and with the associated external steel frame requires a table top ~ 1.2 m square for support and operation. Its size and weight mean that it is restricted to the laboratory and cannot be taken to museums or sites where objects are located, so objects have to be brought to the dome. The 'rule of thumb' is that the dome diameter should be at least three times the diameter of the largest object to be imaged. Thus for the existing dome the largest object would be ~ 34 cm in diameter. If a new dome were to be constructed for digitising a collection of smaller objects, such as coins, it could be much more compact and hence portable. Another possibility would be to mount it on a robot arm

to enable positioning relative to a large object for repeated acquisition. Also it would be feasible to incorporate it into a fixed position above a production line flow for automated mass digitisation. Preliminary discussions have been held with Prof Markku Hauta-Kasari at the University of Joensuu (East Finland) on how a bespoke dome might be incorporated into the production line at the Digitarium (Fig. 8.36).



Figure 8.36: Conveyor belt carrying trays of plant specimens from the Norwegian national herbarium being digitised by a 2D camera setup at the Digitarium in Joensuu, Finland.

8.3.3 Goniometric Adjustment

The angular resolution of the dome, i.e. its ability to sample the angular reflectance distribution of an object surface, is determined by the location and spacing of the lamps over the hemisphere. In the current layout these are spaced at intervals of $\sim 20^\circ$. To achieve an average 10° spacing would require four times the number of lamps, i.e. 256. To achieve an average 1° spacing would require approximately 25,000 lamps, which is clearly impractical. An alternative would be to retain the existing placement of lamps and to tilt the object systematically by mounting it on a double goniometer stage, either by manual adjustment (Fig. 8.37) or motorised under computer control. This would enable the true location of the specular peak to be determined much more accurately and the surrounding reflectance distribution to be sampled at intervals of as little as 0.1° . The drawback is the need to acquire and process a much larger volume of image data.



Figure 8.37: Double goniometer stage.

8.3.4 Anisotropic Surfaces

The modelling of the specular peak in Chapter 6 assumed that the angular reflectance distribution of the specular peak is circularly symmetric, and so could be modelled by a single variable of radial distance from the specular peak. Many real surfaces are anisotropic, however, meaning that they exhibit different reflectance characteristics at different azimuthal angles. Examples are brushed metal, feathers, and woven silk. The power Lorentzian model could be extended by adding two more parameters to represent the anisotropic distribution by a rotated ellipse, replacing the single radial value for scale at each pixel by the semi-major and semi-minor axes of the best-fitting ellipse and its angle of rotation (Fig. 8.38).

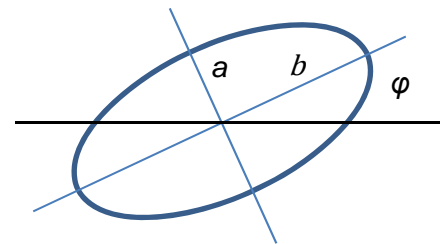


Figure 8.38: Rotated ellipse for modelling anisotropic surfaces.

8.3.5 RTI Algorithms

For palaeography, where the primary requirement is to read characters inscribed or moulded on the surface, the lamps at low angles of elevation ('raking angles') are more efficacious than at high angles, and it would be beneficial to have more lamps at angles less than 15° and relatively fewer at higher angles. A new dome design could combine a sparse pattern of lamps covering the complete hemisphere with a dense pattern of many lamps in a narrow belt just above the equator. New RTI algorithms would need to be developed to represent the intensity distribution as the sum of these two components, i.e. coarse with low angular frequencies over the whole hemisphere and fine with high angular frequencies in the band at low zenith.

8.3.6 High Dynamic Range (HDR) Imaging

Objects with light matte surfaces, such as plaster reliefs, can be photographed well in the existing dome under the individual flash lights, with a camera aperture in the range $f/4$ to $f/8$. This gives a good compromise in the images between exposure, depth of field and geometric distortion. For objects with dark surfaces, however, there is insufficient illuminance from single flash lights and the exposure levels are low, producing an undesirable amount of noise in the shadow regions. For glossy surfaces the problem is exacerbated in two ways: the off-specular angles are darker than Lambertian and the specular peaks are much brighter. Indeed the specular peak intensity may be several thousand times the intensity of a Lambertian surface under the same illumination geometry.

The solution is to make multiple exposures, bracketing the exposure time (not aperture), and to combine them into a single image of extended dynamic range of up to 20 bits. The example in Fig. 8.39 shows ten bracketed images of a gold electro-plated pendant hand-made in Tarquinia in the style of ancient Etruscan jewellery c.500BC. The object was illuminated by a single discrete white LED placed alongside the lens of the Nikon D200 camera mounted in the UCL dome, and the aperture kept constant at $f/4$. There is a 500:1 ratio in exposure time between the endpoints of the series. Such an HDR technique would enable more of the image detail to be captured without under-exposing the shadows or over-exposing the highlights, although in the dome it would require multiple image sets to be acquired (all in pixel register) and pre-

processed to create an HDR version of the image corresponding to each lamp before the usual workflow. Note that this cannot be achieved with flash illumination, so would require a new dome equipped with LEDs.

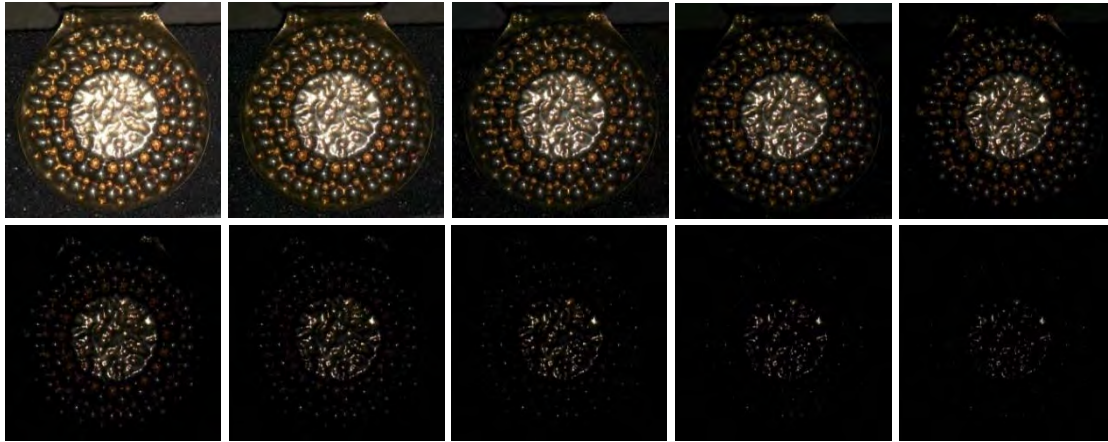


Figure 8.39: A bracketed series of images of an Etruscan-style gold pendant, illuminated by a single LED for ten different exposure times: (top row) 2, 1, 1/2, 1/4, 1/8 sec; (bottom row) 1/15, 1/30, 1/60, 1/125, 1/250 sec.

8.3.7 3D Multiview within Dome

In all dome photography performed for this research programme, the camera was fixed at the 'north pole' of the hemisphere and the object remained in one position on the baseboard. Hence all images were taken from a single viewpoint, which limitation led to the many interesting challenges of photometric 3D reconstruction. Stereo imagery could readily be captured, however, by moving the camera (or using a second camera) to view the object through one of the inspection portholes at an elevation angle of $\sim 50^\circ$. Also multiple image sets of the object could be obtained from the single overhead camera position by moving the object to multiple positions on the baseboard.



Figure 8.40: Demonstration of 3D reconstruction of surface from image pairs by lateral displacement of object in dome: (left) Setup of object on the baseboard of the dome with photogrammetric targets and colour chart; (centre) point cloud of reconstructed surface; (right) map of height differences between reconstruction and 'ground truth' from Arius laser scan.

A pilot study in 2010 in collaboration with Cristina Re from the University of Parma showed that it was feasible to use structure-from-motion (SfM) techniques for dense matching to obtain a digital terrain model. The test object was an Egyptian moulded clay funerary cone from the UCL Petrie Museum UC37585 (Fig. 8.40). The best results were achieved with ring illumination from Tier 3 lamps in the dome and a lateral displacement in the range 4-8 cm of the object between image sets.

8.3.8 Hybrid 3D Reconstruction

The normals derived from photometric stereo can be used to add fine detail in 3D reconstruction to an underlying form derived from other means, such as a laser scanner or dense multiview photogrammetry, as described in Chapter 5. This hybrid approach could be exploited in a systematic way, either within the dome by adding a structured light fringe scanner for example, or in external environments in the laboratory or on site by combining a mobile dome with a laser scanner. There are potentially many applications, both in cultural heritage and in industry, for such a system.

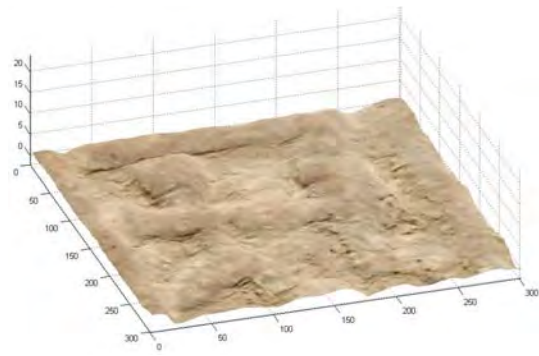


Figure 8.41: 3D reconstruction of a detail of the Egyptian clay funerary cone from the UCL Petrie Museum, UC37585.

8.3.9 Multispectral RTI

In the present dome configuration, objects are illuminated with white light and imaged with an RGB colour camera. Variations in both the illumination and the camera are possible, for example using a monochrome camera, such as the IDS uEye CMOS device used in photogrammetry, and filtering the illumination. This would enable multiple image sets to be acquired, each corresponding to a narrow range of wavelengths, including infrared up to 1100 nm. Such a system would require more complex image processing to correct for wavelength-dependent lateral chromatic aberration in the lens, but would enable characterisation of the angular reflectance distribution of gonio-chromatic materials.



Figure 8.42: uEye monochrome camera fitted with Fujinon 9 mm C-mount lens

8.4 Final Word

This whole research programme was really about the dome imaging system: What is it good for? What are its limitations? What information can be extracted from the image sets? How can it be improved? How can it be used effectively for cultural heritage? How do the results compare with 3D colour laser scanning? Through a series of exploratory technical investigations (metrology, visualisation, rendering, 3D, colour, specularly) a deeper understanding of the strengths and weaknesses of all aspects of the system has been achieved. Through a series of case studies, applying the methods to interesting problems in cultural heritage, the relevance of the system has been demonstrated. There is great potential here for further research and development.

References

- Agrawal A, Chellappa R and Raskar R (2005) An Algebraic Approach to Surface Reconstruction from Gradient Fields, *Proc Tenth IEEE Intl Conf on Computer Vision*, 1:174-181.
- Alman DH (1987) Directional colour measurement of metallic flake finishes, *Proc ISCC Williamsburg Conf on Appearance*, 53-56.
- Ashbaugh JC, Berns RS, Darling BA and Taplin LA (2009) Artist Material BRDF Database for Computer Graphics Rendering, *Proc 17th IS&T/SID Color Imaging Conf*, 62-68.
- Bagher M, Soler C and Holzschuch N (2012) Accurate fitting of measured reflectances using a shifted gamma micro-facet distribution, *Computer Graphics Forum*, 31(4):1509-1518.
- Barsky S and Petrou M (2003) The four-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows, *IEEE Trans on Pattern Analysis & Machine Intell* 25(10):1239-1252.
- Bartesaghi A, Sapiro G, Malzbender T and Gelb D (2005) Three-dimensional shape rendering from multiple images, *Graphical Models*, 67:332-346.
- Beckmann P and Spizzichino A (1963) *The Scattering of Electromagnetic Waves from Rough Surfaces*, Pergamon Press, New York.
- Berns RS (2000) *Billmeyer and Saltzman's Principles of Color Technology*, 3rd Ed, Wiley-Interscience, New York.
- Bernardini F and Rushmeier HE (2002) The 3D Model Acquisition Pipeline, *Computer Graphics Forum* 21(2):149-172.
- Bernardini F, Rushmeier H, Martin IM, Mittleman J and Taubin G (2002) Building a Digital Model of Michelangelo's Florentine Pietà, *IEEE Computer Graphics and Appl* 22(1):59-67.
- Bilbey D and Trusted M (2010) "The Question of Casts" – Collecting and Later Reassessment of the Cast Collections at South Kensington. In *Plaster Casts: Making, Collecting and Displaying from Classical Antiquity to the Present*, ed. R. Frederiksen and E. Marchandin, De Gruyter, Berlin: 465-484.
- Blais F, Taylor J, Cournoyer L, Picard M, Borgeat L, Godin G, Beraldin JA, Rioux M, Lahanier C and Mottin B (2007) More than a poplar plank: the shape and subtle colors of the masterpiece Mona Lisa by Leonardo, *Conf on Videometrics IX, Proc SPIE* 6491-06.
- Blinn, J (1977) Models of light reflection for computer synthesized pictures, *Proc ACM Siggraph*, 11(2):192-198.
- Borgeat L, Godin G, Massicotte P, Poirier G, Blais F and Beraldin JA (2007) Visualizing and Analyzing the Mona Lisa, *IEEE Computer Graphics and Appl* 27(6):60-68.
- Brauers J, Helling S and Aach T (2009) Multispectral Image Acquisition with Flash Light Sources, *J Imaging Science and Technology*, 53(3):031103-(10).
- Brill MH, Finlayson GD, Hubel PM and Thornton WA (1998) Prime Colours and Colour Imaging, *Proc 6th IS&T/SID Color Imaging Conference*, 33-42.
- Brownhill A, Brade R and Robson S (2009) Performance study of non-contact surface measurement technology for use in an experimental fusion device, *Proc 23rd IEEE/NPSS Symp on Fusion Engineering*, SOFE 2009, 1-4.
- Bryan P (2005) User Requirements for Metric Survey, in *Digital Heritage: Applying Digital Imaging to Cultural Heritage*, Ed MacDonald LW, Oxford: Elsevier, Ch 6, 149-173.
- Bryan P, Blake B and Bedford J (2009) *Metric Survey Specifications for English Heritage*, 2nd Ed, York, 126 pp, www.english-heritage.org.uk/publications/metric-survey-specification/

- Buchsbaum G and Gottschalk A (1984) Chromaticity coordinates of frequency-limited functions, *J Opt Soc Am A*, 1(8):885-887.
- Buchsbaum G and Gottschalk A (1983) Trichromacy, opponent colours coding and optimum colour information transmission in the retina, *Proc Royal Society B (London)* 220, 89-113.
- Byczkowski T and Lang J (2009) A Stereo-Based System with Inertial Navigation for Outdoor 3D Scanning, *Proc IEEE Conf on Computer and Robot Vision*, 221-228.
- Cao F, Guichard F and Hornung H (2008) Sensor spectral sensitivities, noise measurements and color sensitivity, *Proc SPIE*, 6817:1-12.
- Carvalho P, Santos A and Martins P (2004) Recovering imaging device sensitivities : A data-driven approach, *Proc IEEE Intl Conf on Image Processing (ICIP)*.
- Cheung V, Westland S and Thomson M (2004) Accurate Estimation of the Nonlinearity of Input/Output Response for Color Cameras, *Color Res and Appl*, 29(6):406-412.
- Cheung V, Westland S, Li C, Hardeberg, J and Connah D (2005) Characterization of trichromatic color cameras by using a new multispectral imaging technique, *J Opt Soc Am A* 22(7):1231-1240.
- Chandraker M, Agarwal S and Kriegman D (2007) Shadow Cuts: Photometric stereo with shadows, *Proc IEEE Conf on Computer Vision and Pattern Recognition (CVPR)*, 1-8.
- Christensen PH and Shapiro LG (1994) Three dimensional shape from colour photometric stereo, *Intl J of Computer Vision* 13(2):213-227.
- Christie JS (1979) Review of geometric attributes of appearance, *J Coatings Tech* 51:64-73.
- Clarke TA, Cooper MAR, Chen J and Robson S (1995) Automated Three Dimensional Measurement Using Multiple CCD Camera Views, *Photogrammetric Record*, 15(85):27-42.
- Clarke PJ (2006) *Surface Colour Measurements*, Good Practice Guide No.96, National Physical Laboratory. <http://publications.npl.co.uk/>
- Coleman EN and Jain R (1982) Obtaining 3-dimensional shape of textured and specular surfaces using four source photometry, *Computer Graphics and Image Processing* 18(4):309-328.
- Cook RL and Torrance KE (1982) A reflection model for computer graphics, *ACM Trans on Graphics*, 1(1):7-24.
- Cooper MAR and Robson S (1996) Theory of close range photogrammetry, in *Close Range Photogrammetry and Machine Vision*, Ed Atkinson KB, Whittles Pub, Caithness. Ch 2:9-51.
- Corsini M, Dellepiane M, Ponchio F and Scopigno R (2009) Image-to-Geometry Registration: a Mutual Information Method exploiting Illumination-related Geometric Properties, *Pac Gr*, 28(7).
- Csuti P and Schanda J (2009) A better description of metameric experience of LED clusters, *Proc CIE Conf on Light and Lighting*, Budapest, May 2009.
- Cyganek B, Siebert JP and Corporation E (2009) *An Introduction to 3D Computer Vision Techniques and Algorithms*, Wiley Online Library.
- Dana K, van Ginneken B, Nayar S and Koenderink J (1999) Reflectance and texture of real-world surfaces, *ACM Trans on Graphics*, 18(1):1-34. www.cs.columbia.edu/CAVE/software/curet/
- Davies J (2006) Microscopy and archival research: interpreting results within the context of historical records and traditional practice, *InFocus*, 3:32-45, Royal Microscopical Society, Oxford.
- Debevec PE and Malik J (1997) Recovering High Dynamic Range Radiance Maps from Photographs, *Proc ACM SIGGRAPH Conf on Computer Graphics* 31:369-378.
- Debevec PE, Hawkins T, Tchou C, Duiker H-P and Sarokin W (2000) Acquiring the reflectance field of a human face, *Proc ACM Siggraph* 24:145-156.

- Denard H (2009) The London Charter for the Computer-Based Visualisation of Cultural Heritage, Draft 2.1, February 2009. www.londoncharter.org
- De Roos H (2004) The Digital Sculpture Project: Applying 3D Scanning Techniques for the Morphological Comparison of Sculptures, *Computer and Information Science* 9(2) 85 pp.
- Devlin K, Chalmers A, Wilkie A and Purgathofer W (2002) *STAR report on tone reproduction and physically-based spectral rendering*. Eurographics.
- Drew M, Hel-Or Y, Malzbender T and Hajari N (2012) Robust estimation of surface properties and interpolation of shadow/specularity components, *Image and Vision Computing*, 30(4):317-331.
- Duffy SM, Bryan P, Earl G, Beale G, Pagi H and Kotouala E (2013) *Multi-Light Imaging for Heritage Applications*, English Heritage, York. www.english-heritage.org.uk/publications
- D'Zmura M (1991) Shading ambiguity: reflectance and illumination. *Computational models of visual processing*, 187-207.
- Earl G, Martinez K and Malzbender T (2010) Archaeological applications of polynomial texture mapping: analysis, conservation and representation, *J Archaeological Science*, 37(8):2040-2050.
- Edwards D, Boulos S, Johnson J, Shirley P, Ashikhmin M, Stark M and Wyman C (2006) The halfway vector disk for BRDF modeling. *ACM Trans on Graphics (TOG)* 25(1):1-18.
- Einarsson P, Hawkins T and Debevec P (2004) Photometric stereo for archeological inscriptions. *Proc ACM Siggraph Sketches* 81.
- Engelhardt K and Seitz P (1993) Optimum color filters for CCD digital cameras, *Appl Opt*, 32(16), 3015-3023.
- Epstein R, Yuille AL and Belhumeur PN (1996) Learning Object Representations from Lighting Variations, *Proc ECCV Workshop on Object Recognition*, 22pp.
- Fairman HS, Brill MH and Hemmendinger H (1997) How the CIE 1931 color-matching functions were derived from Wright-Guild data, *Color Res & Appl* 22(1):11-23.
- Farrell J, Okincha M and Parmar M (2008) Sensor Calibration and Simulation, *Conf on Digital Photography IV, Proc SPIE* 6817:1-9.
- Fattal R, Lischinski D and Werman M (2002) Gradient Domain High Dynamic Range Compression, *ACM Trans on Graphics* 21(3):249-256.
- Faugeras OD and Hebert M (1986) The representation, recognition and locating of 3-D objects, *Intl J Robotics Res* 5(3):27-52.
- Ferwerda JA, Pellacini F and Greenberg DP (2001) A psychophysically-based model of surface gloss perception, *Conf on Human Vision and Electronic Imaging, Proc SPIE* 4299:291-301.
- Filip J and Vávra R (2014) Template-Based Sampling of Anisotropic BRDFs. *Computer Graphics Forum* 33(7):91-99.
- Finlayson GD, Hordley S and Hubel PM (1998) Recovering Device Sensitivities with Quadratic Programming, *Proc 6th IS&T/SID Color Imaging Conf*, 90-95.
- Forstner W and Gulch E (1987) A Fast Operator for Detection and Precise Location of Distinct Points, Corners, and Centres of circular Features, *Proc ISPRS Conf on 'Fast Processing of Photogrammetric Data'*, 281-305.
- Foster DH, Nascimento SMC and Amano K (2004) Information limits on neural identification of coloured surfaces in natural scenes, *Vis Neuroscience* 21:331-336. Images downloaded from: http://personalpages.manchester.ac.uk/staff/david.foster/Hyperspectral_images_of_natural_scenes_04.html
- Frankot RT and Chellappa R (1988) A method for enforcing integrability in shape from shading algorithms, *IEEE Trans on Pattern Analysis & Machine Intell* 10(4) 439-451.

- Fraser CS (2001) Automated Vision Metrology: A Mature Technology for Industrial Inspection and Engineering Surveys, *Australian Surveyor*, 46(1):5-11.
- Freeth T, Bitsakis Y, Moussas X, Seiradakis JH, et al (2006) Decoding the ancient Greek astronomical calculator known as the Antikythera Mechanism, *Nature* 444(7119):587-591.
- Fua P, Gruen A, Plänkner R, D'Apuzzo N and Thalmann D (1998) Human Body Modeling and Motion Analysis from Video Sequences. *ISPRS Archives of Photogrammetry & Remote Sensing* 32:866-873.
- Gardner JL (2006) *Uncertainties in surface colour measurements*, Good Practice Guide No. 95, National Physical Laboratory, London. <http://publications.npl.co.uk/>
- Gautron P, Krivanek J, Pattanaik SN and Bouatouch K (2004). A Novel Hemispherical Basis for Accurate and Efficient Rendering. *Rendering Techniques*, 321-330.
- Germer TA and Asmael CC (1997) A goniometric optical scatter instrument for bidirectional reflectance distribution function measurements with out-of-plane and polarimetry capabilities, *Conf on Scattering and Surface Roughness, Proc SPIE* 3141:220-231.
- Ghosh A, Heidrich W, Achutha S and O'Toole M (2008) A Basis Illumination Approach to BRDF Measurement, *Int. J Computational Visualisation*, Online July 19, 2008.
- Giacometti A, Campagnolo A, MacDonald LW, Mahony S, Terras M, Robson S, Weyrich T and Gibson A (2012) Documenting Parchment Degradation via Multispectral Imaging, *Proc. BCS Conf. on Electronic Imaging and the Visual Arts (EVA)*, London.
- Gouraud H (1971) Continuous Shading of Curved Surfaces, *IEEE Trans Computers*, C-20(6):623-629.
- Granshaw SE (1980) Bundle adjustment methods in engineering photogrammetry, *Photogrammetric Record* 10(56):181-207.
- Grard CL and Trussell HJ (1995) Interpolation and extrapolation for accurate color measurement, *Proc 1995 IEEE Intl Conf on Image Processing*, Washington DC, 350-355.
- Greenberg DP, Torrance KE, Shirley P, Arvo J, Lafortune E, Ferwerda JA and Foo SC (1997) A framework for realistic image synthesis. *Proc ACM Siggraph*, 24:477-494.
- Gregory RL (1973) 'The Confounded Eye' in Gregory RL and Gombrich EH (eds.) *Illusion in Nature and Art*, New York: Charles Scribner's Sons, 49-95.
- Grossberg MD and Nayar SK (2004) Modeling the Space of Camera Response Functions, *IEEE Trans on Pattern Analysis & Machine Intell* 26(10):1272-1282.
- Gruen A and Baltsavias M (1988) Geometrically constrained multiphoto matching, *Photogrammetric Engineering and Remote Sensing*, 54(5):633-641.
- Guild J (1931) The colorimetric properties of the spectrum, *Philos Trans Roy Soc A* 230:149-187.
- Gunawardane P, Wang O, Scher S, Rickards I, Davis J and Malzbender T (2009) Optimized image sampling for view and light interpolation. *Proc 10th Intl Conf on Virtual Reality, Archaeology and Cultural Heritage (VAST)*, Eurographics, 10:93-100.
- Gwinnett AJ and Gorelick L (1993) Beads, Scarabs, and Amulets: Methods of Manufacture in Ancient Egypt. *J American Research Center in Egypt*, 125-132.
- Hammer Ø, Bengtson S, Malzbender T and Gelb D (2002) Imaging Fossils using Reflectance Transformation and Interactive Manipulation of Virtual Light Sources, *Palaeontologia Electronica* 5(4):1-9.
- Hanson AR (2005) *Measurement of Gloss*, Good Practice Guide No. 94, National Physical Laboratory, London. <http://publications.npl.co.uk/>
- Harrison VGW (1945) *The Definition and Measurement of Gloss*, The Printing and Allied Trades Research Association, London.

- Harvey E, Cappo M, Shortis M, Robson S, Buchanan J and Speare P (2003) The accuracy and precision of underwater measurements of length and maximum body depth of southern blue-fin tuna (*Thunnus maccoyii*) with a stereo-video camera system. *Fisheries Res*, 63(3):315-326.
- Hawkins DS and Green P (2008) Spectral characterisation of a digital still camera through a single integrating exposure. *Proc IS&T Eur Conf on Colour in Graphics, Imaging and Vision (CGIV)* 477-480.
- Hel-Or Y, Malzbender T and Gelb D (2003) *Synthesis of Reflectance Function Textures from Examples*, Technical Report HPL-2003-16 (R.1), HP Laboratories, Palo Alto.
- Helmholtz H von (1867) *Handbook of Physiological Optics*, tr Southall JP (1963) Dover Publications, New York.
- Hernández-Andrés J, Lee RL and Romero J (1999) Calculating correlated colour temperatures across the entire gamut of daylight and skylight chromaticities, *Appl Opt*, 38(27):5703-5709.
- Herzog PG, Knipp D, Stiebig H and König F (1999) Colorimetric characterization of novel multiple-channel sensors for imaging and metrology, *J Electronic Imaging*, 8(4) 342-353.
- Hong G, Luo MR and Rhodes PA (2001) A Study of Digital Camera Colorimetric Characterisation Based on Polynomial Modelling, *Color Res Appl*, 26(1), 76-85.
- Horn BKP (1970) *Shape from Shading: A method for obtaining the shape of a smooth opaque object from one view*, AI Technical Report 232, Massachusetts Institute of Technology, Cambridge MA.
- Horn BKP (1990) Height and gradient from shading, *Intl J of Computer Vision*, 5(1):37-75.
- Huang H, Zhang L, Samaras D, Shen L, Zhang R, Makedon F and Pearlman J (2006) Hemispherical harmonic surface description and applications to medical image analysis. *Proc IEEE Third Intl Symp on 3D Data Processing, Visualization and Transmission*, 381-388.
- Hung PC (1993) Colorimetric calibration in electronic imaging devices using a lookup table model and interpolations, *J Electronic Imaging* 2(1), 53-61.
- Hurvich LM and Jameson D (1955) Some Quantitative Aspects of an Opponent-Colors Theory. II. Brightness, Saturation, and Hue in Normal and Dichromatic Vision, *Vision Research* 45(8):602-616.
- Hunt RWG (1982) A model of colour vision for predicting colour appearance, *Color Res and Appl* 7(2):95-112.
- Hunt RWG (1998) *Measuring Colour*, 3rd Ed, Fountain Press, Kingston-upon-Thames.
- Hunter RS and Judd DB (1939) Development of a method of classifying paints according to gloss, *ASTM Bulletin*, 11 March, 118.
- Hunter RS and Harold RW (1987) *The Measurement of Appearance*, 2nd Ed, John Wiley, New York.
- Ikeuchi K and Horn BKP (1981) Numerical Shape from Shading and Occluding Boundaries, *Artificial Intell* 17(1-3):141-184.
- Ji W, Pointer MR, Luo MR and Dakin J (2005) Gloss as an aspect of the measurement of appearance, *J Opt Soc Am A* 23(1):22-33.
- Johnson AJ (1996) Methods for Characterising Colour Scanners and Digital Cameras, *Displays* 16(4):183-190.
- Johnson AJ (1996) *Colour Management in Graphic Arts and Publishing*, PIRA Press, Leatherhead.
- Johnston RM and Stanziola R (1969) Angular color measurement on automotive materials, *Proc Intl Automation Engineering Congress*, Pub. 690241, Soc Automotive Engineers, Warrendale, PA.

- Jolliffe IT (2002) *Principal Component Analysis*, 2nd Ed, Springer Series in Stats, XXIX, New York.
- Judd DB, MacAdam DL and Wyszecki G (1964) Spectral distribution of typical daylight as a function of correlated color temperature, *J Opt Soc Am* 54(8):1031-1040.
- Judd DB (1968) The 1964 CIE Supplementary Observer Applied to the Colorimetry of Rutile and Anatase Forms of Titanium Dioxide, *J Opt Soc Am* 58:1638-1648.
- Julia C, Sappa A, Lumbreras F, Serrat J and López A (2008) Photometric stereo through an adapted alternation approach, *Proc 15th IEEE Intl Conf on Image Processing (ICIP)*, 1500-1503.
- Kang HR (1997) *Color Technology for Electronic Imaging Devices*, SPIE Press, Bellingham WA.
- Kato N, Nakabayashi N, Ito M and Ohno S (1998) Effect of ambient light on the appearance of soft copy images, *J Electronic Imaging* 7:794-806.
- Kautz J and McCool MD (1999) Interactive Rendering with Arbitrary BRDFs using Separable Approximations, *Proc Eurographics Rendering Workshop*.
- Kay G and Caelli T (1995) Estimating the Parameters of an Illumination Model using Photometric Stereo, *Graphical Models and Image Processing*, 57(5):365-388.
- Kim, MH and Kautz, J (2008) Characterization for high dynamic range imaging. *Computer Graphics Forum* 27(2):691-697.
- Kochubey SM and Yatsenko MA (2004) A monitoring system for agricultural crops on chlorophyll basis, *Proc Conf on Remote Sensing for Agriculture, Ecosystems and Hydrology V*, SPIE Vol 5232, 92-99.
- Kovesi P (2005) Shapelets Correlated with Surface Normals Produce Surfaces, *Proc Tenth IEEE Intl Conf on Computer Vision*, 2:994-1001.
- Kuang J, Johnson GM and Fairchild MD (2007) iCAM06: A refined image appearance model for HDR image rendering. *J Visual Communication*, 18(5):406-414.
- Kumar P and MacDonald LW (2006) Measuring gloss by digital photography, *Conf on Digital Photography and Applications, Proc SPIE* 6058.
- Lafortune EP, Foo S, Torrance KE and Greenberg D P (1997) Non-linear approximation of reflectance functions, *Proc ACM Siggraph*, 24:117-126.
- Leloup FB, Forment S, Dutré P, Pointer MR and Hanselaer P (2008) Design of an instrument for measuring spectral bidirectional scatter distribution function, *Appl Opt* 47(29):5454-5467.
- Levoy M and Hanrahan P (1996) Light Field Rendering, *Proc ACM Siggraph* 23:31-42.
- Levoy M, Pulli K, Curless B, et al (2000) The Digital Michelangelo project: 3D Scanning of Large Statues, *Proc ACM Siggraph* 27:131-144. www.graphics.stanford.edu/projects/mich/
- Li H, Foo SC, Torrance KE and Westin SH (2005) Automated three-axis gonioreflectometer for computer graphics applications, *Conf on Advanced Characterization Techniques for Optics, Semiconductors and Nanotechnologies II, Proc SPIE* 5878.
- Luhmann T (2010) Close range photogrammetry for industrial applications, *ISPRS J Photogrammetry and Remote Sensing*, 65(6):558-569.
- Luhmann T, Robson S, Kyle S and Harley I (2008) *Close Range Photogrammetry: Principles, Techniques and Applications*, John Wiley, Chichester.
- MacDonald LW (1996) Developments in colour management systems, *Displays* 16(4):203-211.
- MacDonald LW, Westland S and Shaw J (1999) Colour Image Reproduction: Spectral vs Spatial, *Proc First Intl Symp on Multispectral Imaging and Color Reproduction for Digital Archives*, Chiba, Japan.

- MacDonald LW and Ji W (2002) Colour Characterisation of a High-Resolution Digital Camera, *Proc First Eur. Conf on Colour in Graphics, Imaging and Vision (CGIV)*, Poitiers, France, 433-437.
- MacDonald LW, Ed (2006) *Digital Heritage: Applying Digital Imaging to Cultural Heritage*, Elsevier, Oxford.
- MacDonald LW (2010a) The Limits of Resolution, *Proc BCS Conf on Electronic Visualisation and the Arts (EVA)*, London.
- MacDonald LW (2010b) *3D Imaging of Cultural Objects*, MPhil/PhD Transfer Report, UCL.
- MacDonald LW and Robson S (2010) Polynomial texture mapping and 3D representation, *Proc ISPRS Commission V Symp 'Close Range Image Measurement Techniques'*, Newcastle.
- Makhotkin, OA (1996) Analysis of radiative transfer between surfaces by hemispherical harmonics. *J Quantitative Spectroscopy and Radiative Transfer*, 56(6):869-879.
- Maloney LT (1986) Evaluation of linear models of surface spectral reflectance with small numbers of parameters, *J Opt Soc Am A*, 3(10):1673-1683.
- Malzbender T, Gelb D and Wolters H (2001) Polynomial Texture Maps, *Proc ACM Siggraph* 28:519-528.
- Mann S and Picard R (1995) Being 'Undigital' with Digital Cameras: Extending Dynamic Range by Combining Differently Exposed Pictures, *Proc IS&T 46th Conf*, 422-428.
- Marschner SR, Westin SH, Lafortune EPF and Torrance KE (2000) Image-based bidirectional reflectance distribution function measurement, *Appl Opt-OT*, 39(16):2592-2600.
- Martinez K, Cupitt J, Saunders D and Pillay R (2002) Ten Years of Art Imaging Research, *Proc IEEE*, 90(1):28-41.
- Matusik W, Pfister H, Brand M and McMillan L (2003) Efficient isotropic BRDF measurement, *Proc 14th Eurographics Workshop on Rendering*, 241-247.
- McCamy CS, Marcus H and Davidson JG (1976) A Color-Rendition Chart, *J Applied Photographic Engineering*, 2(3):95-99.
- McCamy CS (1996) Observation and measurement of the appearance of metallic materials: I. Macro appearance, *Color Res & Appl*, 21(4):292-304.
- Mikula M, Čeppan M and Vaško K (2003) Gloss and goniocolorimetry of printed materials, *Color Res & Appl*, 28(5):335-342.
- Mitsunaga T and Nayar SK (1999) Radiometric Self Calibration, *Proc IEEE Conf on Computer Vision and Pattern Recognition*, 2:374-380.
- Mollon JD, Estévez O and Cavanus CR (1993) The two subsystems of colour vision and their roles in wavelength discrimination, Chap 11 in Blakemore C, *Vision: Coding and Efficiency*, CUP, Cambridge, 119-131.
- Moroney N, Fairchild MD, Hunt RWG, Li C, Luo MR and Newman T (2002) The CIECAM02 Color appearance model, *Proc Tenth IS&T/SID Color Imaging Conf*, 23-27.
- Mohammadi M, Nezamabadi M, Berns RS and Taplin LA (2004). Spectral imaging target development based on hierarchical cluster analysis. *Proc 12th IS&T Color Imaging Conf*, 59-64.
- Mudge M, Voutaz JP, Schroer C and Lum M (2005) Reflection Transformation Imaging and Virtual Representations of Coins from the Hospice of the Grand St. Bernard. *Proc 6th Intl Symp on Virtual Reality, Archaeology and Cultural Heritage (VAST)*, Eurographics, 5:29-39.
- Mudge M, Malzbender T, Schroer C and Lum M (2006) New Reflection Transformation Imaging Methods for Rock Art and Multiple-Viewpoint Display, *Proc 7th Intl Symp on Virtual Reality, Archaeology and Cultural Heritage (VAST)*, Cyprus, 6:195-202.

- Mudge M, Malzbender T, Chalmers A, Scopigno R, Davis J, Wang O et al (2008) Image-based empirical information acquisition, scientific reliability, and long-term digital preservation for the natural sciences and cultural heritage. *Eurographics Tutorials*.
- Mudge M and Schroer C (2014) *Reflectance Transformation Imaging: Guide To Highlight Image Processing*, Version 1.4 for RTIBuilder v2.0.2beta. <http://CulturalHeritageImaging.org/Learn/>
- Mueller P, Vereenoooghe T, Vergauwen M, Van Gool L and Waelkens M (2004) Photo-realistic and detailed 3D modeling: the Antonine nymphæum at Sagalassos, *Proc XXXII Conf on Computer Applications and Quantitative Methods to Archaeology (CAA2004)*, Prato, Italy.
- Nascimento SMC, Ferreira FP and Foster DH (2002) Statistics of spatial cone-excitation ratios in natural scenes, *J Opt Soc Am A*, 19(8), 1484-90.
- Nehab D, Rusinkiewicz S, Davis J and Ramamoorthi (2005) Efficiently combining positions and normals for precise 3D geometry. *ACM Trans on Graphics (TOG)* 24(3):536-543.
- Newton, I (1704) *Opticks*, Dover Edition (1979), Book One, Part II, Prop II, 124-125.
- Ng HS, Wu TP and Tang CK (2007) Surface-from-Gradients with Incomplete Data for Single View Modeling, *Proc IEEE 11th Intl Conf on Computer Vision (ICCV)* 1-8.
- Ng HS, Wu TP and Tang CK (2010) Surface-from-Gradients without Discrete Integrability Enforcement: a Gaussian Kernel Approach, *IEEE Trans on Pattern Analysis & Machine Intell* 32(11):2085-2099.
- Ngan A, Durand F and Matusik W (2005) Experimental Analysis of BRDF Models. *Rendering Techniques*, 16th.
- Nicodemus F (1965) Directional reflectance and emissivity of an opaque surface, *Appl Opt* 4(7):767-775.
- Oskoui P and Pirotta E (1999) Determination of Adapted White Points for Various Viewing Environments, *Proc 7th IS&T/SID Color Imaging Conf*, 101-105.
- Padfield J, Saunders D and Malzbender T (2005) Polynomial texture mapping: a new tool for examining the surface of paintings, *Proc ICOM 14th Triennial Meeting*, The Hague, Vol 1:504-510.
- Pakkinen JPS, Hallikainen J and Jaaskelainen T (1989) Characteristic spectra of Munsell colors, *J Opt Soc Am A* 6(2):318-322.
- Peggs GN, Maropoulos PG, Hughes EB, Forbes AB, Robson S, Ziebart M and Muralikrishnan B (2009) Recent developments in large-scale dimensional metrology. *Proc. Inst. Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 223(6):571-595.
- Phong, B-T (1975) Illumination for Computer Generated Pictures, *Comm ACM* 18(6):311-317.
- Plata C, Nieves JL, Valero EM and Romero J (2009) Trichromatic red-green-blue camera used for recovery of albedo and reflectance of rough-textured surfaces under different illumination conditions, *Appl Opt* 48(19):3643-3653.
- Pointer MR, Attridge GG and Jacobson RE (2001) Practical camera characterisation for colour measurement, *Imaging Science J* 49:63-80.
- Pointer MR, Barnes NJ, Clarke PJ, Shaw MJ (2005) A new goniospectrophotometer for measuring gonio-apparent materials, *Coloration Technology*, 121:96-103.
- Pointer MR (2006) *Technical Report: A Framework for the Measurement of Visual Appearance*, CIE Publication 175:2006, Vienna.
- Pottler K, Lüpfert E, Johnston GH and Shortis MR (2004) Photogrammetry: A powerful tool for geometric analysis of solar concentrators and their components, *Proc ASME 2004 International Solar Energy Conf*, 719-726. American Society of Mechanical Engineers.
- Pugsley PC (1975) *Colour Correcting Image Reproducing Methods and Apparatus*, US Pat 3,893,166.

- Quijada MA, Marx CT, Arendt RG and Moseley SH (2004). Angle-of-incidence effects in the spectral performance of the infrared array camera of the Spitzer Space Telescope. *Astronomical Telescopes and Instrumentation*, 244-252.
- Ramamoorthi R and Hanrahan P (2001) An efficient representation for irradiance environment maps. *Proc ACM Siggraph* 28:497-500.
- Reichenbach SE, Park SK and Narayanswamy R (1991) Characterizing digital image acquisition devices, *Optical Engineering* 30(2):170-177.
- Rich D (2002) Instruments and Methods for Colour Measurement, in Green PG and MacDonald LW (Eds) *Colour Engineering: Achieving Device Independent Colour*, John Wiley, Chichester, Ch 2:19-48.
- Robertson AR (1968) Computation of correlated color temperature and distribution temperature, *J Opt Soc Am* 58:1528-1535.
- Robson S, Bucklow S, Woodhouse N and Papadaki H (2004) Periodic photogrammetric monitoring and surface reconstruction of a historic wood panel painting for restoration purposes, *Proc ISPRS Archives of Photogrammetry and Remote Sensing*. Zurich 30(5W1):379-384.
- Robson S, Brewer A, Cooper MAR, Clarke TA, Chen J, Setan HB and Short T (1995) Seeing the Wood from the Trees - An Example of Optimised Digital Photogrammetric Deformation Detection, *Proc ISPRS Intercom Workshop "From Pixels to Sequences"*, Zurich, IAPRS, 30, Pt 5W1.
- Robson S and Setan HB (1996) The Dynamic Digital Photogrammetric Measurement and Visualisation of a 21m Wind Turbine Rotor Blade undergoing Structural Analysis, *ISPRS Archives of Photogrammetry and Remote Sensing* 31(B5):493-499.
- Rushmeier H and Bernardini F (1999) Computing Consistent Normals and Colors from Photometric Data, *Proc IEEE Second Intl Conf on 3-D Imaging and Modeling (3DIM)*.
- Rusinkiewicz SM (1998) A new change of variables for efficient BRDF representation, *Proc 9th Eurographics Workshop on Rendering Techniques*, Vienna.
- Rössler G (1990) Multigeometry color measurement of effect surfaces, *Die Farbe*, 37:111-121.
- Sadlo F, Weyrich T, Peikert R and Gross M (2005) A Practical Structured Light Acquisition System for Point-Based Geometry and Texture, *Proc Eurographics Symp on Point-Based Graphics*, 89-98.
- Sansoni G, Corini S, Lazzari S, Rodella R and Docchio F (1997) Three dimensional imaging based on Gray-code light projection, *Appl Opt* 36(19):4463-4472.
- Sato Y, Wheeler M and Ikeuchi K (1997) Object Shape and Reflectance Modeling from Observation, *Proc ACM Siggraph* 23:379-387.
- Saunders D (1998) High-Quality Imaging at the National Gallery: Origins, Implementation and Applications, *Computers and the Humanities*, 31(3):153-167.
- Saunders D and Cupitt J (1993) Image Processing at the National Gallery: The VASARI Project, *National Gallery Technical Bulletin*, 14:72-85.
- Schlüns K and Witting O (1993) Photometric Stereo for Non-Lambertian Surfaces Using Color Information, *Proc Fifth Intl Conf Computer Analysis of Images and Patterns*, 444-454.
- Schenk T (2005) *Introduction to Photogrammetry*. Dept of Civil and Environmental Engineering and Geodetic Science, Ohio State University, 79-95.
- Sebe N, Lew MS, Huijsmans DP (2000) Towards improved ranking metrics, *IEEE Trans on Pattern Analysis & Machine Intell*, 22(10):1132-1143.
- Shafer SA (1985) Using Color to Separate Reflection Components, *Color Res & Appl* 10(4):210-218.

- Sharpe LT, Stockman A, Jagla W and Jägle H (2005) A luminous efficiency function, $V^*(\lambda)$, for daylight adaptation, *J of Vision* 5(11) 948-968. <http://jovvision.org/5/11/3/article.aspx>
- Shortis MR, Clarke TA and Robson S (1995) Practical testing of the precision and accuracy of target image centering algorithms, *Proc SPIE Conf on Videometrics IV*, Vol. 2598, 65.
- Shortis MR and Harvey ES (1998) Design and calibration of an underwater stereo-video system for the monitoring of marine fauna populations. *ISPRS Archives of Photogrammetry and Remote Sensing* 32(5):792-799.
- Shortis MR, Ogleby CL, Robson S, Karalis EM and Beyer HA (2001) Calibration modelling and stability testing for the Kodak DC200 series digital still camera, *Proc SPIE Conf on Videometrics and Optical Methods for 3D Shape Measurement*, *Proc SPIE* 4309:148-153.
- Shortis MR (2001) *What is VMS? Documentation in VMS Help File*, Geomsoft.
- Sloan PP, Kautz J and Snyder J (2002) Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. *ACM Trans on Graphics (TOG)* 21(3):527-536.
- Stokes M, Anderson M, Chandrasekar S and Motta R (1996) A Standard Default Color Space for the Internet - sRGB Version 1.10. <http://www.w3.org/Graphics/Color/sRGB.html>
- Sun Y (2002) A physically-based reflection model for glossy appearance, *Proc 10th IS&T/SID Color Imaging Conf*, 172-177.
- Taylor J, Beraldin J-A, Godin G, Cournoyer L, Baribeau R, Blais F, Rioux M and Domey J (2003) NRC 3D imaging technology for museum and heritage applications, *J Visualization and Computer Animation*, 14(3):121-138.
- Thomson M and Westland S (2001) Colour-Imager Characterization by Parametric Fitting of Sensor Responses, *Color Res & Appl*, 26(6):442-449.
- Thornton WA (1971) Luminosity and color-rendering capability of white light, *J Opt Soc Am* 61:191-194.
- Thornton WA (1972) Three-color visual response, *J Opt Soc Am*, 62:457-459.
- Thornton WA (1999) Spectral sensitivities of the normal human visual system, color matching functions and their principles, and how and why the two sets should coincide, *Color Res and Appl*, 24(2):139-156.
- Thornton WA (2000) Suggested Optimum Primaries and Gamut in Color Imaging, *Color Res and Appl*, 25(2):148-150.
- Tominaga S, Ujike H and Horiuchi T (2010) Surface reconstruction of oil paintings for digital archiving. *Proc IEEE Southwest Symposium on Image Analysis & Interpretation (SSIAI)*, 173-176.
- Tommaselli AMG and Reiss M (2005) A Photogrammetric Method for Single Image Orientation and Measurement, *Photogrammetric Engineering & Remote Sensing* 71:727-732.
- Torrance KE and Sparrow EM (1967) Theory for off-specular reflection from roughened surfaces, *J Opt Soc Am*, 57(9):1105-1114.
- Trezza TA and Krochta JM, Specular reflection, gloss, roughness and surface heterogeneity of biopolymer coatings, *J Applied Polymer Science* 79(12):2221-2229.
- Triantaphillidou T, Jacobson RE and Attridge GG (2002) A case study in the digitisation of a photographic collection, *Imaging Science J*, 50:97-115.
- Trowbridge T and Reitz K (1975) Average irregularity representation of a roughened surface for ray reflection, *J Opt Soc Am* 65(5):531-536.
- Van Haasteren AJP and Frankena HJ (1994) Real time displacement measurement using a multi-camera phase stepping speckle interferometer, *Appl Opt*, 33(19):4137-4142.
- Verbiest F, Willems G and Van Gool L (2006) Image-based rendering for photo-realistic visualization. *Virtual and Physical Prototyping*, 1(1):19-30.

- Vrhel MJ, Gershon R and Iwan LS (1994) Measurement and analysis of object reflectance spectra. *Color Research & Application* 19(1):4-9.
- Vos JJ (1978) Colorimetric and photometric properties of a 2-deg fundamental observer, *Color Research & Application* 3(2):125-128.
- Ward GJ (1992) Measuring and modeling anisotropic reflection, *Proc ACM Siggraph* 16(2):265-272.
- Ward GJ (1994) The RADIANCE Lighting Simulation and Rendering System, *Proc ACM Siggraph*, 18:459-472.
- Ward R (ed) (2014) *Court and Craft: A Masterpiece of Northern Iraq*, Exhibition Catalogue, The Courtauld Gallery, Paul Holberton Publishing, London, 95-97.
- Watson R and Raven P (2001) Comparison of measured BRDF data with parameterized reflectance models, *Proc SPIE Conf on Targets and Backgrounds VII*, 4370:159-168.
- Westin SH, Arvo JR and Torrance KE (1992) Predicting reflectance functions from complex surfaces *Proc ACM Siggraph* 26(2):255-264.
- Weyrich T, Deng J, Barnes C, Rusinkiewicz S and Finkelstein A (2007) Digital Bas-Relief from 3D Scenes, *Proc ACM Trans Graphics*, 26(3):32-38.
- Weyrich T, Lawrence J, Hendrik P, Lensch A, Rusinkiewicz S and Zickler T (2008) *Principles of Appearance Acquisition and Representation*, Class Notes, ACM SIGGRAPH.
- Winter J, Maines CA and Dickerson JH (1996) Applications of Image Processing and Analysis in Research on Works of Art, in *Imaging the Past*, British Museum Occasional Paper 114, 35-48.
- Wolf PR and Dewitt BA (2000) *Elements of Photogrammetry with Applications in GIS*, 3rd Ed, McGraw-Hill.
- Woodham RJ (1980) Photometric method for determining surface orientation from multiple images, *Opt Eng* 19(1):139-144.
- Wright WD (1929) A re-determination of the trichromatic coefficients of the spectral colours, *Trans Opt Soc* 30:141-164.
- Wu W and Allebach JP (2000) Imaging Colorimetry Using a Digital Camera, *J Imaging Sci and Tech* 44(4):267-279.
- Wu TP and Tang CK (2010) Photometric stereo via expectation maximization, *IEEE Trans on Pattern Analysis & Machine Intell* 32(3):546-560.
- Wyble DR and Rich DC (2007a) Evaluation of Methods for Verifying the Performance of Color-Measuring Instruments. Part I: Repeatability, *Color Res & Appl* 32(3):166-175.
- Wyble DR and Rich DC (2007b) Evaluation of Methods for Verifying the Performance of Color-Measuring Instruments. Part II: Inter-Instrument Reproducibility, *Color Res & Appl* 32(3):176-194.
- Zhang M and Drew MS (2014) Efficient robust image interpolation and surface properties using polynomial texture mapping. *EURASIP J Image and Video Processing*, 1:1-19.
- Zhang R, Tsai P-S, Cryer JE and Shah M (1999) Shape from shading: A survey, *IEEE Trans on Pattern Analysis & Machine Intell*, 21(8):690-706.

Standards

ASTM (1997) E430, *Standard test method for measurement of gloss of high-gloss surfaces by goniphotometry*, American Society for Testing and Materials, West Conshohocken, PA.

ASTM (1999) D523, *Standard test method for specular gloss*, American Society for Testing and Materials, West Conshohocken, PA.

ASTM (2002) E284, *Standard terminology of appearance*, American Society for Testing and Materials, West Conshohocken, PA.

ASTM (2004) E1767, *Standard Practice for Specifying the Geometry of Observations and Measurements to Characterize the Appearance of Materials*, American Society for Testing and Materials, West Conshohocken, PA.

CIE Publication 13.3 (2006), *Method of specifying and measuring colour rendering properties of light sources*, 3rd Ed, Commission Internationale de l'Éclairage, Vienna.

CIE 15.3:2004, *Colorimetry*, 3rd Ed, Commission Internationale de l'Éclairage, Vienna.

ICC.1:2004-10 *Image technology colour management — Architecture, profile format, and data structure*, International Color Consortium, www.color.org/ICC1v42_2006-05.pdf

IEC 61966-9:2004 *Multimedia systems and equipment — Colour measurement and management — Part 9: Digital cameras*, European Committee for Electrotechnical Standardization, Brussels (CENELEC),

ISO 3664:2000 *Viewing conditions — graphic technology and photography*, International Organization for Standardization, Geneva.

ISO 7589:2002 *Photography — Illuminants for sensitometry — Specifications for daylight, incandescent tungsten and printer*, International Organization for Standardization, Geneva.

ISO 9845-1:1992 *Solar energy — Reference solar spectral irradiance at the ground at different receiving conditions — Part 1: Direct normal and hemispherical solar irradiance for air mass 1,5*, International Organization for Standardization, Geneva.

ISO 11664-1:2007, *Colorimetry — Part 1: CIE standard colorimetric observers*, International Organization for Standardization, Geneva.

ISO 12233:2000 *Photography — Electronic still-picture cameras — Resolution measurements*, Geneva: International Organization for Standardization.

ISO 12641:1997 *Graphic technology — Prepress digital data exchange — Colour targets for input scanner calibration*, International Organization for Standardization, Geneva.

ISO 14524:1999 *Photography — Electronic still-picture cameras — Methods for measuring opto-electronic conversion functions (OECFs)*, International Organization for Standardization, Geneva.

ISO 15469:2004, *Spatial distribution of daylight — CIE standard general sky*, International Organization for Standardization, Geneva.

ISO 15739:2003 *Photography — Electronic still-picture imaging — Noise measurements*, International Organization for Standardization, Geneva.

ISO 16067-1:2003 *Photography — Spatial resolution measurements of electronic scanners for photographic images — Part 1: Scanners for reflective media*, International Organization for Standardization, Geneva.

ISO 17321-1:2006 *Graphic technology and photography — Colour characterisation of digital still cameras (DSCs) — Part 1: Stimuli, metrology and test procedures*, International Organization for Standardization, Geneva.

Glossary of Abbreviations

BRDF	Bidirectional Reflectance Distribution Function
BTF	Bidirectional Texture Function
BM	British Museum
CCD	Charge-Coupled Device
CHI	Cultural Heritage Imaging
CIE	Commission Internationale de l'Éclairage
CMM	Coordinate Measuring Machine
FFT	Fast Fourier Transform
HDR	High Dynamic Range
HSH	Hemi-Spherical Harmonics
ISO	International Standards Organisation
LED	Light Emitting Diode
LMA	London Metropolitan Archives
LMS	Least Median of Squares
LUT	Look-Up Table
MTF	Modulation Transfer Function
NEF	Nikon Electronic Format
NCS	Natural Colour System
NPL	National Physical Laboratory
OECF	Opto-Electronic Conversion Function
PCA	Principal Components Analysis
PTM	Polynomial Texture Map
RAM	Random Access Memory
RGB	Red, Green, Blue
RMS	Root Mean Square
RTI	Reflectance Transform Imaging
SPD	Spectral Power Distribution
SVD	Singular Value Decomposition
TIFF	Tagged Image File Format
UCL	University College London
UCS	Uniform Colour Space
VMS	Vision Measurement System
XRF	X-Ray Fluorescence

Appendix I – Procedure for image capture from dome

The sequence of operations for capturing a set of images in the dome is as follows:

1. Select the lens (17-55 mm zoom or 105 mm macro or 200 mm macro) according to the object size, and engage onto the bayonet mount on the Nikon D200 camera body;
2. Check on the camera menu that the USB mode is set to 'P' (program), not 'M' (memory). Check that focus mode switches on both camera body and lens are set to AF;
3. Mount the camera and lens on the plate above the dome, attached by 1/4-inch screw, and ensure that lens front is aligned with the circular hole at top of dome;
4. Connect ribbon cable from the flash control box to dome flash board #1, switch on power and confirm that focus lights can be switched on;
5. Connect three cables to the camera: mains power supply, USB from camera to computer, and flash trigger from camera to flash control box;
6. Start *Nikon Camera Control Pro II* software on the host PC and check that the camera is connected and operational.
7. Lift the dome by chrome handle at front and use the black steel bar to prop it open, like the bonnet of a car. Place object at the centre of the baseboard on black cardboard;
8. Close the dome and switch on LED focus lights (flash boards in Tier 4). Open one of the portholes on the left side of the dome, illuminate the object with a torch, and half-press the camera shutter button to autofocus the camera on the surface of the object. The focus should not be changed again throughout the image sequence;
9. Via the Nikon GUI on the computer, set exposure time to 1/60 s, aperture to *f*/8 and speed to ISO100. Select RAW (or NEF) image format. Edit the directory, file name and number for the location of the downloaded images;
10. On the flash control box, set the flash mode to 1 (ring), select Tier 3, and take an image using the 'Manual' button on the GUI (not AF);
11. Check composition of image, and adjust zoom on lens if necessary for magnification. If X-Y positioning is necessary to centre the object in the image frame, open dome and move the black card and/or object as required. Repeat steps 8-11 until composition is correct;
12. On the control box, switch off the Focus LED lights, and set the flash mode to 0 (single). Press the green 'LOAD' button and then the red 'FLASH' button to check that a single flash light fires (light #1 is on bottom Tier, at the rear of dome).
13. Via the Nikon GUI, set the lens aperture according to the surface gloss of the object. The typical range is from *f*/4 for a completely matte surface to *f*/16 for a mirror-like surface. Consideration should also be given to depth of field needed to span the range of relief on the object surface;
14. Switch off the room lights and any other sources of stray light (lamps, etc.) Reduce the luminance of the computer screen. Capture a sequence of 64 images under single flash

lighting, pressing the blue 'INC' button on the flash controller to increment the flash after each, at intervals of approximately 10 seconds;

15. Take one additional image, without flash, as the 'black' ambient image;
16. Open the dome, remove the object and the black cardboard sheet. Close the dome and set the flash mode to 1, and select Tier 5 (four lamps at top of dome near camera). Without changing the aperture or focus settings on the lens, take an image of the retroreflective targets on the dome baseboard.

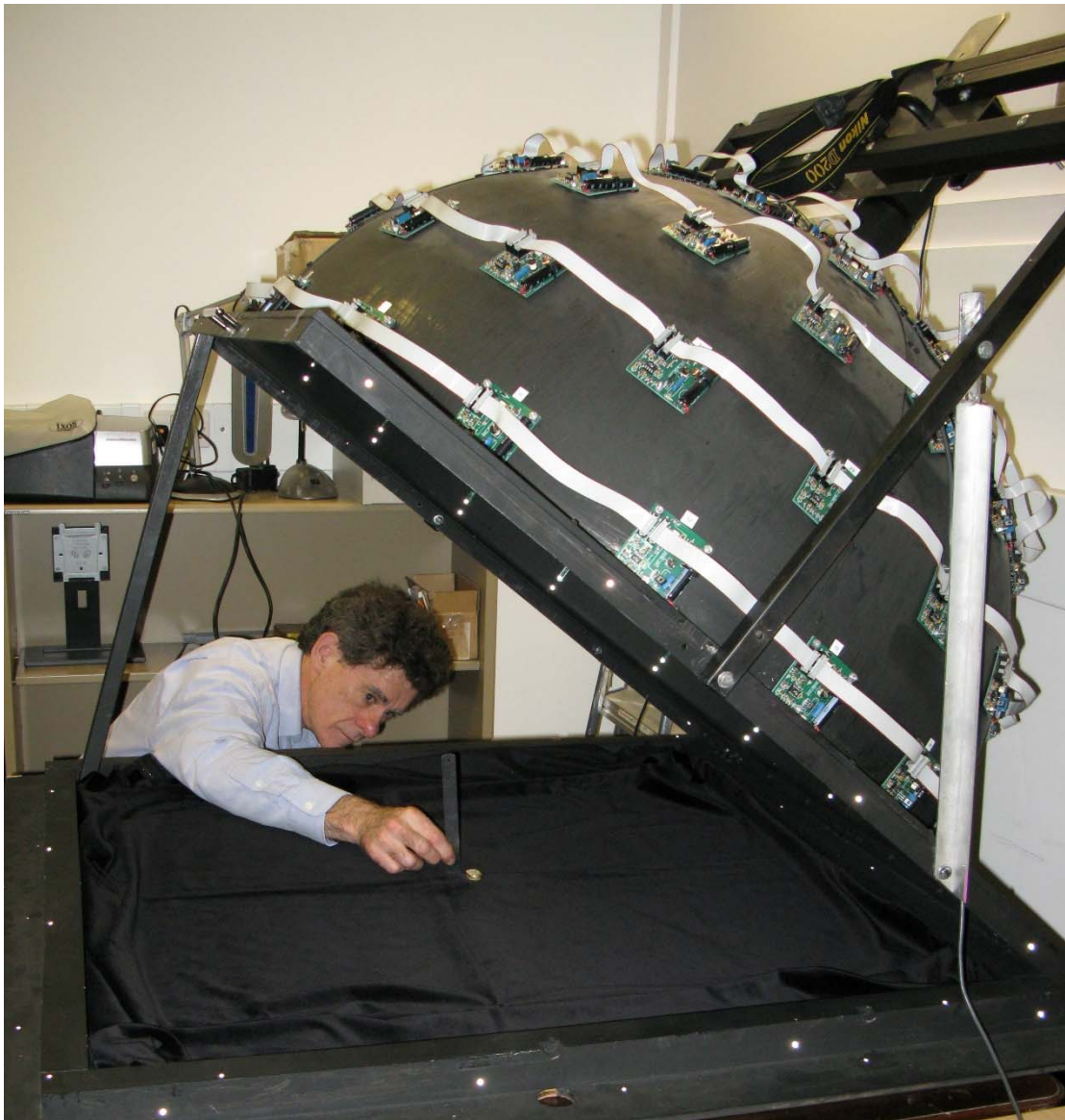


Figure A.1.1: Setting up for photography of the Egyptian scarab (Sec. 8.1.2)

Alternatively, eliminating r from Eq. 3-2 and Eq. A-2 gives:

$$\begin{aligned}
 R^2 - z^2 &= s'^2 \left(\frac{z+v}{h} - 1 \right)^2 = \frac{s'^2}{h^2} (z+v-h)^2 = \frac{s'^2}{h^2} (z^2 + 2(v-h)z + (v-h)^2) \\
 \therefore \left(\frac{s'^2}{h^2} + 1 \right) z^2 + 2 \frac{s'^2}{h^2} (v-h)z + \frac{s'^2}{h^2} (v-h)^2 - R^2 &= 0 \\
 \therefore z &= \frac{-2 \frac{s'^2}{h^2} (v-h) \pm \sqrt{4 \left(\frac{s'^2}{h^2} (v-h) \right)^2 - 4 \left(\frac{s'^2}{h^2} + 1 \right) \left(\frac{s'^2}{h^2} (v-h)^2 - R^2 \right)}}{2 \left(\frac{s'^2}{h^2} + 1 \right)} \\
 \therefore z &= \frac{-\left(\frac{s'}{h} \right)^2 (v-h) \pm \sqrt{\left(\left(\frac{s'}{h} \right)^2 (v-h) \right)^2 - \left(\left(\frac{s'}{h} \right)^2 + 1 \right) \left(\left(\frac{s'}{h} \right)^2 (v-h)^2 - R^2 \right)}}{\left(\frac{s'}{h} \right)^2 + 1} \quad (\text{A-3})
 \end{aligned}$$

As a check, if $v = 0$, Eq. A-3 simplifies to the same as Eq. 3-4:

$$z = \frac{\left(\frac{s'}{h} \right)^2 h \pm \sqrt{\left(\left(\frac{s'}{h} \right)^2 h \right)^2 - \left(\left(\frac{s'}{h} \right)^2 + 1 \right) (s'^2 - R^2)}}{\left(\frac{s'}{h} \right)^2 + 1} \quad (\text{A-4})$$

If $v = h$, Eq. A-3 simplifies to:

$$z = \frac{\pm \sqrt{(s'^2 + h^2) R^2}}{(s'^2 + h^2)/h} = \frac{Rh}{\sqrt{(s'^2 + h^2)}} \quad (\text{A-5})$$

For the measured uplift of $v = 55$ mm, the lamp coordinates calculated by the pin shadow method are set out in Table A.1. Fig. A.2.2 shows the solutions of equations A-2 and A-3 as a function of the length of the shadow cast by the two pins used in the experiment.

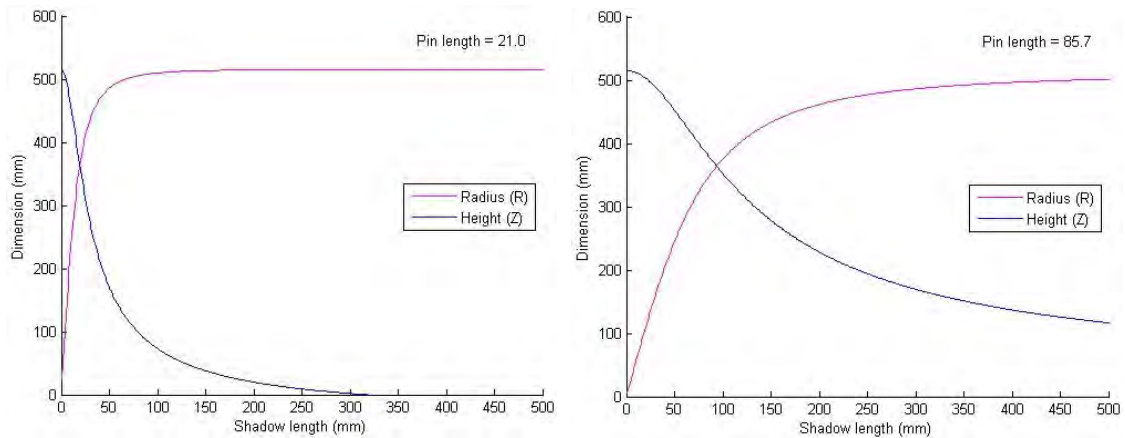


Figure A.2.2: Calculated radial distance and height of lamp as a function of shadow length for two pin lengths.

Table A.1: Lamp coordinates (X,Y,Z) relative to centre of baseboard (pin shadow values in grey).

Lamp #	X	Y	Z	X	Y	Z	Diff 3D
1	-51.98	501.95	42.26	-46.26	505.87	45.44	10.95
2	-246.42	442.07	40.94	-243.60	445.43	43.15	8.19
3	-398.38	318.54	42.28	-395.86	317.48	50.90	9.45
4	-497.93	132.86	43.17	-490.67	130.51	47.89	5.83
5	-515.62	-61.60	42.84	-503.83	-62.83	47.69	8.86
6	-456.57	-251.67	42.26	-443.31	-246.59	43.55	8.51
7	-328.94	-408.64	40.60	-317.27	-396.58	46.72	13.54
8	-149.20	-503.38	38.61	-141.02	-488.33	41.77	12.82
9	43.76	-523.03	36.70	45.03	-506.43	39.68	13.19
10	237.65	-466.62	39.46	232.82	-451.71	42.89	11.46
11	393.58	-340.32	40.18	383.74	-331.81	43.32	9.24
12	489.77	-162.90	43.02	481.75	-161.31	43.63	3.88
13	510.34	33.84	42.76	507.20	30.14	43.12	3.26
14	456.91	222.60	43.08	458.18	219.25	45.44	6.53
15	327.62	385.01	41.41	333.90	382.55	47.58	11.71
16	145.51	482.63	41.99	153.56	483.19	45.54	12.14
17	47.54	472.57	191.26	50.54	470.48	190.28	3.40
18	-140.84	453.03	192.54	-137.70	451.71	192.58	3.06
19	-307.77	365.62	192.36	-301.46	363.73	192.12	3.09
20	-429.04	218.81	196.88	-417.38	217.67	196.20	7.73
21	-486.03	37.71	195.51	-470.37	40.85	192.78	11.63
22	-468.28	-150.76	190.19	-452.52	-140.86	188.39	13.16
23	-377.83	-317.76	187.51	-365.82	-302.18	186.97	15.23
24	-233.80	-436.27	190.13	-226.64	-416.11	188.60	17.15
25	-53.75	-495.14	189.45	-51.31	-471.50	187.53	19.12
26	135.12	-479.99	190.27	128.62	-455.48	189.98	20.77
27	303.16	-389.14	196.96	292.18	-369.70	195.08	18.26
28	427.80	-243.19	191.43	412.85	-232.20	188.96	13.97
29	483.47	-59.68	193.70	468.95	-55.85	192.47	10.48
30	465.35	123.50	195.13	453.82	123.06	195.13	6.37
31	376.02	293.15	196.41	371.48	290.39	193.36	1.29
32	226.19	417.50	193.92	226.70	413.20	193.82	3.28
33	113.32	366.96	338.72	113.37	366.54	335.99	3.23
34	-39.39	379.70	338.77	-38.51	379.34	338.69	3.70
35	-183.91	335.60	338.66	-179.47	336.45	338.64	5.94
36	-305.55	237.96	337.78	-299.06	239.19	336.80	5.56
37	-378.32	103.42	339.87	-368.68	108.99	335.12	8.38
38	-393.43	-48.03	342.68	-377.30	-39.12	340.88	15.98
39	-351.89	-193.74	335.40	-343.28	-180.23	331.30	13.80
40	-252.64	-312.10	337.42	-246.23	-292.13	337.82	17.95
41	-120.35	-388.60	333.87	-118.47	-369.60	330.84	15.41
42	33.06	-406.24	338.19	31.77	-383.27	333.97	19.24
43	183.79	-363.56	337.12	175.03	-342.98	333.41	19.55
44	303.30	-263.91	341.04	291.41	-246.13	338.54	18.37
45	375.96	-129.31	343.49	363.43	-118.17	337.72	13.65
46	389.03	23.31	347.62	373.07	30.75	346.34	13.19
47	347.85	170.44	343.01	337.77	171.62	341.39	7.38
48	246.02	297.07	341.03	240.52	293.98	339.46	3.69

49	76.30	235.51	447.03	73.59	237.30	445.38	5.33
50	-59.53	238.72	445.48	-59.01	241.94	445.08	6.86
51	-175.12	173.31	446.32	-172.07	178.35	445.74	7.89
52	-243.18	60.15	448.44	-238.78	67.83	445.49	9.00
53	-252.72	-67.10	446.04	-243.04	-53.11	443.52	13.24
54	-186.24	-196.02	441.83	-182.68	-181.25	440.28	13.39
55	-82.72	-253.76	443.26	-81.55	-237.41	443.96	13.61
56	56.40	-263.38	445.82	53.55	-245.51	443.80	16.78
57	171.85	-203.15	447.71	163.48	-187.22	443.98	16.24
58	243.97	-91.23	449.15	235.42	-78.95	445.49	13.61
59	251.63	43.82	450.66	241.54	52.38	446.10	11.92
60	187.52	162.42	450.35	182.78	166.82	445.94	6.72
61	-43.46	63.26	506.19	-45.75	73.29	502.61	10.21
62	-77.30	-55.70	503.26	-76.04	-42.43	502.50	13.25
63	47.14	-90.76	506.88	42.99	-75.48	502.55	15.03
64	-51.98	501.95	42.26	75.48	41.31	502.71	11.27

Table A.2: Lamp radius (in X-Y plane), theta, phi (pin shadow values in grey).

Lamp	R	θ	ϕ	R	θ	ϕ
1	503.63	95.91	3.79	507.98	95.22	5.11
2	506.11	119.14	3.63	508.17	118.77	3.85
3	510.08	141.35	3.74	507.44	141.27	5.73
4	515.35	165.06	3.79	507.73	165.11	5.39
5	519.28	186.81	3.72	507.73	187.11	5.37
6	521.34	208.86	3.63	508.15	209.03	3.90
7	523.59	231.17	3.43	507.87	231.34	5.26
8	525.02	253.49	3.21	508.28	253.89	3.70
9	525.86	273.77	3.99	508.43	275.08	3.46
10	523.66	296.99	3.31	508.18	297.27	3.82
11	521.07	319.22	3.41	508.06	319.23	3.99
12	516.15	341.60	3.76	508.04	341.49	5.02
13	511.46	3.79	3.78	508.09	3.40	3.96
14	508.25	25.98	3.84	507.94	25.57	5.11
15	505.54	49.60	3.68	507.77	48.89	5.35
16	503.09	73.22	3.76	507.96	72.40	5.12
17	473.96	83.26	21.93	473.18	83.87	21.91
18	473.42	107.27	22.09	472.23	106.95	22.19
19	477.92	130.09	21.92	472.42	129.65	22.13
20	481.61	152.98	22.23	470.73	152.46	22.63
21	487.50	175.56	21.85	472.14	175.04	22.21
22	491.95	197.85	21.14	473.94	197.29	21.68
23	493.68	220.06	20.80	473.49	219.56	21.51
24	495.45	241.71	20.99	473.83	241.42	21.70
25	498.05	263.80	20.83	473.28	263.79	21.57
26	498.64	285.72	20.89	473.29	285.77	21.87
27	493.91	308.01	21.74	471.22	308.32	22.49
28	492.09	330.38	21.26	473.67	330.64	21.75
29	487.14	352.96	21.68	472.26	353.21	22.17
30	481.71	13.98	22.05	471.17	15.14	22.50
31	477.41	38.04	22.36	471.52	38.02	22.40
32	473.84	61.55	22.32	471.30	61.25	22.46
33	383.35	72.70	41.39	383.67	72.81	41.21
34	381.73	95.92	41.59	381.29	95.80	41.61
35	382.69	118.72	41.51	381.32	118.08	41.61
36	387.28	142.09	41.09	382.95	141.35	41.33
37	392.47	163.57	40.89	383.45	163.53	41.08
38	397.34	186.94	40.78	379.32	185.92	41.95
39	402.18	208.96	39.83	387.72	207.70	40.51
40	401.53	231.01	40.04	382.06	229.87	41.48
41	406.81	252.79	39.46	388.12	252.23	40.44
42	407.58	273.65	39.68	383.58	273.74	41.06
43	408.27	296.75	39.55	385.06	297.04	40.97
44	402.80	319.07	40.25	381.44	319.82	41.59
45	397.58	341.02	40.83	382.15	341.99	41.47
46	389.79	3.58	41.73	373.33	3.71	42.78
47	387.36	26.10	41.61	378.87	26.93	42.02
48	385.71	50.37	41.48	380.61	50.81	41.73

49	247.56	72.05	61.02	248.45	72.77	60.85
50	246.03	103.00	61.09	249.04	103.71	60.77
51	246.38	135.30	61.10	247.82	133.97	60.93
52	251.48	166.16	60.72	248.23	163.14	60.87
53	261.47	193.87	59.62	249.96	192.50	60.65
54	270.39	226.46	58.53	257.34	223.78	59.69
55	266.90	251.95	59.00	251.02	251.04	60.52
56	270.32	282.04	58.77	251.29	282.30	60.48
57	266.85	310.09	59.20	249.21	311.30	60.75
58	261.40	339.58	59.80	248.30	341.46	60.87
59	255.41	9.88	60.46	247.15	12.24	61.01
60	248.08	40.90	61.15	247.47	42.39	60.97
61	77.58	123.07	81.29	86.39	121.97	80.25
62	95.28	215.78	79.30	87.08	209.16	80.17
63	102.28	297.45	78.59	86.87	299.67	80.19
64	83.98	21.39	80.53	86.05	28.69	80.29

Appendix 3 – Principal component analysis of spectra

This appendix repeats the analysis of Sec. 7.3.2 for eight datasets of reflectance spectra in colour atlases and natural scenes.

Berns-Taplin Pigment Set

The dataset contains reflectance spectra for 100 samples of acrylic paint, cut from draw-downs and measured with an X-rite eye-one Pro spectrophotometer in the Munsell Color Science Laboratory at Rochester Institute of Technology (RIT). The spectra were recorded at 10 nm intervals from 380 to 730 nm (36 values). The data were interpolated (by cubic spline) to 5 nm intervals, giving 71 reflectance values 380-730 nm.

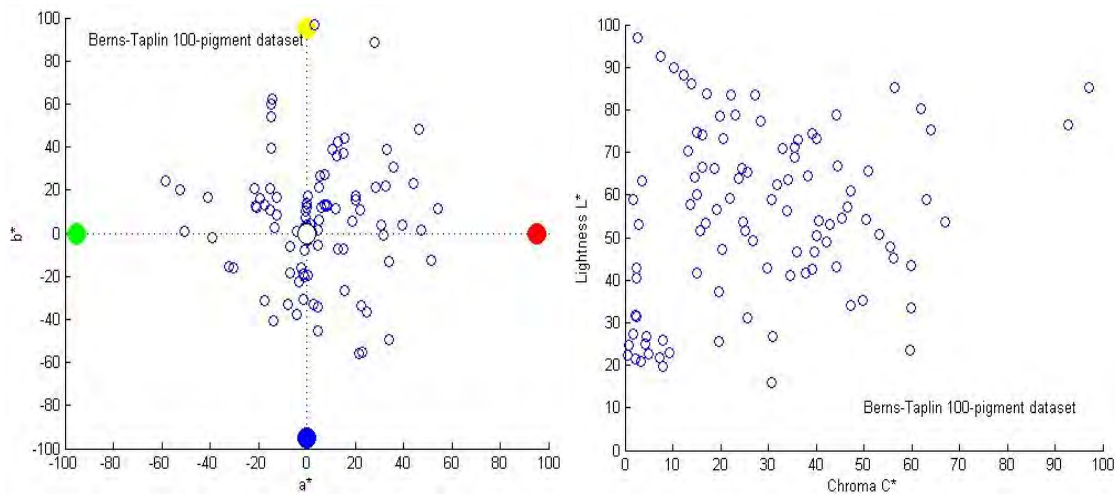


Figure A.3.1: Colours of 100 patches of Berns-Taplin acrylic paint dataset, plotted in CIELAB space for D65: (left) projection in a^*-b^* plane; (right) projection onto L^*-C^* plane.

The colours corresponding to the 100 samples are shown in CIELAB coordinates in Fig. A.3.1. The chroma of all colours is less than 70, except for two yellows. The spectra of the samples are shown in Fig. A.3.2 (left), together with the mean (in red). The first five principal components are shown in Fig. A.3.2 (right).

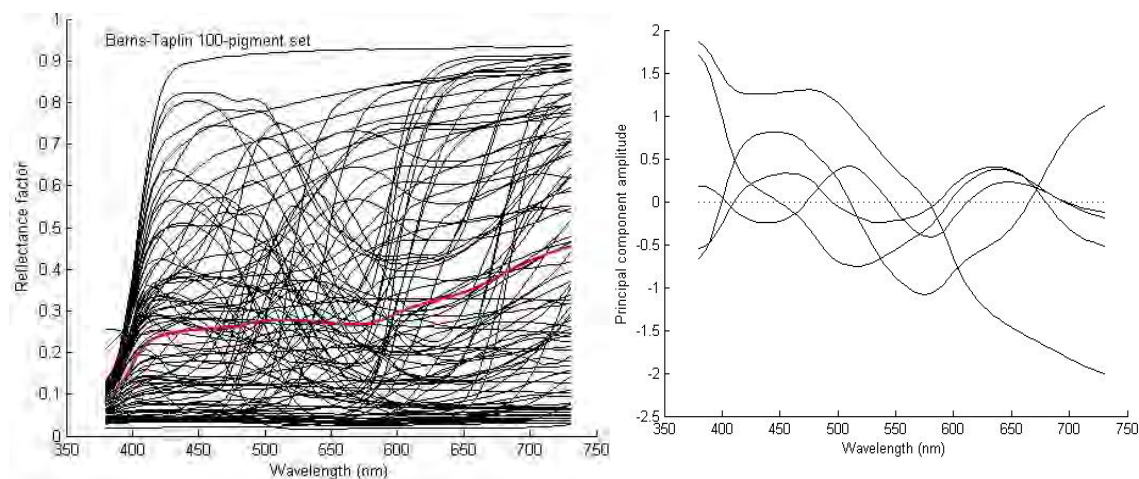


Figure A.3.2: (left) reflectance spectra of 100 samples of Berns-Taplin dataset, with mean (red); (right) spectra of first five principal components.

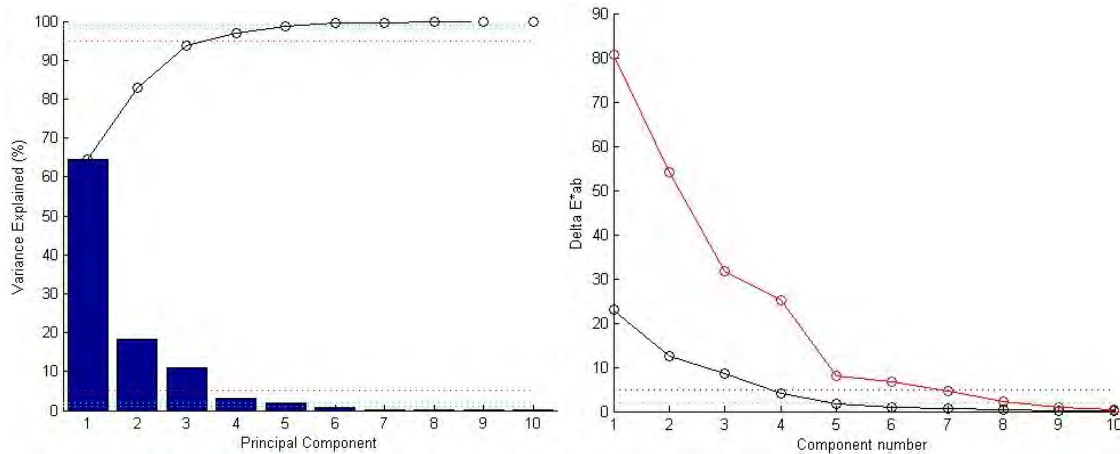


Figure A.3.3: (left) variance explained by first ten principal components, with cumulative distribution; (right) maximum (red) and mean (blue) colour error (ΔE^*_{ab}) corresponding to the difference between the true spectrum of each sample and that reconstructed from increasing numbers of principal components.

The variance and colour errors associated with each component and the cumulative variance are illustrated in Fig. A.3.3. The first five components account for 98.8% of the variance, and over the whole set of samples the spectra reconstructed with five components have a mean colour error (ΔE^*_{ab}) of 1.64, with a maximum error of 7.98. The ΔE^*_{ab} values for reconstruction with three components are plotted in Fig. A.3.4. Errors tend to be greater for higher levels of chroma C^* but with not much dependence on L^* or h_{ab} .

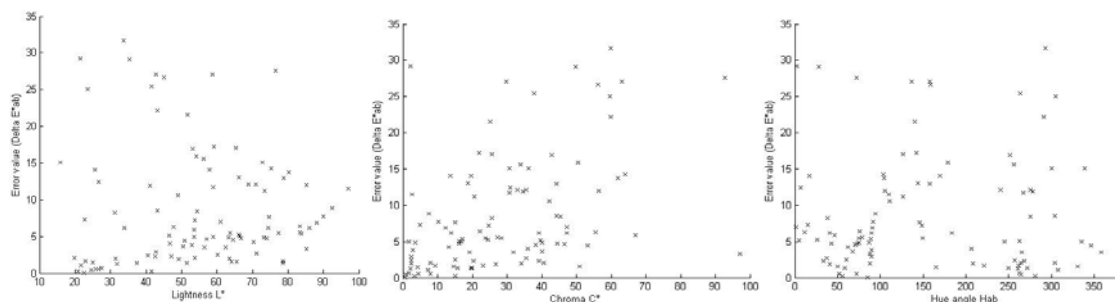
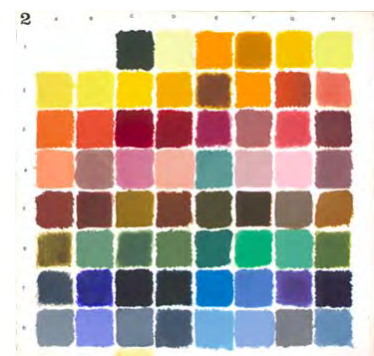


Figure A.3.4: Distribution of errors between true and reconstructed spectra, plotted against: (left) lightness, (centre) chroma, and (right) hue angle in CIELAB coordinates.

National Gallery Paint Chart

The chart contains $8 \times 8 = 64$ patches overall, and was hand-painted in 1990 as a test target for characterisation of the VASARI scanner, using traditional art pigments employed by restorers at the National Gallery, London (Saunders and Cupitt, 1993). The chart was measured at the National Gallery with a Monolight spectrophotometer, and the spectra recorded at 2 nm intervals from 380 to 780 nm (201 values). The data were resampled (by cubic spline) to 5 nm intervals and the last ten values were discarded, giving 71 reflectance values 380-730 nm. The colours corresponding to the 64 patches are shown in CIELAB coordinates in Fig. A.3.5. The majority of the hues are between red and yellow, because fine artists use more pigments in this range (including browns).



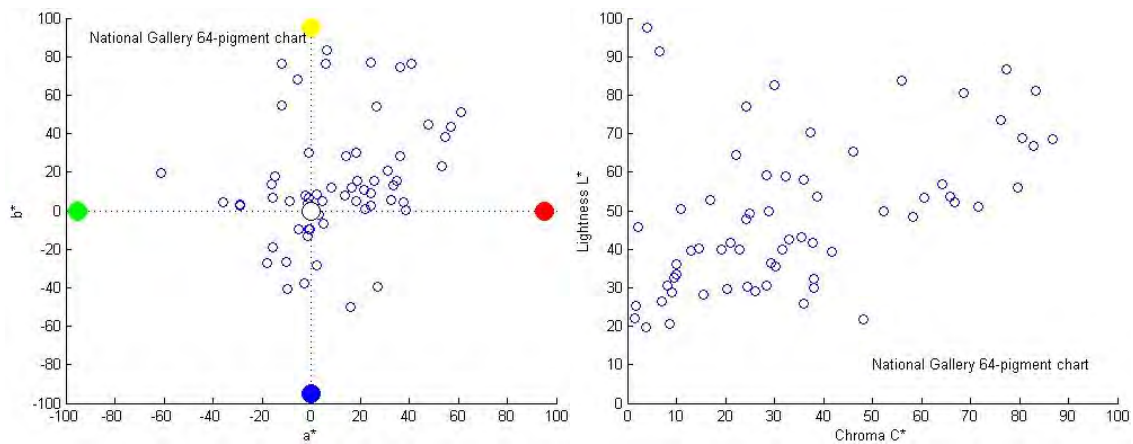


Figure A.3.5: Colours of 64 patches of National Gallery chart, plotted in CIELAB space for D65: (left) projection in a^*-b^* plane; (right) projection onto L^*-C^* plane.

The spectra of all patches are shown in Fig. A.3.5 (left), together with mean (in red). The first five principal components are shown in Fig. A.3.5 (right).

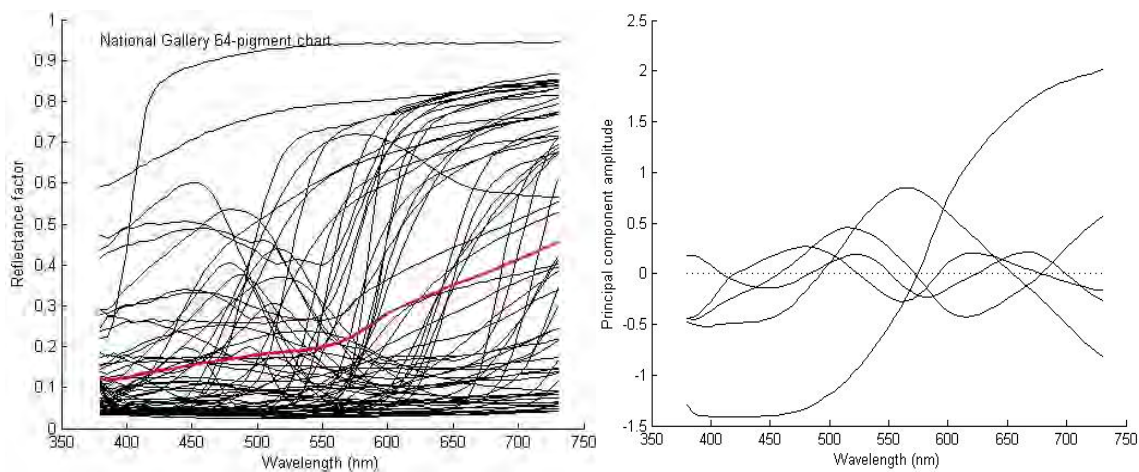


Figure A.3.6: (left) reflectance spectra of 64 patches of the National Gallery chart, with mean (red); (right) spectra of first five principal components.

The variance and colour errors associated with each component and the cumulative variance are illustrated in Fig. A.3.6. The first five components account for 99.0% of the variance, and over the whole set of samples the spectra reconstructed with five components have a mean colour error (ΔE^*_{ab}) of 2.05, with a maximum error of 11.73.

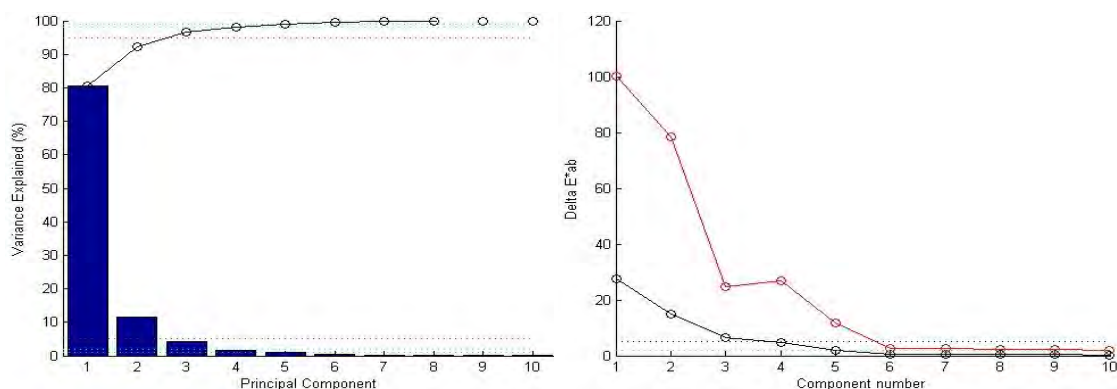


Figure A.3.6: (left) variance explained by first ten principal components, with cumulative distribution; (right) maximum (red) and mean (blue) colour error (ΔE^*_{ab}) corresponding to the difference between the true spectrum of each sample and that reconstructed from increasing numbers of principal components.

Natural Materials

The dataset contains reflectance spectra of 494 samples of natural organic materials such as leaves, petals and grass. They were collected in the park surrounding the Colour and Imaging Institute at Kingsway House, Derby, by Vien Cheung in 2002. The samples were measured with the Macbeth 7000A spectrophotometer with diffuse geometry and specular component included. The spectra were recorded at 10 nm intervals from 400 to 700 nm (31 values). The data were interpolated (by cubic spline) to 5 nm intervals, giving a vector of 61 reflectance values 400-700 nm.

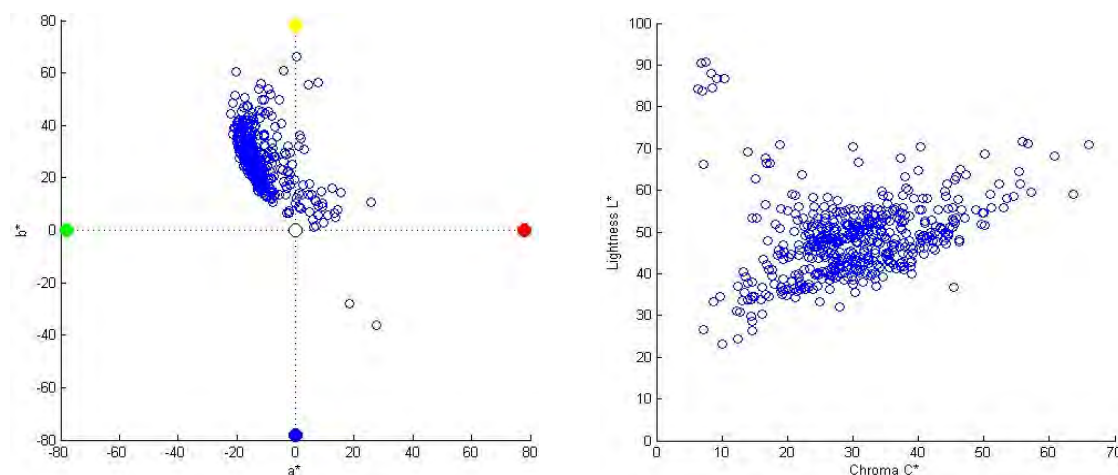


Figure A.3.7: Colours of 494 samples of natural materials, plotted in CIELAB space for D65: (left) projection in a^*-b^* plane; (right) projection onto L^*-C^* plane.

The colours corresponding to the 494 samples are shown in CIELAB coordinates in Fig. A.3.7. The great majority occupy a small region of the colour space between greenish yellow and brown. There are no dark colours in the set (the lightness of all colours is greater than 20 in Fig. A.3.7 right), probably because the measurements from glossy surfaces were made with ‘specular included’ geometry.

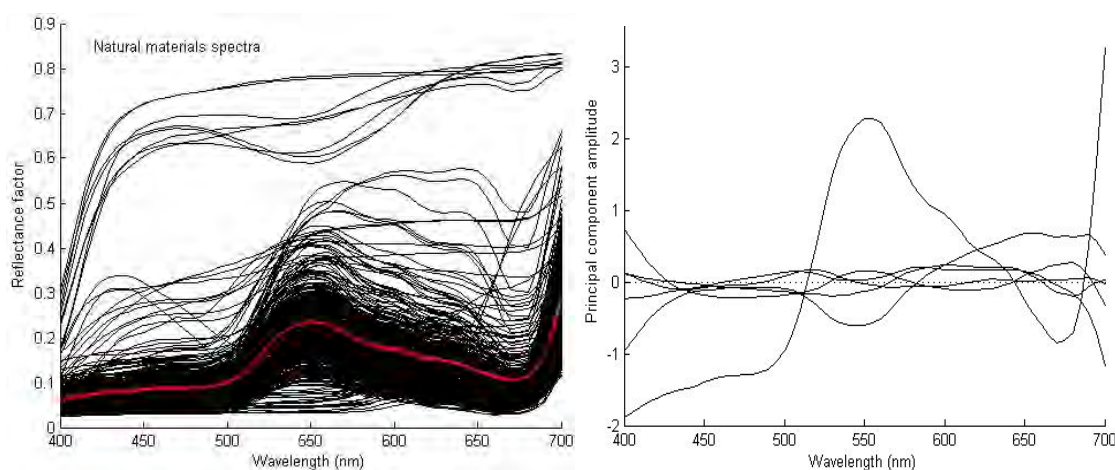


Figure A.3.8: (left) reflectance spectra of 494 samples of natural materials, with mean (red); (right) spectra of first five principal components.

The spectra of all patches are shown in Fig. A.3.8 (left), together with the mean (in red). The predominant shape of the spectra is the characteristic signature of chlorophyll, with a peak at

550 nm, a minimum at 670 nm and then a sharp rise in reflectance above 680 nm into the infrared. This growth region is known as the ‘plant spectrum reflectance red edge’ and is used as a diagnostic signal for the detection of chlorophyll in vegetation, for example in terrestrial remote sensing from space (Kochubey and Yatsenko, 2004). The first five principal components are shown in Fig. A.3.9 (right) and the amplitude of the second and subsequent components is rather low, because the spectra are mostly so similar.

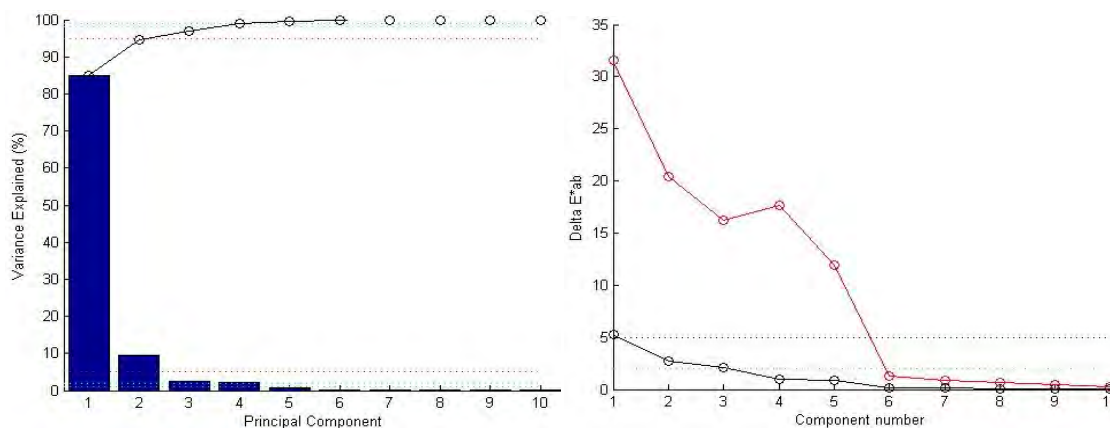


Figure A.3.9: (left) variance explained by first ten principal components, with cumulative distribution; (right) maximum (red) and mean (blue) colour error (ΔE^*_{ab}) corresponding to the difference between the true spectrum of each sample and that reconstructed from increasing numbers of principal components.

The variance and colorimetric errors associated with each component and the cumulative variance are illustrated in Fig. A.3.9. The first two components already account for 94.4% of the variance and the first four components account for 98.9%. Over the whole set of samples the spectra reconstructed with only four components have a mean colour error (ΔE^*_{ab}) of 0.97, but the maximum error remains high until six components are employed.

Chinese CNCS Dyed Textiles

The dataset contains reflectance spectra of 899 samples of polyester treated with synthetic dyes. They were measured by staff in the laboratory of the Chinese Natural Colour System (CNCS) in Shanghai, using a Datacolor 600 spectrophotometer with diffuse (d/8) geometry, specular component excluded and UV excluded. The spectra were recorded at 10 nm intervals from 400 to 700 nm (31 values). The data were interpolated (by cubic spline) to 5 nm intervals, giving 61 reflectance values 400–700 nm.



The colours corresponding to the 899 samples are shown in CIELAB coordinates in Fig. A.3.10. They cover the full range of hues, but the chroma values are noticeably larger for yellows ($+b^*$). The hues are roughly clustered in twenty angles around the full circle, at intervals of average 18° . The lightness values (L^*) are clustered at seven intervals spaced by approximately 10 units. Chroma is clustered in approximately seven concentric rings. There are no really dark colours in the set (the minimum lightness is 19.8 in Fig. A.3.10 right), probably because a diffuse measuring geometry (d/8) was used with an integrating sphere.

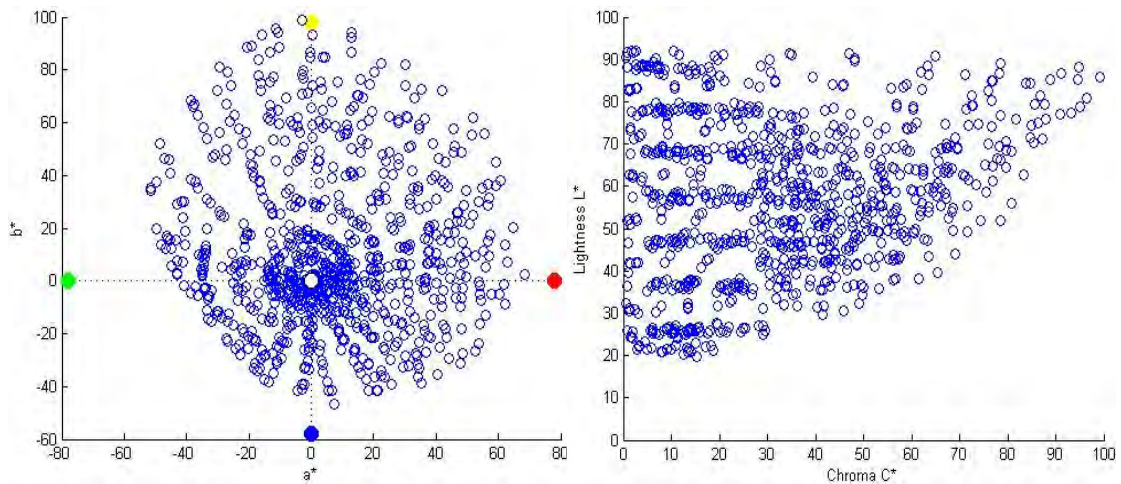


Figure A.3.10: Colours of 899 samples of dyed textile samples, plotted in CIELAB space for D65: (left) projection in a^*-b^* plane; (right) projection onto L^*-C^* plane.

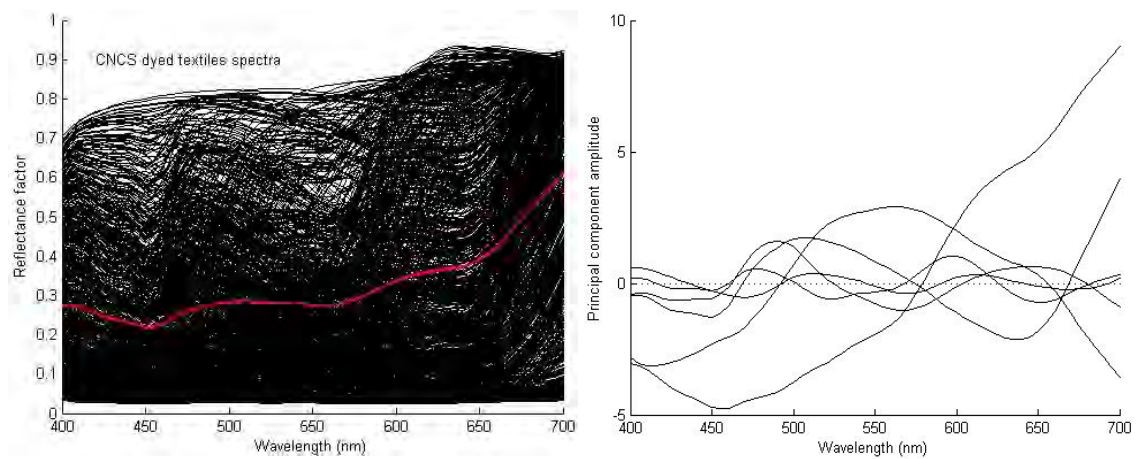


Figure A.3.11: (left) reflectance spectra of 899 dyed textile samples, with mean (red); (right) spectra of first six principal components.

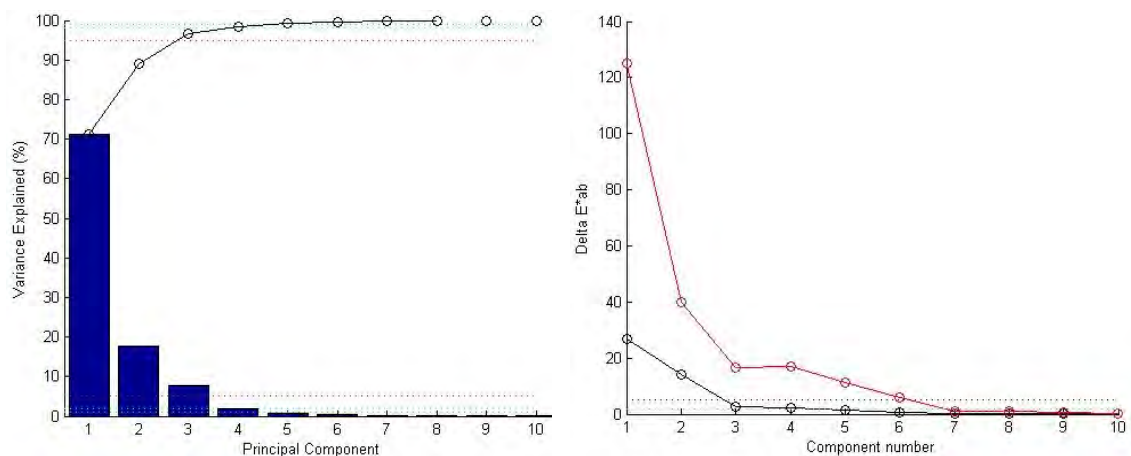


Figure A.3.12: (left) variance explained by first ten principal components, with cumulative distribution; (right) maximum (red) and mean (blue) colour error (ΔE^*_{ab}) corresponding to the difference between the true spectrum of each sample and that reconstructed from increasing numbers of principal components.

The spectra of all patches are shown in Fig. A.3.11 (left), together with the mean (in red). The overall trend is upward with greater reflectance at longer wavelengths. The first six principal

components are shown in Fig. A.3.11 (right). The variance and colorimetric errors associated with each component and the cumulative variance are illustrated in Fig. A.3.12. The first three components account for 96.6% of the variance and the first five components for 99.3%. Over the whole set of samples the spectra reconstructed with five components have a mean colour error (ΔE^*_{ab}) of 1.32, but it needs seven components to bring maximum error down to 1.15.

Swedish NCS Colour Atlas

The dataset contains reflectance spectra of 1950 painted samples, representing 50 colours on each of 40 hue angles in the NCS colour atlas. They were measured by staff in the Swedish NCS laboratory, using a Gretag Spectrolino spectrophotometer with 45/0 geometry. The spectra were recorded at 10 nm intervals from 380 to 730 nm (36 values) and Lagrange interpolated by NCS to 5 nm intervals (71 values). The colours corresponding to 1950 samples are shown in CIELAB coordinates in Fig. A.3.13. They cover the full range of hues, but the chroma values are larger for yellows ($+b^*$). Lightness values range from 96.3 (white) to 9.3 (black).

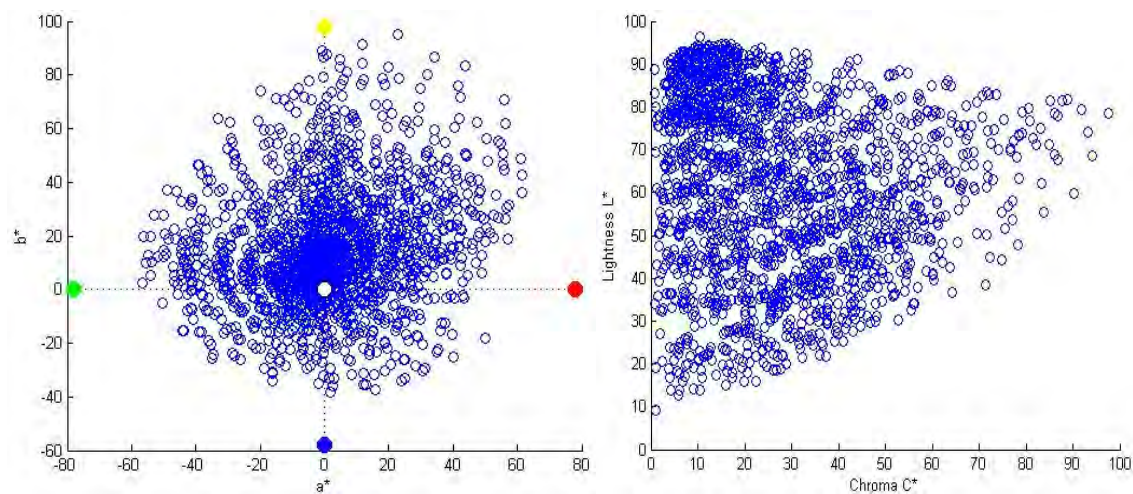


Figure A.3.13: Colours of 1950 NCS samples, plotted in CIELAB space for D65: (left) projection in a^*-b^* plane; (right) projection onto L^*-C^* plane.

The spectra of all patches are shown in Fig. A.3.14 (left), together with the mean (in red). The overall trend is upward with greater reflectance at longer wavelengths. The first six principal components are shown in Fig. A.3.14 (right). The variance and colorimetric errors associated with each component and the cumulative variance are illustrated in Fig. A.3.15. The first three components account for 96.9% of the variance and the first five components for 99.4%. Over the whole set of samples the spectra reconstructed with five components have a mean colour error (ΔE^*_{ab}) of 0.79, but it needs seven components to bring maximum error down to 1.72.

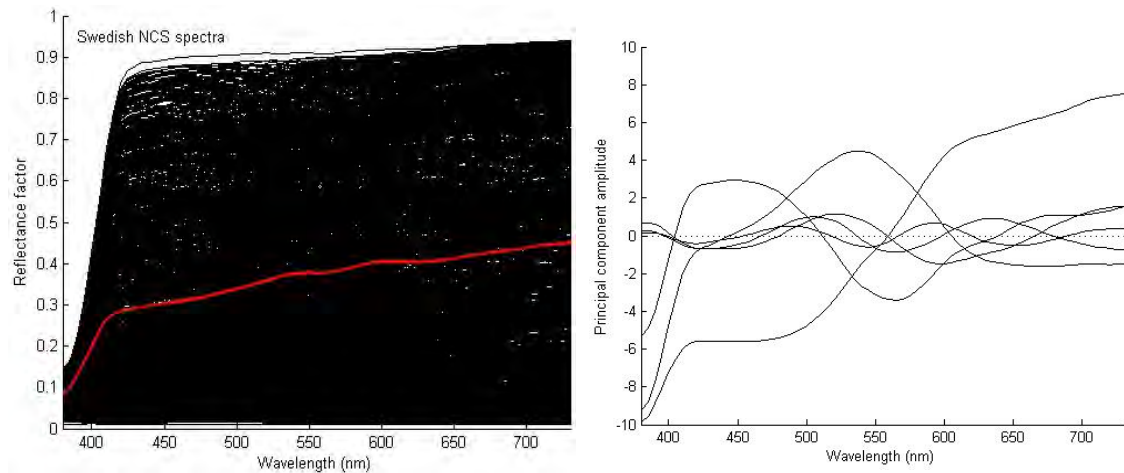


Figure A.3.14: (left) reflectance spectra of 1950 NCS colour samples, with mean (red); (right) spectra of first six principal components.

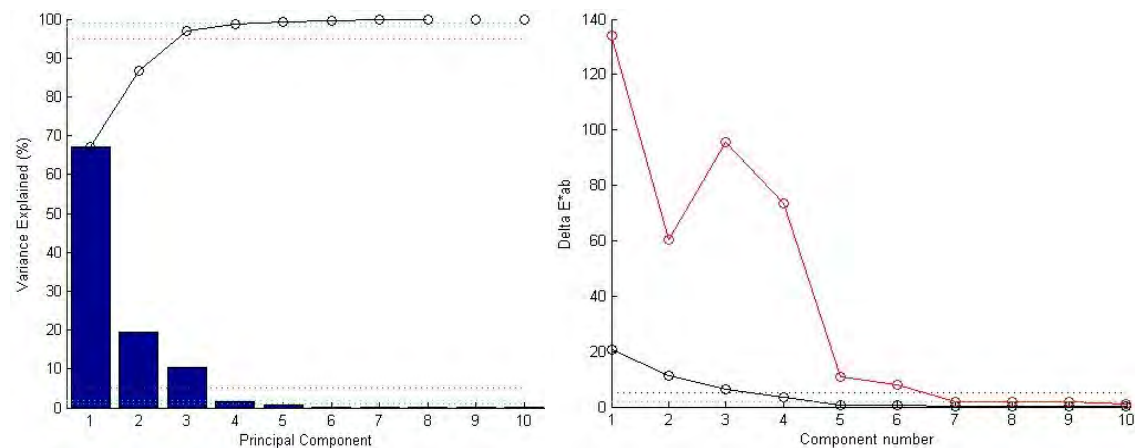


Figure A.3.15: (left) variance explained by first ten principal components, with cumulative distribution; (right) maximum (red) and mean (blue) colour error (ΔE^*_{ab}) corresponding to the difference between the true spectrum of each sample and that reconstructed from increasing numbers of principal components.

Munsell Matte Colour Atlas

The dataset contains reflectance spectra of 1269 matte colour chips, representing an average of approximately 32 colours on each of 40 hue angles (5 primaries with eight sub-divisions of each) in the Munsell colour atlas. They were measured by Jouni Hiltunen in the multispectral laboratory at Joensuu University, using a Perkin-Elmer lambda 9 UV/VIS/NIR spectrophotometer with 8/diffuse geometry. The spectra were recorded at 1 nm intervals from 380 to 800 nm (421 values). The 1 nm data for the CIE 1931 Standard Observer and D65 illuminant were used, from the CVRL database.



The colours corresponding to the 1269 samples are shown in CIELAB coordinates in Fig. A.3.14. The regular sampling of perceptual colour space is clear, with the hues falling onto well-defined radii from the centre and the lightness in intervals of 10 units. The chroma values are compressed on the blue axis (negative values of b^*).

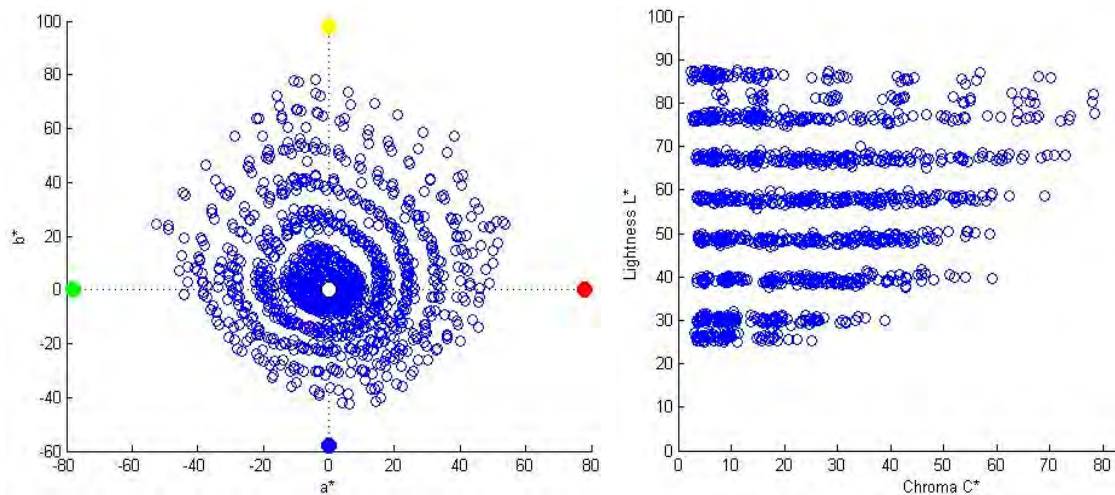


Figure A.3.14: Colours of 1269 Munsell matte samples, plotted in CIELAB space for D65: (left) projection in a^*-b^* plane; (right) projection onto L^*-C^* plane.

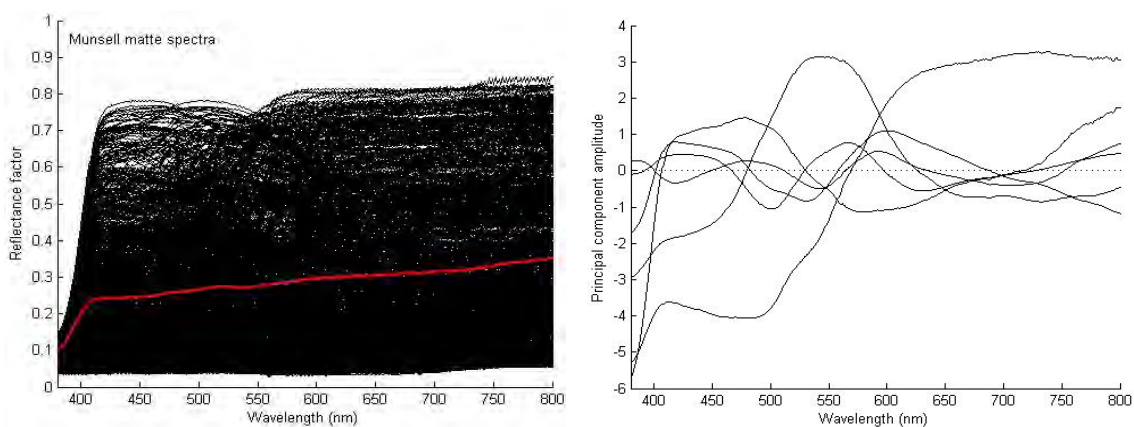


Figure A.3.15: (left) reflectance spectra of 1269 Munsell matte colour samples, with mean (red); (right) spectra of first six principal components.

The spectra of all patches are shown in Fig. A.3.15 (left), together with the mean (in red). The overall trend is upward with greater reflectance at longer wavelengths. The first six principal components are shown in Fig. 7.77 (right). The variance and colorimetric errors associated with each component and the cumulative variance are illustrated in Fig. A.3.16. The first three components account for 94.2% of the variance and the first five components for 98.6%. Over the whole set of samples the spectra reconstructed with six components have a mean colour error (ΔE^*_{ab}) of 0.61, but it needs nine components to bring the maximum error down to 1.91.

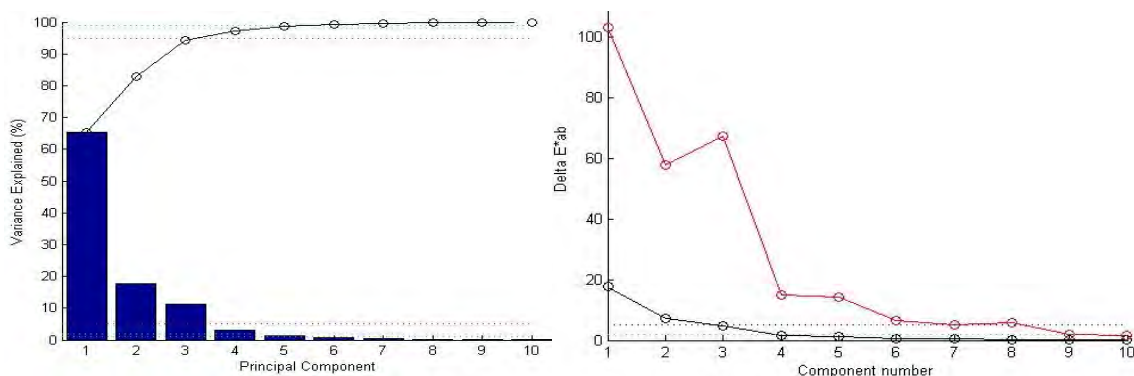


Figure A.3.16: (left) variance explained by first ten principal components, with cumulative distribution; (right) maximum (red) and mean (blue) colour error (ΔE^*_{ab}) corresponding to the difference between the true spectrum of each sample and that reconstructed from increasing numbers of principal components.

Munsell Gloss Colour Atlas

The dataset contains reflectance spectra of 1600 glossy colour chips, representing an average of 40 colours on each of 40 hue angles (5 primaries with eight sub-divisions of each) in the Munsell colour atlas. They were measured by Joni Orava in the multispectral laboratory at Joensuu University, using a Perkin-Elmer lambda 18 UV/VIS/NIR spectrophotometer with d/8 geometry. The spectra were recorded at 1 nm intervals from 380 to 780 nm (401 values).

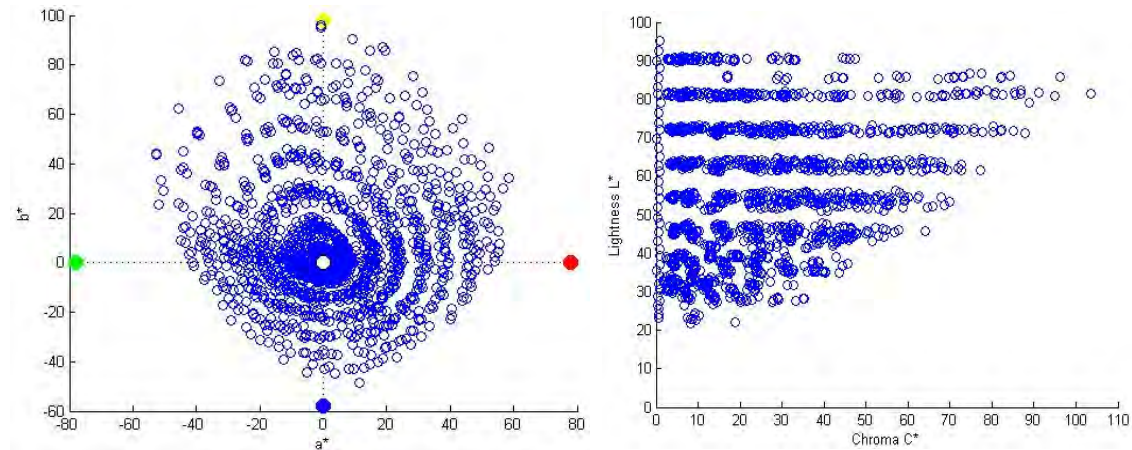


Figure A.3.17: Colours of 1600 Munsell glossy samples, plotted in CIELAB space for D65: (left) projection in a^*-b^* plane; (right) projection onto L^*-C^* plane.

The colours corresponding to the 1600 samples are shown in CIELAB coordinates in Fig. A.3.17. The regular sampling of perceptual colour space places the hues falling onto well-defined radii from the centre and the lightness in intervals of 10 units. Chroma values are compressed on the blue axis (negative values of b^*). The colour gamut is larger than for the matte samples (Fig. A.3.14), with maximum colourfulness increased to 103.49 and minimum lightness reduced to 21.65. The spectra of all patches are shown in Fig. A.3.18 (left), together with the mean (in red). The overall trend is upward with greater reflectance at longer wavelengths. The first six principal components are shown in Fig. A.3.18 (right).

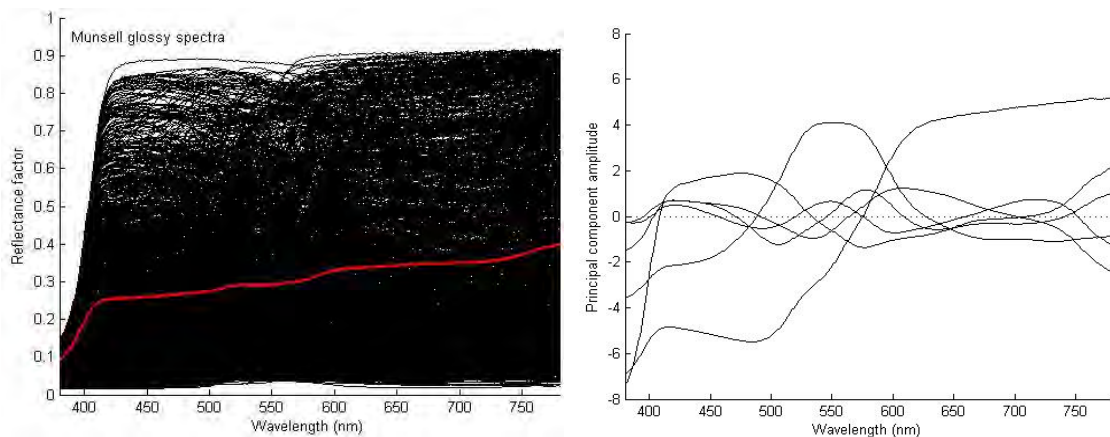


Figure A.3.18: (left) reflectance spectra of 1600 Munsell glossy colour samples, with mean (red); (right) spectra of first six principal components.

The variance and colorimetric errors associated with each component are illustrated in Fig. A.3.19. The first three components account for 94.7% of the variance and the first five components for 98.4%. Over the whole set of samples the spectra reconstructed with six components have a mean colour error (ΔE^*_{ab}) of 0.70, but it needs ten components to bring the maximum error down to 1.90. The distribution has a surprisingly long tail.

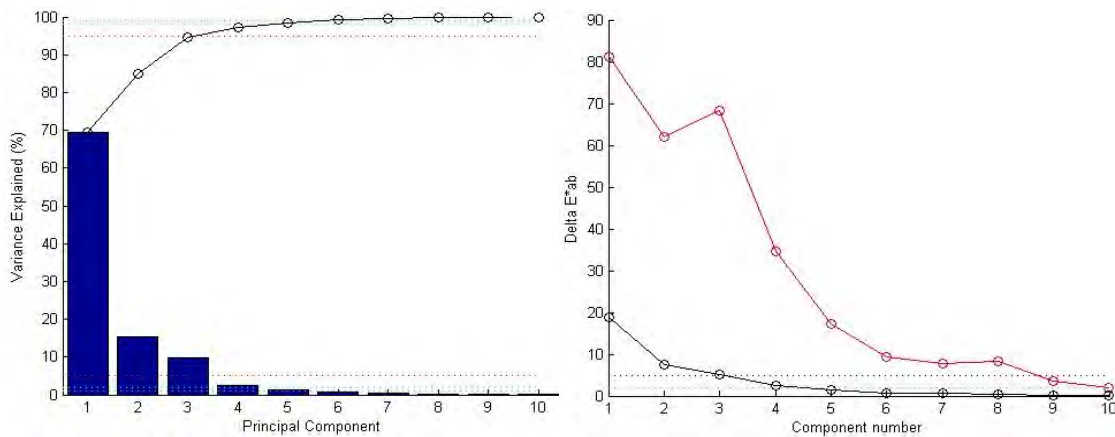


Figure A.3.19: (left) variance explained by first ten principal components, with cumulative distribution; (right) maximum (red) and mean (blue) colour error (ΔE^*_{ab}) corresponding to the difference between the true spectrum of each sample and that reconstructed from increasing numbers of principal components.

Hyperspectral Images

A set of hyperspectral images of natural scenes was downloaded from David Foster's site¹. These had been created under carefully controlled conditions from a camera with a tunable liquid crystal filter (Foster *et al*, 2004). Five images were selected from the set (Fig. A.3.21), and 100 points randomly sampled from each. This produced 100 reflectance spectra for each image and a dataset of 500 spectra in total, each with 33 values at 10 nm intervals in the range 400–720 nm. The spectra of the complete set of 500 patches are shown in Fig. A.3.20 (left), together with the mean (in red). The first five principal components are shown in Fig. A.3.20 (right).

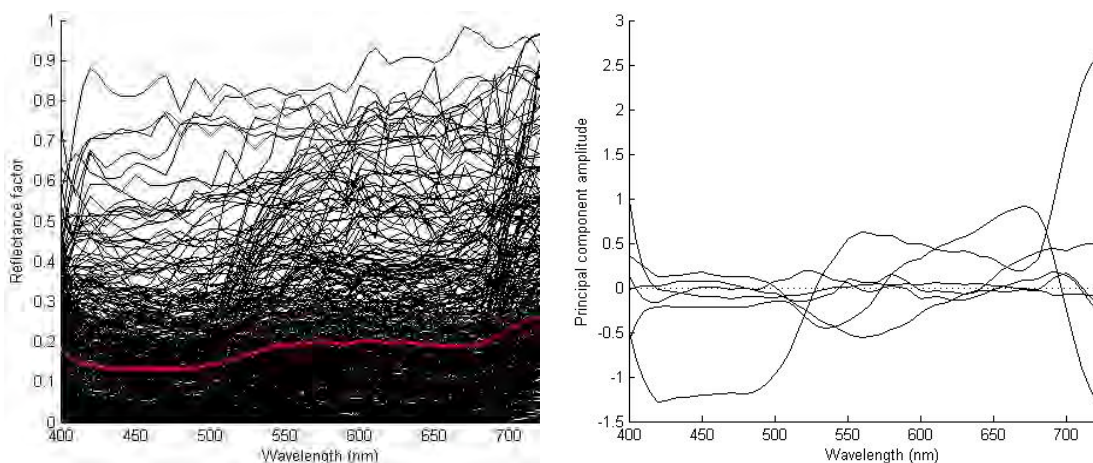


Figure A.3.20: (left) reflectance spectra of 500 samples selected from five hyperspectral images of natural scenes, with mean (red); (right) spectra of the first five principal components.

¹ http://personalpages.manchester.ac.uk/staff/david.foster/Hyperspectral_images_of_natural_scenes_04.html

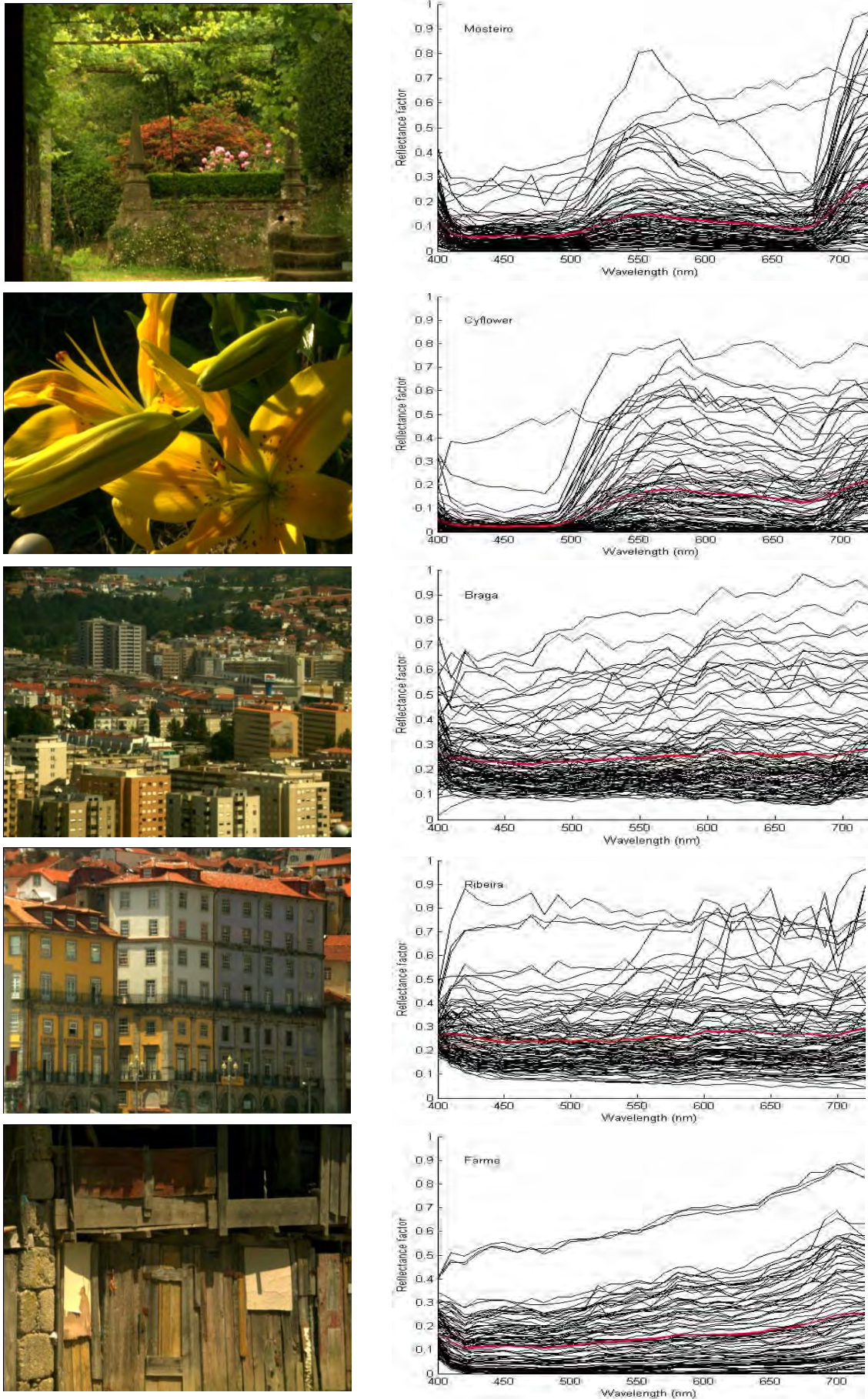


Figure A.3.21: Five hyperspectral images, with reflectance spectra 400-720 nm sampled at 100 points in each.

In the samples corresponding to green leaves (top two images of the monastery garden and the yellow lilies in Fig. A.3.21), can be seen the characteristic spectral signature of chlorophyll, with a peak at 550 nm, a minimum at 670 nm and then a sharp rise in reflectance above 680 nm into the infrared. This was also observed in the spectra of natural materials (Fig. A.3.8). The spectra for each image were normalised to reflectance factors in the range $[0, 1.0]$ as follows. The spectra of ten thousand points were randomly sampled from the image, and the 99th percentile value at each wavelength determined. The maximum value of the pseudo-spectrum was taken as the normalising factor. 100 spectra were selected from the set of 10,000 on condition that the scaled maximum did not exceed 1.0 (i.e. any point with a spectrum that would have exceeded 1.0 at any wavelength was discarded). The data were interpolated (by cubic spline) to 5 nm intervals, giving a vector of 65 reflectance values, for colorimetric calculation with 5 nm data for D65 illumination and the Judd-Vos standard observer.

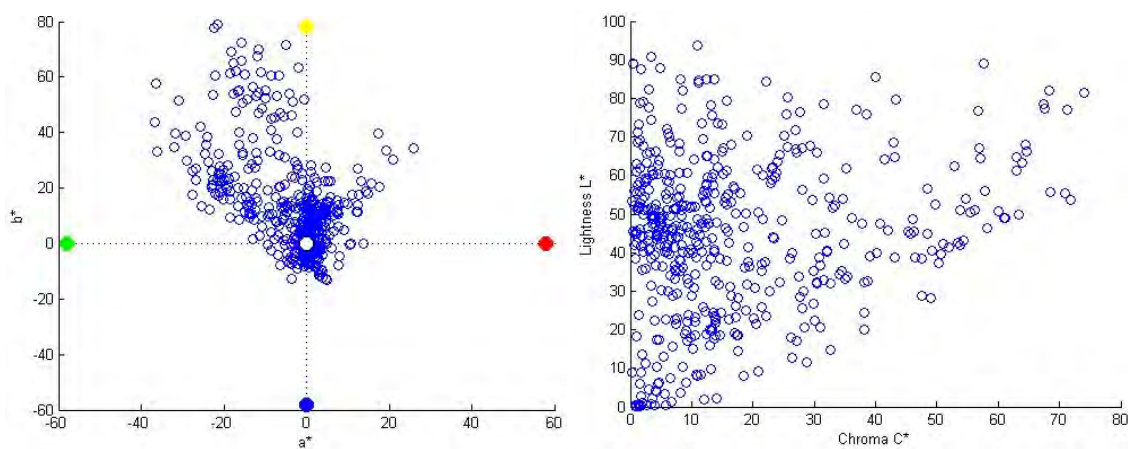


Figure A.3.22: Colours of 500 spectra of natural scenes, plotted in CIELAB space for D65: (left) projection in a^*-b^* plane; (right) projection onto L^*-C^* plane.

The colours corresponding to the 500 hyperspectral image samples are shown in CIELAB coordinates in Fig. A.3.22. They occupy a small range of hues between yellow-green and orange, but with a wide range of lightness and chroma. The variance and colorimetric errors associated with each component and the cumulative variance are illustrated in Fig. A.3.23. The first four components account for 96.2% of the variance, but additional components contribute only small amounts so that ten components are required before 99% is accounted for.

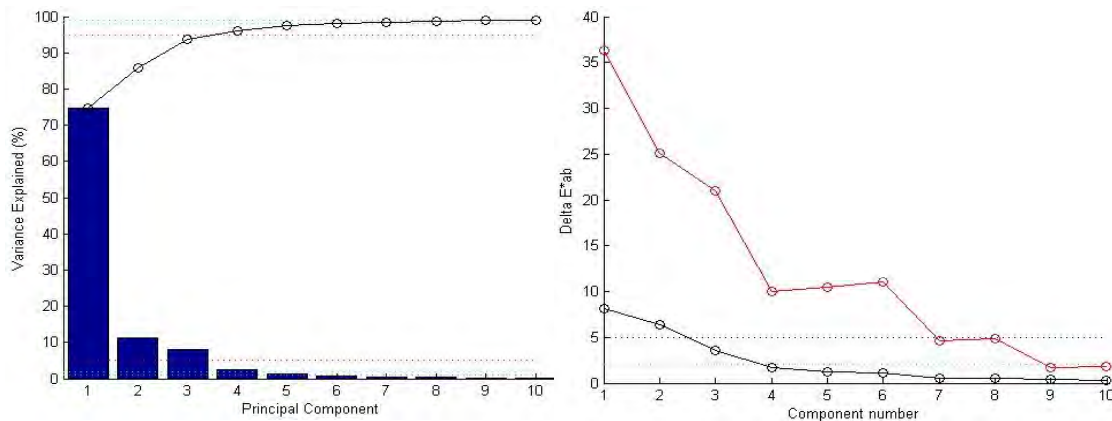


Figure A.3.23: (left) variance explained by first ten principal components, with cumulative distribution; (right) maximum (red) and mean (blue) colour error (ΔE^*_{ab}) corresponding to the difference between the true spectrum of each sample and that reconstructed from increasing numbers of principal components.

Over the whole set of samples from the hyperspectral images the spectra reconstructed with only four components have a mean colorimetric error (ΔE_{ab}^*) of 1.72 and a maximum error of 9.96, but addition of the fifth and sixth components produces little improvement in the mean and actually increases the maximum. This behaviour is quite different from the spectra of other materials examined, which in general show a steady reduction of mean and maximum colorimetric error with increasing numbers of components. The reason can be seen in the individual spectra, which exhibit an unnatural jitter from one wavelength to the next. Fig. A.3.24 shows the successive reconstruction of one reflectance spectrum from ten principal components and the error between the 'true' spectrum (black line) and the successive approximations (blue lines). The figure compares the hyperspectral image data (top) with the direct spectrophotometer measurement (bottom) for green leaves (in the Cheung dataset) with very similar reflectance spectra showing the distinctive chlorophyll signature. The jitter in the hyperspectral image data prevents the reconstruction from converging with successive principal components in the way that it does for the much smoother measured waveform.

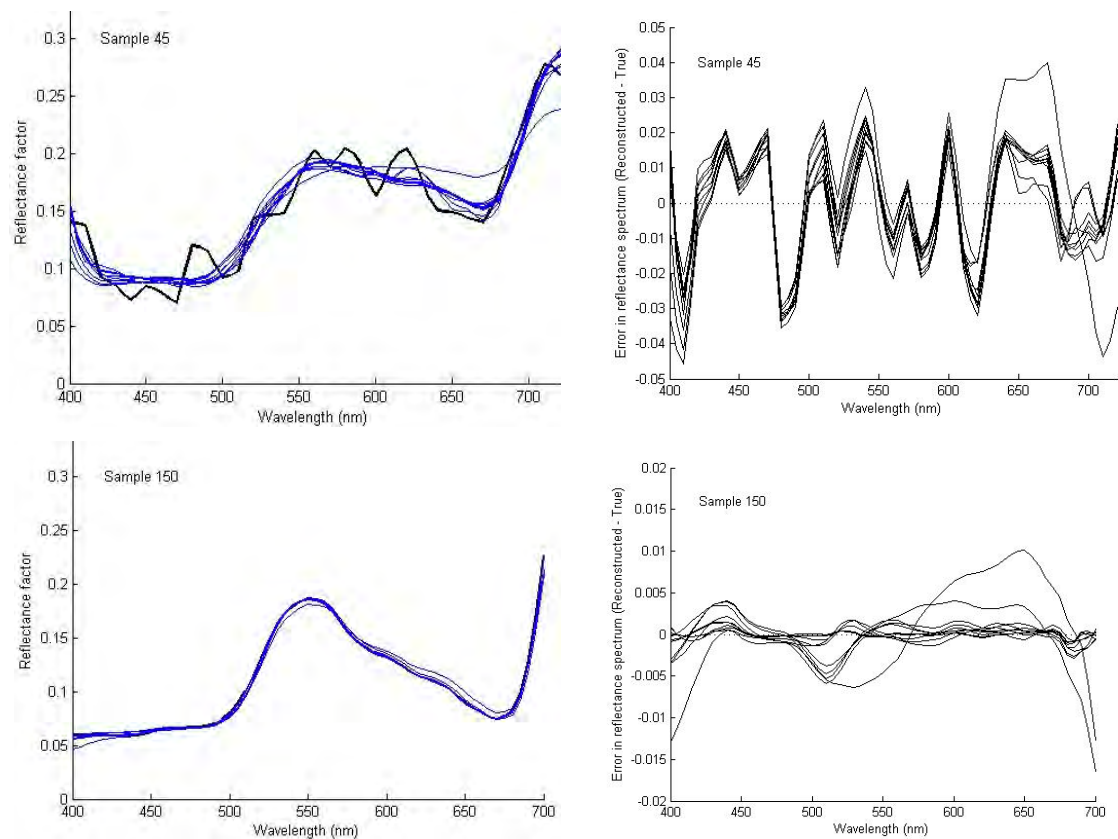


Figure A.3.24: Comparison of spectral reconstruction from first ten principal components for (top) hyperspectral image reflectance sample and (bottom) measured reflectance sample: (left) reconstruction of reflectance spectrum; (right) error between the true spectrum and successive refinement.

The jitter in the hyperspectral image spectra is also evident in the distribution of second-order differences, i.e. the changes in slope between adjacent pairs of points in the spectra, calculated in Matlab as `diff(diff(R))`. The measured reflectance spectra have a standard deviation of 0.0191 vs 0.0298 for the hyperspectral spectra (Fig. A.3.25 left) and are much smoother throughout the range of wavelengths (Fig. A.3.25 right). Real reflectance spectra are band-limited and therefore exhibit a good correlation between the reflectance values at successive wavelengths. The jitter in the spectra of the hyperspectral images must have been caused by

errors in calibration of the acquisition process, for example by small changes in the scene illumination between exposures through successive wavelength settings of the liquid crystal filter, or by non-repeatable variations in the transmittance of the filter.

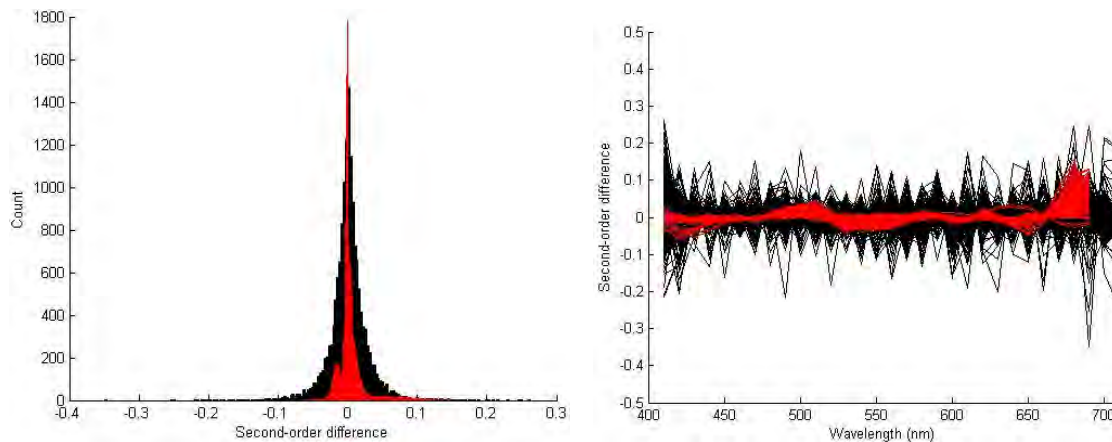


Figure A.3.25: Comparison of distributions of second-order differences between hyperspectral image spectra (black) and measured natural spectra (red): (left) histogram of occurrence; (right) magnitude vs wavelength.

It thus seems to me,
that in repeating an argument that I have learned,
I could readily have discovered it myself;
or rather, even if this is an illusion,
if I were not clever enough to have created it myself,
I rediscover it myself, to the extent that I repeat it.

– Jules Henri Poincaré
In Ayoub, R.G. (2004)
Musings of the Masters: An anthology of mathematical reflections,
Part I: Mathematics and the Intellect, p.22
Mathematical Association of America, Washington.