Spectrum Representation, Photometric Invariants and Shape Recovery in Imaging Spectroscopy

Cong Phuoc Huynh

April 24, 2012

A thesis submitted for the degree of

Doctor of Philosophy

of the Australian National University
To my dear parents, Huynh Thang and Diep Thi Minh Hai, and brother Huynh Cong Nghia, for their unconditional love, understanding and support.
Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of the authors knowledge, it contains no material previously published or written by another person, except where due reference is made in the text.

______________________________

Cong Phuoc Huynh
April 24, 2012
Acknowledgements

First of all, I am deeply grateful for the immense love and support of my mother Hai, my father Thang and my brother Nghia, who have always been with me throughout this scholastic endeavour. Their constant support has helped me build a strong mentality to overcome difficult times during these years. Their love, care and encouragement have drawn me close to them despite our geographical distance.

I am greatly indebted to Dr. Antonio Robles-Kelly for his supervision and guidance throughout the course of this thesis. I owe him for the knowledge and the scholastic experience I have received over the past few years. It has been a worthwhile and enjoyable studentship with him. I also thank him for sharing his wisdoms on various aspects of life.

I am grateful for the opportunity to collaborate with Professor Edwin Hancock on the work concerning polarisation imaging. Indeed this work has contributed a substantial part to this thesis. I thank him for his suggestions and advice during our meetings over the last few years.

I would also express my gratitude to the members of the Spectral Imaging and Source Mapping research group at National ICT Australia (NICTA) for their technical support. In particular, my thanks go to Dr. Ran Wei whose has put camera systems together for the acquisition of the data used in my thesis.

Also, I would like to thank staff and student fellows at the College of Engineering and Computer Science and National ICT Australia (NICTA) for creating a friendly and helpful professional environment. Specifically, my many thanks go to Marie Katselas for her great effort in administering the paper work related to my PhD candidature.

I would like to acknowledge the important role of my friends who have made the chapter of my life in Canberra enjoyable and memorable. My appreciation goes to the
family of Harry Hoang, Hannah Tran and Charlie Hoang for their warm hospitality and true friendship. I would also like to thank Debdeep Banerjee, John Lim, Akshay Asthana, Novi Quadrianto, Abhinav Dhall and Jyoti Joshi for their companion during trips and meals and their entertaining stories. I also feel fortunate to form a close bond with Khoa Pham, Anh Pham and Vu Pham, who treat me like a family member. My life would have been less meaningful without these experiences. In addition, I appreciate my friendship with fellow residents at Fenner Hall, Graduate House and the University House of the Australian National University (ANU).

Lastly, I would like to acknowledge the financial support of the Australian Government and National ICT Australia (NICTA) in providing me with an Australian Postgraduate Award and other funds to undertake this study.
Abstract

This thesis addresses the issues of spectrum representation, photometric invariants and shape recovery in multispectral and hyperspectral imagery. We focus on the development of algorithms to recovery shape and photometric invariants from spectral images captured from a single view and the design of a compact spectrum representation for the purpose of recognition, reconstruction and storage of spectra.

Our first contribution is a novel representation of reflectance spectra for imaging spectroscopy, based on the control points resulting from the interpolation of B-Spline curves to multispectral and hyperspectral reflectance data. Since this interpolation is based upon a knot removal scheme in the parameter domain, the representation can exploit the local support of Splines so as to recover a compact representation robust to noise and local perturbation of reflectance data. In addition, this representation permits the manipulation of image data by the use of numerically stable algorithms and methods commonly used to model smooth curves in graphics. The representation presented here also allows pattern recognition and computer vision tasks to be effected on spectra of dissimilar lengths. Furthermore, our Spline-based representation can be applied not only to single spectral reflectance curves but also multispectral and hyperspectral images by providing a common basis for spatially varying spectral reflectance over various materials.

Secondly, we address the problem of photometric invariance in multispectral imaging making use of an optimisation approach based upon the dichromatic model. In this manner, we cast the recovery of the illuminant spectra and the surface reflectance spectra, the shading and the specular factors in a structural optimisation setting. To facilitate the recovery, we make use of the spectral information provided by multispectral imaging and the structure of image patches to formulate an objective function combin-
ing the dichromatic error and the smoothness priors for the surfaces under study. The objective function is quite general, allowing the enforcement of alternative smoothness constraints; and the optimisation framework can be extended in a straightforward manner to trichromatic settings. Moreover, the objective function is convex with respect to the subset of variables to be optimised in each alternating step of the minimisation strategy. This gives rise to an optimal closed-form solution for each of the iterations in our algorithm.

Finally, we address the simultaneous recovery of the shape and refractive index of an object from a spectro-polarimetric image captured from a single view. Here, we focus on the diffuse polarisation process occurring at dielectric surfaces due to sub-surface scattering and transmission from the object surface into the air. The diffuse polarisation of the reflection process is modelled by the Fresnel transmission theory. We present a method for estimating the azimuth angle of surface normals from the spectral variation of the phase of polarisation. Moreover, we estimate the zenith angle of surface normals and refractive index simultaneously in a well-posed optimisation framework. We achieve well-posedness by introducing additional constraints to the problem, including the surface integrability and the material dispersion equation. This yields closed-form solutions to both the zenith angle and the refractive index in each iteration.
## Contents

### Abstract

<table>
<thead>
<tr>
<th>1 Introduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Contributions</td>
</tr>
<tr>
<td>1.1.1 Spectrum Representation</td>
</tr>
<tr>
<td>1.1.2 Photometric Invariants</td>
</tr>
<tr>
<td>1.1.3 Shape and Refractive Index Recovery</td>
</tr>
<tr>
<td>1.2 Organisation of the Thesis</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2 Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Spectral Imaging and Applications</td>
</tr>
<tr>
<td>2.2 Pattern Recognition in Spectral Imagery</td>
</tr>
<tr>
<td>2.3 Photometric Invariants</td>
</tr>
<tr>
<td>2.3.1 Colour Constancy</td>
</tr>
<tr>
<td>2.4 Reflectance Models</td>
</tr>
<tr>
<td>2.4.1 Physics-based Models</td>
</tr>
<tr>
<td>2.4.2 Phenomenological Models</td>
</tr>
<tr>
<td>2.4.3 Empirical models</td>
</tr>
<tr>
<td>2.5 Shape and Photometric Invariants from Polarisation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3 Spectral Imagery</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Radiometric Definitions</td>
</tr>
<tr>
<td>3.2 Image Formation</td>
</tr>
<tr>
<td>3.2.1 The Bidirectional Reflectance Distribution Function (BRDF)</td>
</tr>
</tbody>
</table>
3.2.2 Reflectance Modelling ................................. 37
3.2.3 Image Irradiance ..................................... 40
3.2.4 Colour Responses .................................... 40
3.3 Trichromatic Visualisation of Spectral Images .......... 41
  3.3.1 Planckian Illuminant Colour .......................... 43
  3.3.2 MacBeth Colour Simulation ............................ 46
  3.3.3 Colour Rendering of Spectral Images ................. 49
3.4 Concluding Remarks ..................................... 52

4 B-Spline Spectrum Representation .......................... 53
  4.1 Spline-based Representation ............................. 57
    4.1.1 B-Splines ........................................... 57
    4.1.2 Representation of a Reflectance Spectrum ............... 59
    4.1.3 Cost-Optimal B-Spline Representation .................. 60
    4.1.4 Representation of a Collection of Spectra ............... 63
  4.2 Implementation Issues ................................ 65
    4.2.1 Knot Removal ....................................... 66
    4.2.2 Classification Features .............................. 70
  4.3 Experiments ............................................. 71
    4.3.1 Reconstruction of Reflectance Spectra ................. 71
    4.3.2 Skin Recognition .................................... 79
    4.3.3 Skin Biometrics ..................................... 90
  4.4 Concluding Remarks ..................................... 96

5 Reflection Parameter Recovery .............................. 97
  5.1 Recovery of the Reflection Model Parameters ............ 99
    5.1.1 The Dichromatic Reflection Model .................... 100
    5.1.2 Objective Function .................................. 102
    5.1.3 Homogeneous Patch Selection ......................... 103
    5.1.4 Optimisation Procedure ............................. 105
  5.2 Imposing Smoothness Constraints ........................ 112
    5.2.1 Robust Shading Smoothness ......................... 114
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.2</td>
<td>Curvature Consistency</td>
<td>114</td>
</tr>
<tr>
<td>5.3</td>
<td>Adaptation to Trichromatic Imagery</td>
<td>116</td>
</tr>
<tr>
<td>5.4</td>
<td>Experiments</td>
<td>117</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Illumination Spectrum Recovery</td>
<td>118</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Skin Recognition and Material Clustering</td>
<td>133</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Specularity Removal</td>
<td>139</td>
</tr>
<tr>
<td>5.5</td>
<td>Concluding Remarks</td>
<td>143</td>
</tr>
<tr>
<td>6</td>
<td>Shape and Refractive Index Recovery</td>
<td>145</td>
</tr>
<tr>
<td>6.1</td>
<td>Preliminaries</td>
<td>147</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Polarisation of Light</td>
<td>147</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Polarisation Upon Diffuse Reflection</td>
<td>152</td>
</tr>
<tr>
<td>6.1.3</td>
<td>Fresnel Transmission Ratio and Phase of Polarisation</td>
<td>154</td>
</tr>
<tr>
<td>6.2</td>
<td>Shape and Refractive Index Recovery</td>
<td>155</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Decomposing Polarisation Images</td>
<td>157</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Azimuth Angle and Phase of Diffuse Polarisation</td>
<td>159</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Zenith Angle and Refractive Index</td>
<td>162</td>
</tr>
<tr>
<td>6.3</td>
<td>Optimal Zenith Angle and Refractive Index</td>
<td>167</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Recovery of the Zenith Angle</td>
<td>168</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Recovery of the Refractive Index</td>
<td>172</td>
</tr>
<tr>
<td>6.4</td>
<td>Experiments</td>
<td>176</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Synthetic Data</td>
<td>177</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Real-world Imagery</td>
<td>188</td>
</tr>
<tr>
<td>6.5</td>
<td>Concluding Remarks</td>
<td>195</td>
</tr>
<tr>
<td>7</td>
<td>Conclusion</td>
<td>197</td>
</tr>
<tr>
<td>7.1</td>
<td>Summary of Contributions</td>
<td>197</td>
</tr>
<tr>
<td>7.1.1</td>
<td>Spectrum Representation</td>
<td>198</td>
</tr>
<tr>
<td>7.1.2</td>
<td>Reflection Parameter Recovery</td>
<td>199</td>
</tr>
<tr>
<td>7.1.3</td>
<td>Shape and Refractive Index from Polarisation</td>
<td>200</td>
</tr>
<tr>
<td>7.2</td>
<td>Future Work and Discussion</td>
<td>202</td>
</tr>
<tr>
<td>A</td>
<td>List of Publications</td>
<td>205</td>
</tr>
</tbody>
</table>
CONTENTS

Bibliography 207
List of Figures

2.1 A pair of OKSI visible and near-infrared hyperspectral cameras. ................. 9
2.2 A schematic diagram of the structure of a spectral image. ......................... 10

3.1 The solid angle subtended by a patch $dA$ at a particular point is the projection of the patch onto a unit sphere centred at that point. ......................... 30
3.2 The amount of energy transmitted by a patch into an infinitesimal region of solid angle $d\omega$ along the direction $(\theta, \phi)$ is proportional to the radiance, the patch area $dA$, the angle $\theta$ between the normal vector $\vec{N}$ and the propagation direction, the solid angle and the transmission time. ......................... 32
3.3 Notation in Section 3.2. ................................................. 34
3.4 Reflection geometry. The incident light travels from a source with a solid angle $d\omega_i$ from the direction $\vec{L}$, impinging upon the surface which reflects it in the viewing direction $\vec{V}$. The right-handed local coordinate system has the origin located at the surface point under study and the surface normal aligned with the $z$-axis. In this coordinate system, the incident and viewing directions are parameterised by the polar coordinates $(\theta_i, \phi_i)$ and $(\theta_o, \phi_o)$, respectively. $\theta_i$ and $\theta_o$ are the zenith angles while $\phi_i$ and $\phi_o$ are the azimuth angles of these directions. .......................................................... 36
3.5 Color matching functions and camera spectral sensitivity responses ....... 43
3.6 Estimated colour temperature for the range $[1000^\circ K, 20000^\circ K]$ corresponding to the three camera models including Canon 10D, Nikon D70 and Kodak DCS420, and the human eye as measured by Stiles and Burch’s colour matching functions. .......................................................... 44
3.7 Reflectance spectra of five colour patches. ........................................... 46
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.8</td>
<td>The colours observed by the three cameras Canon 10D, Nikon D70 and Kodak DCS420 and the human eye as measured by Stiles and Burch.</td>
<td>47</td>
</tr>
<tr>
<td>3.9</td>
<td>CIE-XYZ plots for the five colours under study.</td>
<td>48</td>
</tr>
<tr>
<td>3.10</td>
<td>Spectral radiance of the sun and spectral reflectance of human skin and a leaf.</td>
<td>49</td>
</tr>
<tr>
<td>3.11</td>
<td>Rendering of a head with the skin reflectance of a Caucasian male under bright sunlight</td>
<td>50</td>
</tr>
<tr>
<td>3.12</td>
<td>Rendering of a fern with the reflectance of a leaf under sunlight.</td>
<td>51</td>
</tr>
<tr>
<td>4.1</td>
<td>Notation in Chapter 4.</td>
<td>56</td>
</tr>
<tr>
<td>4.2</td>
<td>Left panel: A third-degree B-Spline curve (in blue) is interpolated to the discrete data points ( Q_i, i = 0, \ldots, 5 ). The control points of the curve are depicted as ( P_i, i = 0, \ldots, 5 ). Note that the control points at both ends coincide with the first and last data points. Right panel: The basis functions ( N_{i,3}, i = 0, \ldots, 5 ), of the B-Spline curve in the left panel are plotted with respect to the independent parameter ( t ). The horizontal line below the graph shows the knot positions, which determine the local support of the basis functions, i.e. where the basis functions assume non-zero values.</td>
<td>58</td>
</tr>
<tr>
<td>4.3</td>
<td>B-Spline curves (in green) with 34 knots and 30 control points resulting from our knot removal algorithm performed on the discrete spectral reflectance samples of human skin (left) and a leaf (right), which is depicted as red scattered dots. The original data contain 88 samples per spectrum, in the range of 400–750 nm.</td>
<td>69</td>
</tr>
<tr>
<td>4.4</td>
<td>Plots of sample StellarNet reflectance spectra. From left to right: the original and reconstructed reflectance of human skin, a leaf and plastic. From top to bottom: the reconstruction results for 12, 18, 24 and 30 control points of the B-Spline representation, as compared to 4, 6, 8 and 10 Gaussian components, respectively.</td>
<td>73</td>
</tr>
<tr>
<td>4.5</td>
<td>Absolute reconstruction error (ARE) for the StellarNet spectrometer reflectance spectra versus the representation length for both the Spline representation and the Gaussian Mixture in [3].</td>
<td>76</td>
</tr>
</tbody>
</table>
4.6 Absolute and relative reconstruction error for sample hyperspectral images. First column: pseudo colour images. Second and third columns: Absolute (ARE) and relative reconstruction error (RRE) for our Spline-based representation, where pure white in the error maps corresponds to $2 \times 10^{-2}$ and 10% in the second and third columns, respectively. Fourth and fifth columns: Absolute (ARE) and relative reconstruction error (RRE) yielded by the Gaussian Mixture descriptor, where pure white in the error maps corresponds to $10 \times 10^{-2}$ and 30% in the fourth and fifth columns, respectively. 

4.7 The skin spectrum recognition accuracy for the reflectance spectra captured by a StellarNet spectrometer with respect to the descriptor length of our descriptor and the alternatives.

4.8 Sample face images used to train an SVM classifier for skin segmentation, shown at 670 nm.

4.9 The overall skin classification rates (CR), the correct skin detection rates (CDR) and the false detection rates (FDR) achieved by several spectral descriptors on the multispectral image database, for various descriptor’s lengths. The Spline-based descriptor is recovered by enforcing a common knot vector and the common wavelength coordinates of control points between the training and testing reflectance spectra.

4.10 Skin probability maps of a sample subject. First column: the original input image rendered in pseudo trichromatic colour. The remaining columns, from left to right, correspond to the skin probability maps delivered by the Spline-based descriptor, the Gaussian Mixture descriptor, LDA, PCA and the input reflectance spectra as discriminant features. The brightness on the skin probability maps indicates the predicted likelihood of being skin. The rows, from top to bottom, correspond to the descriptor lengths of 12, 18, 24 and 30.
4.11 Skin probability maps for sample representative subjects of different ethnic
groups. Column (a): input images rendered in pseudo colour. Columns
(b)–(e): skin probability maps delivered by 12-element descriptors (Spline-based, Gaussian Mixture, LDA and PCA, respectively). Column (f): the skin probability maps delivered by the input reflectance spectra consisting of 30 bands. Columns (g)–(j): the skin probability maps delivered by 30-element descriptors (Spline-based, Gaussian Mixture, LDA and PCA respectively). The brightness on the skin probability maps indicates the predicted skin likelihood.

4.12 Skin type association maps of three subjects belonging to the Caucasian, Indian and Oriental groups, listed from top to bottom. The red, green and blue values correspond to the likelihood that each pixel belongs to the Caucasian, Indian and Oriental groups, respective. First column: input images rendered in pseudo colour; Remaining columns, from left to right: Skin type association maps for the Spline-based descriptor, the Gaussian Mixture, LDA, PCA and the raw reflectance spectra.

4.13 Visualisation of the ethnic group association of test subjects in a 3D coordinate system. The sub-figures show the plots of the ethnicity association probabilities yielded by (a) the third-degree B-Spline descriptor with 19 knots and 15 control points; (b) the Gaussian Mixture descriptor with 5 components; (c) the 15 principal LDA components of reflectance spectra; (d) the 15 principal PCA components of reflectance spectra; (e) raw reflectance spectra.

5.1 Notation in Chapter 5

5.2 The pseudo-colour rendering of a spectral image of a human subject captured in the visible range (430–720nm). The image is overlaid with the rectangular boundaries (coloured in red) of all the homogeneous patches detected by the patch selection strategy described in Section 5.1.3.

5.3 Accuracy versus number of patches used of our illuminant estimation method on the multispectral face image database, in degrees, compared to Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s method. The results for both the visible (left) and near-infrared ranges (right) are shown.
5.4 Ground truth illuminant spectra and those estimated by our method and the
alternatives from the image of a human subject in the dataset illuminated
by a high-oblique light direction (from above and to the left of the camera
center). Here we show the estimated illuminant spectra in both the visible (left
column) and near-infrared (right column). From top to bottom: the spectra
estimated using 1, 5, 10, 40 and 50 patches from the image. 122

5.5 Accuracy versus the number of patches used for our illuminant estimation
method on the multispectral outdoor image database, in degrees, compared to
Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s method. The
results for both the visible (left) and near-infrared ranges (right) are shown. 124

5.6 Ground truth illuminant spectra and those estimated by our method and the
alternatives from the image of an outdoor scene. Here we show the estimated
illuminant spectra in both the visible (left column) and the near-infrared (right
column) ranges. From top to bottom: the spectra estimated using 1, 10, 20,
30 and 40 patches from the image. 125

5.7 Accuracy of the estimated illuminant spectra versus the standard deviation
of Gaussian noise. The vertical axis shows the angular deviation of the esti-
mated spectra from the corresponding ground-truth, while the horizontal axis
shows the standard deviation of Gaussian noise as the percentage of maxi-
mum brightness of the imagery. The performance for the visible spectrum
is shown in the left-hand image, while that corresponding to the infra-red
spectrum is shown on the right-hand panel. 126

5.8 Accuracy versus number of image patches for our illuminant estimation method
with $\alpha = 10000$ on the RGB images synthesized from the multi-spectral face
imagery, in degrees, compared to Finlayson and Schaefer’s method (F & S)
and Sunshine et al.’s approach. The figures include results for simulated RGB
images as captured with (a) Stiles and Burch’s colour matching functions (b)
a Canon 10D camera sensor (c) a Nikon D70 camera sensor. 129
5.9 The accuracy of our illuminant estimation method with $\alpha = 100$ on the RGB images synthesized from the multi-spectral outdoor image database, in degrees, compared to Finlayson & Schaefer’s method (F & S) and Sunshine et al.’s method, versus the number of patches used. The figures include results for simulated RGB images as captured with (a) Stiles and Burch’s colour matching functions [34] (b) a Canon 10D camera sensor and (c) a Nikon D70 camera sensor. ................................................................. 131

5.10 Accuracy of our illuminant estimation method with $\alpha = 1000$ for the Mondrian and specular datasets with 8-bit and 16-bit dynamic ranges described in [10]. Our method is compared to that of Finlayson and Schaefer (F & S) and Sunshine et al. ................................................................. 132

5.11 Skin segmentation results. Top row: sample training images for skin recognition, with labelled skin regions (with red borders) and non-skin regions (in blue borders). Second row: the test images of the corresponding subjects, captured under a different illumination direction. The third to last row are the skin probability map obtained using different features. Third row: obtained using the reflectance estimated given the ground-truth illuminant spectrum. Fourth row: yielded using the reflectance estimated given the estimated illuminant spectrum. Fifth row: yielded using the reflectance obtained by normalising radiance spectra by the ground-truth illuminant spectrum. Sixth row: yielded using the raw radiance of the input images. Seventh row: recovered making use of the top 20 LDA components of the raw radiance spectra. .... 135

5.12 Material clusters, with each material marked by a different shade of gray. Top row: A band of the input images, shown at 670 nm. Second row: material clustering maps, obtained with the clustering feature being the reflectance estimated given the ground-truth illuminant. Third row: material clustering maps resulting from the use of the reflectance feature recovered with an estimated illuminant spectrum as input. Fourth row: material clustering maps, using the reflectance obtained by normalising the input radiance image by the ground-truth illuminant spectrum. .................. 139

5.14 A comparison between the highlight removal results of our method and that by Mallick et al. [108]. First column: the input image. Second and third columns: the diffuse and specular components resulting from our method. Fourth and fifth columns: the diffuse and specular components yielded by Mallick et al.’s method.

6.1 Notation used throughout Chapter 6.

6.2 The polarisation of an electromagnetic wave. (a) The polarisation of light is caused by the rotation of its electric field (\(\vec{E}\)) along its propagation direction \(z\). The electric field can be decomposed into the two orthogonal harmonic components \(\vec{E}_x\) and \(\vec{E}_y\) with magnitudes \(E_{0x}\) and \(E_{0y}\). (b) The projection of the trajectory for the front of \(\vec{E}\) as projected onto the \(x-y\) plane perpendicular to the propagation direction.

6.3 (a) The superposition of a unpolarised wave \(E_{un}^\alpha\) with an elliptically polarised one \(E_{pol}^\alpha\). The elliptically polarised wave can be further decomposed into two component vectors \(\vec{E}_x\) and \(\vec{E}_y\) whose lengths oscillate sinusoidally with magnitudes \(E_{0x}\) and \(E_{0y}\). The shape and orientation of the major and minor axes of the ellipse are determined by the ratio \(\frac{E_{0x}}{E_{0y}}\) and the phase difference between \(\vec{E}_x\) and \(\vec{E}_y\). (b) The transmitted wave through a polariser with the transmission axis direction \(\vec{P}\) oriented at an angle \(\theta\) with respect to the \(x\)-axis.
6.4 (a) Polarisation upon diffuse reflection from a dielectric surface with a normal vector \( \vec{N} \) and refractive index \( \eta \). The incident light waves penetrate the surface, scatter inside the dielectric body and finally refract through the material-air boundary. The electric fields of the waves incident on and transmitted through the material-air interface are represented by vectors rotating in a plane perpendicular to the propagation direction. These electric field vectors, in turn, can be decomposed into two orthogonal components which are parallel and perpendicular to the plane of emittance, which contains both the surface normal and the emittance direction. The vectors \( \vec{E}_{i\parallel} \) and \( \vec{E}_{i\perp} \) are the components of the internal incident field on the material-air surface boundary before refraction. The vectors \( \vec{E}_{T\parallel} \) and \( \vec{E}_{T\perp} \) are the components of the wave emerging from the surface after refraction. (b) The emitted electric field, where the surface normal \( \vec{N} \) is aligned with the y-axis. \( \vec{E}_{T\parallel} \) vibrates along the y-axis because it is in the plane of reflection. The perpendicular component \( \vec{E}_{T\perp} \) lies in the x-axis and \( \vec{P} \) is the orientation of the polariser’s transmission axis.

6.5 The transmitted radiance through a linear polariser varies sinusoidally with respect to the polariser orientation angle.

6.6 The Fresnel transmission ratio for several refractive indices and zenith angles in the range of \([0, \frac{\pi}{2}]\). (a) The true Fresnel transmission ratio function. (b) The approximating function.

6.7 The approximation error for the Fresnel transmission ratio plotted for refractive indices between 1 and 3 and zenith angles in the range of \([0, \frac{\pi}{2}]\).

6.8 Synthetic images of our test shapes rendered with the refractive index of Polystyrene under five different illumination conditions. The rows, from top to bottom, correspond to the combinations of illumination directions \( L_3, L_4, L_5, L_2 + L_4 \) and \( L_1 + L_3 \). The images are rendered in pseudo trichromatic colours synthesized from the multispectral radiance of the polarisation component oriented at 45° with respect to the horizontal axis of the image coordinate system.
6.9  Pseudo trichromatic rendering of synthetic spectral-polarimetric images of eight shapes synthesized with the refractive index of Polystyrene under the frontal light source direction. Each row, from top to bottom, shows the polarisation component oriented at 0°, 30°, 45°, 60°, 90°...

6.10  The estimated needle maps of eight different shapes as compared to the corresponding groundtruth. First row: ground truth needle maps. Second, fourth and sixth rows: The needle maps recovered from images illuminated under the illumination directions L₃, L₄ and L₅, respectively. Third, fifth and seventh rows: The error of the estimated needle maps for the illumination directions L₃, L₄ and L₅, respectively. ...

6.11  The estimated shading maps of eight different shapes as compared to the corresponding groundtruth. First row: The ground truth shading maps. From the second to the fourth row: The shading maps recovered from images illuminated under the illumination directions L₃, L₄ and L₅, respectively. ...

6.12  The estimated surface depths of eight different shapes as compared to the corresponding groundtruth. First row: The ground truth depth. From the second to the fourth row: The depth maps recovered from images illuminated under the illumination directions L₃, L₄ and L₅, respectively. ...

6.13  Shape estimation and rendering for real-world images. First column: The 45° polarisation component images rendered in trichromatic pseudo-colours. Second column: Needle maps. Third column: Shading maps. Fourth and fifth columns: Two-view stereo rendering of the input shapes under L₃. In the fourth row, the view direction is 10° to the left of the original view. In the fifth row, the view direction forms an angle of 10° to the right of the original view. ...

6.14  A novel view of the objects shown in Figure 6.13, rendered under the light directions L₃ (first row) and L₂ + L₄ (second row). ...
6.15 First and third columns: Pseudo trichromatic colour rendering of real-world objects from the frontal view under $L_3$ and $L_2 + L_4$, respectively. Second and fourth columns: The trichromatic colour rendering of ground-truth unpolarised multispectral images corresponding to the lighting conditions in the first and third columns.

6.16 Refractive index spectra of a number of selected regions in the input images. First column: The $45^\circ$ polarisation component images rendered in pseudo-colour, with the selected regions indicated by rectangular bounding boxes. Second to fifth columns: The mean and standard deviation of the estimated refractive index spectra over the pixels in each selected region, plotted in colours matching those of the region boundaries depicted in the first column.
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Mean absolute reconstruction error (ARE) for the StellarNet dataset.</td>
<td>75</td>
</tr>
<tr>
<td>4.2</td>
<td>The absolute and relative reconstruction error for our hyperspectral reflectance imagery.</td>
<td>77</td>
</tr>
<tr>
<td>4.3</td>
<td>The mean and standard deviation of the classification accuracy over 20 random trials on the dataset of spectra acquired by a StellarNet spectrometer for various feature combinations from the 34 knots and 30 control points of a third-degree B-Splines descriptor.</td>
<td>80</td>
</tr>
<tr>
<td>4.4</td>
<td>The skin spectrum recognition accuracy (in %) on the Spectrometer dataset for several spectral reflectance descriptors, with respect to the variation in the descriptor length. The descriptors under consideration are the $y$-coordinates of the B-Splines representing individual training reflectance spectra (Spline-1); the $y$-coordinates of the B-Splines sharing a common knot vector and common $x$-coordinates, that have been extracted from the entire training set of reflectance spectra (Spline-2) ; the Gaussian Mixture descriptor proposed in [3] (Gaussian Mix.,); principal components recovered by performing LDA and PCA on the spectra; and the raw reflectance spectra.</td>
<td>82</td>
</tr>
<tr>
<td>4.5</td>
<td>The means of the overall skin classification rates (CR), the correct skin detection rates (CDR) and the false detection rates (FDR) achieved by several spectral descriptors, over seven trials on the multispectral image database.</td>
<td>85</td>
</tr>
</tbody>
</table>
4.6 Percentage of individuals classified into the Caucasian (C), Indian (I) and Oriental (O) ethnic groups. The diagonal elements for each sub-table spanning columns 3–5 present the overall classification rates of ethnic groups at the individual level, whereas the off-diagonal elements show the “confusion” between these groups per descriptor under study.

4.7 Percentage of image pixels classified into the Caucasian (C), Indian (I) and Oriental (O) ethnic groups. The diagonal elements for each sub-table spanning columns 3–5 present the overall classification rates for ethnic groups at the pixel level, whereas the off-diagonal elements show the “confusion” between these groups per descriptor under study.

5.1 Robust kernels and influence functions.

5.2 Accuracy versus the number of patches used for our illuminant estimation method on the multispectral face image database captured under both the visible and near-infrared spectra, in degrees, compared to Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s method.

5.3 Accuracy, in degrees, versus the number of patches used for our illuminant estimation method on the multispectral outdoor image database captured under both the visible and near-infrared spectra compared to Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s approach.

5.4 Accuracy of our illuminant estimation method with $\alpha = 10000$ on the synthesized RGB face image database, in degrees, compared to Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s approach. We show the results on RGB face images synthesized from the multispectral face imagery for the Stiles and Burch’s colour matching functions (rows 2–4), the spectral sensitivity response functions of a Canon 10D (rows 5–7) and a Nikon D70 camera (rows 8–10).
5.5 Accuracy of our illuminant estimation method with $\alpha = 100$ on the synthesized RGB outdoor image database, in degrees, compared to Finlayson & Schaefer’s method (F & S) and Sunshine et al.’s approach. We show the results on RGB outdoor images synthesized from the multispectral outdoor imagery for the Stiles and Burch’s colour matching functions (rows 2–4), the spectral sensitivity response functions of a Canon 10D (rows 5–7) and a Nikon D70 camera (rows 8–10).  130

5.6 Accuracy of our illuminant estimation method with $\alpha = 1000$ on the Mondrian and specular datasets reported in [10], in degrees, compared to Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s approach. The accuracy is measured in degrees.  132

5.7 Accuracy of several skin pixel recognition methods, each using a different reflectance-based feature as input for classification.  138

6.1 The accuracy of the recovered surface normals for several illumination directions computed as the average absolute angular difference (in degrees) between the estimated surface normal directions and the ground truth.  185

6.2 The accuracy of surface depth reconstruction for several illumination directions. This is computed as the mean absolute difference between the estimated and the ground-truth depth, with both the depth maps being normalised to the range between 0 and 1.  186

6.3 Means and standard deviations of the Euclidean angular difference (in degrees) between the estimated refractive indices and the ground truth, across all the reported materials.  187

6.4 Means and standard deviations of the mean squared error between the estimated absolute values of the refractive indices and the ground truth, across all the reported materials.  187

6.5 The angular deviation (in degrees) between the spectral reflectance images rendered for the frontal viewing direction and the ground truth images. The mean and standard deviation of these errors across pixels are reported for each image.  195
Chapter 1

Introduction

The development of spectral image sensing technologies has made it possible to capture wavelength-indexed image data in hundreds or thousands of bands across a broad range of the electromagnetic spectrum. This is due to the ability of multispectral and hyperspectral sensing devices, filter-sets and multiplexed lights to provide a means of capturing an information-rich representation of reflectance and radiance data. This large volume of data offers opportunities to develop computational algorithms for scene and object understanding based on the reflection geometry and wavelength-dependent properties of the physical world, such as reflectance, colour, refractive index and other photometric invariants.

Originally developed as a remote sensing tool, multispectral and hyperspectral imaging have been traditionally used in aerial imaging [72, 152, 159]. More recently, spectral imaging has found many opportunities and challenges for terrestrial applications thanks to its capability of capturing a large amount of information across a broad range of wavelengths, including the visible and near infrared segments of the spectrum. This technology has tremendous potential for a wide range of applications in colour reproduction [140], material recognition [56], shape recovery [81], medicine, agriculture, manufacturing, forensics and an ever increasing array of other applications. The wavelength-dependency captured by this imaging modality adds dimensions of colour information to the traditional primary values captured by trichromatic sensors. Moreover, narrow-band spectral values of illuminants and surface reflectance allow greater accuracy for colour representation than traditional trichromatic imagery. For
CHAPTER 1. INTRODUCTION

This reason, spectral imaging has been employed in the areas of digital archiving of art works [12, 17, 63], textile production [74, 154], printing [189] and high dynamic range imaging [25, 64]. In addition, the ability of spectral imaging to distinguish metamers, i.e. matching apparent colours with different spectral power distributions, has benefited applications in material identification [98, 153].

1.1 Contributions

This thesis is concerned with both the study of shape and photometric invariant recovery from spectral imagery and the development of a compact spectrum representation. The work presented here has potential applications in areas of shape analysis, image rendering and material and scene recognition. For instance, the recovered shape could be treated as input to three-dimensional shape and object recognition algorithms and benefit rendering applications with novel material appearance under various illumination conditions. In addition, with the recovered photometric invariants in hand, the statistical distribution of their discriminative features can be learnt. These spectral features may be extracted and represented using our spectrum representation for the purpose of scene analysis based on material classification.

1.1.1 Spectrum Representation

To take full advantage of the information captured by spectral imagery, we need to address the issue of representing its high-dimensional spectral domain for efficient recognition and storage. A number of approaches to modelling reflectance spectra attempted to fit a mixture of basis functions to discrete spectral samples [3, 111]. Although these approaches can represent spectra with a few basis functions within a reasonable fitting error, they do not contemplate the utility of these representations for recognition tasks.

Thus, we commence by presenting a compact representation of spectra that captures their overall shape. This is ever more important since multispectral and hyperspectral imaging systems can produce vast amounts of image data whose efficient processing and storage can become burdensome. In this respect, it is necessary to strive for compactness to gain advantages of efficiency and storage. Further, to facilitate
the analysis and comparison of spectra acquired by different instruments, we aim to
develop a common representation of spectra captured by spectral sensors of various
spectral resolutions and ranges. Moreover, such a representation should aim to be effi-
cient in terms of its descriptive power with respect to dimensionality. With the above
properties, the spectral representation to be developed shall allow effective and effi-
cient pattern recognition procedures to be performed and accurate reconstruction of a
continuous spectrum from discrete spectral samples.

1.1.2 Photometric Invariants

With the spectrum representation above, one may wish to extract meaningful patterns
from spectral photometric invariants. However, the observed image of a scene is a
composition of factors including object geometry, reflectance and illumination [13]. In
many cases, machine vision employs only a subset of these factors as a cue for pattern
recognition. Since these factors coexist and collectively influence the input imagery,
there is an obvious need for separating the intrinsic characteristics of the material, ge-
ometry and illumination for higher-level scene analysis. For material identification and
segmentation, we wish to extract photometric properties that provide discriminative
information of the observed material and are invariant to changes in the illumination
condition and the object shape. Examples of photometric invariants include reflectance
ratio [122], moment invariants [151], image ratio [84], chromaticity distribution [105]
and chromaticity [103].

Hence, we address the simultaneous recovery of the illumination power spectrum
and the material reflectance, the shading factor and the specular coefficient in the scene
from a single spectral image. The analysis is performed on individual images, there-
fore avoiding the need for cumbersome acquisition of multiple images. We under-
take a purely physics-based approach based on the dichromatic reflection model [147].
Specifically, we cast the recovery of the above parameters as a minimisation of the
rendering error of the dichromatic reflection model with respect to the input spectral
image. Making use of the geometric regularity within local image regions, we can
reformulate the optimisation problem with fewer variables and solve it by an iterative
approach. We render the problem well-posed by including a surface smoothness error
term as a regulariser. Since the objective function can be shown to be quadratic, we can derive a closed-form solution which is iteration-wise optimal.

1.1.3 Shape and Refractive Index Recovery

The third contribution of this thesis aims to simultaneously recover the surface normals and the refractive index from spectral images. To make this problem well-posed, we exploit polarisation in addition to the spatial and spectral information of spectral imagery. The objective of this work is closely related to the recovery of the dichromatic reflection model parameters described above. Specifically, the surface normals are related to the object shape, which in turn determines the shading factor. Both of refractive index and reflectance are examples of a broad range of photometric invariants.

Although the polarisation of light has long been well-studied in astronomy [62], optics [18, 110] and crystallography, it has just been introduced as an imaging modality over the last two decades. The physics foundation of polarisation imaging is based on the wave theory developed by Huygens (1629–1695) to explain the rectilinear, omnidirectional propagation of light. The creation of polarisation capture devices was initially inspired by biological mechanisms for detection and navigation tasks performed by animals such as Mantis shrimps, fish and bees. With recent advances in camera technologies, these polarisation effects can be captured by polarisation cameras [182, 183, 186]. The development of portable, low-cost and fast polarisation camera sensors has found the application of polarisation imaging in areas such as target detection [60, 142] and material property recovery [184].

Given the attention given to polarisation in the Computer Vision community, it is somewhat surprising that there is little work on the use of polarisation methods for the simultaneous estimation of surface orientation and photometric invariants, such as the index of refraction. In fact, most of the methods for surface orientation recovery rely on the assumption that the index of refraction is known [4, 6, 7, 8, 118, 119, 120, 143]. On the other hand, material classification methods [32, 184, 180, 179] have often neglected the wavelength-dependence of the index of refraction. In our work, we propose a new constraint to render the recovery problem well-posed, based on the spectral variation of spectro-polarimetric images captured from a single viewpoint. In contrast to
alternative methods reported elsewhere in the literature, our work concentrates on the use of Fresnel reflection theory, the material dispersion equations and the surface integrability constraint so as to estimate the shape and refractive index simultaneously from polarisation images acquired from a single viewpoint. The work presented here assumes minimal knowledge of the illumination conditions, the material under study or the object surface orientation.

To the best of our knowledge, this is the first time a method for estimating both surface normals and refractive index from spectro-polarimetric imagery has been proposed in the Computer Vision literature. Refractive index is an important surface characteristic which can be used to quantify subsurface structure. Our study makes this information available without the need for carefully calibrating optical bench measurements. Moreover, the setup is simple, requiring only a single view. More importantly, the method does not assume any knowledge of the illumination condition and index of refraction of the surface under study. This is a major advantage of the method over existing approaches which require either accurate measurements of shape or refractive index, or complicated instrument setups. Here, we propose an optimisation framework that permits the recovery of surface normals and refractive index simultaneously. To do this, we render the problem well-posed by drawing on the integrability constraint and the physics of both Fresnel reflection theory and the material dispersion. This permits the use of an iterative procedure to find an approximately optimal solution to the above optimisation problem. Further, our iterative approach is computationally efficient due to the use of closed-form solutions for the recovery of the zenith angle of surface normals and the refractive index.

1.2 Organisation of the Thesis

The remainder of this thesis is organised as follows. Chapter 2 reviews literature relevant to pattern recognition, shape recovery and photometric invariants for spectral imagery. This chapter focuses on the relationship between spectral imaging, photometric invariants, reflectance modelling, shape recovery and polarisation. Chapter 3 presents an overview of the image formation model of spectral imagery and links it to trichro-
matic imaging and colour reproduction. In Chapter 4, we present our novel method to represent the reflectance and radiance spectra in a compact form. This representation is not only applicable to individual spectra, but also extendable to multispectral and hyperspectral images. The work in this chapter is of particular relevance to Chapters 5 and 6 since this representation can be applied to the photometric invariants recovered by the methods discussed in these latter chapters for pattern recognition purposes. Subsequently, Chapter 5 presents an approach to spectral photometric invariant recovery, aiming to simultaneously estimate the material-intrinsic reflectance, the illumination power spectrum, the surface geometry and the specularity from a single spectral image. Chapter 6 studies the problem of simultaneous recovery of shape and refractive index of object materials from spectro-polarimetric imagery as input. The recovery approach is applicable to a wide range of materials by making use of the dispersion of the material refractive index over wavelengths. Furthermore, the method is insensitive to the illumination and does not make prior assumptions on the shape under observation. It is worth noting that the formulations in each of the Chapters 3, 4, 5 and 6 employ a separate set of notation. In each Chapter, we provide a summary of symbols for reference purposes. Lastly, Chapter 7 summarises the contributions of this thesis and discusses possible future research directions of the reported work.
Chapter 2

Literature Review

In this chapter, we aim to review the literature relevant to this thesis. In Section 2.1, we commence with an overview of the recent developments of spectral imaging devices and their applications. Since this thesis aims to make sense of the physics involved in the image formation process, including the illumination condition, the shape and the material properties of objects in the scene captured by these sensors, the review focuses on prior work regarding the recovery and recognition of these properties. In Section 2.2, we review techniques to extract features from the rich information inherent in spectral imagery for pattern recognition tasks. In relation to this, we also consider methods for representing spectral images as descriptors for image reconstruction and classification purposes. Subsequently, in Section 2.3, we review previous work on using photometric invariants, i.e. features that are invariant to illumination conditions and surface geometry, for object and material recognition in trichromatic and multispectral imagery. In relation to photometric invariants, we review Color Constancy methods, which aim to recover material-intrinsic reflectance from imagery captured under varying illumination conditions. In Section 2.4, we consider reflectance models related to shape analysis and recognition tasks. Finally, in Section 2.5, we provide an overview of polarimetric imaging and review methods for recovering shape and photometric invariants from polarisation information captured in both trichromatic and spectral images.
2.1 Spectral Imaging and Applications

Multispectral and hyperspectral imaging sensors were originally designed and used in airborne and spaceborne vehicles for aerial imaging applications. In remote sensing and computer vision, much of the earlier effort concentrated on automatic detection and classification of materials, such as earth resources. The work by Chang [31] addresses an array of techniques aiming at the recognition and analysis of material spectral signatures for remote sensing applications. Healey and Slater [72] developed a maximum likelihood algorithm for material identification that is invariant to illumination and atmospheric conditions. In [152], the authors synthesized the radiance spectra of a particular material over a wide range of conditions to classify the material of 3D objects. In a related work, Suen and Healey [159] presented an optimisation approach to the recovery and detection of linear mixtures of materials.

A variety of multispectral and hyperspectral image sensing methods have been developed over the past decades. The majority of passive sensors fall into three categories, including Liquid Crystal Tunable Filters based on birefringence, Acousto-Optical Tunable Filters based on diffraction, and Fabry-Perot filters based on optical interference. Miller and Hoyt [115] introduced the Liquid Crystal Tunable Filter (LCTF) as an optical instrument to filter spectral bands electronically with no moving parts. Subsequently, the LCTF was employed by Hardeberg et al. [65] to acquire device-independent multispectral color images. Using a different technique, Acousto-Optical Tunable Filters (AOTF) select a specific wavelength of light by applying radio-frequency (RF) acoustic waves to compress and relax an optically anisotropic crystal [38, 61, 66]. As a consequence, this mechanism tunes the refractive index of the crystal, creates diffraction and separates a specific wavelength from the broad spectrum of incident light. In the last category, Fabry-Perot devices [85] commonly involve a Fabry-Perot interferometer, which consists of two parallel planar surfaces whose inner sides are coated with highly reflective materials. Light enters the cavity and undergoes multiple reflections between its surfaces, creating constructive interference if the transmitted beams are in phase or destructive interference otherwise. As a result, certain wavelengths are transmitted with much more power than the others, subject to a relationship between the incident angle, the thickness of the material between the sur-
2.1. SPECTRAL IMAGING AND APPLICATIONS

Figure 2.1: A pair of OKSI visible and near-infrared hyperspectral cameras.

faces and its refractive index. As opposed to passive sensing methods, active sensing approaches capture spectral images of scenes by controlling and altering the illumination spectrum. Hauta-Kasari et al. [68] manipulated the light source spectrum by controlling the transmittance pattern on a Liquid Crystal panel. In [129], a method to capture hyperspectral images through multiplexed light is presented.

These research efforts have gradually been transformed into commercial ground-based multispectral and hyperspectral imaging systems. For instance, the Hyperspectral Image Intensified Camera System \(^1\) of Opto-Knowledge Systems Inc. (OKSI) can capture samples at intervals of 10 nm in the visible and near infrared range. Figure 2.1 depicts the components of a pair of OKSI hyperspectral cameras equipped with Liquid Crystal Tunable Filters (LCTFs), one of which is able to filter wavelengths in the visible range of 400 – 720nm while the other operates in the near-infrared range of 650 – 1100nm. The camera system acquires images using a pair of 12-bit monochrome Charge-Coupled Device (CCD) sensors. Each camera also includes an interchangable C-mount front lense and relay optics to avoid vignetting.

This is a typical example of a multispectra/hyperspectral camera system that cap-

\(^1\)http://www.techexpo.com/WWW/opto-knowledge/prodhiicsi.html
The ability to represent illumination and surface reflectance as a spectral signature allows greater accuracy and flexibility to interpret and distinguish colours than traditional trichromatic imagery. This is due to the robustness of spectral signatures to metamerism, i.e. trichromatic matches between materials which may be very different. The utility of spectral imaging for accurate colour reproduction in applications such as digital art archiving and analysis has been investigated. Baronti et al. analysed the colour pigments of art works using a multispectral imaging system [12]. Haneishi et al. studied the optimal filter design for accurate digital archiving of art works [63]. In Computational Photography, hyperspectral imaging has been identified as a promising tool to reveal chemical and biological features of surfaces for realistic rendering and
2.2 Pattern Recognition in Spectral Imagery

As mentioned in Section 2.1, due to its ability to capture the reflectance variation across hundreds or thousands of wavelengths, spectral signatures of materials are more robust to metamerism than trichromatic colours. Even though trichromatic imagery can be used to classify materials based on colours with a certain degree of success, it is often error-prone when dealing with metamerism. As a result, spectral reflectance has been used in machine vision for material identification as an alternative to trichromatic colours. Hence, we review approaches to representing features extracted from spectral photometric invariants in general, and spectral reflectance in particular, for recognition purposes.

In spectral imagery, each pixel is associated with a spectrum which can be viewed as a vector in a high dimensional space. Due to this high dimensional nature of spectral data, a large number of samples are required to sufficiently capture the class statistics. When the number of sample spectra is limited, the classification of spectral imagery suffers from the well-known “curse of dimensionality” phenomenon. For this reason, it is desirable to reduce the dimensionality of hyperspectral data by mapping it onto a low-dimensional feature space without losing separability information [86].

To fulfill this purpose, many classical algorithms in pattern recognition and machine learning have been naturally borrowed and adapted so as to perform feature extraction and classification [98]. Commonly used techniques such as Principle Component Analysis (PCA) [88], Linear Discriminant Analysis (LDA) [58], Projection Pursuit [87], and decision boundary feature extraction [99] treat the full spectra as input vectors in a high-dimensional space, where the dimensionality is given by the number of bands. Principal-component analysis assumes that the input data take the form of a normal distribution, whose parameters are described by the mean vector and the covariance matrix of the distribution. This assumption is problematic because it treats the entire dataset as belonging to the same distribution. Meanwhile, Linear Discriminant Analysis [58] finds a linear projection that maximises the ratio of
between-class to within-class scatter. Further, projection pursuit aims to deal with the circumstance in which there are only a limited number of training samples. Instead of performing computation at full dimensionality, it selects potentially interesting sub-space linear projections by maximising or minimising a function called the projection index. This projection index is the measure of class separability and is normally based on a Bhattacharyya distance among the classes. Lee and Landgrebe [99] proposed a feature extraction method based directly on decision boundaries. Overall, several of the above methods have a high computational demand since they operate on the full dimensionality of the input data. In addition, linear subspace projections do not cope well with the case of nonlinearly separable classes.

To overcome these issues, kernel-based classifiers [41] view the classification task as an implicit nonlinear mapping from the input space onto a feature space, often with reduced dimensionality, via the optimisation of a given criterion which leads to statistically optimal solutions. Further, feeding raw reflectance spectra as input to classifiers such as Support Vector Machines (SVM) using a kernel mapping function achieves better performance than that yielded by using lower-dimensional features [55, 148]. This is not surprising, since feature extraction may potentially lose discriminant information, whereas the direct application of SVM to raw spectra can be viewed as an implicit feature selection effected through kernel mapping. Although dimensionality reduction techniques such as Principle Component Analysis (PCA) [88], Linear Discriminant Analysis (LDA) [58] and Projection Pursuit [87] may mitigate the curse of dimensionality, these methods are not able to accurately represent and reconstruct the original reflectance and radiance spectra. This is further exacerbated by the fact that spectra acquired with different sensors are often sampled at disparate spectral resolutions and spectral ranges. As a result, the captured spectra may vary in length and resolution, depending on the acquisition hardware.

An alternative to the use of raw reflectance spectra as a means for classification and recognition is the use of a reflectance descriptor robust to changes in illumination, noise and object surface geometry. Recognition and identification methods have benefited from the recovery of properties which are invariant to geometric and photometric effects. For instance, Nayar and Bolle [122] presented an algorithm to compute re-
reflectance ratios from a single brightness image in just two raster scans. Subsequently, they showed that the reflectance ratio of the foreground to the background is invariant to illumination and optical parameters, from which a hash table of these ratios is built for object recognition and localisation [123]. Dror et al. [40] described a vision system that learnt the relationship between surface reflectance and certain statistics computed from grayscale images. Based on this relationship, the system was then able to classify the material of novel objects under novel illumination conditions. Slater and Healey [151] used a set of Gaussian filters to derive moment invariants for recognition. Jacobs et al. [84] employed image ratios for comparing images under variable illumination. Lin and Lee [105] used an eigenspace of chromaticity distribution to obtain illumination direction, illumination color and specularity-invariants for three-dimensional object recognition. Lenz et al. [103] used perspective projections in the canonical space of colour signals to separate intensity from chromaticity so as to recover a three-dimensional colour descriptor.

There has also been growing interests in the utility of a spectral image reflectance descriptor for the purposes of recognition and image reconstruction. Within this context, Angelopoulou [1] computed a descriptor based on the gradient of reflectance spectra of the materials in the scene. In another work, Angelopoulou et al. [3] used spectrophotometric data to model skin colour using several sets of basis functions including Gaussians and their first derivatives, wavelets and PCA. The aim of these representations is to capture the prominent “W” pattern with a central wavelength at around 560 nm in the reflectance spectrum of human skin. In [111], Marimont and Wandell computed a set of linear basis functions which yields the minimal approximating error for surface reflectance spectra and illuminant spectra. Prior to this work, Maloney had validated the use of low-dimensional linear models for representing reflectance spectra through a number of evaluations on empirical measurements of surface reflectance [109]. More recently, Lansel et al. [155] proposed a dictionary-based sparse representations of reflectance which can outperform linear models in reflectance spectrum recovery.
2.3 Photometric Invariants

One of the pre-processing steps toward obtaining illumination invariants is to detect and remove specularity from images. In this area, there has been a large bulk of work which analyses the specular effects exhibited by shiny or rough surfaces in trichromatic imaging. A number of these methods are based on the dichromatic reflection model [147], which describes the observed colour of an object as a linear mixture of the diffuse (body) component and the specular (highlight) component. Based on this model, Klinker et al. [91] visualised the colour distribution observed on surfaces made of the same material as a skewed T-shape with two limbs in the RGB colour space. One limb represents purely diffuse points while the other corresponds to highlight points. Using this model, their reflection component separation algorithm fits a skewed T-shape to the colour distribution of each image area, thus determining the diffuse and specular vectors for the area. Subsequently, it computes the diffuse and specular component at each pixel by means of projection onto the corresponding limb. By assuming a linear model of surface reflectance, Lin and Shum [106] separated the diffuse and specular reflection components of the dichromatic model using two photometric images. For textured surfaces, Tan and Ikeuchi [161] employed the maximum chromaticity to separate the dichromatic components using a single image without the need for colour segmentation. The method, however, may cause a diffuse colour shift due to the arbitrary choice of the diffuse maximum chromaticity.

There have also been several attempts to remove specularity from images of non-Lambertian surfaces. For instance, Brelstaff and Blake [26] used a thresholding strategy to identify specularity on real-world surfaces. Ragheb and Hancock [135] developed a probabilistic framework for specularity subtraction which used the Torrance and Sparrow model to account for the distribution of specular reflection and highlights. Note that most of the work mentioned above either require colour segmentation [91], make assumptions of the surface material [106] or rely on parameter tuning [26, 135, 161]. Moreover, they are not readily applicable to spectral images, as there can be tens or even hundreds of bands for each pixel. Thus, any local, pre-processing or postprocessing operation must be exercised with caution and in relation to neighbouring spectral bands so as to prevent spectral signature distortion.
In Computer Vision, works have been often devoted to the derivation of photometric invariants from images of matte or Lambertian surfaces. In [33], Chen et al. proposed an analytical probability distribution of the image gradient as a function of the surface geometry and reflectance. From this distribution, they observed that the direction of image gradient is invariant to the illumination direction. Wolff and Fan [185] also utilised the direction of the image intensity gradient as a photometric invariant under varying illumination to identify elliptical, hyperbolic and parabolic points in the diffuse reflection image of smooth objects. This work permits the segmentation of a smooth object surface into the three above types of curvature, independent of the pose of the object with respect to the viewpoint and the surface albedo. In [93], Koenderink and van Doorn found a number of structural angles on the isophotes, i.e. contours of equal brightness, of Lambertian surfaces, which are invariant to the light source position. Slater and Healey [150] used physical models of image formation to derive local color invariants that capture information about the spectral reflectance distribution of an object. Being intrinsic to the surface of interest, these invariants provide substantial discriminative power for identifying surfaces in cluttered scenes. Weiss [178] formulated a maximum-likelihood estimation problem to separate the reflectance image from the illumination image given an image sequence captured under varying illumination. More recently, Zickler et al. [192] introduced an illumination-dependent colour space through a rotation of the original RGB colour space. Using this colour space and the dichromatic reflection model [147], they obtained specularity and shading invariants for Lambertian surfaces. However, this approach assumes prior knowledge of the illuminant power spectrum and mostly applies to only homogeneous, dichromatic surfaces.

There has also been research on the development of photometric invariants for non-Lambertian reflectance surfaces. Narasimhan et al. [121] formulated a scene radiance model for the class of “separable” Bidirectional Reflectance Distribution Functions (BRDFs), which can be used to separate the model into material, object shape and lighting terms. Based on this model, they derived geometric and material invariants as rational functions of the material, shape and lighting parameters recovered from several image acquisitions. Despite being effective, the application of this method to
multispectral imagery is somewhat limited since they are either constrained to trichromatic imagery or rely on the closed form of the Bidirectional Reflectance Distribution Function (BRDF). Belhumeur and Kriegman [15] stated and proved a proposition on the subspace formed by monochrome images of an object with an arbitrary shape and reflectance distribution function, seen under all possible lighting conditions. Although this subspace structure suggests approaches to illumination invariant recognition, its construction requires a large number of lighting conditions and it does not lend itself to physically meaningful properties.

Specific to multispectral imagery, Healey et al. [72, 151, 159] addressed the use of photometric invariants for material classification and mapping in aerial imaging in the presence of photometric artifacts induced by atmospheric effects and changing solar illumination. These works concentrated on the utility of spectral reflectance as a photometric invariant for the sake of material analysis. In [158], a method was presented for hyperspectral edge detection. The method is robust to photometric effects, such as shadows and specularities. In [1], a photometrically invariant approach was proposed based on the derivative analysis of spectra, which is intrinsic to the surface albedo. Nonetheless, since the analysis in [1] was derived from the Lambertian reflection model, it is not applicable to specular reflections. In [2], the author derived a method to detect specular highlights in multispectral images by making use of the spectral derivative of the Fresnel reflection coefficient. In [57], an orthogonal subspace projection method for specularity-invariant representation for hyperspectral images was presented. Fu and Robles-Kelly [55] proposed a method to select the most discriminant band ratios for multi-class material classification. In another work [56], the authors employed spectral absorption features as an invariant to shading and specularity for material identification in both terrestrial and remotely sensed hyperspectral imaging data. Their approach selects the most discriminant absorption spectral segments between the materials under study based on Canonical Correlation Analysis (CCA) and the maximisation of the $\alpha$-divergence [141].
2.3. PHOTOMETRIC INVARIANTS

2.3.1 Colour Constancy

Being invariant to the illuminant power spectrum, spectral reflectance is directly related to a broad class of photometric invariants. Since the irradiance image captures both the material-intrinsic reflectance and the illuminant colour, the former is often confounded by the latter and other factors such as shadows. Thus, it is desirable to separate the material reflectance and the illuminant power spectrum from the irradiance image. In this respect, the matter of estimating the illuminant power spectrum and the material-intrinsic reflectance image is closely related to the large body of work in Colour Constancy. In machine vision, Colour Constancy is the ability to resolve the intrinsic material colour from trichromatic images captured under varying illumination conditions. The research on computational Colour Constancy branches in two main trends, one of them relies on the statistics of illuminant and material reflectance, the other is drawn upon the physics-based analysis of local shading and specularity of the surface material.

In the statistics-based approaches, the colour of input images is often correlated against a collection of known illuminant chromaticities, such as those of Planckian light sources or black-body radiators. A few of these employ Bayes rule [22, 23] to compute the best estimate from a posterior distribution by standard methods such as maximum a posteriori (MAP), minimum-mean-squared error (MMSE) or maximum local mass (MLM) estimation. The illuminant and surface reflectance spectra typically take the form of a finite linear model with a Gaussian basis. A well-known instance of this category is the Colour by Correlation method [11, 47], where a correlation matrix is built for a set of known plausible illuminants to characterise all the possible image chromaticities that can be observed. From this correlation matrix, one can easily obtain the maximum-likelihood solution for the illuminant colour corresponding to a given colour image. Gamut mapping methods [45, 48, 52], instead, gather the statistics of surface colours illuminated by a reference light source by taking the convex hull of the observed image colours. The rationale behind gamut mapping is to establish a linear mapping from the colour gamut of a given image to a canonical one. As a result, the illuminant colour can be recovered by the inverse mapping.

Note that all the above methods require training on a variety of material and il-
luminant colours. Apart from those, a number of Colour Constancy algorithms can be applied to single images with no pre-processing steps and no prior colour statistics gathered from a large number of images. In an early work, the Grey-World hypothesis [28] assumes that the spatial average of surface reflectances in a scene is achromatic, i.e. illuminant spectra can be estimated by taking the average of the sensor responses in the image. Meanwhile, the White-Patch approach [114] assumes the existence of a white material in the scene. To generalise the previous two methods, Finlayson and Trezzi [51] hypothesized that the average scene reflectance is a shade of gray. On the other hand, the Grey-Edge hypothesis [170] states that the average edge derivative in an image is achromatic.

In contrast to statistics-based approaches, physics-based colour constancy analyses the physical processes by which light interacts with matter for the purpose of illuminant and surface reflectance estimation. The two famous corner stones of physics-based colour constancy are Land’s Retinex theory [96, 97] and the dichromatic reflection model [147]. Land’s Retinex theory, through a series of psychophysical measurements of human colour sensation, has shown that this sensation is attributed to three different kinds of cones in the retinal-cortical system, which respond to the short, medium and long wavelengths in the visible spectrum. Each of these kinds of cones forms a separate lightness image of observed colours to enable a colour processing mechanism independent of the spectral composition of the incident light. Although some of the early implementations of the theory did not meet the performance of human colour constancy [24], the Retinex theory has inspired several related computational approaches to colour constancy [90].

On the other hand, the dichromatic model describes reflected light as a combination of body reflection and specular reflection, therefore treating the illumination estimation problem as an analysis of highlights from shiny surfaces [91, 92, 106, 162]. Based on this theory, the colours of all the pixels with the same material span a two-dimensional subspace, i.e. a dichromatic plane, of the colour space. One of the basis vectors of the dichromatic plane is the illuminant colour vector, whereas the other one is the material diffuse colour vector. In a scene composed of two or more different dichromatic materials, all of the induced dichromatic planes intersect at the illuminant
colour vector. Making use of this property, several authors have proposed illumination estimation techniques by computing the intersection of dichromatic planes [49, 164] or introducing additional constraints such as assumed chromaticities of common light sources [50].

2.4 Reflectance Models

In computer vision, the modelling of surface reflectance is a topic of pivotal importance for purposes of shape analysis and pattern recognition. In relation to photometric invariants, a number of physically meaningful parameters of reflectance models are material-intrinsic and therefore potentially useful for pattern recognition tasks. In shape from shading, a classical approach usually relies on the minimisation of a data error term with respect to an image irradiance equation based on a reflectance model.

Reflectance models are designed to capture the relationship between the light energy reflected from a surface, the illumination conditions, the surface geometry and the material properties under study. In other words, the reflectance, which is the fraction of incident light reflected from the surface, is only dependent on the geometric and material properties. The geometric properties include, but are not limited to, the roughness scale, the distribution, structure and shape of the constituent micro-facets of the surface. Furthermore, the spectral distribution of surface reflectance is characteristic of the material and could provide useful information for material classification.

Traditionally, reflectance models can be classified into the following three broad categories, including empirical, physics-based and phenomenological (semi-empirical). The first category comprises models that are efficient for the purpose of image synthesis and rendering in computer graphics [70, 130]. At the other extreme of the spectrum, physics-based models employ light scattering theories for modelling surface reflectance. In this category, Kirchhoff’s theory of the scattering of electromagnetic waves was first adopted by Beckmann to develop physics-based reflectance models for slightly rough or very rough surfaces [14]. Lastly, semi-empirical or phenomenological models are developed to explain optical phenomena, such as the off-specular peak reflection [166] or Fresnel reflection theory [181].
2.4.1 Physics-based Models

The Kirchhoff’s electromagnetic theory allows a class of physics-based reflectance models to be derived. The earliest work on modelling surface reflectance as a wave scattering process based on this theory was undertaken by Beckmann [14]. Here, the Kirchhoff integral of scalar diffraction theory relates the wave scattering of light to a Fresnel reflection coefficient, the local slope and the position of the reflecting point. The theory models surface geometry with the variance of the surface height distribution and the autocorrelation length, which is the spacing between surface peaks. However, since the model relies on the evaluation of the Kirchoff integral, it is intractable in closed form. Further, it breaks down when either the surface roughness or the scattering angle is large.

Subsequent work in reflectance modelling aimed to overcome the failure of Beckmann-Kirchhoff model at large scattering angles due to energy absorption, self-shadowing and multiple scattering effects. Several authors [171, 67] attempted to modify the geometric factor in Beckmann’s model to improve the model prediction with respect to experimental scattering data for rough surfaces with large incident and scattering angles. Ogilvy [126] added the Lambert’s cosine of the incident angle to the original geometric factor of the Beckmann-Kirchhoff model. Nieto and Garcia [125] suggested a method to avoid the prediction of the Beckmann-Kirchhoff model from becoming infinite at grazing viewing angles. He et al. [70] related directional and uniform diffuse reflection to a Fresnel term that depends on the wavelength, the incident angle, the surface roughness and the refractive index of the material under study. Recently, Ragheb and Hancock [136] replaced the geometric term in the Beckmann-Kirchhoff model with a Fresnel correction term to provide accurate predictions for moderately rough surfaces.

2.4.2 Phenomenological Models

One of the simplest and most widely used reflectance models in Computer Vision and Computer Graphics is perhaps the Lambertian model [95]. Although being simple, this model has successfully predicted the reflectance of a number of smooth surfaces. Being known as Lambert’s cosine law, the diffuse reflection from a surface only de-
2.4. REFLECTANCE MODELS

pends on the incident angle between the light direction and the surface normal, but not the viewing angle. In addition, the reflected radiance is proportional to the illuminant power and the material reflectance. Mathematically, the reflected diffuse radiance \( I(u, \lambda) \) at a location \( u \) and a wavelength \( \lambda \) is given by

\[
I(u, \lambda) = \frac{1}{\pi} L(\lambda) S(u, \lambda) \cos \theta_i d\omega_i
\]  

(2.1)

where \( S(\lambda) \) is the wavelength-dependent surface reflectance, which quantifies the fraction of incident light reflected by the surface, \( L(\lambda) \) is light source radiance, \( \theta_i \) is the incident angle and \( d\omega_i \) is the solid angle of the light source when viewed from the location \( u \). Note that the material reflectance \( S \) and the illuminant power \( L \) are modelled as wavelength-dependent terms so as to address the various colours of materials and light sources.

Suppose that the unit light and normal vectors at the location \( u \) are \( \overrightarrow{L} \) and \( \overrightarrow{N(u)} \), respectively. The incident angle \( \theta_i \) is therefore related to these two vectors as

\[
\cos \theta_i = \langle \overrightarrow{L}, \overrightarrow{N(u)} \rangle
\]  

(2.2)

So far, the Lambertian model of diffuse reflection has been incorporated into more complex reflectance models. These latter models serve to explain phenomena absent from the Lambertian model such as specularity and shadow [147, 166]. In practice, other models such as the Oren-Nayar [128] and Wolff [181] models generalise the Lambertian model for rough or non-Lambertian surfaces.

Perhaps one of the simplest reflection models that account for both diffuse and specular reflection is the dichromatic model introduced by Shafer [147]. This model assumes uniform illumination across the spatial domain of the observed scene. According to this model, surface radiance is decomposed into a diffuse component and a specular component. Consider an object with surface radiance \( I(u, \lambda) \) at pixel location \( u \) and wavelength \( \lambda \) be illuminated by an illuminant with a power spectrum \( L(\lambda) \). With these ingredients, the dichromatic model then becomes

\[
I(u, \lambda) = g(u)L(\lambda)S(u, \lambda) + k(u)L(\lambda)
\]  

(2.3)

In Equation 2.3, the first term on the right-hand side accounts for the diffuse reflection component. We note the resemblance of this term to the Lambert’s cosine law.
The shading factor $g(u)$ governs the proportion of diffuse light reflected from the object and depends solely on the surface geometry. On the other hand, the factor $k(u)$ models the irregularities of the microfacet structure that cause specularities in the scene.

In fact, for a purely Lambertian surface, the shading factor is related to the incident angle $\theta_i$ as

$$g(u) = \frac{d\omega_i}{\pi} \cos \theta_i$$

(2.4)

As before, $d\omega_i$ is the solid angle of the light source. Due to the fact that $\theta_i$ is the angle between the surface normal $\mathbf{N}(u)$ and the light direction $\mathbf{L}$, the shading factor is governed by the surface orientation with respect to the light source.

Oren and Nayar [128] generalised the Lambertian model to address the view-dependent intensity of light reflected from rough surfaces. This model fundamentally differs from the Lambertian model in that the reflected radiance of illuminated surfaces increases as the viewing direction approaches the source direction. In this model, the rough surface structure is viewed as a collection of long symmetric V-cavities to account for the shadowing, masking and interreflections between particles on the surfaces.

In an attempt to explain the non-Lambertian behaviour of a number of dielectrics such as plastic, ceramic and rubber, Wolff [181] developed a diffuse reflectance model that accounted for multiple subsurface scattering of light within the dielectric body. The reflectance model incorporates the Fresnel attenuation coefficient and the Snell law of refraction to explain the scattering process within the dielectric body and the refraction through the boundary between material body and the air. The model also explains the peak diffuse reflection effect observed when the angle between the light source and the viewing directions is larger than $50^\circ$, which deviates significantly from the Lambertian model.

A well-known model of the off-specular spike phenomenon was formulated by Torrance and Sparrow, who observed the maximum reflection at a viewing angle larger than the specular angle [166]. The model assumes that the surface is composed of small, randomly disposed, mirror-like facets. With this surface structure, the off-specular peak phenomenon is explained in terms of mutual masking, shadowing, multiple reflections and internal scattering between the microfacets. Similar to the Wolff
2.4.3 Empirical models

Although empirical reflectance models may not be based on a physics theory, they aim to fit the empirical measurements observed on real-world data. In 1924, Opik [127] designed an empirical model to estimate the reflection behaviour of the moon. In 1941, Minnaert [116] modified Opik’s model to obtain a reflectance function that was dependent on the polar angles of incidence and reflection and the surface roughness. This function was designed to obey Helmholtz’s reciprocity constraint [169], but did not originate from a theory in physics. Instead, it aimed to predict the light reflection behaviour of realistic non-Lambertian surfaces such as the moon. Phong’s model [130] achieved rendering realism by an interpolation over the vertex surface normals of polygons constituting the surface being rendered. Ward [176] introduced a physically valid, yet computationally simple model for the rendering of surfaces with anisotropic reflectance.

2.5 Shape and Photometric Invariants from Polarisation

The polarisation of light is a concept describing the distribution of its electromagnetic field at different oscillation orientations in the plane perpendicular to the propagation direction. It has long been a well-studied subject in astronomy [62], optics [18, 110] and crystallography. Although the human vision system is oblivious to polarisation, a number of animals, such as Mantis shrimps, naturally possess a polarisation vision system [113]. Biology researchers have also noticed evidence for biophysical mechanisms of polarisation coding in fish [69]. With recent advances in camera technologies, these polarisation effects can be captured by devices such as polarimeters and more recently, polarisation cameras [182, 183, 186]. In [182, 183, 186], Wolff et al. developed a liquid crystal polarization video camera with Twisted Nematic liquid crystals being
electro-optically controlled to replace the need for a mechanical rotation of linear polarisers. The development of these portable, low-cost and fast polarisation camera sensors potentially extends the applications of polarisation imaging to areas such as target detection and segmentation [60, 142] and material property recovery [184].

In image analysis, polarisation provides an important source of information for the shape and material composition of the object being observed. Shape and material properties such as refractive index, are usually co-existing factors that influence the appearance of an object to an observer and the polarisation properties of the emitted radiation. In an early work, Torrance, Sparrow and Birkebak [167] measured the specular reflectance distribution of rough surfaces for different polarisation orientations. The reflectance model attributes polarisation to specular reflection from a collection of small, randomly disposed, mirror-like facets that constitute the surface area. To fulfill this purpose, it includes a specular reflection component based on the Fresnel reflection theory and a microfacet distribution function. Reflectance models such as the Torrance-Sparrow model [166] and the Wolff model [181] are motivated by the Fresnel reflection theory. As a result, these reflectance models consider the reflected light as a combination of polarisation components parallel and perpendicular to the plane of reflection, and are applicable to polarised light sources. In these models, the material properties and the geometry of the reflection process are expressed in a single equation with multiple degrees of freedom. As a result, the simultaneous recovery of the photometric and shape parameters becomes an underconstrained problem.

To remedy the ill-posedness nature of the problem, Miyazaki et al. [120] assumed that the histogram of zenith angles for a given object was similar to that for a sphere and used this property to recover a mapping from the degree of polarisation to the zenith angle. Despite being effective, their approach is limited to only surfaces with a uniform distribution of surface normal directions. Moreover, the mapping is not necessarily consistent across different material refractive indices. The related work in [120] and [119] employed the degree of polarisation in the visible and far-infrared regions to resolve the surface orientation of transparent objects. The main drawback of the method in [119] is the need for an omni-directional diffuse illumination source, which limits its applicability in real-world settings. Moreover, the method requires measure-
ments of the energy emitted in the far-infrared spectrum, which is susceptible to thermal noise due to the heating of the object under study. Using a similar experimental setup involving a spherical optical diffuser, Saito et al. [143] were able to recover the shape of transparent objects with known refractive index. Rahmann [137] presented a method for computing the light source position and the orientation of rough surfaces through the polarisation of specular highlights. However, because the method in [137] relies on the existence of highlights reflected from flat surfaces, it cannot be applied to objects which exhibit purely diffuse reflectance. In a subsequent development, Rahmann [138] employed level sets to reconstruct the surface of an object from a single polarisation image. This was done under the assumption of weak perspective camera projection. Atkinson and Hancock [4] recovered the surface orientation from the measured diffuse polarisation of smooth dielectric surfaces. However, in their work, they assumed a known refractive index in order to estimate the zenith angle of surface normals from the degree of polarisation. Overall, methods for inferring surface orientation from polarisation images captured from a single view either assume a known refractive index or a known surface orientation distribution, or require complex instrument setup.

Despite these efforts, the recovery of object shape from a single view still remains a challenging task due to the presence of photometric artifacts and object surface discontinuities. To overcome these difficulties, the Computer Vision community has turned its attention to the use of multiple images. Rahmann and Canterakis [139] proposed a polarisation imaging method to recover the shape of specular surfaces. Their approach made use of the correspondences between the polarisation phases recovered from multiple views. They showed that three views were sufficient for surface reconstruction. Atkinson and Hancock [7] also made use of the link between the phase and the degree of polarisation for shape recovery. Using two views, they proposed a method to disambiguate the two candidates for the azimuth angle of surface normals. They also resolved the mapping from the degree of specular polarisation to the zenith angle of surface normals. However, the method in [7] assumes the homography between the views is known in order to match points across the imagery. This work was later extended in [8], where robust statistics were used to refine the correspondence estimates between the two views of an object. Miyazaki et al. [118] disambiguated the two pos-
sible zenith angles which yield the same degree of specular polarisation by physically
tilting the observed object by a small angle.

As an alternative to the use of multiple views, polarisation imaging can also be
combined with photometric stereo. Drbohlav and Šára [39] showed how to disam-
biguate surface orientations from uncalibrated photometric stereo images by altering
the polarisation angles of the incident and emitted light. Their method uses two pro-
jections of the object normals onto two planes perpendicular to the viewing direction
and the illumination direction. When combined with the integrability constraint, this
yields a method that can cope with both the bas-relief [16] and convex/concave ambi-
guities. In [6], Atkinson and Hancock disambiguated the surface normals by combin-
ing polarisation data with photometric stereo using three known light source positions.
Thilak et al. [163] presented a nonlinear least-squares estimation algorithm to extract
the complex index of refraction and the zenith angle of the surface normals from mul-
tiple images illuminated by unpolarised light sources. However, their method requires
prior knowledge of the light source positions relative to the observer. Furthermore,
it employs a polarimetric Bidirectional Reflectance Distribution Function (BRDF) of
light scattering that is limited to the case where the light source direction, the surface
normal direction and the view direction are co-planar.

It is also noted that polarisation has proven to be an effective tool in revealing ma-
terial properties from images. An early work was carried out by Wolff and Boult [184]
on classifying image regions into metallic and dielectric materials. In this work, the
ratio between the two Fresnel reflection components, i.e. the Fresnel ratio, was deemed
to characterise the material’s relative electrical conductivity, and was therefore treated
as a discriminant feature for classification. Later, the same authors developed a Fresnel
reflectance model to predict the magnitude of a polarisation component with an arbi-
trary orientation in the image plane [179, 180]. Using this model, the Fresnel ratio can
be estimated for the classification of dielectrics. More recently, Chen and Wolff [32]
employed the phase angle of polarisation for distinguishing conducting metals from
dielectric materials. Their approach hinges on the physics principle that, upon surface
reflection, metals alter the phase difference between polarisation components [18],
whereas this is not the case for dielectrics.
Chapter 3

Preliminaries on Spectral Image Formation

Color is but a sensation and has no existence outside the nervous system of living beings.

Nicholas Ogden Rood
American Physicist, 1831–1902

Before proceeding to the topics of spectrum representation (in Chapter 4), illuminant invariant recovery (in Chapters 5), and shape and photometric invariant recovery from polarisation (in Chapter 6), it is necessary to provide a background on spectral imagery and its formation process. This is because the formalism in the subsequent chapters departs from the fundamental concepts and the image formation process described in this chapter. Furthermore, it is important to provide a link between spectral imagery and trichromatic imagery. This relationship further clarifies how the physical aspects of the scene, which can be captured and recovered from spectral imagery, affect the colour perceived by a human observer or that output by a trichromatic sensor. A number of concepts involved in modelling the scene reflectance, such as light scattering and Fresnel reflection theory, are of particular relevance to the physics of polarisation (mentioned in Chapter 6) and illumination invariants (mentioned in Chapters 5). The material in this chapter also aims to clearly point out the separation between elements of the physical world such as geometry, illuminants and materials from the subjective
CHAPTER 3. SPECTRAL IMAGERY

colour sensation of a human observer and a trichromatic sensor. Moreover, the ability to produce a sensor-dependent colour image from a corresponding spectral image is a highly utilised tool in the subsequent chapters for the purposes of data generation and image display.

Therefore, the focus of this chapter is to establish an understanding of the formation process of spectral imagery and its relationship with colour imaging. To commence, we introduce relevant concepts and formulations from the areas of radiometry, photogrammetry, colorimetry, spectral imaging and reflectance modelling. Subsequently, we show how camera simulation and comparison can be effected with respect to the CIE-1931 colour standard. Here, we use the spectral response of the camera to recover RGB colours corresponding to known illuminant and material reflectance spectra. We have compared these RGB values to those computed making use of the colour matching functions proposed by Stiles and Burch [156, 157].

In colorimetry, the aim is to capture and reproduce colours to achieve perceptual accordance between the scene image and the observation by the viewer. The simulation and evaluation of this information is important to the understanding of the relation between the scene and the camera image. Moreover, the accurate capture and reproduction of colours as acquired by digital camera sensors are an active area of research which has applications in colour correction [21, 46, 77, 174], camera simulation [107] and sensor design [42].

However, the capture and reproduction of colours that are perceptually consistent with human observations is not a straightforward task. Digital cameras comprise three kinds of spectral broad-band colour sensors which provide responses for the three colour channels, i.e. red (R), green (G) and blue (B). In practice, the RGB values recorded by these spectral broadband sensors are not necessarily a linear combination of the CIE colour matching functions of human vision [34]. Rather, colours, as acquired by digital cameras, are device dependent and can be highly non-linear with respect to the CIE-XYZ colour gamut [188].

While colorimetry focuses on the accuracy of the colours acquired by the camera, spectroscopy studies aspects of the physical world such as the spectrum of light absorbed, transmitted, reflected or emitted by objects and illuminants in the scene. In
3.1. **RADIOMETRIC DEFINITIONS**

contrast with trichromatic sensors, multispectral and hyperspectral sensing devices can acquire wavelength-indexed reflectance and radiance data in tens of hundreds of bands across a broad spectral range. Making use of photogrammetry and spectroscopy techniques, it is possible to recover the spectral response of the camera under study [172]. This spectral response, together with material reflectance and illuminant radiance measurements, can be used to perform colorimetric simulations and comparisons.

The organisation of this chapter is as follows. We commence with an introduction to radiometric definitions in Section 3.1 for purposes of quantifying the amount of light energy transferred from light sources to surfaces, between surfaces and from surfaces to image sensors. In Section 3.2, we present the image formation process for spectral imagery based on the radiometric definitions introduced in the previous section. Lastly, in Section 3.3, we present the trichromatic simulation of spectral imagery for three different commercially available cameras and the human vision system through three experimental vehicles. The evaluation demonstrates that colour images can be reproduced from spectral imagery with a high level of accuracy, while being dependent on the spectral sensitivity response of trichromatic sensors.

### 3.1 Radiometric Definitions

The imagery captured by an imaging sensor is the result of the arrival of energy or photons at the sensor. Before quantifying light energy, we commence with concepts involved in the field of light measurement, or also known as radiometry. Radiometry defines terms and units that quantify the energy transfer from light sources to surface areas and from surface areas to image sensors and observers. Much of this process is due to the physical phenomena that occur during and after the impingement of photons upon surfaces.

#### 3.1.1 Radiometric Definitions

The amount of energy received by a surface depends on both the foreshortened area and the radiant energy of the light source. Here, foreshortening refers to how big a surface area looks from the point of view of the source. It accounts for the fact that
surfaces receive different amounts of light from different illumination directions and may radiate different amounts of light in different outgoing directions. Specifically, a surface patch tilting away from the source looks smaller than one that faces perpendicular to the illumination direction. As a result, surfaces tilting away from the source, i.e. receive a less amount of energy than those perpendicular to the direction of light propagation, due to their smaller areas of light reception.

### Foreshortening and Solid Angle

Foreshortening can be quantified as the area of a surface patch viewed from a light source. We capture this quantity by the notion of the solid angle subtended by the surface patch at the light source position. This notion is based on the intuition that patches that “look the same” on the unit sphere subtend the same solid angle. We illustrate the solid angle using the diagram in Figure 3.1. Formally, the solid angle subtended by a surface patch at a point is the area of the region obtained by projecting the patch onto the unit sphere centred at that point [53].

In Figure 3.1, we consider an infinitesimal patch with an area $dA$ and a distance $r$ from the centre of the unit sphere. In addition, its surface normal forms an angle $\Delta \theta$ with the line of sight. The infinitesimal solid angle $d\omega$ it subtends at the centre of the

Figure 3.1: The solid angle subtended by a patch $dA$ at a particular point is the projection of the patch onto a unit sphere centred at that point.
3.1. RADIOMETRIC DEFINITIONS

The imaging process is concerned with the distribution of light in space as a function of position and direction. To this end, we refer to radiance as the measurement of the distribution of light in space. It is defined as the power of light travelling at a point in a specified direction per unit surface area perpendicular to the travel direction, per unit solid angle [53]. Radiance is measured in Watts per square meter per steradian (W.m\(^{-2}\).sr\(^{-1}\)).

Based on this definition, we can quantify the amount of energy transmitted through a region along a particular direction. Figure 3.2 shows a planar surface patch with area \(dA\) radiating energy in a direction forming an angle \(\theta\) with the surface normal \(\vec{N}\) and a tilt angle \(\phi\) with respect to a reference axis on the patch. Assuming that the radiance from the patch is \(L(\theta, \phi)\), the amount of energy \(E_o\) transmitted through an infinitesimal region of a solid angle \(d\omega\) around the direction of propagation in time \(dt\) is

\[
E_o = L(\theta, \phi) \cos \theta dA d\omega dt
\]

Irradiance

Image sensors are often designed to measure the amount of energy arriving at the image plane as a result of reflection or scattering from different locations and directions in the scene. The amount of incoming energy to the sensor is termed irradiance. By definition, irradiance is the power of incident light per unit surface area not foreshortened, and is measured in Watts per square meter (W.m\(^{-2}\)).
Figure 3.2: The amount of energy transmitted by a patch into an infinitesimal region of solid angle $d\omega$ along the direction $(\theta, \phi)$ is proportional to the radiance, the patch area $dA$, the angle $\theta$ between the normal vector $\vec{N}$ and the propagation direction, the solid angle and the transmission time.

In fact, the amount of incoming energy is proportional to the source radiance. Specifically, we consider incident illumination with radiance $L_i(\theta, \phi)$ coming from a region of solid angle $d\omega$ at angles $(\theta, \phi)$ with respect to the local coordinate system of the patch as illustrated in Figure 3.2. The irradiance arriving at a surface patch illuminated by this incident flux is

$$E_i = L_i(\theta, \phi) \cos \theta d\omega$$  \hspace{1cm} (3.4)

### 3.2 Image Formation

We now proceed to study the process by which light is transmitted, reflected and scattered upon arrival at a surface before impinging on image sensors. Real-world surfaces exhibit a combination of these effects. A reflectance model describes the fraction of radiance energy reflected from a surface illuminated by a light source. Conventionally, this fraction is dependent on the reflection geometry, i.e. the illuminant and viewing directions and the surface normal, the power spectrum of the illuminant and the material properties such as roughness, albedo, shininess and subsurface structure. In our analy-
3.2. IMAGE FORMATION

sis, we assume that no light energy is absorbed by the surface material, to concentrate our modelling on the transmission, reflection and scattering of light at the surface and subsurface levels.

The photometric process is summarised as follows. A fraction of the light incident on the surface being observed is reflected towards the camera. This fraction depends on the scene geometry and the surface material reflectance. Subsequently, the reflected light passes through the camera lens and is focussed onto the image plane of the camera. Finally, the R, G, B values for each pixel in the image are determined by the spectral sensitivity of the R, G, B receptors of the camera to the incoming light.

We adopt the following notation for our formalism. Since the surface reflectance and the illuminant power spectrum are wavelength-dependent, we represent them as functions $S(\lambda)$ and $L(\lambda)$ with respect to the wavelength $\lambda$. In addition, let us denote the spectral sensitivity of the red(R), green(G) and blue(B) sensors with respect to the wavelength $\lambda$ as $C_R(\lambda)$, $C_G(\lambda)$ and $C_B(\lambda)$, respectively. For mnemonic purposes, in Figure 3.3, we provide the list of notation used to model the image formation process.

The remainder of this Section is organised as follows. Firstly, we describe a general model of local reflection in Section 3.2.1. Further, in Section 3.2.2, we present a reflectance model which accounts for both diffuse and specular reflection. Subsequently, we elaborate on how to integrate this model into the image irradiance equation in Section 3.2.3 and how to simulate object colour using camera sensor responses in Section 3.2.4.

3.2.1 The Bidirectional Reflectance Distribution Function (BRDF)

The most general model of local reflection is the Bidirectional Reflectance Distribution Function (BRDF). To formulate the reflectance models herein, let us consider a reflection geometry defined with respect to a right-handed local coordinate system whose origin is the surface location under study and whose z-axis aligns with the local surface normal, as shown in Figure 3.4. In this coordinate system, the illuminant and viewing directions are described by two pairs of angles $(\theta_i, \phi_i)$ and $(\theta_s, \phi_s)$, where $\theta_i$ and $\theta_s$ are the zenith angles while $\phi_i$ and $\phi_s$ are the azimuth angles of the respective directions. The solid angle of the cone of incident light rays when viewed from the
<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S(\lambda) )</td>
<td>Material spectral reflectance at wavelength ( \lambda ).</td>
</tr>
<tr>
<td>( L(\lambda) )</td>
<td>Illuminant spectral power at wavelength ( \lambda ).</td>
</tr>
<tr>
<td>( (\theta_i, \phi_i) )</td>
<td>Polar and azimuth angles of the incoming light direction.</td>
</tr>
<tr>
<td>( (\theta_s, \phi_s) )</td>
<td>Polar and azimuth angles of the outgoing direction.</td>
</tr>
<tr>
<td>( E_i(\theta_i, \phi_i, \lambda) )</td>
<td>Source irradiance.</td>
</tr>
<tr>
<td>( E_o(\theta_s, \phi_s, \lambda) )</td>
<td>Outgoing radiance.</td>
</tr>
<tr>
<td>( f(\theta_i, \phi_i, \theta_s, \phi_s, \lambda) )</td>
<td>Bidirectional Reflectance Distribution Function (BRDF)</td>
</tr>
<tr>
<td>( d\omega_i )</td>
<td>Solid angle of the source viewed from the surface location.</td>
</tr>
<tr>
<td>( g )</td>
<td>Surface roughness.</td>
</tr>
<tr>
<td>( P(\theta_i, \phi_i, \theta_s, \phi_s) )</td>
<td>Mean scattered power from the surface according to the Beckmann-Kirchoff model.</td>
</tr>
<tr>
<td>( D(\theta_i, \phi_i, \theta_s, \phi_s) )</td>
<td>Diffuse scattering power according to the Beckmann-Kirchoff model.</td>
</tr>
<tr>
<td>( G_{FC}^2 )</td>
<td>The attenuation term of the modified Beckmann-Kirchoff reflectance model.</td>
</tr>
<tr>
<td>( F(\theta_i, \eta) )</td>
<td>Fresnel reflection coefficient of a material with refractive index ( \eta ).</td>
</tr>
<tr>
<td>( I_{im}(\lambda) )</td>
<td>Spectral image irradiance.</td>
</tr>
<tr>
<td>( C_c(\lambda) )</td>
<td>Spectral sensitivity of the trichromatic sensor ( c ), where ( c \in { R, G, B } ).</td>
</tr>
<tr>
<td>( I_c )</td>
<td>Colour response for the trichromatic sensor ( c ).</td>
</tr>
</tbody>
</table>

Figure 3.3: Notation in Section 3.2.
3.2. IMAGE FORMATION

surface is annotated as \(d\omega_i\) in Figure 3.4.

Since the patch is an infinitesimal area, the source radiance \(L(\lambda)\) is almost constant over this patch. Thus, the energy flux per unit area perpendicular to the light direction is proportional to the source radiance and the solid angle, \(i.e. L(\lambda)d\omega_i\). Since the source irradiance \(E_i\) arriving at the surface is measured per unit area perpendicular to the local surface normal, we have

\[
E_i(\theta_i, \phi_i, \lambda) = L(\lambda) \cos \theta_i d\omega_i \tag{3.5}
\]

where \(\cos \theta_i\) accounts for the ratio of a foreshortened area in the incident light direction \((\theta_i, \phi_i)\) to its projection onto the surface tangent plane at the point under study.

The Bidirectional Reflectance Distribution Function (BRDF) \(f(\theta_i, \phi_i, \theta_s, \phi_s, \lambda)\) quantifies the fraction of incident light that is reflected from the surface into the air in various directions. Let the light source energy arriving at the surface in the direction \((\theta_i, \phi_i)\) be \(E_i(\theta_i, \phi_i, \lambda)\), and the radiance at the same surface point, seen from the direction \((\theta_s, \phi_s)\) be \(E_o(\theta_s, \phi_s, \lambda)\). The BRDF of the surface is the ratio of the outgoing radiance to the incoming irradiance

\[
f(\theta_i, \phi_i, \theta_s, \phi_s, \lambda) = \frac{E_o(\theta_s, \phi_s, \theta_i, \phi_i, \lambda)}{E_i(\theta_i, \phi_i, \lambda)} \tag{3.6}
\]

Note that the source irradiance, the surface radiance and the BRDF are dependent on the wavelength \(\lambda\).

Since the interaction of light with objects may incur complex phenomena, we make the following assumptions to ensure the applicability of a Bidirectional Reflectance Distribution Function (BRDF). Firstly, the radiance leaving a surface point is contributed by only light irradiance arriving at the same point on the surface. This assumption avoids the case of subsurface scattering encountered in translucent or multi-layered surfaces, where light may arrive, penetrate and re-emerge from the surface at different points. Secondly, the surface under study is not fluorescent, \(i.e.\) it does not absorb light in the short wavelength range and emit light in a longer wavelength. In other words, the reflected or scattered light at a given wavelength is attributed to incident light at the same wavelength. Thirdly, we assume that surfaces do not themselves emit internal energy and treat the light source energy separate from the surface reflectance.
Figure 3.4: Reflection geometry. The incident light travels from a source with a solid angle \( d\omega_i \) from the direction \( \vec{L} \), impinging upon the surface which reflects it in the viewing direction \( \vec{V} \). The right-handed local coordinate system has the origin located at the surface point under study and the surface normal aligned with the \( z \)-axis. In this coordinate system, the incident and viewing directions are parameterised by the polar coordinates \((\theta_i, \phi_i)\) and \((\theta_o, \phi_o)\), respectively. \( \theta_i \) and \( \theta_o \) are the zenith angles while \( \phi_i \) and \( \phi_o \) are the azimuth angles of these directions.

The BRDF in Equation 3.6 depends on four angular variables and is therefore cumbersome to directly model the characteristics of a surface. For isotropic surfaces, where the BRDF remains constant as the illuminant and viewing directions are rotated about their surface normals by the same tilt (azimuth) angle, the BRDF only depends on three variables, including the polar angles \( \theta_i \) and \( \theta_s \) and the difference of azimuth angles \( \phi_s - \phi_i \), i.e. the BRDF can be rewritten as \( f(\theta_i, \phi_i, \theta_s, \phi_s, \lambda) = f^*(\theta_i, \theta_s, \phi_s - \phi_i, \lambda) \). On the other hand, the Helmholtz reciprocity principle [169] states that the energy travelling from the light source to the surface and being reflected to the observer is equal that travelling in the reverse direction. Mathematically, this principle is expressed as

\[
f(\theta_i, \phi_i, \theta_s, \phi_s, \lambda) = f(\theta_s, \phi_s, \theta_i, \phi_i, \lambda) \tag{3.7}
\]

For isotropic surfaces, this also means that \( f^*(\theta_i, \theta_s, \phi_s - \phi_i, \lambda) = f^*(\theta_i, \theta_s, \phi_s - \phi_i, \lambda) \) or \( f(\theta_i, \phi_i, \theta_s, \phi_s, \lambda) = f^*(\theta_i, \theta_s, |\phi_s - \phi_i|, \lambda) \).
3.2. **IMAGE FORMATION**

In addition, we note that the formulation in Equation 3.6 is wavelength-dependent. Here, we have omitted the wavelength from the list of arguments for the sake of brevity. As we will discuss in Section 3.2, irradiance and radiance terms are generally wavelength-dependent, thus determining the colour of illuminants and objects in the scene.

A disadvantage of using a non-parametric BRDF model is that irradiance and radiance have to be measured for every pair of illuminant and viewing directions. To reduce the effort for data acquisition, we resort to parametric reflectance models which provide analytical expressions relating the reflection geometry, the illuminant spectrum and the material properties to the observed radiance. In Section 3.2.2, we discuss such a parametric subsurface scattering and reflection model for the sake of colour reproduction using spectral images.

### 3.2.2 Reflectance Modelling

An important part of the image formation process is the reflection and scattering of light at the surface boundary. To describe the physical phenomena involved, we adopt a parametric reflectance model to relate the diffuse scattering and specular reflection to various material properties and the reflection geometry. In this section, we are interested in simulating the appearance of surface colours as a result of surface and subsurface scattering of light.

In practice, light scattered from many natural multi-layered surfaces, including human skin and plant tissue, is composed of a large subsurface scattering component due to reflection at the bottom surface boundaries and refraction through translucent layers. Therefore, a physics-based reflectance model that accounts for refractive attenuation of incident light through several surface layers is the most appropriate model for estimating the colour of these surfaces. To model light transmission and scattering within the material body and light reflection at the surface level, we have chosen to incorporate into our framework the variant of the Beckmann-Kirchoff model proposed by Ragheb and Hancock [136]. This model combines a wave scattering theory with a Fresnel correction term [181] to represent the refraction of light through the top surface boundary.
The model in [136] makes use of two physical parameters. The first of these is the standard deviation $\sigma$ of the height variation with respect to a mean surface level. The second parameter is the surface correlation length $\tau$, which characterises the random nature of the surface profile in terms of the relative horizontal spacing between peaks and valleys of the surface.

Before further formalism, we refer to the reflection geometry depicted in Figure 3.4. As before, the reference local coordinate system is defined with its z-axis aligned to the local surface normal and its origin being the point of interest on the surface. The incident light and viewing directions are defined by the polar coordinates $(\theta_i, \phi_i)$ and $(\theta_s, \phi_s)$, respectively. For the sake of simplicity, we assume that the incident light is always in the $xz$-plane, i.e. $\phi_i = \pi$.

The physical mechanisms involved in the scattering, transmission and reflection of light are broadly attributed to two different reflection phenomena, which are specular and diffuse reflection. In [136], the authors described the contribution of these two phenomena to the mean scattered power from a surface as

$$P(\theta_i = \pi, \theta_s, \phi_s) = \rho_0^2 e^{-g} + D(\theta_i = \pi, \theta_s, \phi_s)$$

where the first term on the right-hand side represents the scattering component in the specular direction, and the second term is the diffuse scattering component.

First we draw attention to the terms contributing to specular component. Let us consider a rectangular surface patch surrounding the point of interest with an area of $A = XY$, where $X$ and $Y$ are the width and length of the patch. With these ingredients, the scattering coefficient $\rho_0$ of the surface is given by

$$\rho_0 = sinc(v_x X) sinc(v_y Y)$$

where $sinc(x) = sin(x)/x$, $v_x = k(sin\theta_i - sin\theta_s \cos \phi_s)$, $v_y = -k \sin \theta_s \sin \phi_s$. Here, $k$ is the propagation rate of the incident light, related to its wavelength $\lambda$ through the equation $k = 2\pi/\lambda$.

In addition, the term $g$ is the roughness of the surface, which is dependent on the standard deviation $\sigma$ of the local surface height as $g = \sigma^2 v_z^2$, where the quantity $v_z$ is defined as $v_z = -k(cos\theta_i + cos\theta_s)$. 


Now we focus on the modelling of the diffuse scattering component. To model the microstructure of the surface, we assume that the surface correlation function is Gaussian. Given this assumption, it has been shown that the diffuse component is

\[ D(\theta_i, \phi_i = \pi, \theta_s, \phi_s) = \frac{\pi \tau^2}{A} G^2_{FC} e^{-g} \sum_{l=1}^{\infty} \frac{g^l}{l!} e^{-\frac{\tau^2 v^2_{xy}}{4}} \]  

(3.10)

where \( v^2_{xy} = v^2_x + v^2_y \) and \( G_{FC} \) is a geometric attenuation term.

Next, we turn our attention to the modelling of the geometric term \( G_{FC} \). Later in this section, we elaborate further on the roughness factor \( \sigma \). Here, we adopt the approach of Ragheb and Hancock \[136\], where the original geometric attenuation factor in the Beckmann-Kirchoff model is replaced with that derived by Wolff \[181\]. Thus, the geometric attenuation term becomes

\[ G^2_{FC} = (1 - F(\theta_i, \eta)) (1 - F(\arcsin \left( \frac{\sin \theta_i}{\eta} \right), \frac{1}{\eta})) \cos \theta_i \]  

(3.11)

where \( F(\cdot, \cdot) \) is the Fresnel reflection coefficient given by

\[ F(\theta_i, \eta) = \frac{1}{2} \left( \frac{a - \cos \theta_i}{a + \cos \theta_i} \right)^2 \times \left( 1 + \frac{a - \sin \theta_i \tan \theta_i}{a + \sin \theta_i \tan \theta_i} \right) \]  

(3.12)

where \( a = (\eta^2 - \sin^2 \theta_i)^{\frac{1}{2}} \). In the expression above, \( \eta \) is the refractive index of the surface material.

For the sake of simplicity and computational efficiency, we adopt the approach taken in \[136\] and divide the computation of \( D(\cdot) \) into three different cases dependent on the value of the roughness term \( g \). These cases include slightly rough surfaces, i.e. \( g \ll 1 \), moderately rough surfaces, i.e. \( g \approx 1 \) and very rough surfaces, i.e. \( g \gg 1 \). When the surface is rough or very rough, i.e. \( g \gg 1 \), \( D(\cdot) \) is approximated by

\[ D(\theta_i, \phi_i = \pi, \theta_s, \phi_s) = \frac{\pi \tau^2 G^2_{FC} g}{A \sigma^2 v^2_{z}} \exp \left( -\frac{\tau^2 v^2_{xy}}{4 \sigma^2 v^2_{z}} \right) \]  

(3.13)

Note that, if the surface is moderately rough, i.e. \( g \approx 1 \), \( D(\theta_i, \phi_i = \pi, \theta_s, \phi_s) \) can be approximated with a finite number of terms of the infinite sum in Equation 3.10. Further, for cases in which the surface is slightly rough, the diffuse component can be approximated by the first term of the summation in Equation 3.10, i.e.

\[ D(\theta_i, \phi_i = \pi, \theta_s, \phi_s) = \frac{\pi \tau^2 G^2_{FC} g}{A} \exp \left( -\frac{\tau^2 v^2_{xy}}{4} \right) \]
CHAPTER 3. SPECTRAL IMAGERY

3.2.3 Image Irradiance

Having modelled the reflection and scattering from the surface under study, we now derive the image irradiance captured by the sensor. We commence by relating the Bidirectional Reflectance Distribution Function (BRDF) to the mean scattered power formulated in Section 3.2.2. Indeed, the BRDF \( f(\theta_i, \phi_i, \theta_s, \phi_s, \lambda) \) can be viewed as the product of the surface reflectance \( S(\lambda) \) and the mean scattered power \( P(\theta_i, \phi_i = \pi, \theta_s, \phi_s) \)

\[
f(\theta_i, \phi_i = \pi, \theta_s, \phi_s, \lambda) = P(\theta_i, \phi_i = \pi, \theta_s, \phi_s)S(\lambda)
\]

The surface reflectance \( S(\lambda) \) can be measured with a setting in which the incident light and viewing directions are aligned to the surface normal. In this setting, the diffuse scattering of light is governed solely by the surface geometry. Furthermore, the mean scattering power \( P(\theta_i, \phi_i, \theta_s, \phi_s) \) established by Equation 3.8 is the geometric attenuation factor of light energy scattered in the viewer's direction.

Combining Equations 3.5, 3.6 and 3.14 yields the surface radiance

\[
E_o(\theta_i, \pi, \theta_s, \phi_s, \lambda) = P(\theta_i, \pi, \theta_s, \phi_s)S(\lambda)L(\lambda) \cos \theta_i d\omega_i
\]

We assume that the camera lens has unit transmittance, i.e., the flux radiated from the surface is transmitted through the lens without any loss of energy. According to Horn [78], the irradiance reaching the image plane is given by

\[
I_{im}(\lambda) = E_o(\theta_i, \pi, \theta_s, \phi_s)S(\lambda)L(\lambda) \frac{\pi}{4} \left( \frac{d}{z} \right)^4 \cos^4 \alpha
\]

\[
= mP(\theta_i, \pi, \theta_s, \phi_s)S(\lambda)L(\lambda) \cos \theta_i \cos^4 \alpha d\omega_i
\]

where \( m = \frac{\pi}{4} \left( \frac{z}{d} \right)^4 \), \( d \) is the lens diameter, \( z \) is the focal length of the lens and \( \alpha \) is the angle between the optical axis of the camera and the line-of-sight from the surface patch to the center of the lens.

3.2.4 Colour Responses

We now proceed to model the RGB responses of the camera sensors. Since we are interested in the chromaticity rather than the intensity of the light reflected from the surface, we focus our attention on the relative spectral distribution of radiance across
3.3. TRICROMATIC VISUALISATION OF SPECTRAL IMAGES

wavelengths as an alternative to the absolute radiance value. Since the term \( m \) in Equation 3.16 only depends on the camera geometry, it is constant with respect to the spectral and angular variables of the image irradiance equation. Also, note that the derivation of trichromatic responses applies equally to each of the three colour channels. Therefore, we consider a general formulation for a particular colour channel \( c \), where \( c \in \{R, G, B\} \). Recall that the spectral sensitivity functions of the red(R), green(G) and blue(B) sensors are denoted as \( C_R(\lambda) \), \( C_G(\lambda) \) and \( C_B(\lambda) \), respectively. With this notation, the response for the colour sensor is given by

\[
I_c = \kappa_c \int_W C_c(\lambda)I_{im}(\lambda)d\lambda
\]

where \( W = [380\,nm, 780\,nm] \) is the human visible spectrum.

In Equation 3.16, the value of \( \kappa_c \) corresponds to the colour balance factor of the camera against a predetermined reference. If a white reference is used, these become the white balance factors. By balancing, the values of \( I_B, I_G \) and \( I_R \) are scaled such that a smooth surface, with unit reflectance \( S(\lambda) = 1 \) and unit scattering \( P(\cdot) = 1 \), placed perpendicular to the camera axis, \( \text{i.e.} \ \theta_i = \theta_s = \alpha = 0 \), presents a colour perceptually consistent with the reference colour \( s_c \). Thus, the primary of the colour reference is given by

\[
s_c = m\kappa_c d\omega_i \int_W C_c(\lambda)L(\lambda)d\lambda
\]

By solving the equation above for \( \kappa_c \) and substituting into Equation 3.16, we get

\[
I_c = s_c \cos \theta_i \cos^4 \alpha \times \frac{\int_W C_c(\lambda)P(\theta_i, \pi, \theta_s, \phi_s)S(\lambda)L(\lambda)d\lambda}{\int_W C_c(\lambda)L(\lambda)d\lambda}
\]

3.3 Trichromatic Visualisation of Spectral Images

Now we demonstrate the formalism of the relationship between spectral irradiance and its colour responses for a number of trichromatic sensors. In the following experiments, we consider the trichromatic sensor responses of the human eye and three camera models, including the Canon 10D, the Nikon D70 and the Kodak DCS420 models. The human eye trichromatic responses were measured as 10-degree colour matching
functions in [156, 157]. Note that the Kodak DCS420 is a discontinued model dating from the mid-nineties. Therefore, its capacity to reproduce colour with respect to the human eye is expected to be outperformed by the Nikon D70 and the Canon 10D. Nonetheless, we have included the Kodak DCS420 as a matter of comparison with the work on spectral sensitivity function measurement in [172, 173].

We commence by estimating the colour of a Planckian blackbody radiator [102] heated at various temperatures. Subsequently, we examine the simulation of standard MacBeth colour indices, which are commonly used for colour correction and calibration in digital photography. Finally, we illustrate how the image formation model above can be used for rendering purposes.

To assess the fidelity of the colours produced by these cameras, we compare their simulated output with the standard XYZ colours recovered using the colour matching functions in [156, 157], the spectral surface reflectance and the light source radiance measured with a StellarNet spectrometer. As a common practice, for comparison purposes, we plot the colours yielded by the camera spectral sensitivity functions and the colour matching functions on the CIE 1931 colour gamut [34]. To this end, we convert the RGB values into the CIE-XYZ colour space using the following linear transformation

\[
\begin{pmatrix}
X \\
Y \\
Z
\end{pmatrix} =
\begin{pmatrix}
0.4124 & 0.3576 & 0.1805 \\
0.2126 & 0.7152 & 0.0722 \\
0.0193 & 0.1192 & 0.9505
\end{pmatrix}
\begin{pmatrix}
R \\
G \\
B
\end{pmatrix}
\]  

(3.19)

To recover the spectral sensitivity function of the cameras under study, we use a procedure akin to that in [172, 173]. Narrow band illumination is obtained by passing light from a calibrated tungsten source through an OKSI Liquid Crystal Tunable filter. This narrow band light is used to image a LabSphere \(^1\) Spectralon white reference. For photometric calibration purposes, we measure the spectrum of the narrow band illuminant on the white reference making use of a StellarNet spectrometer \(^2\) whose probe is placed at a similar geometric configuration with respect to the white reference as that of the camera used in [172, 173]. In Figure 3.5 we show the colour matching functions and the spectral sensitivity functions of the Canon 10D, the Nikon D70 and

\(^1\)http://www.labsphere.com/
\(^2\)http://www.stellarnet-inc.com
3.3. TRICROMATIC VISUALISATION OF SPECTRAL IMAGES

(a) Stiles and Burch RGB colour matching functions

(b) Canon 10D

(c) Nikon D70

(d) Kodak DCS420

Figure 3.5: (a) Stiles and Burch RGB colour matching functions; (b)-(d): The RGB sensitivity functions for the Canon 10D, the Nikon D70 and Kodak DCS420 camera sensors, respectively.

the Kodak DCS420 cameras. It is worth noting that the spectral sensitivity functions for the Kodak DCS420 camera are consistent with those reported in [172, 173].

3.3.1 Planckian Illuminant Colour

As mentioned earlier, we first examine the colour temperature characteristic of a Planckian illuminant, which is given by its chromaticity as compared with a heated black body radiator. The temperature, in Kelvin degrees, at which the chromaticity of the black body radiator is equivalent to that of the illuminant under consideration is the illumi-
nant colour temperature. Colour temperature has applications in photography and is relevant to white balancing of digital cameras. Further, it has been used as a standard benchmark in colorimetry [188].

The spectral radiant exitance per unit surface area per unit solid angle for a black body radiator is governed by Planck’s law. As a function of wavelength, Planck’s law is given by

\[ E_{bb}(\lambda, T) = \frac{c_1}{\lambda^5 \left(e^{c_2/\lambda T} - 1\right)} \]  

(3.20)

where \( T \) is the temperature in Kelvin, \( \lambda \) is the wavelength variable as before and \( c_1 = 3.74183 \times 10^{-16} \) Wm \(^2\) and \( c_2 = 1.4388 \times 10^{-2} \) mK are constants.

Figure 3.6: Estimated colour temperature for the range \([1000^\circ K, 20000^\circ K]\) corresponding to the three camera models including Canon 10D, Nikon D70 and Kodak DCS420, and the human eye as measured by Stiles and Burch’s colour matching functions.
3.3. TRICROMATIC VISUALISATION OF SPECTRAL IMAGES

Making use of the spectral radiant exitance, the colour temperature captured by the camera can be computed by using \( E_{bb}(\lambda, T) \) as an alternative to the image irradiance \( I_{im}(\lambda) \). Note that the colour temperature is determined by the illuminant chromaticity. As a result, we can use Equation 3.18 to determine the colour temperature by assuming that the RGB spectral sensitivity functions of the camera are normalized to unity, \( i.e. s_R = s_G = s_B = 1 \), against a white colour reference with unit reflectance \( S(\lambda) \equiv 1 \). Moreover, if the incident light and viewing directions are aligned with the surface normal and the optical axis of the camera, \( i.e. \theta_i = \alpha = 0 \), we can simplify Equation 3.18 so as to obtain the colour of the blackbody radiator at temperature \( T \) by

\[
\begin{align*}
I_B(T) &= \frac{\int W C_B(\lambda) E_{bb}(\lambda, T) d\lambda}{\int W C_B(\lambda) d\lambda} \\
I_G(T) &= \frac{\int W C_G(\lambda) E_{bb}(\lambda, T) d\lambda}{\int W C_G(\lambda) d\lambda} \\
I_R(T) &= \frac{\int W C_R(\lambda) E_{bb}(\lambda, T) d\lambda}{\int W C_R(\lambda) d\lambda}
\end{align*}
\]

(3.21)

In Figure 3.6 we show a plot on the CIE-XYZ gamut of the colour temperatures for a black body radiator heated at temperatures ranging from \( 1000^\circ K \) to \( 20000^\circ K \). Each curve corresponds to the colour temperature yielded by each camera as compared with those obtained using the colour matching functions corresponding to the human eye’s spectral sensitivity responses.

Note that, from the figure, the curves in the temperature range of \([5000^\circ K, 6200^\circ K]\) for the three cameras are in good accordance to that yielded by the colour matching functions. This implies that all the cameras under evaluation are able to perform appropriate white balancing against natural sunlight. Further, the length of each curve across the gamut indicates the dynamic range of the corresponding camera. In particular, the dynamic range of Kodak DCS420 camera is shorter than those of the others. The second observation is that the Canon 10D and Nikon D70 curves are closer to that for the human eye than the Kodak DCS420 curve. In the cases where the colour temperature is close to its extreme values, the colour produced by Kodak DCS420 is considerably different from those produced by the other two cameras.
CHAPTER 3. SPECTRAL IMAGERY

3.3.2 MacBeth Colour Simulation

Our second experimental vehicle focuses on the estimation of colours from the MacBeth Colour Checker. This is a chart that provides a set of known colours commonly used for colour correction and white balance. Here we have used a subset of colours from the chart. These are Red (C1), Green (C2), Blue (C3), White (C4) and Orange Yellow (C5). For purposes of simulation, we measure the reflectance of each of the colour patches under study making use of a StellarNet spectrometer so as to obtain the spectra in the range between 380 and 780 nm. The reflectance spectra of the colours under study are shown in Figure 3.7.

![Reflectance spectra of five colour patches. From left-to-right: Red (C1), Green (C2), Blue (C3), White (C4) and Orange Yellow (C5).](image)

Given the reflectance spectrum for each colour patch, we estimate the RGB response for the sensors in each camera and compare them with those yielded by the colour matching functions [156, 157]. To estimate the intrinsic surface colour, we set the illumination spectrum to a uniform one across wavelengths, \( i.e. \ L(\lambda) = 1, \) for all \( \lambda \in [380 \text{ nm}, 780 \text{ nm}] \). With similar assumptions as those used in the previous section,
Figure 3.8: The colours observed by the three cameras Canon 10D, Nikon D70 and Kodak DCS420 and the human eye as measured by Stiles and Burch.

The RGB values of the colour patch under study are estimated by the equation

\[
I_B = \frac{\int W_C B(\lambda) S(\lambda) d\lambda}{\int W_C B(\lambda) d\lambda}
\]

\[
I_G = \frac{\int W_C G(\lambda) S(\lambda) d\lambda}{\int W_C G(\lambda) d\lambda}
\]

\[
I_R = \frac{\int W_C R(\lambda) S(\lambda) d\lambda}{\int W_C R(\lambda) d\lambda}
\]

(3.22)

where \( S(\lambda) \) is the measured spectral reflectance of the colour patch. Finally, the estimated RGB colour is transformed into the CIE-XYZ colour space by Equation 3.19.

In Figure 3.8, we show the estimated colours in the form of a chart indexed to trichromatic sensor and colour of reference. Note that the estimated RGB value for the patch C4 (White) is the same for the three cameras and the colour matching functions. This is not surprising and follows from the equations above. If the reflectance spectra is unity across the visible spectrum, i.e. \( S(\lambda) = 1 \) for all \( \lambda \in [380nm, 780nm] \), the RGB values are unity. In Figure 3.8, the Canon 10D and Nikon D70 cameras achieve a colour estimate close to the human eye for patch C1 (Red) and patch C5 (Orange Yellow). Also, the resulting Green (C2) and Blue (C3) colours yielded by these cameras are perceptually distinguishable from that for the human eye. The Kodak
DCS420 camera produces colours which are distinct from those for the other cameras and the human eye.

Further, we plot in Figure 3.9 the simulated colours under study for the three camera models and the human eye on the CIE-XYZ gamut. Each cluster in the figure corresponds to a different colour. Note that the qualitative results in Figure 3.8 agree with those in Figure 3.9. The Red (C1) and Orange Yellow (C5) points for the Canon 10D and Nikon D70 cameras are close to those yielded by the colour matching functions, while the Green (C2) and Blue (C3) coordinates for these two cameras are further apart from those for the human eye. This result suggests that the Canon 10D and Nikon D70 cameras can produce Red and Orange Yellow shades more accurately than Green and Blue ones. Also, the Kodak DCS420 response curves produce colour point-coordinates.
3.3. TRICHROMATIC VISUALISATION OF SPECTRAL IMAGES

(a) Spectral radiance of the sun  (b) Human skin reflectance  (c) Leaf reflectance

Figure 3.10: The radiance of the sun (3.10(a)), the spectral reflectance of human skin (3.10(b)) and a plant leaf (3.10(c)).

that are well separated from the others in each of the C1, C2, C3 and C5 clusters. This agrees with both our qualitative results in Figure 3.8 and those yielded by our colour temperature simulation.

3.3.3 Colour Rendering of Spectral Images

In this section, we demonstrate the utility of the colour estimation method presented in Section 3.2.4 for purposes of simulating the colour appearance of 3D meshes for the digital cameras under study. To this end, we have used a fern and a human head. These meshes can be viewed as a collection of planar polygons whose RGB colour is determined by Equation 3.18.

For purposes of rendering, the illuminant in the scene is considered to be a point light source. Here, the radiance $L(\lambda)$ of the illuminant is given by the sunlight spectrum. The surface reflectance spectra used here are those of human skin for the head and a plant leaf for the fern. We measured these spectra using a StellarNet spectrometer and plotted them in Figure 3.10(a). For white balancing, we have assumed a light source colour equivalent to the CIE D55 colour temperature (5500° K). At the colour temperature of 5500° K, the $s_R : s_G : s_B$ proportion for the computation in Equation 3.18 is $1:1.1040:0.9870:0.8233$.

In our implementation, the mean scattered power $P(\cdot)$ in Equation 3.18 is computed for the mesh polygons adjacent to the vertex of reference. The tangent plane at each vertex is the one perpendicular to the local surface normal. This plane is used as the
mean surface level to compute the standard deviation $\sigma$ of the height variation. For a specific wavelength $\lambda$ of incident light, the roughness level of the local area around a vertex is determined by the value of $g = \sigma^2 v_z^2$, where $v_z = -\frac{2\pi}{\lambda} (\cos \theta_i + \cos \theta_s)$. The approximation presented in Section 3.2.2 is used to compute the mean scattered power of the surface patch.

Figure 3.11 presents the rendering of the head yielded using the camera spectral sensitivity functions and the Stiles and Burch’s colour matching functions. The Canon 10D and the Nikon D70 cameras produce results that closely resemble the one computed using the colour matching functions. The colour achieved by the Kodak DCS420 cameras.

Figure 3.11: Rendering of a head with the skin reflectance of a Caucasian male under bright sunlight. Results yielded by (a) the Stiles and Burch’s colour matching functions, (b) the Canon 10D, (c) Nikon D70 and (d) Kodak DCS420 cameras.
3.3. TRICROMATIC VISUALISATION OF SPECTRAL IMAGES

Figure 3.12: Rendering of a fern with the reflectance of a leaf under sunlight. Rendering yielded by, (a) the Stiles and Burch’s colour matching functions, (b) the Canon 10D, (c) Nikon D70 and (d) Kodak DCS420 cameras.

camera has a higher saturation and tends further to the white, which makes the head appear slightly brighter.

Similarly, in Figure 3.12, we show the rendering for the fern plant with the reflectance of a leaf. The surface colour rendered with the Kodak DCS420 response curves is closer to the white reference than those produced for the Canon 10D and Nikon D70 cameras. The rendering for the latter cameras indeed has a stronger green component, which looks more natural than that for the former. However, as compared to the results in Figure 3.11, the colours for the former two cameras are perceptually different from that computed with the RGB colour matching functions.
3.4 Concluding Remarks

In this chapter, we have described the image formation process for the purpose of understanding spectral imagery. The core component of this process is the illumination and reflectance modelling as related to the interaction of light with the surface under study and the transformation of scene radiance energy into image values. An important point to stress is that narrow-band spectral images are designed to capture the physics of illuminants and materials, and the scene geometry. In contrast, broad-band trichromatic receptors built in commercial cameras transform the scene radiance spectrum into device-dependent colour values. This relationship between spectral and trichromatic images has been demonstrated by colour simulation results produced for several trichromatic digital camera sensors using techniques from photogrammetry, colorimetry and spectroscopy. The method is quite general in nature and applies to a wide variety of trichromatic digital cameras commercially available as well as the human vision system. Specifically, we have presented results involving Planckian illuminants and the Macbeth colour chart. We have also rendered surfaces using spectroscopic measurements. These results are relevant to areas such as white balancing, colour correction and photorealistic rendering.
Chapter 4

A B-Spline Representation of Spectral Signatures

I procured me a triangular glass prism to try therewith the celebrated phenomena of colours.

Isaac Newton

English physicist, mathematician, astronomer and philosopher

1643 - 1727

Previously, we provided an introduction to spectral imagery in Section 2.1 and discussed a spectral image formation model in Chapter 3. In relation to spectral imagery, we also reviewed different feature representations and classification methods for recognition purposes based on spectral photometric invariants in general and on spectral reflectance in particular. The goal of this Chapter is to develop a spectral reflectance representation that is robust to noise and local perturbation of spectra for pattern recognition purposes. In addition, the representation aims to be compact in order to facilitate efficient computation, classification and storage. We strive for an expressive representation that is capable of reconstructing a continuous spectrum akin to the original shape of the discretely-sampled spectra. The representation presented here permits analytical operations, such as derivative analysis, on the continuous spectral domain for feature extraction purposes. Moreover, this representation can be applied
equally to other spectral photometric invariants rather than reflectance, to derive compact invariant features for recognition tasks.

Thus, in this chapter, we focus on the development of a free-form representation for arbitrary wavelength-dependent spectral data. The representation presented here can reduce densely sampled reflectance spectra, which could potentially consist of hundreds of samples, to an efficient and compact form for reconstruction and classification. To this end, we turn our attention to the properties of Splines curves \[132\]. It is worth noting in passing that, although Splines are widely used in Computer-Aided Design and graphics, they do not appear to have been used for the representation of reflectance spectra. This is somewhat surprising since Splines allow an analytical representation of free-form curves using a set of control points with local support, i.e. the control points have a bounded effect on the span of the curve. By treating reflectance spectra as 2D curves, Splines become a natural choice of basis functions for the representation of arbitrary shapes of spectra.

There has been an extensive range of techniques of representing spectra that involve the decomposition of the original spectrum into frequency components. Examples of these operations include the wavelet transform, the Fourier transform and its special cases, the sine and cosine transform \[59\]. In these transforms, each basis function represents the presence of a certain frequency in the input spectrum. Albeit being effective for image compression, these transforms are cumbersome to represent imaging spectra in a compact form. This is because such a transform would lead to a large number of basis functions corresponding to various scales and positions of the peaks and valleys in the input spectrum. Using these transforms, individual spectra need to be decomposed separately, which adds overhead to the processing of spectral images with a large number of pixels. In addition, long feature vectors formed by numerous basis coefficients do not confer an efficiency advantage for recognition tasks. Further, the range of constituent frequencies may vary per spectrum, causing difficulties in enforcing a consistent basis across spectra.

As opposed to the above choices of basis functions, Splines possess a number of properties that benefit the representation of spectral data. Due to the local support property of Splines, noise corruption in the spectra only affects the representation lo-
cally, without any propagation to other wavelength-indexed sections of the spectrum. As a result, the descriptor presented here is robust to noise and local perturbations in the spectra. Hence, the local support property of Splines offers an advantage over other choices of basis functions to represent reflectance spectra, such as Gaussian functions or wavelets [3]. On the other hand, the flexibility to control the degree of the Splines and the number of basis functions delivers a good trade-off between goodness of fit and representation dimensionality.

Moreover, the descriptor permits the interpolation at any point in the spectral domain via a set of numerically stable algorithms and the continuous representation and reconstruction of spectra of dissimilar lengths captured by different sensors. This provides the added benefit of allowing operations to be effected on the continuous representation using tools from functional analysis rather than discrete approximation. This is particularly useful for absorption band detection. So far, absorption features have been traditionally used as signatures for the classification of chemicals [168] and their concentration [160]. These are detected making use of the second derivative of the spectrum with respect to wavelength index [29, 37, 43]. In these methods, derivatives are evaluated using techniques similar to finite differences, which are often inaccurate and noise-sensitive. With a continuous closed-form representation of reflectance spectra, derivative analysis can be performed analytically at any point in the spectral domain.

Our B-Spline representation involves the parameterisation of reflectance and wavelength with respect to an independent parameter. To arrive at a compact Spline-based representation, we employ a knot removal scheme in the parameter domain, subject to the goodness of fit and the representation length, \(i.e\). dimensionality. In addition, the representation presented here can be used to reconstruct the original spectra. Thus, we achieve a compact representation of image reflectance that yields high accuracy in the reconstruction of spectra. We also demonstrate the utility of this representation for material recognition and classification.

The chapter is organised as follows. Section 4.1 provides a detailed formulation of the Spline-based descriptor. We describe our knot removal scheme, which aims at minimising the reconstruction error so as to arrive at a compact representation of re-
### Notation Description

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{i,p}(t)$</td>
<td>The $i^{th}$ B-Spline basis function of degree $p$ with respect to an independent variable $t$.</td>
</tr>
<tr>
<td>$U = {u_0, \ldots, u_m}$</td>
<td>The knot vector of a B-Spline.</td>
</tr>
<tr>
<td>$(x_i, y_i)$</td>
<td>The wavelength $(x_i)$ and reflectance coordinates $(y_i)$ of the $i^{th}$ control point $P_i$, $i = 0, \ldots, n$ of a B-Spline with $n + 1$ control points.</td>
</tr>
<tr>
<td>$\lambda(t)$ and $R(t)$</td>
<td>Wavelength and reflectance interpolation functions with respect to an independent parameter $t$.</td>
</tr>
<tr>
<td>$\lambda_k, R_k$</td>
<td>The wavelength and reflectance of the $k^{th}$ sample from the input reflectance spectrum.</td>
</tr>
<tr>
<td>$l$</td>
<td>The number of samples on the input reflectance spectrum.</td>
</tr>
<tr>
<td>$t_k$</td>
<td>The independent parameter value corresponding to the $k^{th}$ wavelength-reflectance sample from the input spectrum.</td>
</tr>
<tr>
<td>$\mathbf{Y} \triangleq [y_0, y_1, \ldots, y_n]^T$</td>
<td>Our spectral reflectance descriptor for a single spectrum.</td>
</tr>
<tr>
<td>$R_{v,k}$</td>
<td>Measured reflectance at wavelength $\lambda_k$, $k = 1, \ldots, l$, of the spectrum with index $v$.</td>
</tr>
<tr>
<td>$\mathbf{R}_v$</td>
<td>The vector representation of samples from the input spectrum with index $v$, where $\mathbf{R}<em>v \triangleq [R</em>{v,1}, R_{v,2}, \ldots, R_{v,l}]^T$.</td>
</tr>
<tr>
<td>$\mathcal{R}_v(t)$</td>
<td>Reflectance interpolation function corresponding to the input spectrum with index $v$.</td>
</tr>
<tr>
<td>$(x_i, y_{v,i})$</td>
<td>The coordinates of the $i^{th}$ control point $P_{v,i}$, $i = 0, \ldots, n$, of the B-Spline curve corresponding to the input spectrum with index $v$.</td>
</tr>
</tbody>
</table>

Figure 4.1: Notation in Chapter 4.
4.1. SPLINE-BASED REPRESENTATION

Reflectance spectra. Moreover, we include a strategy to enforce a common knot vector for the representation of a collection of reflectance spectra and those in a multispectral or hyperspectral image. Section 4.2 discusses implementation issues pertaining the computation of our Spline representation. For reference purposes, we summarise in Figure 4.1 the list of symbols commonly used in this chapter. Finally, Section 4.3 presents experimental results for the tasks of spectrum reconstruction and skin recognition. Here, we explore the use of our representation for biometrics by investigating its utility for ethnic group recognition. In all the experiments, we provide performance comparison with a number of alternative descriptors elsewhere in the literature.

4.1 Spline-based Representation

In this section we present the theoretical foundation of our compact reflectance spectrum descriptor. To do this, we commence with an introduction to B-Spline curves and relate them to an interpolation operation performed on the spectrum. We then formulate the reflectance descriptor through performing knot removal so as to minimise a cost function in the parameter domain of the Spline curve. Using our knot removal strategy, we aim to reduce the number of knots and control points while fixing the B-Spline degree.

4.1.1 B-Splines

To commence, we introduce the concept of a B-Spline function. It is a polynomial function of an independent variable, with two or more dimensions. The function has support with respect to degree, smoothness and domain partition. These properties make B-Splines a flexible tool for fitting arbitrary shapes and data. The smoothness property makes the interpolating curve robust to noise. The local support property permits the modification of the curve over a given wavelength range while keeping the rest of the Spline unaffected. The degree, on the other hand, permits the representation of complex shapes by allowing the use of polynomial basis functions of different orders.

First, we require some formalism. Since reflectance spectra can be treated as functions of wavelength, we restrict our analysis to the two-dimensional case. A $p$-degree
CHAPTER 4. B-SPLINE SPECTRUM REPRESENTATION

(a) Fitting a B-Spline through discrete points

(b) B-Spline basis functions and knot vector

Figure 4.2: Left panel: A third-degree B-Spline curve (in blue) is interpolated to the discrete data points $Q_i, i = 0, \ldots, 5$. The control points of the curve are depicted as $P_i, i = 0, \ldots, 5$. Note that the control points at both ends coincide with the first and last data points. Right panel: The basis functions $N_{i,3}, i = 0, \ldots, 5$, of the B-Spline curve in the left panel are plotted with respect to the independent parameter $t$. The horizontal line below the graph shows the knot positions, which determine the local support of the basis functions, i.e. where the basis functions assume non-zero values.

B-Spline curve $C$ in $\mathbb{R}^2$ composed of $n$ segments is a function in the parameter domain $\mathcal{U}$ of the univariate variable $t$ given by the linear combination $C(t) = \sum_{i=0}^{n} N_{i,p}(t)P_i$, where $P_i = (x_i, y_i)$ are the 2D control points, and $N_{i,p}(t)$ are the $p$-degree B-Spline basis functions defined on the parameter domain [132]. The coordinates $(x, y)$ of a point on the curve are expressed in the parametric form

$$x(t) = \sum_{i=0}^{n} N_{i,p}(t)x_i$$  \hspace{1cm} (4.1)

$$y(t) = \sum_{i=0}^{n} N_{i,p}(t)y_i$$  \hspace{1cm} (4.2)

In Figure 4.2(a), we illustrate the interpolation of a B-Spline curve with degree $p = 3$ to six 2D points $Q_i, i = 0, \ldots, 5$. The control points, denoted $Q_i, i = 0, \ldots, 5$, do not lie exactly on, but are distributed along the curve to govern its overall shape. Here, we note that the control points at both ends coincide with the first and last data points.

A B-Spline is characterised not only by the control points but also by a knot vector.
4.1. SPLINE-BASED REPRESENTATION

\[ U = \{u_0, \ldots, u_m\}, \text{ where } m = n + p + 1. \]  

With these ingredients, we can define the \(i^{th}\) B-Spline basis function \(N_{i,p}(t)\) of degree \(p\) as follows

\[
N_{i,0}(t) = \begin{cases} 
1 & \text{if } u_i \leq t < u_{i+1} \\
0 & \text{otherwise}
\end{cases}
\]

\[
N_{i,p}(t) = \frac{t-u_i}{u_{i+p} - u_i} N_{i,p-1}(t) + \frac{u_{i+p+1} - t}{u_{i+p+1} - u_{i+1}} N_{i+1,p-1}(t)
\]

In Figure 4.2(b), we show the graph of the B-Spline basis functions \(N_{i,p}(t), i = 0, \ldots, n,\) of the curve in Figure 4.2(a) with respect to the independent parameter \(t\). We note that each basis function \(N_{i,p}(t)\) is a piecewise polynomial assuming non-zero values only in the interval \([u_i, u_{i+p+1})\). Therefore, it only affects the shape of the Spline in this local support. The line below the graph shows the distribution of the knot vector in the parameter domain. Note that the local support of each basis function is bounded between a pair of knot locations.

4.1.2 Representation of a Reflectance Spectrum

To formulate our B-Spline representation, we treat a spectrum as a collection of spectral samples with two coordinates \((\lambda_k, R_k)\), where \(R_k\) is the \(k^{th}\) reflectance sample at the wavelength \(\lambda_k\). Thus, the parametric form of a B-Spline curve through the spectral data points is obtained by representing the wavelength and reflectance as two functions of an independent parameter \(t\), which we denote \(\lambda(t)\) and \(R(t)\), respectively. The descriptor is then governed by a knot vector and a control point-set which minimise a cost function defined by squared differences between the measured reflectance \(R_k\) and the reflectance \(R(t)\) reconstructed according to the parametric form of the B-Spline. We depart from an initial interpolation of the input reflectance spectrum so as to arrive at a B-Spline curve that minimises the cost function through a knot removal algorithm.

Once the control points and knot vector that minimise the cost function are in hand, we proceed to construct the descriptor. We do this by selecting the features from the control point-set and the knot vector that are the most characteristic of the input reflectance spectrum. We observe that the knots and \(x\)-coordinates of each control point, \(i.e.\) the \(x_i\) variables in Equation 4.1, vary in a specific neighbourhood of the
parameter and wavelength domain where they govern the local support of the Spline. As a result, their variations across spectra may not be as related to the spectrum shape as the \( y \)-coordinates of the control points, *i.e.* the \( y_i \) variables in Equation 4.2. In other words, the variation of the \( y \)-coordinates of the control points primarily determines the general shape of the spectrum. Thus, it is expected that the \( y \)-coordinates should provide better discrimination between spectra than the knots and \( x \)-coordinates. To minimise the complexity of the descriptor in terms of vector length and computational cost, we limit our descriptor \( Y \) to the \( y \)-coordinates of the control points, *i.e.* \( Y = [y_0, y_1, \ldots, y_n]^T \).

### 4.1.3 Cost-Optimal B-Spline Representation

To optimise the choice of the interpolating B-Spline curve for reflectance spectra, we devise a cost function. Our cost function aims to approximate the input data with the lowest approximation error using the least number of control points. Ideally, given a material with reflectance \( R_k \) at each wavelength \( \{\lambda_k\}, k = 1, \ldots, l \), we aim to recover an interpolating B-Spline curve through every sampled point on the spectrum. Assuming that the curve has a control point set \( \{P_i = (x_i, y_i), i = 0, \ldots, n\} \) and a knot vector \( U \), the interpolation implies

\[
\lambda(t_k) = \sum_{i=0}^{n} N_{i,p}(t_k)x_i \tag{4.3}
\]

\[
R(t_k) = \sum_{i=0}^{n} N_{i,p}(t_k)y_i \tag{4.4}
\]

where the parameter \( t_k \in U \) corresponds to the \( k^{th} \) wavelength, *i.e.* \( \lambda_k = \lambda(t_k) \forall k \).

The cost of interpolating the points \( (\lambda_k, R_k), k = 1, \ldots, l \), using the B-Spline curve above is then given by

\[
K = \alpha \sum_{k=1}^{l} (\mathcal{R}(t_k) - R_k)^2 + (1 - \alpha)|U| \tag{4.5}
\]

where \( |.| \) denotes the length of the vector argument and \( \alpha \) is a constant between zero and unity.

Note that the first term in Equation 4.5 above is the weighted sum of squared errors in the two-dimensional space defined by the wavelength-reflectance pairs, whereas the
second term is the weighted number of knots. Thus, the optimal interpolating curve minimises the sum of squared distances \((R(t_k) - R_k)^2\), while penalising a large number of knots. This imposes a balance in the resulting curve between describing the general shape of reflectance spectra, while minimising the number of knots required to describe the spectra. This trade-off is governed by a balancing factor \(\alpha\). A small value of \(\alpha\) favours a short descriptor over one that tightly fits the shape of the original data. Also, note that, as the number of knots decreases, the interpolating curve becomes smoother. Thus, an appropriate choice of \(\alpha\) makes the descriptor less susceptible to noise while preventing over-smoothing and loss of detail.

To recover the B-Spline representation, in practice, we depart from an initial approximation of the curve to the reflectance spectrum under study. To do this, we apply the curve interpolation algorithm in [132], which employs the centripetal method of Lee [100] to recover parameter values for every control point. We then apply knot removal so as to minimise the interpolation cost introduced in Equation 4.5.

Once the initial approximation is in hand, we proceed to remove knots sequentially using a knot removal method akin to that in [177]. This knot removal strategy reduces the number of knots and control points of the interpolating B-Splines while fixing their degree. The algorithm proceeds in two passes. In the first pass, removable knots are identified. In the second pass, knots are sequentially removed and new control points are computed. Note that, although being effective, the algorithm in [177] does not automatically determine the best knot to remove in each pass, but rather assumes the candidate knot for removal to be designated as input. As a result, our knot removal procedure comprises the computation of the potential cost reduction contributed by removable knots. Once these contributions to the cost reduction are evaluated, we apply Tiller’s knot removal algorithm [177] to the selected knots.

Thus, our algorithm selects amongst the candidate knots the one that yields the maximum cost reduction. We note that, following the strategy above, the parameter \(t_k\) should be recovered for every wavelength \(\lambda_k\). This is not a straightforward task since the function \(\lambda(t_k)\) is expressed as a linear combination of the basis functions \(N_{i,p}\) given in Equation 4.3. Unfortunately, the equation \(\lambda(t_k) = \lambda_k\) cannot be solved for \(t_k\) analytically. To overcome this problem, we adopt a numerical approach in order to find
Algorithm 1 KnotRemoval\((Q, p, \alpha)\)

Require: \(Q, p, \alpha\)

\(Q\): A single spectrum with wavelength-reflectance samples denoted as \(Q = \{(\lambda_k, R_k), k = 1, \ldots, l\}\)

\(p\): The degree of basis functions

\(\alpha\): The balance factor

\(U_0, P_0\): the returned knots and control points

1: \((U_0, P_0) \leftarrow \text{Interpolate}(Q, p)\)

2: while true do

3: \(\text{Reduction}_{\text{max}} \leftarrow -1\) //Highest cost reduction

4: for all \(u \in U_0\) do

5: \((\text{flag}, U_1, P_1) \leftarrow \text{RemoveKnot}(u)\) \(\text{//flag is true if knot} \ u \text{ is removable}\)

6: if flag then

7: \(\text{SSE}_{\text{old}} \leftarrow \text{SSE}(U_0, P_0, Q)\)

8: \(\text{SSE}_{\text{new}} \leftarrow \text{SSE}(U_1, P_1, Q)\)

9: \(\text{Reduction} \leftarrow (1 - \alpha) + \alpha(\text{SSE}_{\text{old}} - \text{SSE}_{\text{new}})\)

10: if Reduction > Reduction\(_{\text{max}}\) then

11: \(\text{Reduction}_{\text{max}} \leftarrow \text{Reduction}\)

12: \(\text{candidate} \leftarrow u\)

13: end if

14: end if

15: end for

16: if Reduction\(_{\text{max}} < 0\) then

17: return \(U_0, P_0\)

18: end if

19: \((\text{flag}, U_0, P_0) \leftarrow \text{RemoveKnot}(\text{candidate})\)

20: end while

21: return \(U_0, P_0\)
an approximate solution with a reduced computational cost. In fact, it is always reasonable to assume that the wavelength $\lambda(t_k)$ is an increasing function in the parameter domain. Therefore, for a given wavelength $\lambda_k$, we can perform a binary search for $t_k$ such that $\lambda_k \sim \lambda(t_k)$.

The knot removal process can be viewed as a greedy approach which is reminiscent of a gradient descent method. In practice, this is an iterative method in which, at every iteration, we locate the knot that maximises the reduction of the cost $K$ in Equation 4.5. The knot removal algorithm is summarised in Algorithm 1, where the sum of squared errors before and after knot removal are denoted as $SSE_{old}$ and $SSE_{new}$, respectively. The $\text{RemoveKnot}(\cdot)$ procedure implements the knot removal algorithm reported in [177].

As a result of the local support property of B-Spline curves, the removal only affects the curve partition in the neighbouring sections of the removed knot. To take advantage of the local support of Splines and to improve computational efficiency, we compute the change in sum of squared errors as the change within the neighbourhood of the removal candidate $u$. To achieve this, we use the spans of the B-Spline basis functions [132] and employ lists to back-track their effect across the Spline. The knot removal algorithm terminates when removing any knot cannot further reduce the interpolation cost.

### 4.1.4 Representation of a Collection of Spectra

We now note that, so far, the representation above applies to a single spectrum. In this section, we extend the B-Spline representation described in Section 4.1.2 to represent a collection of spectra or a spectral image. For this representation to be effective for pattern recognition purposes, the data under study should be represented by compact features that lie in a feature space with a common basis. This motivates the idea that the knot vector should be common to all the pixels in the image. Hence, we now proceed to extend the developments in the previous sections to the computation of a B-Spline representation for a given collection of reflectance spectra, with a common knot vector across all the spectra.

The purpose of enforcing a common knot vector for our reflectance representation
is threefold. Firstly, this provides a common B-Spline basis to represent and compare reflectance spectra of different materials in the scene. The continuity, differentiability and local support properties of these basis functions enable further functional analysis of spectral signatures. Secondly, it allows the recognition of materials based on compact Spline features with a common basis in the wavelength domain across the image. Thirdly, it opens up the possibility of effecting tasks such as compression, whereby each pixel in the image is represented by a number of control points and a set of knots which act as global indexes in the wavelength domain. In this manner, a hyperspectral image, which may comprise tens or hundreds of bands can be reduced to much fewer control points with a shared knot vector.

The formulation of such a Spline-based descriptor with a shared knot vector across the image builds upon that for a single spectrum presented earlier. Furthermore, it is required that all the reflectance spectra under study are represented by the same B-Spline basis functions \( N_{i,p}(t), i = 0, \ldots, n \). This implies that, for a collection of spectra signatures sampled at the same set of wavelengths, their Spline representations share the same knot vector and parameter values corresponding to the wavelengths.

Let the input image be treated as a collection of reflectance spectra \( \{ R_v | R_v = [R_{v,1}, R_{v,2}, \ldots, R_{v,l}]^T \} \), where \( v \) denotes the pixel-index in the spectral image and \( R_{v,k} \) is the measured reflectance of the spectrum \( v \) at the wavelength \( \{ \lambda_k \}, k = 1, \ldots, l \). Note that the feature which renders each spectrum different from one another is the control point-coordinates in the reflectance domain. Formally, the problem is formulated as follows. We aim to recover a B-Spline curve \( C_v \) with control points \( P_{v,i} = (x_i, y_{v,i}), i = 0, \ldots, n \), for each spectrum index \( v \), and a knot vector \( U \) common to all the spectra such that the coordinates of the curve are functions of an independent parameter \( t \) as

\[
\lambda(t) = \sum_{i=0}^{n} N_{i,p}(t)x_i \quad (4.6)
\]

\[
R_v(t) = \sum_{i=0}^{n} N_{i,p}(t)y_{v,i} \quad (4.7)
\]

Note that in Equations 4.6 and 4.7, the basis functions \( N_{i,p}(t) \) and the wavelength interpolation function \( \lambda(t) \) are the same for all the given spectra. In contrast, the reflectance interpolation function \( R_v(t) \) and its associated control point coordinates \( y_{v,i} \) vary with each spectrum. Thus, the representation of a collection of spectral signatures
in the image effectively consists of the knot vector $U$, the wavelength coordinates $[x_0, \ldots, x_n]^T$ and the reflectance coordinates $[y_{v,0}, \ldots, y_{v,n}]^T$ of the control points of the interpolating B-Spline curves.

Comparing Equation 4.7 to Equation 4.4 suggests the reformulation of the cost function in Equation 4.5. To do this, we consider the set of wavelength-indexed reflectance spectra $\{R_v| R_v = [R_v,1, R_v,2, \ldots, R_v,l]^T\}$. Let the parameter $t_k \in U$ correspond to the $k$th wavelength, i.e. $\lambda_k = \lambda(t_k), \forall k = 1, \ldots, l$. It then becomes straightforward to rewrite the cost function in Equation 4.5 as follows

$$K = \frac{1}{N} \sum_v \sum_{k=1}^l (R_v(t_k) - R_v,k)^2 + (1 - \alpha)|U| \quad (4.8)$$

where $N$ is the number of spectral signatures, $|.|$ denotes the length of the vector argument and $\alpha$ is a constant between zero and unity.

The cost function in Equation 4.8 mainly differs from that in Equation 4.5 in the first term, which has now become the weighted sum of squared errors in the two-dimensional space of wavelength-reflectance pairs, averaged over all the spectra. Therefore, as long as we enforce a common knot vector across the spatial dimension of the image, this cost function can be minimised using the knot removal procedure akin to that presented in Algorithm 1.

### 4.2 Implementation Issues

As mentioned in Section 4.1.3, the minimisation of the cost function departs from an initial interpolation of the B-Spline curves to the image reflectance spectra. For the minimisation of the cost function presented in Section 4.1.4, we depart from the curve interpolation algorithm described by Piegl and Tiller in [132]. Since the method in [132] aims at a single Spline, it is not readily applicable to the interpolation of B-Splines to multiple spectra. As mentioned in Section 4.1.4, it is necessary to enforce a common knot vector and common basis functions across the input reflectance spectra. A strategy to satisfy this constraint is to impose it on the initial interpolating B-Spline curves for the input reflectance spectra. Thus, we extend the interpolation method in [132] to include this constraint and describe its implementation in Section 4.2.1.
In addition, since classification and recognition are supervised processes, we need to extract feature vectors that can be used to perform training and testing on the input spectra. In Section 4.2.2, we present a procedure to extract features for classification purposes.

4.2.1 Knot Removal

In practice, a common knot vector and wavelength coordinates of control points are imposed upon the whole image as follows. In Algorithm 2, we present a method of interpolating B-Splines to the input image reflectance. Note that this procedure is the first step of the knot removal algorithm, as shown in Line 1 of Algorithm 1. At the beginning of Algorithm 2, we employ the centripetal method in [100] (Lines 4–7) to compute parameter values $t_{v,k}$ separately for each reflectance spectrum. Note that our choice of the centripetal method is not exclusive and other methods may be used to compute these parameter values. After obtaining the average $t_k$ of these parameter values across all the spectral signatures, a common knot vector across all the input spectra is computed in Lines 9–11. The $FindSpan(l, p, t_k, U)$ procedure in Line 14 obtains the knot span index of $t_k$ using a binary search method. Subsequently, the $BasisFunctions(span, t_k, p, U)$ procedure in Line 15 evaluates the basis functions at the parametric point $t_k$. Both the previously mentioned procedures have been thoroughly described in [132]. The control point coordinates are then the solutions to the linear equations presented in Lines 17 and 19.

Once a knot is removed, an interpolation step is applied so as to iteratively reduce the length of the B-Spline representation, i.e. the number of control points. It is also worth stressing that the target number of knots and control points should be imposed as a hard constraint since each B-Spline representation should be normalised in length, i.e. all the spectra should have the same number of control points at output. To achieve this, we apply the knot removal method presented in Algorithm 1 recursively on resampled wavelength-reflectance values. To resample wavelength and reflectance values, we evaluate the wavelength and reflectance interpolation functions in Equations 4.6 and 4.7 with respect to sampled values of the independent parameter $t$. Note that the sampling operation may cause over-smoothing and distortion of the...
Algorithm 2 \textit{Interpolate}(Λ, [R_v], p)

\textbf{Require:} Λ, R, p

Λ: The vector of sampled wavelengths, where Λ = [λ_1, \ldots, λ_l]^T

{R_v}: A collection of reflectance spectra, where \( R_{v,k} \) is the measured reflectance at the \( k^{th} \) band of the spectrum with index \( v, k = 1, \ldots, l \).

p: The degree of the basis functions.

1: \( l \leftarrow \) the number of bands (wavelengths)
2: \( m \leftarrow l + p + 1 // l \) is also the number of control points per interpolating curve
3: \( N \leftarrow \) the number of spectra
4: \textbf{for} reflectance spectrum with index \( v \) \textbf{do}
5: \( d_v \leftarrow \sum_{k=2}^{l} ((R_{v,k} - R_{v,k-1})^2 + (\lambda_k - \lambda_{k-1})^2)^{\frac{1}{4}} \)
6: \text{Compute the parameters values } \tilde{t}_{v,k}, \text{ where } k = 1, \ldots, l, \text{ such that } \tilde{t}_{v,1} = 0, \tilde{t}_{v,l} = 1, \text{ and } \tilde{t}_{v,k} = \tilde{t}_{v,k-1} + \frac{((R_{v,k} - R_{v,k-1})^2 + (\lambda_k - \lambda_{k-1})^2)^{\frac{1}{4}}}{d_v}, \text{ where } k = 2, \ldots, l,
7: \textbf{end for}
8: \( t_k \leftarrow \frac{1}{N} \sum v \tilde{t}_{v,k} \forall k = 1 \ldots l \)
9: \text{Compute the common knot vector for all spectra}
10: \( u_k = 0 \forall k = 1, \ldots, p + 1, \quad u_k = 1 \forall k = m - p, \ldots, m \)
11: \( U \leftarrow [u_1, \ldots, u_m]^T // \text{The knot vector} \)
12: Initialise A to a zero-valued array of size \( l \times l \)
13: \textbf{for } k \leftarrow [1 \ldots l] \textbf{do}
14: \text{span} \leftarrow \text{FindSpan}(l, p, t_k, U)
15: \( A[k][\text{span} - p \ldots \text{span}] \leftarrow \text{BasisFunctions(span, } t_k, p, U) \)
16: \textbf{end for}
17: \( x \leftarrow \) the solution to \( Ax = \Lambda \)
18: \textbf{for} reflectance spectrum with index \( v \) \textbf{do}
19: \( y_v \leftarrow \) the solution to \( Ay_v = R_v, \text{ where } R_v = [R_{v,1}, R_{v,2}, \ldots, R_{v,l}]^T // \text{y}_v: \text{ the reflectance-coordinates of the spectrum with index } v. \)
20: \textbf{end for}
21: \textbf{return} U, x, y_v \forall v
Algorithm 3 \textit{IterKnotRemoval}(Q, p, \alpha, \text{target})

\textbf{Require:} \( Q, p, \text{target} \)

- \( Q \): The set of input sampled wavelength-reflectance pairs \( Q = \{(\lambda_k, R_k)\}, k = 1, \ldots, l \)
- \( p \): The degree of basis functions
- \( \text{target} \): The target number of knots
- \( \alpha \): The balance factor

\( U, P \): The final knots and control points

1: \( \text{samples} \leftarrow Q \)
2: \textbf{while} true \textbf{do}
3: \( (U, P) \leftarrow \text{KnotRemoval}(\text{samples}, p, \alpha) \)
4: \( h \leftarrow \text{the number of remaining control points per spectrum} \)
5: Select parameters \( (t_1, \ldots, t_h) \)
6: \( \text{samples} \leftarrow \text{Resample}(p, U, P, t_1, \ldots, t_h) \)
7: \textbf{if} \( |U| \leq \text{target} \) \textbf{then}
8: \textbf{return} \( U, P \)
9: \textbf{end if}
10: \textbf{end while}
4.2. IMPLEMENTATION ISSUES

Figure 4.3: B-Spline curves (in green) with 34 knots and 30 control points resulting from our knot removal algorithm performed on the discrete spectral reflectance samples of human skin (left) and a leaf (right), which is depicted as red scattered dots. The original data contain 88 samples per spectrum, in the range of 400–750nm.

original spectra when the sampled values of \( t \) are close to the midpoints of the knot spans \([u_i, u_{i+1}), i = 1, \ldots, n + p\). Thus, we select the values of \( t \) close to the knot positions remaining in each iteration to preserve the shape of the original reflectance curves. Overall, we found that this resampling operation allows further knots to be removed by reducing the number of curve sections without changing the distribution of the control points with respect to the original reflectance data. The pseudocode for the recursion on the knot removal algorithm is shown in Algorithm 3.

To illustrate the behaviour of our algorithms, in Figure 4.3, we show the Spline curves obtained by removing knots from the reflectance spectra of human skin and a leaf. The original reflectance data, shown as red scattered dots, are sampled every 4nm, consisting of 88 samples in the range of 400–750nm. For this illustration, we employed third-degree B-Spline basis functions. This yielded 34 knots and 30 control points per curve. Note that, while this number of knots is considerably less than the number of data points on the input spectra, the resulting B-Splines, plotted in green solid curves, are in good accordance with the original data. Our cost-optimal knot removal algorithm is, in effect, performing dimensionality reduction on the control points while respecting the global shape of the reflectance spectra.
4.2.2 Classification Features

Now we focus on the procedure for extracting the B-Spline descriptor used for classification purposes. We do this by enforcing a common knot vector and wavelength-coordinates between the training and testing reflectance spectra. Let the training set of reflectance spectra be \( R_{\text{train}}^{v} = [R_{v,1}, R_{v,2}, \ldots, R_{v,l}]^T \), where \( v \) denotes the spectrum index and \( R_{v,k} \) is the measured reflectance of the training spectrum at the wavelength \( \lambda_k, k = 1, \ldots, l \). The minimisation of the cost function in Equation 4.8 results in a shared knot vector \( U \), basis functions \( \{N_i, p(t)\}, i = 0, \ldots, n \} \), and the wavelength-coordinates \( [x_0, \ldots, x_n]^T \), which are common to all the training spectra. In addition, each reflectance spectrum \( R_{\text{train}}^{v} \) is represented by a vector of reflectance-coordinates of the control points \( Y_{\text{train}}^{v} = [y_{\text{train}}^0, \ldots, y_{\text{train}}^n]^T \).

Let a testing reflectance spectrum be given by \( R_{\text{test}} = [R_{\text{test}}^1, R_{\text{test}}^2, \ldots, R_{\text{test}}^l]^T \), where \( R_{\text{test}}^k \) is the discrete sample reflectance at the wavelength \( \lambda_k \). Our goal is to find an interpolating B-Spline curve for the test reflectance spectrum, with the same knot vector \( U \) and wavelength-coordinates \( [x_0, \ldots, x_n]^T \) as those of the training data. Note that the B-Spline basis functions of the test spectrum, which are determined by the knot vector, are also the same as those of the training spectra, \( i.e. \{N_i, p(t), i = 0, \ldots, n\} \). As a result, the interpolation to the test spectrum amounts to finding the \( y \)-coordinates of the test spectrum, \( Y_{\text{test}} = [y_{\text{test}}^0, \ldots, y_{\text{test}}^n]^T \), such that

\[
\lambda_k = \sum_{i=0}^{n} N_{i,p}(t_k)x_i \quad (4.9)
\]

\[
R_{\text{test}}^k = \sum_{i=0}^{n} N_{i,p}(t_k)y_i^{\text{test}} \quad (4.10)
\]

where \( t_k \) is an independent parameter corresponding to the wavelength \( \lambda_k \).

Note that with the values of \( x_i \) and \( \lambda_k \) recovered from the training data, Equation 4.9 can be solved for the set of independent parameters \( t_k \)'s, \( k = 1, \ldots, l \) as a binary search strategy as discussed in Section 4.1.3. Subsequently, the basis functions \( N_{i,p}(t_k) \) can be evaluated at the obtained values of \( t_k \)'s, \( k = 1, \ldots, l \).

With the values of \( \{N_i, p(t), i = 0, \ldots, n\} \) in hand, Equation 4.10 for \( k = 1, \ldots, l \) becomes an over-determined linear system of the unknown vector \( Y_{\text{test}} = [y_{\text{test}}^0, \ldots, y_{\text{test}}^n]^T \) when the number of control points \( \hat{n} \) does not exceed the number of wavelengths, \( i.e. \)}
\hat{n} = n + 1 \leq l. The coordinate vector \( \mathbf{Y}_{test} \), once obtained, can be viewed as a descriptor of the test reflectance spectrum which satisfies Equation 4.10 in the least-squares sense.

### 4.3 Experiments

In this section, we demonstrate the utility of our Spline-based reflectance representation in three experimental vehicles. In Section 4.3.1, we use spectrometer data and hyperspectral images as a means to validate the accuracy of reconstruction with respect to the input reflectance spectra. In Section 4.3.2, we explore the use of the our Spline-based representation as a descriptor for skin recognition purposes. Lastly, in Section 4.3.3, we explore the use of this descriptor for biometrics applications, where we perform ethnic group classification based on human skin reflectance spectra.

#### 4.3.1 Reconstruction of Reflectance Spectra

To provide a quantitative study on the reconstruction accuracy for reflectance spectra represented using our B-Spline method, we perform our experiments on two data-sets. The first of these consists of 297 reflectance spectra acquired in house using a StellarNet spectrometer, which we name the StellarNet dataset. These spectra correspond to nine material categories, including cloth, different kinds of paint, human skin, leaves from a number of plant varieties, metals, papers of different colours, plastic, porcelain and wood. In our dataset there are 157 spectra of human skin and 140 of the other materials above. All the spectra have been sampled in 0.5 nm intervals between 430–720 nm with reflectance normalised between [0, 1]. The second set of data consists of a multispectral face image database captured under 10 artificial light sources with varying illumination directions and spectral power. The database comprises frontal views of 51 human subjects acquired with cluttered backgrounds. The images have been captured using a hyperspectral camera system which acquires wavelength-resolved images in 10 nm intervals over the visible spectrum, i.e. 30 bands between 430 nm and 720 nm. To obtain the ground-truth reflectance of these images, we normalised the raw radiance images by the illumination power spectrum of each image as measured on a white
The two data-sets above permit the application of our method to the reconstruction of reflectance data acquired with several sensing devices whose spectral sampling resolution may be quite dissimilar from one another. Here, we utilise the knot vectors and control points extracted from the spectral reflectance samples as a means of representation and data reduction in the spectral domain. Furthermore, we compare the reconstruction performance of our Spline representation with that of the descriptor recovered through Gaussian Mixture Regression reported in [3]. Recall that, in [3], a linear combination of $M$ Gaussian basis functions is fitted to each reflectance spectrum. Thus, the descriptor in [3] is similar to our method in the sense that it treats the reflectance $R(\lambda)$ as a function of the wavelength $\lambda$, where the normalised reflectance spectrum can be reconstructed as

$$R(\lambda) = \sum_{k=1}^{M} \frac{\beta_k}{\sqrt{2\pi}\sigma_k} \exp\left(-\frac{(\lambda - \mu_k)^2}{2\sigma_k^2}\right)$$

(4.11)

where $M$ is the number of Gaussian mixture components and the $k^{th}$ Gaussian component is associated with a mean $\mu_k$, a standard deviation $\sigma_k$ and a mixture coefficient $\beta_k$. The fitting problem in Equation 4.11 can be treated as a nonlinear least-squares optimisation one, which may be solved numerically using a Levenberg-Marquardt optimisation procedure [112].

To further demonstrate the advantage of our representation, in Figure 4.4, we present sample plots of the reconstructed reflectance spectra using the two methods overlaid on top of the original data as measured by the StellarNet spectrometer. In each column from left to right, we present the reflectance spectra of human skin, a leaf and a plastic material with various representation lengths. The rows show the reconstruction results produced with descriptor lengths of 12, 18, 24 and 30. The curves reconstructed from the Splines and Gaussian Mixture descriptors are plotted as dot-dashed lines and dashed-lines, respectively, whereas the original data is plotted as a solid trace. The general trend of these plots shows that the Spline-based representation presented here fits the spectra better as its length increases. This is not necessarily the case for the alternative in [3]. In the plots, we notice that the Spline representation tends to capture absorption bands much better than the Gaussian Mixture one. This typical effect can
4.3. EXPERIMENTS

(a) Human Skin  
(b) Leaf  
(c) Plastic

Figure 4.4: Plots of sample StellarNet reflectance spectra. From left to right: the original and reconstructed reflectance of human skin, a leaf and plastic. From top to bottom: the reconstruction results for 12, 18, 24 and 30 control points of the B-Spline representation, as compared to 4, 6, 8 and 10 Gaussian components, respectively.
be easily observed on the reflectance spectra of the skin sample at around 550\textit{nm}, the leaf sample at approximately 680\textit{nm} and the plastic material at 440\textit{nm}. Moreover, the Spline representation is able to reproduce absorption bands with as few as 12 control points (top row). In contrast, the Gaussian mixture descriptor tends to smooth out the original curves and does not conform to the overall shape of the original spectra even with 10 Gaussian components, as seen in the bottom-row plots with a length of 30. This is due to the smooth, global support nature of the Gaussian mixture approach in [3], which faces difficulties with fitting non-smooth curves with sharp peaks and valleys. In contrast, our algorithm departs from an initial curve interpolation with a data error term of zero, performing cost-based knot removal so as to reduce the descriptor length. This procedure assures a low reconstruction error.

To provide a more quantitative study, we use two measurements for the reconstruction error. The first of these is the mean of the absolute difference between the reconstructed and the original reflectance spectra per band, \textit{i.e.} the average absolute reconstruction error (ARE). The second error measurement used here pertains the relative reconstruction error (RRE), which we quantify as the ratio of the absolute error to the original reflectance value. In this way, the mean absolute error allows the assessment of the reconstruction results across each data set, whereas the relative reconstruction error provides a means of quantifying the proportion of the errors with respect to the magnitude of the reflectance data. It is worth noting in passing that, for the reflectance data acquired using spectrometers, its customary to normalise the spectrum to unity. Thus, the absolute and relative reconstruction errors for each spectrum are proportional up to a multiplicative constant. This is not the case for hyperspectral and multispectral imagery, where the spectra in an image cannot be normalised individually per pixel. Thus, later in the section, we solely focus on the absolute reconstruction for the StellarNet dataset while providing both measures for our hyperspectral image dataset.

We first turn our attention to examining the reconstruction results for the reflectance samples measured by the StellarNet spectrometer. We commence by assessing the effect of varying the representation length. The purpose here is to provide a quantitative evaluation of the representation accuracy as a function of the “compactness” of
Table 4.1: Mean absolute reconstruction error (ARE) for the StellarNet dataset.

<table>
<thead>
<tr>
<th>Length</th>
<th>Spline</th>
<th>Gaussian Mixture</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>(1.60 ± 2.46) × 10⁻²</td>
<td>(2.11 ± 2.66) × 10⁻²</td>
</tr>
<tr>
<td>12</td>
<td>(1.57 ± 2.39) × 10⁻²</td>
<td>(1.79 ± 2.30) × 10⁻²</td>
</tr>
<tr>
<td>15</td>
<td>(1.45 ± 2.19) × 10⁻²</td>
<td>(1.57 ± 2.09) × 10⁻²</td>
</tr>
<tr>
<td>18</td>
<td>(1.29 ± 1.98) × 10⁻²</td>
<td>(1.35 ± 1.76) × 10⁻²</td>
</tr>
<tr>
<td>21</td>
<td>(1.08 ± 1.73) × 10⁻²</td>
<td>(1.28 ± 1.67) × 10⁻²</td>
</tr>
<tr>
<td>24</td>
<td>(0.90 ± 1.52) × 10⁻²</td>
<td>(1.21 ± 1.61) × 10⁻²</td>
</tr>
<tr>
<td>27</td>
<td>(0.77 ± 1.30) × 10⁻²</td>
<td>(1.20 ± 1.58) × 10⁻²</td>
</tr>
<tr>
<td>30</td>
<td>(0.68 ± 1.17) × 10⁻²</td>
<td>(1.19 ± 1.61) × 10⁻²</td>
</tr>
</tbody>
</table>

our representation. To this end, we recover a Spline for all these reflectance spectra, with a common knot vector and common wavelength coordinates of control points as presented in Section 4.1.4. Thus, the length of our Spline-based representation per reflectance spectrum is given by $|U| + \hat{n} + N\hat{n}$, where $N$ is the number of reflectance spectra, $\hat{n} = n + 1$ is the number of control points and $|U|$ is the length of the knot vector. Here, we employ Splines with a degree $p = 3$. To represent the “W” pattern typically found in skin reflectance spectra [3], we employed at least three Gaussian basis functions, one of which to model the peak of this pattern and the others to represent the low and high ends of the spectrum. Further, the Gaussian representation consists of triplets of the mixture coefficient, the mean and the standard deviation of Gaussian components. To facilitate comparison with the Gaussian representation, the descriptor length under study is a multiple of three. Therefore, we vary the number of control points between 9 and 30 in increments of 3. Since the dataset includes $N = 297$ reflectance spectra, $\hat{n} \leq 30$ and $|U| = \hat{n} + p + 1 \leq 34$, the fraction $\frac{|U| + \hat{n}}{N}$ can be neglected, i.e. the length of our Spline representation can be approximated by the number of control points. The above range of the Spline representation length, i.e. between 9 and 30 is equivalent to 3 to 10 Gaussian components of the Gaussian descriptor in [3].
StellarNet spectrometer reflectance spectra with respect to the representation length, for our Spline representation and the Gaussian Mixture descriptor. From the figure, we can conclude that our Spline representation consistently yields a lower absolute reconstruction error (ARE) than the alternative with the smallest performance difference between the two methods occurring when six Gaussian Mixture components are used, \(i.e.\) when the Spline representation comprises 18 control points. Note that, with increasing length, the ARE becomes more stable for our representation, with a decreasing standard deviation. Another important observation is that increasing the number of Gaussian components beyond six does not significantly reduce the reconstruction error. On the other hand, our Spline-based representation shows a strong and consistent reduction in reconstruction error as the number of control point increases. These trends imply that our representation outperforms the alternative in both representation accuracy and compactness.

In Table 4.1, we show the values of the error measures plotted in Figure 4.5. Note that the absolute errors are in the order of \(10^{-2}\), which is negligible for normalised reflectance values in the range \([0, 1]\). As shown before, when 30 control points are
4.3. EXPERIMENTS

<table>
<thead>
<tr>
<th></th>
<th>Spline (18 control points)</th>
<th>Gaussian Mixture (6 components)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute error</td>
<td>((0.35 \pm 0.43) \times 10^{-2})</td>
<td>((2.72 \pm 2.85) \times 10^{-2})</td>
</tr>
<tr>
<td>Relative error</td>
<td>3.89 ± 4.52(%)</td>
<td>16.61 ± 11.53(%)</td>
</tr>
</tbody>
</table>

Table 4.2: The absolute and relative reconstruction error for our hyperspectral reflectance imagery.

used, the absolute reconstruction error of our Spline-based representation can be as low as half of that of the Gaussian Mixture descriptor.

Now we present results on multispectral image representation. Here, we apply the knot removal algorithm described in Section 4.1.4 to the collection of reflectance spectra at all the pixels in a multispectral image so as to arrive at a per-image common knot vector. As a result, the control points for the image reflectance can be viewed as a “lossy” representation of the imagery in our dataset. In this case, the data reduction ratio is quantified as \(\frac{|U| + \hat{n} + N\tilde{h}}{Nl}\), where \(|U|\) is the length of the knot vector, \(\hat{n} = n + 1\) is the number of control points, \(N\) is the number of image pixels and \(l\) is the number of bands in the multispectral image. For an image with a sufficiently large number of pixels \(N\), this ratio can be approximated by \(\frac{\hat{n}}{l}\).

In this experiment, we focus on the reconstruction error yielded by our algorithm and the alternative when applied to the hyperspectral face image database introduced earlier. In the previous reconstruction experiments on the StellarNet spectral curves, both representations are the most comparable when the representation in [3] is set to comprise six Gaussian components. Thus, in Table 4.2, we compare the reconstruction performance of the two representations on the hyperspectral image dataset when their length is set to 18, \(i.e.\) the Spline representation has a degree \(p = 3\) and consists of 22 knots and 18 control points. As shown in the table, our Spline-based representation yields an absolute reconstruction error nearly eight times lower than that delivered by the Gaussian Mixture descriptor. Likewise, the relative reconstruction error of the Splines is over four times lower than that of the alternative.

In Figure 4.6, we illustrate the quantitative results above by showing the reconstruction error, graded in shades of gray, for four sample subjects in our dataset. The first
Figure 4.6: Absolute and relative reconstruction error for sample hyperspectral images. First column: pseudo colour images. Second and third columns: Absolute (ARE) and relative reconstruction error (RRE) for our Spline-based representation, where pure white in the error maps corresponds to $2 \times 10^{-2}$ and 10% in the second and third columns, respectively. Fourth and fifth columns: Absolute (ARE) and relative reconstruction error (RRE) yielded by the Gaussian Mixture descriptor, where pure white in the error maps corresponds to $10 \times 10^{-2}$ and 30% in the fourth and fifth columns, respectively.
white corresponds to 10% for our method and 30% for the alternative in [3]. In the Figure, the absolute error yielded by our representation is smaller than that of the alternative by an order of magnitude. Likewise, the relative error of the Spline is lower than that of the Gaussian Mixture descriptor. Note that, as expected, the regions with large relative errors are those with low reflectance in the original images.

4.3.2 Skin Recognition

In this section we show the utility of our Spline-based representation as a descriptor for skin recognition. We commence by performing skin spectra recognition on the Stellar-Net dataset. Later, we study the behaviour of such a descriptor on multispectral images. Throughout the section, we compare the recognition accuracy yielded by our method with that delivered by a number of alternative descriptors. The first of these is the Gaussian Mixture descriptor [3] used in the previous section. To construct the Gaussian Mixture feature for recognition purposes, we concatenate the triplet $(\beta_k, \mu_k, \sigma_k)$, i.e. the mixture coefficient, the mean and standard deviation for every component in the descriptor. We have done this since, in our experiments, this triplet consistently delivers better classification performance than any other tuple combinations involving $\beta_k, \mu_k$ and $\sigma_k$.

Other alternative spectral reflectance descriptors include the principal components resulting from applying Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) to wavelength-indexed spectral reflectance vectors and the raw spectral reflectance vector. As before, we study the recognition accuracy of these descriptors with respect to their lengths. To model skin reflectance spectra, we employed at least three Gaussian basis functions to represent the “W” pattern the skin reflectance spectra under study, in a similar manner to the work performed by Angelopoulou et al. [3]. Since the length of the Gaussian Mixture descriptor is always a multiple of 3, we perform skin recognition experiments with descriptor lengths between 9 and 30 with an increment of 3.
CHAPTER 4. B-SPLINE SPECTRUM REPRESENTATION

<table>
<thead>
<tr>
<th>Feature combination</th>
<th>Spline-1 Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$–coordinates</td>
<td>90.08 ± 3.19</td>
</tr>
<tr>
<td>$x$–coordinates</td>
<td>51.27 ± 5.63</td>
</tr>
<tr>
<td>Knots</td>
<td>55.69 ± 8.31</td>
</tr>
<tr>
<td>$x$ and $y$–coordinates</td>
<td>85.99 ± 2.45</td>
</tr>
<tr>
<td>Knots and $y$–coordinates</td>
<td>85.56 ± 2.79</td>
</tr>
<tr>
<td>Knots and $x$–coordinates</td>
<td>56.19 ± 8.17</td>
</tr>
<tr>
<td>Knots, $x$ and $y$–coordinates</td>
<td>84.04 ± 2.53</td>
</tr>
</tbody>
</table>

Table 4.3: The mean and standard deviation of the classification accuracy over 20 random trials on the dataset of spectra acquired by a StellarNet spectrometer for various feature combinations from the 34 knots and 30 control points of a third-degree B-Splines descriptor.

**Spectrometer Data**

For our recognition experiments on the StellarNet dataset we effected 20 trials as follows. For each trial, we randomly selected 50 skin and 50 non-skin spectra for purposes of training. The remaining spectra were used for testing. The spectra were sub-sampled at a 5nm sampling interval in the wavelength domain, resulting in 59 bands per spectrum. The performance of the descriptors involved in this experiment was computed by the average performance over the 20 trials. The skin spectra recognition task was effected using a soft-margin Support Vector Machines (SVM) classifier [19, 36] with a third-degree polynomial kernel function. The other parameter values of the SVM classifier were selected by a 4-fold cross validation procedure.

In this experiment, we adopted two different approaches to extracting the Spline-based descriptor for the skin recognition task. In the first variant, which we denote Spline-1, the training and test features include the optimal control points extracted from each individual reflectance spectrum. The optimal Spline curve recovery process from a single reflectance spectrum has been described in Section 4.1.2. The other option used here, which we denote Spline-2, employs the features composed of the $y$-coordinates of the control points extracted from the whole set of training reflectance spectra using a common knot vector as described in Sections 4.1.4.
We also assess the discriminative power of the \( y \)-coordinates of control points in the variant Spline-1, which we denote as \( \mathbf{Y} = [y_0, y_1, \ldots, y_n]^T \), in comparison with the \( x \)-coordinates and the knot vector. For the results shown here, we represent each original wavelength-indexed spectral reflectance vector separately using a third-degree B-Spline curve with 34 knots and 30 control points. In Table 4.3, the recognition rate of the descriptor \( \mathbf{Y} \) is compared with those of other feature combinations of knots and control points. As shown, the choice of the descriptor \( \mathbf{Y} \) achieves the highest and most stable recognition rate of all the feature combinations involving the knot vector and control points of the resulting Spline curves. From this Table, we note that combining the knot vector and \( x \)-coordinates with the \( y \)-coordinates of the control points does not necessarily improve the classification performance, despite the added computational complexity. This is consistent with the understanding that the knot vector and the \( x \)-coordinates of all spectra vary within the same range of values in the parameter and wavelength domains, respectively. Thus, these two features and their combination deliver an almost random classification result. Meanwhile, the variation of the \( y \)-coordinates is in good accordance with the shape of reflectance spectra, therefore yielding high classification performance. Thus, in our experiments from now on, we favour the formulation \( \mathbf{Y} = [y_0, y_1, \ldots, y_n]^T \) as the Spline-based descriptor over the \( x \)-coordinates and the knot vector. Our choice of \( \mathbf{Y} \) as a descriptor for our experiments stems from the fact that it yields better performance with lower computational cost.

We now evaluate the performance of the Spline-based descriptor variants on the data set focusing solely on the \( y \)-coordinates, \textit{i.e.} reflectance-coordinates, of the control points of the spectra. As before, the descriptor length varies between 9 and 30, in increments of 3, corresponding to 3 to 10 components of the Gaussian Mixture descriptor. Here, to obtain the Spline-1 descriptor variant, we compacted the original spectral samples into third-degree B-Spline curves. Meanwhile, the Spline-2 variant is represented by B-Spline curves of first-degree.

In Table 4.4, we report the skin recognition accuracy for the above mentioned spectral reflectance descriptors. In the second and third columns, we show the mean and standard deviation yielded by our Spline-based descriptors over the 20 trials. In the remaining columns, we present the recognition rates of the Gaussian Mixture, Linear
Table 4.4: The skin spectrum recognition accuracy (in %) on the Spectrometer dataset for several spectral reflectance descriptors, with respect to the variation in the descriptor length. The descriptors under consideration are the $y$-coordinates of the B-Splines representing individual training reflectance spectra (Spline-1); the $y$-coordinates of the B-Splines sharing a common knot vector and common $x$-coordinates, that have been extracted from the entire training set of reflectance spectra (Spline-2); the Gaussian Mixture descriptor proposed in [3] (Gaussian Mix.); principal components recovered by performing LDA and PCA on the spectra; and the raw reflectance spectra.

<table>
<thead>
<tr>
<th>Length</th>
<th>Spline-1</th>
<th>Spline-2</th>
<th>Gaussian Mix.</th>
<th>LDA</th>
<th>PCA</th>
<th>Raw</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>87.79 ± 2.49</td>
<td>92.28 ± 1.76</td>
<td>90.66 ± 2.11</td>
<td>86.78 ± 2.21</td>
<td>93.53 ± 2.54</td>
<td>_</td>
</tr>
<tr>
<td>12</td>
<td>87.56 ± 2.96</td>
<td>92.64 ± 2.27</td>
<td>81.98 ± 2.93</td>
<td>95.94 ± 1.27</td>
<td>95.79 ± 2.06</td>
<td>_</td>
</tr>
<tr>
<td>15</td>
<td>88.12 ± 2.57</td>
<td>92.79 ± 2.39</td>
<td>83.43 ± 4.05</td>
<td>96.12 ± 1.44</td>
<td>94.01 ± 2.27</td>
<td>_</td>
</tr>
<tr>
<td>18</td>
<td>89.04 ± 2.03</td>
<td>93.96 ± 2.26</td>
<td>81.65 ± 3.25</td>
<td>96.90 ± 1.39</td>
<td>94.42 ± 1.98</td>
<td>_</td>
</tr>
<tr>
<td>21</td>
<td>89.49 ± 2.65</td>
<td>93.55 ± 2.44</td>
<td>83.55 ± 3.30</td>
<td>97.51 ± 1.40</td>
<td>95.03 ± 1.67</td>
<td>_</td>
</tr>
<tr>
<td>24</td>
<td>89.72 ± 2.60</td>
<td>93.93 ± 2.63</td>
<td>83.60 ± 2.95</td>
<td>98.27 ± 0.99</td>
<td>95.08 ± 2.55</td>
<td>_</td>
</tr>
<tr>
<td>27</td>
<td>90.71 ± 2.55</td>
<td>94.06 ± 2.39</td>
<td>82.13 ± 4.28</td>
<td>98.45 ± 0.85</td>
<td>94.62 ± 1.94</td>
<td>_</td>
</tr>
<tr>
<td>30</td>
<td>91.19 ± 3.18</td>
<td>94.26 ± 2.31</td>
<td>85.99 ± 2.47</td>
<td>98.40 ± 1.90</td>
<td>93.50 ± 2.79</td>
<td>_</td>
</tr>
<tr>
<td>59</td>
<td>_</td>
<td>_</td>
<td>_</td>
<td>_</td>
<td>_</td>
<td>94.44 ± 2.72</td>
</tr>
</tbody>
</table>

Discriminant Analysis (LDA), Principal Component Analysis (PCA) and the Raw input reflectance feature. Note that the Raw reflectance feature consists of 59 spectral bands. To provide a visual illustration of these numerical results, in Figure 4.7, we plot the recognition performance of the descriptors as functions of the descriptor length, with a trace corresponding to each descriptor.

In Table 4.4, although LDA delivers the best performance at lengths greater or equal to 12, it is outperformed by our Spline descriptor obtained with a common knot vector (Spline-2) for length 9. The high performance of LDA and PCA can, hence, be explained by the fact that, when applied to the spectra, they may remove sampling noise. Of the descriptors presented, our Spline-based representation and the Gaussian mixture reported in [3] are those that permit the spectra to be reconstructed. Amongst
4.3. EXPERIMENTS

Figure 4.7: The skin spectrum recognition accuracy for the reflectance spectra captured by a StellarNet spectrometer with respect to the descriptor length of our descriptor and the alternatives.

these two, our descriptor obtained with a common knot vector (Spline-2) consistently delivers better results. Furthermore, the Spline-2 descriptor steadily increases in performance with respect to the increase of descriptor length. This implies that this descriptor can be effected with a short length, to achieve a high recognition accuracy with a low computational overhead. In addition, the Spline-2 descriptor does not suffer from fluctuation of recognition accuracy as the descriptor length varies. With the Gaussian Mixture descriptor and PCA, the recognition accuracy may degrade with respect to increasing representation length. This can be understood as the effect of overfitting, where additional basis functions tend to capture noise rather than the discriminant features. The stability of the Spline-2 descriptor is illustrated in Figure 4.7, where its recognition rate approaches that of the Raw reflectance spectra feature near the high end of the length range, i.e. 30.
CHAPTER 4. B-SPLINE SPECTRUM REPRESENTATION

Figure 4.8: Sample face images used to train an SVM classifier for skin segmentation, shown at 670 nm.

**Multispectral Images**

Following our experiments on the StellarNet spectrometer data, we now present results on our hyperspectral image database of human subjects. The database contains three ethnic groups, with 22 of the subjects identifying themselves as Caucasian, 8 as Indian, 18 as Oriental and 3 as others. To acquire training data, we randomly selected a subject representative of each of the main ethnic groups. On the images of these subjects, we randomly sampled pixels from skin and non-skin regions, each consisting of 685 skin and 2968 non-skin spectra. We performed the skin recognition on the images that were acquired under the frontal illumination direction. To obtain the image reflectance, we simply normalised the raw radiance images by the ground truth illumination power, which had been measured on a white calibration target, i.e. a Spectralon. In practical situations, where a measurement of the illumination power spectrum is unavailable, the image reflectance can still be estimated from the original image as a result of illumination spectrum recovery [82].

On the hyperspectral image database, we show results over seven trials, each of these effected with different randomly selected training spectra as described above. Examples of skin and non-skin regions on sample training images are shown as rectangles in Figure 4.8, with red borders for skin regions and blue ones for non-skin regions in the image. To obtain the Spline-based reflectance features used for training purposes, we compute a cost-optimal set of B-Spline curves with respect to the whole set of training reflectance spectra, as described in Sections 4.1.4. The y-coordinates (reflectance coordinates) of the resulting Spline curves are then employed as feature for recognition.

Similarly to our previous experiments, we use a soft-margin Support Vector Ma-
4.3. EXPERIMENTS

chines classifier (SVM) [19, 36] for purposes of classification. For our image data, we use a Radial Basis Function kernel as an alternative to the polynomial one. The parameter values of the SVM classifier are, again, selected by a 4-fold cross validation procedure. Subsequently, skin recognition is performed by applying the trained classifier to the spectral reflectance descriptor at each image pixel for all the testing imagery, i.e. that corresponding to the 48 subjects remaining at each trial after excluding the three training images.

<table>
<thead>
<tr>
<th>Descriptor’s length</th>
<th>9</th>
<th>12</th>
<th>15</th>
<th>18</th>
<th>21</th>
<th>24</th>
<th>27</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spline</strong> CR</td>
<td>96.70</td>
<td>97.02</td>
<td>97.04</td>
<td>97.15</td>
<td>97.25</td>
<td>97.21</td>
<td>97.24</td>
<td>97.24</td>
</tr>
<tr>
<td>CDR</td>
<td>75.27</td>
<td>79.87</td>
<td>79.43</td>
<td>79.13</td>
<td>80.75</td>
<td>80.59</td>
<td>80.34</td>
<td>80.39</td>
</tr>
<tr>
<td>FDR</td>
<td>1.78</td>
<td>1.80</td>
<td>1.76</td>
<td>1.62</td>
<td>1.62</td>
<td>1.65</td>
<td>1.61</td>
<td>1.62</td>
</tr>
<tr>
<td><strong>Gauss Mix.</strong> CR</td>
<td>96.79</td>
<td>96.63</td>
<td>96.72</td>
<td>96.70</td>
<td>97.17</td>
<td>97.34</td>
<td>97.32</td>
<td>96.56</td>
</tr>
<tr>
<td>CDR</td>
<td>74.53</td>
<td>73.50</td>
<td>73.57</td>
<td>72.80</td>
<td>73.68</td>
<td>74.29</td>
<td>73.00</td>
<td>72.02</td>
</tr>
<tr>
<td>FDR</td>
<td>1.70</td>
<td>1.77</td>
<td>1.68</td>
<td>1.63</td>
<td>1.21</td>
<td>1.10</td>
<td>1.01</td>
<td>1.75</td>
</tr>
<tr>
<td><strong>LDA</strong> CR</td>
<td>96.64</td>
<td>96.62</td>
<td>97.02</td>
<td>97.23</td>
<td>97.10</td>
<td>97.10</td>
<td>97.03</td>
<td>97.01</td>
</tr>
<tr>
<td>CDR</td>
<td>76.05</td>
<td>78.71</td>
<td>79.83</td>
<td>80.91</td>
<td>80.94</td>
<td>80.80</td>
<td>80.52</td>
<td>80.47</td>
</tr>
<tr>
<td>FDR</td>
<td>1.92</td>
<td>2.13</td>
<td>1.80</td>
<td>1.64</td>
<td>1.78</td>
<td>1.76</td>
<td>1.82</td>
<td>1.84</td>
</tr>
<tr>
<td><strong>PCA</strong> CR</td>
<td>97.53</td>
<td>97.64</td>
<td>97.81</td>
<td>97.72</td>
<td>97.69</td>
<td>97.69</td>
<td>97.51</td>
<td>97.38</td>
</tr>
<tr>
<td>CDR</td>
<td>79.00</td>
<td>80.32</td>
<td>80.10</td>
<td>80.40</td>
<td>82.10</td>
<td>81.57</td>
<td>82.05</td>
<td>82.70</td>
</tr>
<tr>
<td>FDR</td>
<td>1.21</td>
<td>1.18</td>
<td>0.98</td>
<td>1.10</td>
<td>1.25</td>
<td>1.21</td>
<td>1.43</td>
<td>1.62</td>
</tr>
<tr>
<td><strong>Reflectance</strong> CR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>97.92</td>
</tr>
<tr>
<td>CDR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>78.07</td>
</tr>
<tr>
<td>FDR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Table 4.5: The means of the overall skin classification rates (CR), the correct skin detection rates (CDR) and the false detection rates (FDR) achieved by several spectral descriptors, over seven trials on the multispectral image database.

In Table 4.5 we show the performance of our method and the alternatives, i.e. the Gaussian Mixture, LDA, PCA and the raw reflectance spectra. Here we quantify the
skin recognition accuracy in terms of the overall classification rate (CR), the correct detection rate (CDR) and false detection rate (FDR). Here, we aim to compare descriptors with the same length, rather than classifiers with various parameter settings. Therefore, we favour these metrics over the Receiver Operating Characteristic (ROC) curve, which often varies a function of classifier parameters. The correct detection rate is the percentage of skin pixels correctly classified. The false detection rate is the percentage of non-skin pixels incorrectly classified. The classification rate is the overall percentage of skin and non-skin pixels classified accurately. The Table shows the mean accuracy measures over the 7 trials effected using different randomly selected training sets. The rows of the Table correspond to the different spectral reflectance features described earlier.

As expected, the descriptor composed of the $y$-coordinates of the B-Spline achieves a comparable result to that delivered by the SVM when LDA and PCA is used to reduce the dimensionality of the image reflectance. Moreover, it delivers a superior skin detection rate (CDR) to that yielded by the Gaussian mixture, with a comparable false detection rate (FDR). This is consistent with the qualitative results on the StellarNet spectrometer data presented earlier. Furthermore, the overall performance difference between the LDA, PCA and our method is negligible. In fact, our Spline-based representation achieves lower false detection rates than those achieved by LDA. In addition, it delivers slightly higher skin detection rates (CDR) than the raw reflectance feature, when the descriptor is composed of 12 or more elements. Compared to the Gaussian Mixture descriptor, our B-Spline descriptor achieves a higher CDR and a comparable FDR. Lastly, it is noticed that the hyperspectral images used in this experiment contain much more non-skin pixels than the skin ones. Hence, although the descriptors above may deliver significantly different CDR and FDR, the large proportion of non-skin pixels results in similar overall classification rates (CR) across different descriptors.

In Figure 4.9, we plot the means and standard deviations of the recognition rates as functions of the descriptor length. As shown in Figure 4.9(a), the B-Spline descriptor achieves a comparable skin detection rate (CDR) to that yielded by the LDA and PCA features. Figure 4.9(b), on the other hand, indicates that the raw reflectance spectra delivers the lowest FDR of all the descriptors involved. Our descriptor is comparable
4.3. EXPERIMENTS

(a) Correct detection rates (CDR)  

(b) False detection rates (FDR)  

(c) Overall classification rates (CR)  

Figure 4.9: The overall skin classification rates (CR), the correct skin detection rates (CDR) and the false detection rates (FDR) achieved by several spectral descriptors on the multispectral image database, for various descriptor’s lengths. The Spline-based descriptor is recovered by enforcing a common knot vector and the common wavelength coordinates of control points between the training and testing reflectance spectra.
Figure 4.10: Skin probability maps of a sample subject. First column: the original input image rendered in pseudo trichromatic colour. The remaining columns, from left to right, correspond to the skin probability maps delivered by the Spline-based descriptor, the Gaussian Mixture descriptor, LDA, PCA and the input reflectance spectra as discriminant features. The brightness on the skin probability maps indicates the predicted likelihood of being skin. The rows, from top to bottom, correspond to the descriptor lengths of 12, 18, 24 and 30.

to the remaining ones in the low range of descriptor’s length (between 9 and 18) while being more stable than LDA in terms of FDR and the overall CR. Lastly, our descriptor’s performance grows steadily with the increase of the descriptor’s length. This is evident in Figure 4.9(c), where our descriptor improves its performance with length, whereas this is not the case for the descriptors other than the raw reflectance. Indeed, the B-Spline descriptor closely approaches the overall performance of PCA with 30 control points.

Figure 4.10 presents the skin probability maps for a sample subject, which result from the employment of the above spectral descriptors with various lengths. The rows of the Figure, from top to bottom, correspond to increasing descriptor lengths of 12, 18, 24 and 30. The first column shows the original input image rendered in pseudo colour. The remaining columns, from left to right, correspond to skin probability maps obtained by the B-Spline descriptor, the Gaussian Mixture descriptor, LDA, PCA and
4.3. EXPERIMENTS

Figure 4.11: Skin probability maps for sample representative subjects of different ethnic groups. Column (a): input images rendered in pseudo colour. Columns (b)–(e): skin probability maps delivered by 12-element descriptors (Spline-based, Gaussian Mixture, LDA and PCA, respectively). Column (f): the skin probability maps delivered by the input reflectance spectra consisting of 30 bands. Columns (g)–(j): the skin probability maps delivered by 30-element descriptors (Spline-based, Gaussian Mixture, LDA and PCA respectively). The brightness on the skin probability maps indicates the predicted skin likelihood.
the raw reflectance spectra. As before, the skin probability maps shown here are, indeed, consistent with the observations drawn from Table 4.5 and Figure 4.9. In fact, the Gaussian Mixture delivers noisier skin maps than the other features due to overfitting. On the other hand, the B-Spline descriptor produces similar skin maps to PCA, LDA and the raw reflectance. Note that the skin maps produced by the B-Spline descriptor do not vary significantly with respect to the descriptor length. Therefore, in addition to its reconstruction capability, the B-Spline descriptor offers competitive recognition rates with lower complexity, as compared to the full reflectance spectra.

In Figure 4.11, we show the skin probability maps for representatives of several ethnic groups. Here, we observe the same trend as in Figure 4.10, where the Gaussian Mixture descriptor produces a much noisier skin maps than the others, in both the skin and non-skin regions. Meanwhile, the Spline-based descriptor delivers similar skin segmentation maps to PCA, LDA and the raw spectral reflectance feature, irrespective of the descriptor length. Furthermore, this trend is consistent across different ethnic groups.

### 4.3.3 Skin Biometrics

Having shown results on skin recognition, we now examine the last of our experimental vehicles. In this section, we illustrate the utility of our representation for purposes of ethnic group identification from skin reflectance. Here, we perform our experiments on the imagery used in the previous sections and obtain our representation using, as before, a third-degree B-Spline curve and fixed the descriptor length to 15 control points, which corresponds to 19 knots.

Following our skin recognition experiments, we recovered a descriptor vector from the optimal knots and control points, according to the Equations 4.6 and 4.7. For the classification of subjects in our dataset into the main ethnic groups (Caucasian, Indian and Oriental), we used training data from two typical representatives randomly selected from each ethnic group. To this end, we used frontally-illuminated images of these training subjects and select 5 rectangular skin regions from each image, with an average size of \(25 \times 17\) pixels. This gave rise to three training classes (one for each ethnic group) which we used to train three Support Vector Machines (SVM) classifiers.
4.3. EXPERIMENTS

Figure 4.12: Skin type association maps of three subjects belonging to the Caucasian, Indian and Oriental groups, listed from top to bottom. The red, green and blue values correspond to the likelihood that each pixel belongs to the Caucasian, Indian and Oriental groups, respective. First column: input images rendered in pseudo colour; Remaining columns, from left to right: Skin type association maps for the Spline-based descriptor, the Gaussian Mixture, LDA, PCA and the raw reflectance spectra.

According to a one-versus-all strategy. Each of these employed a first-degree polynomial kernel with its parameters tuned through a 4-fold cross validation procedure. During the test phase, the three SVM classifiers assigned an ethnic group association probability to each skin pixel in the test image. The overall ethnic identity associated with the test subjects was determined by the majority voting of the ethnic group association across their skin pixels.

For comparison purposes, we consider alternative reflectance spectra descriptors. The first of these consists of the coefficients, means and standard deviations of a mixture of 5 Gaussian components fitted to the given reflectance spectra [3]. We also compare the performance of the B-Spline descriptor with that delivered by the raw reflectance spectra and by their 15 principal LDA and PCA components.

In Figure 4.12, we present the association probability maps of several subjects with respect to the three skin groups. Each row of the figure shows the probability maps of a subject representing each skin group, namely the Caucasian, Indian and Oriental groups, from top to bottom. On the maps, the values of the red, green and
Figure 4.13: Visualisation of the ethnic group association of test subjects in a 3D coordinate system. The sub-figures show the plots of the ethnicity association probabilities yielded by (a) the third-degree B-Spline descriptor with 19 knots and 15 control points; (b) the Gaussian Mixture descriptor with 5 components; (c) the 15 principal LDA components of reflectance spectra; (d) the 15 principal PCA components of reflectance spectra; (e) raw reflectance spectra.
4.3. EXPERIMENTS

blue channels of a pixel are proportional to the association probability of that pixel to the Caucasian, Indian and Oriental groups, respectively. As shown by the Figure, the Spline-based descriptor, the PCA and the raw reflectance feature produce similar skin group association maps, which correctly identify the expected ethnic groups of the subjects. Their results significantly outperform the Gaussian Mixture descriptor and LDA. While the Gaussian Mixture descriptor correctly recognises the skin groups of the Indian and Oriental subjects, it suffers from confusion between the Caucasian and Oriental skin groups when being used to recognise the ethnic group of the Caucasian subject, as shown in the first row of Column (c). Meanwhile, the LDA feature experiences greater confusion between the three skin groups. In particular, its degree of confusion results in the incorrect classification of the Indian subject as Oriental. In contrast, all the other spectral features were able to clearly separate the Indian subject from the other skin groups.

To provide a summary of the skin association probability distribution of the hyperspectral image database, in Figure 4.13, we visualise the ethnic group association probabilities of the test subjects as points in a 3-D coordinate system, where the three axes represent the main ethnic groups in our dataset. The ground-truth ethnic groups of the test subjects are colour-coded, as red, green and blue for Caucasian, Indian and Oriental, respectively. In all the plots except that for the LDA feature, the Indian skin cluster is well-separated from the other groups. This is expected as the Indian skin reflectance spectra from our database have “flatter” slopes and less absorption bands than the other two groups. It is noticeable that the B-Spline descriptor provides the highest degree of separation between the Caucasian and the Oriental groups, with the least number of confounded examples, followed by PCA and the raw reflectance spectra, which yield similar distributions. We also notice that the skin groups obtained with the B-Spline descriptor are clustered most tightly, thus allowing more accurate classification of the skin groups. For the B-Spline descriptor, PCA and the Raw reflectance spectra, the Caucasian and Oriental groups are fairly separable from each other with a few examples confounded. This is probably due to image saturation and shadow. It is also noted that each ethnic group is correctly located within the proximity of the extreme point corresponding to the pure skin example of the group. The distributions
Table 4.6: Percentage of individuals classified into the Caucasian (C), Indian (I) and Oriental (O) ethnic groups. The diagonal elements for each sub-table spanning columns 3–5 present the overall classification rates of ethnic groups at the individual level, whereas the off-diagonal elements show the “confusion” between these groups per descriptor under study.

<table>
<thead>
<tr>
<th></th>
<th>Classified</th>
<th>Caucasian(C)</th>
<th>Indian(I)</th>
<th>Oriental (O)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spline</td>
<td>Caucasian(C)</td>
<td>90</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Indian(I)</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Oriental(O)</td>
<td>6.25</td>
<td>0</td>
<td>93.75</td>
</tr>
<tr>
<td>Gauss Mix.</td>
<td>Caucasian(C)</td>
<td>60</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Indian(I)</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Oriental(O)</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>LDA</td>
<td>Caucasian(C)</td>
<td>45</td>
<td>0</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Indian(I)</td>
<td>0</td>
<td>66.67</td>
<td>33.33</td>
</tr>
<tr>
<td></td>
<td>Oriental(O)</td>
<td>6.25</td>
<td>0</td>
<td>93.75</td>
</tr>
<tr>
<td>PCA</td>
<td>Caucasian(C)</td>
<td>85</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Indian(I)</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Oriental(O)</td>
<td>6.25</td>
<td>0</td>
<td>93.75</td>
</tr>
<tr>
<td>Reflectance</td>
<td>Caucasian(C)</td>
<td>85</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Indian(I)</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Oriental(O)</td>
<td>6.25</td>
<td>0</td>
<td>93.75</td>
</tr>
</tbody>
</table>

We evaluate the numerical accuracy of the ethnicity identification task using the Spline descriptor as opposed to the alternatives. In Table 4.6, we report the percentage of individuals classified into each ethnic group as a confusion matrix between ethnicities. Overall, the Spline-based descriptor outperforms the others in recognising the three ethnic groups. Although the Gaussian Mixture descriptor yields a higher recognition rate for the Oriental group, its overall performance is significantly lower than the
Table 4.7: Percentage of image pixels classified into the Caucasian (C), Indian (I) and Oriental (O) ethnic groups. The diagonal elements for each sub-table spanning columns 3–5 present the overall classification rates for ethnic groups at the pixel level, whereas the off-diagonal elements show the “confusion” between these groups per descriptor under study.

Spline descriptor due to its bias toward the Oriental group, i.e. 40% of the Caucasian subjects are classified as Oriental using the Gaussian Mixture descriptor. In addition, the B-Spline descriptor outperforms the LDA components in separating the Oriental group from the other two. This implies that the general shape of spectra, which the B-Spline descriptor conforms to, plays a pivotal role in discriminating spectra of similar kinds, i.e. skin spectra in this case. Following this hypothesis, the B-Spline descriptor is expected to deliver similar recognition rates to the raw reflectance spectra, which is consistent with the numerical results reported in the Table. Indeed, the B-Spline descriptor recognises the Caucasian group slightly better than PCA and the raw re-
To elaborate further on the quantitative results of this experiment, in Table 4.7, we report the percentage of pixels classified into each ethnic group as a confusion matrix between ethnicities. Again, the B-Spline descriptor outperforms the Gaussian Mixture descriptor and LDA, especially in separating the Caucasian and Oriental groups. While being better at recognising the Caucasian group, the B-Spline descriptor is comparable to PCA and the raw reflectance at recognising the other ethnic groups. However, it is worth stressing that, although PCA and the raw reflectance spectra deliver comparable performance to our method, the B-Spline representation also allows for reconstruction in addition to identification and recognition. In this regard, the raw spectra can be computationally costly for recognition purposes due to its high dimensionality.

4.4 Concluding Remarks

In this chapter, we have presented a novel B-Spline representation of reflectance spectra. This representation is quite general in nature and permits reconstruction, segmentation and recognition to be effected based on the control points resulting from the interpolation of B-Spline curves to individual spectra or imaging spectroscopy data. We have presented a knot removal scheme in the parameter domain which takes advantage of the local support property of Splines. We have also presented a method for enforcing a common knot vector and common basis functions on the interpolating B-Splines across reflectance spectra at all the pixels in the image. This delivers a representation which is not only highly descriptive but also efficient in terms of representation complexity. This procedure also provides a stable representation robust to local perturbation in the spectra, which can be used for pattern recognition and computer vision tasks. We have provided a quantitative study on the reconstruction capability of our representation for real-world spectra and hyperspectral images. We have also demonstrated the utility of our B-Spline representation for the purposes of skin segmentation and biometrics, where we have shown results on ethnic group recognition.
Chapter 5

Reflection Parameter Recovery from a Single Spectral Image

Colours have their greatest significance for us in so far as they are properties of bodies and can be used as marks of identification of bodies. Hence in our observations with the sense of vision we always start out by forming a judgment about the colours of bodies, eliminating the differences of illumination by which a body is revealed to us.

Hermann Ludwig Ferdinand von Helmholtz
German physician and physicist, 1821 - 1894
Treatise on Physiological Optics, Volume II, pages 286–287 [169]

In Chapter 4, we presented a compact representation of spectral signatures that offers efficient recognition and storage and accurate reconstruction capabilities. The representation described previously is quite general and can be applied to other spectral phometric invariants. To benefit from such a representation, we now turn our attention to the recovery of reflection parameters from a single spectral image.

Here, we address the simultaneous recovery of the illumination power spectrum, the material reflectance, the shading and the specularity in the scene from a single multispectral image. We defer the study of other photometric invariants such as refractive index until Chapter 6. Our proposed method assumes that the scene is uniformly illuminated. This assumption is common and valid for a wide range of situations, e.g.
where the scene is illuminated by natural sunlight or a distant light source. Based on the dichromatic reflection model [147], we pose the recovery problem in a structural optimisation setting. Making use of the spectral information provided by multispectral imaging and the structure of automatically selected image patches, we recover the dichromatic parameters of the scene. Since the objective function is convex with respect to each variable subset to be optimised upon, we can recover a closed-form solution which is iteration-wise optimal. We employ a surface smoothness error and other alternatives as regularisers for the optimisation problem. Subsequently, we show the successful application of our method to the tasks of illumination recovery and reflectance-based recognition. Lastly, the recovered reflection parameters can also be employed for specularity removal.

In contrast to prior literature on colour constancy, the work presented here integrates the recovery of the illuminant, photometric invariants, \textit{i.e.} the material reflectance, the shading and specularity factors from a single multispectral image in a unified optimisation framework. Not only the work extends the colour constancy problem from trichromatic to multispectral and hyperspectral imagery, but also confers several advantages. By optimising the data closeness to the dichromatic model, the method is generally applicable to surfaces exhibiting both diffuse and specular reflection. In addition, our method makes no assumption on the parametric form or prior knowledge of the illuminant and surface reflectance spectra. This is in contrast to other approaches where assumptions are made on the chromaticities of common light sources or the finite linear model of illuminants and surface reflectance. Compared to other methods which make use of the dichromatic model [49, 50, 164], our approach is able to perform well even on a small number of different material reflectance spectra. Furthermore, unlike the dichromatic plane-based methods for trichromatic imagery, our method does not require pre-segmented images as input. Instead, we employ an automatic dichromatic patch selection process to determine the uniform-albedo patches to be used for illuminant estimation. The noise perturbation analysis described in Section 5.4 shows that our illumination estimation method is more accurate than the alternatives and stable with respect to the number of surface patches used.
5.1. RECOVERY OF THE REFLECTION MODEL PARAMETERS

Moreover, the optimisation framework presented here is flexible and general in the sense that any regulariser that quantifies that spatial smoothness of the input image can be incorporated into the method. In Section 5.2, we present two instances of robust regularisers for the smoothness of the shading field. The utility of regularisers has been a common practice in early vision problems [133] and particularly in Shape-from-Shading [79], where regularisation together with occluding boundaries add supplementary constraints to render the underconstrained problem of inferring shape from shading well-posed [83]. Further, our objective function generalises prior colour constancy work [49, 144, 164] based on least-squares optimisation of the dichromatic model by controlling the surface smoothness through the use of regularisers. It is worth noting in passing that the shading factor in the dichromatic model reflects the angle between the incoming light direction and surface normals. Thus, the recovery of the shading factor by our optimisation method can be regarded as a pre-processing step for Lambertian Shape-from-Shading problems with spatially varying surface reflectance.

The chapter is structured as follows. In Section 5.1, we formulate an objective function to address the reflection parameter recovery problem. We elaborate further on the optimisation approach adopted here for the recovery of the parameters of the dichromatic reflection model. In Section 5.2 we show how smoothness constraints may be imposed upon the optimisation process. In Section 5.3, we show the applicability of our method, which is originally designed for hyperspectral images, to trichromatic imagery. In Section 5.4 we illustrate the utility of the method for the purposes of illuminant spectrum recovery, skin recognition, material clustering and specularity removal. This section mainly focuses on illumination recovery with supporting results from the skin recognition and material clustering experiments. In addition, it presents results for specularity removal purposes.

5.1 Recovery of the Reflection Model Parameters

Here, we present a structural approach based upon the processing of smooth surface-patches whose spectral reflectance is uniform over all the pixels they comprise. As mentioned earlier, the process of recovering the photometric parameters is based on
an optimisation method which aims at reducing the difference between the estimate yielded by the dichromatic model and the input image. In this section, we commence by providing an overview of the dichromatic model as presented by Shafer [147]. Subsequently, we formulate a target minimisation function with respect to the model in [147] and derive an optimisation strategy based upon the structure of smooth image patches with uniform reflectance. Throughout the section, we also present our strategy for selecting patches to be included in the objective function and describe in detail the coordinate descent optimisation procedure. This optimisation strategy is based upon interleaved steps aimed at recovering the light spectrum, the surface shading and surface reflectance properties so as to recover the optima of the dichromatic reflection parameters. Before further formalism, in Figure 5.1, we provide a summary of the main notation used throughout this chapter. These are defined in the order of appearance.

5.1.1 The Dichromatic Reflection Model

Throughout the chapter, we employ the dichromatic model introduced by Shafer [147] so as to relate the captured surface radiance to the light spectral power, surface reflectance and the scene geometry. Let us recall from Chapter 2 that the model assumes uniform illumination across the spatial domain of the observed scene. Consider an object with surface radiance $I(u, \lambda)$ and surface reflectance $S(u, \lambda)$ at pixel-location $u$ and wavelength $\lambda$ under an illumination spectrum $L(\lambda)$. With these ingredients, the dichromatic model is formulated as

$$I(u, \lambda) = g(u)L(\lambda)S(u, \lambda) + k(u)L(\lambda)$$  \hspace{1cm} (5.1)

In Equation 5.1, the shading factor $g(u)$ governs the proportion of diffuse light reflected from the object and depends solely on the surface geometry. For a purely Lambertian surface, $g(u)$ is proportional to the cosine of the angle between the surface normal $\vec{N}(u)$ and the light direction $\vec{L}$. On the other hand, the factor $k(u)$ models the proportion of light directly reflected at the air-material interface. To be precise, this factor is related to the Fresnel reflection coefficients, which in turn depends on the material refractive index. Although the refractive index is strictly wavelength-dependent,
its variation within the visible and near-infrared regions is generally slow for a wide variety of materials. This permits the assumption that the specularity coefficient is constant across the spectrum. Using this model, we aim to recover the shading factor

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I(u, \lambda)$</td>
<td>Scene radiance at pixel $u$ and wavelength $\lambda$.</td>
</tr>
<tr>
<td>$I(u)$</td>
<td>Spectral radiance vector at image pixel $u$, $I(u) \triangleq [I(u, \lambda_1), \ldots I(u, \lambda_n)]^T$.</td>
</tr>
<tr>
<td>$L(\lambda)$</td>
<td>Illuminant power at wavelength $\lambda$.</td>
</tr>
<tr>
<td>$L$</td>
<td>Spectral power vector of the illuminant, $L \triangleq [L(\lambda_1), \ldots L(\lambda_n)]^T$.</td>
</tr>
<tr>
<td>$S(u, \lambda)$</td>
<td>Reflectance at pixel $u$ and wavelength $\lambda$.</td>
</tr>
<tr>
<td>$S(u)$</td>
<td>Spectral reflectance vector at image pixel $u$, $S(u) \triangleq [S(u, \lambda_1), \ldots S(u, \lambda_n)]^T$.</td>
</tr>
<tr>
<td>$S_P$</td>
<td>The common spectral reflectance vector for a patch $P$ with a uniform-albedo, $S_P \triangleq [S_P(\lambda_1), \ldots S_P(\lambda_n)]^T$.</td>
</tr>
<tr>
<td>$S_P(\lambda)$</td>
<td>Spectral reflectance of a patch $P$ at wavelength $\lambda$.</td>
</tr>
<tr>
<td>$D_P$</td>
<td>Diffuse spectral radiance vector for a patch $P$ with a uniform-albedo, $D_P \triangleq L \cdot S_P$.</td>
</tr>
<tr>
<td>$D_P(\lambda)$</td>
<td>Diffuse radiance of a patch $P$ at wavelength $\lambda$.</td>
</tr>
<tr>
<td>$g(u)$</td>
<td>Shading factor at pixel location $u$.</td>
</tr>
<tr>
<td>$g_P$</td>
<td>Shading map of all pixels in patch $P$, $g_P = [g(u_1), \ldots g(u_l)]^T$ with $u_1, \ldots u_l$ being all the pixels in the patch $P$.</td>
</tr>
<tr>
<td>$g$</td>
<td>Shading map of all the patches in the image, $g = [g_{P_1}^T, \ldots g_{P_r}^T]^T$ where $P_1, \ldots P_r$ are all the patches in the image.</td>
</tr>
<tr>
<td>$k(u)$</td>
<td>Specularity coefficient at pixel location $u$.</td>
</tr>
<tr>
<td>$k_P$</td>
<td>Specularity map of all pixels in patch $P$, $k_P = [k(u_1), \ldots k(u_l)]^T$.</td>
</tr>
<tr>
<td>$k$</td>
<td>Specularity map of all the patches in the image, $k = [k_{P_1}^T, \ldots k_{P_r}^T]^T$.</td>
</tr>
<tr>
<td>$R(u)$</td>
<td>Regulariser, i.e. smoothness error at pixel $u$.</td>
</tr>
<tr>
<td>$C_c(\lambda)$</td>
<td>Spectral sensitivity of the trichromatic sensor $c$, where $c \in {R, G, B}$.</td>
</tr>
</tbody>
</table>

Figure 5.1: Notation in Chapter 5
g(u), the specular coefficient k(u), the light spectrum L(\lambda) and the spectral reflectance S(u, \lambda) at location u and wavelength \lambda from the spectral radiance I(u, \lambda) of the image.

### 5.1.2 Objective Function

With the dichromatic model above, we proceed to define our target function for purposes of optimisation. Our algorithm takes as input a multispectral image whose pixel values correspond to the measurements of the spectral radiance I(u, \lambda) indexed to the wavelengths \lambda \in \{\lambda_1, \ldots, \lambda_n\}. As mentioned previously, our goal is fitting the observed data to the dichromatic model to recover the parameters g(u), k(u) and S(u, \lambda). In general, here we view the dichromatic cost function of a multispectral image \mathcal{I} as the weighted sum of its dichromatic error and a regularisation term \mathcal{R}(u) for each image location. This is

\[
F(\mathcal{I}) = \sum_{u \in \mathcal{I}} \left[ \sum_{i=1}^{n} [I(u, \lambda_i) - L(\lambda_i)(g(u)S(u, \lambda_i) + k(u))]^2 + \alpha \mathcal{R}(u) \right]
\]

In equation 5.2, \alpha is a constant that acts as a balancing factor between the dichromatic error and the regularisation term \mathcal{R}(u) on the right-hand side. The wavelength-independent regularisation term \mathcal{R}(u) is related to the surface shading and will be elaborated upon later.

For now, we focus our attention on the solution space of Equation 5.2. Note that minimising the cost \(F(\mathcal{I})\) without further constraints is an underdetermined problem. This is due to the fact that, for an image with n spectral bands containing m pixels, we would have to minimise over \(2m + n + m \times n\) variables while having only \(m \times n\) terms in the summation of Equation 5.2. However, we notice that this problem can be further constrained if the model is applied to smooth surfaces made of the same material, i.e. the albedo is uniform across the patch or image region under consideration. This imposes two constraints. Firstly, all locations on the surface share a common diffuse reflectance. Therefore, a uniform albedo surface \(P\) is assumed to have the same reflectance for each pixel \(u \in P, S(u, \lambda_i) = S_P(\lambda_i)\). Note that this constraint significantly reduces the number of unknowns \(S(u, \lambda_i)\) from \(m \times n\) to \(N \times n\), where \(N\) is the number of surface albedos in the scene. In addition, the smooth variation of the patch geometry allows us to formulate the regularisation term \(\mathcal{R}(u)\) in equation 5.2 as a function of the
5.1. RECOVERY OF THE REFLECTION MODEL PARAMETERS

shading factor $g(u)$. In brief, smooth, uniform albedo surface patches naturally pro-
vide constraints so as to reduce the number of unknowns significantly while providing
a plausible formulation of the regularisation term $R(u)$.

Following the rationale above, we proceed to impose constraints on the minimisation
problem. For a smooth, uniform-albedo surface patch $P \in I$, we consider the
following cost function

$$ F(P) = \sum_{u \in P} \left[ \sum_{i=1}^{n} [I(u, \lambda_i) - L(\lambda_i)(g(u)S_{P}(\lambda_i) + k(u))]^2 + \alpha R(u) \right] $$

As before, we have $S(u, \lambda_i) = S_{P}(\lambda_i)$, for all $u \in P$. Furthermore, the smooth-
ness constraint on the patch implies that the shading factor $g(u)$ should vary smoothly
across the pixels in $P$. This constraint can be effectively formulated by minimising
the variation of gradient magnitude of the shading map. This, effectively, precludes
discontinuities in the shading map of $P$ via the regularisation term

$$ R(u) = \left[ \frac{\partial g(u)}{\partial x(u)} \right]^2 + \left[ \frac{\partial g(u)}{\partial y(u)} \right]^2 $$

(5.3)

where the variables $x(u)$ and $y(u)$ are the column and row coordinates, respectively, for
pixel location $u$.

Thus, by making use of the set $\mathcal{P}$ of uniform-albedo patches in the image $I$, we
can recover the dichromatic model parameters by minimising the target function

$$ F^*(I) = \sum_{P \in \mathcal{P}} F(P) $$

$$ = \sum_{P \in \mathcal{P}} \sum_{u \in P} \left[ \sum_{i=1}^{n} [I(u, \lambda_i) - L(\lambda_i)(g(u)S_{P}(\lambda_i) + k(u))]^2 + \alpha R(u) \right] $$

(5.4)

as an alternative to $F(I)$.

5.1.3 Homogeneous Patch Selection

In the previous section, we formulated the recovery of the dichromatic model param-
eters as an optimisation procedure over the surface patch-set $\mathcal{P}$. In this section, we
describe our method for automatically selecting uniform-albedo surface patches for
the minimisation of the cost function in Equation 5.4. The automatic patch selec-
tion method presented here allows the application of our method to arbitrary images.
CHAPTER 5. REFLECTION PARAMETER RECOVERY

Figure 5.2: The pseudo-colour rendering of a spectral image of a human subject captured in the visible range (430–720 nm). The image is overlaid with the rectangular boundaries (coloured in red) of all the homogeneous patches detected by the patch selection strategy described in Section 5.1.3.

It is worth noting that this contrasts with other methods elsewhere in the literature [49, 50, 164, 165], which are only applicable to pre-segmented images.

Our patch selection strategy is performed as follows. We first subdivide the image into patches of equal size in a lattice-like fashion. For each patch, we fit a two-dimensional hyperplane to the radiance vectors of the pixels in the patch. Next, we note that, in perfectly dichromatic patches, the wavelength-indexed radiance vector of each pixel lies perfectly in this hyperplane, *i.e.* the dichromatic plane. To allow for noise effect, we regard dichromatic patches as those containing a percentage of at most $t_p$ pixels whose radiance vectors deviate from their projection given by the Singular Value Decomposition (SVD) in [164]. We do this by setting a threshold $t_a$ on the angular deviation from the dichromatic plane, where $t_p$ and $t_a$ are global parameters.

In Figure 5.2, we show all the patches of uniform albedo in a spectral image of a human subject. The threshold values for patch selection are set at $t_a = 2^\circ$ and $t_p = 0.95$. The image was captured in the visible range (430–720 nm) and rendered in pseudo colour to emulate the colour matching function of the human eye as measured by Stiles and Burch [156, 157]. In the figure, we colour the rectangular boundaries of the patches in red.

However, not all these patches are useful for purposes of illumination spectrum
5.1. RECOVERY OF THE REFLECTION MODEL PARAMETERS

recovery. This is due to the fact that perfectly diffuse surfaces do not provide any information regarding the illuminant spectrum. The reason being that, a spectral radiance vector space for this kind of surfaces is one-dimensional, spanned only by the wavelength-indexed diffuse radiance vector. On the other hand, the dichromatic model implies that the specularities have the same spectrum as the illuminant, where the specular coefficient can be viewed as a scaling factor solely dependent on the surface shading.

Thus, for the recovery of the dichromatic model parameters, we only use highly specular patches by selecting regions with the highest contrast amongst those deemed to have a uniform albedo. We recover the contrast of each patch by computing the variance of the mean radiance over the spectral domain. These highly specular patches provide a means to the recovery of the light spectrum. This is due to the fact that, for highly specular surface patches with uniform albedo, the surface diffuse radiance vector and the illuminant vector span a hyperplane in the radiance vector space. This is a well known property in colour constancy, where a number of approaches [71, 92, 101] have employed subspace projection for purposes of light power spectrum recovery.

5.1.4 Optimisation Procedure

We now present the optimisation procedure employed in our method. Here, we adopt an iterative approach so as to find the global light power spectrum $L$, the reflectance $S_P$, the shading factor $g_P$ and the specular coefficient $k_P$ for each patch $P$, which yield the minimum of the cost function in Equation 5.4. At each iteration, we minimise the cost function with respect to $L$ and the triplet $\{g_P, k_P, S_P\}$ in separate steps.

The procedure presented here is, in fact, a coordinate descent approach [20] which aims at minimising the cost function. The step sequence of our minimisation strategy is summarised in the pseudocode of Algorithm 4. The coordinate descent approach comprises two interleaved minimisation steps. At each iteration, we index the dichromatic variables to iteration number $t$ and optimise the objective function, in interleaved steps, with respect to the two subsets of variables $\{g_P, k_P, S_P\}, \{L\}$. Once the former variables are in hand, we can obtain optimal values for the latter ones. We iterate between these two steps until convergence is reached.
The algorithm commences by initialising the unknown light spectrum $L(\lambda)$ to an unbiased uniform illumination spectrum, as indicated in Line 1 of Algorithm 4. It terminates once the illuminant spectrum does not change, in terms of angle, by an amount beyond a preset global threshold $t_L$ between two successive iterations. In the following two subsections we show that the two optimisation steps above can be employed to obtain the optimal values of the dichromatic parameters in closed form.

**Algorithm 4** Estimate dichromatic variables from a set of homogeneous patches

**Require:** Image $I$ with radiance $I(u, \lambda)$ for each band $\lambda \in \{\lambda_1, \ldots, \lambda_n\}$ and location $u$ and the collection of homogeneous patches $P$

**Ensure:** $L, S_P, g, k$, where

- $L$: the estimated illuminant spectrum.
- $S_P$: the diffuse reflectance of each surface patch $P$.
- $g, k$: the diffuse and specular reflection coefficients at all locations.

1: $t \leftarrow 1; L^0 \leftarrow I^T$
2: while true do
3: for all $P \in I$ do
4: $[g'_P, k'_P, S'_P] \leftarrow \arg\min_{g_P, k_P, S_P} F(P)|_{L^{t-1}}$
5: end for
6: $[L'] \leftarrow \arg\min_{S_{P_1}, \ldots, S_{P_r}} \sum_{P \in P} F(P)|_{g', k'}$
7: if $\angle(L', L^{t-1}) < t_L$ then
8: break
9: else
10: $t \leftarrow t + 1$
11: end if
12: end while
13: return $L', g', k', S'_{P_1}, \ldots, S'_{P_r}$

**Recovery of the Patch Reflectance, Shading and Specular Coefficients**

In the first step, we estimate the optimal surface reflectance and shading given the light spectrum $L^{t-1}$ recovered at iteration $t-1$. This corresponds to Lines 3–5 in Algorithm 4.
5.1. Recovery of the Reflection Model Parameters

Note that, at iteration $t$, we can solve for the unknowns $g_t^p, k_t^p$ and $S_t^p$ separately for each surface patch $P$. This is because, for each patch, these variables appear in a separate term in Equation 5.4. This step is, therefore, reduced to minimising

$$F(P)|_{L^{-1}} = \sum_{u \in P} \left[ \|I(u) - g(u)D_{P}^{-1} - k(u)L_{P}^{-1}\|^2 + \alpha R(u) \right]$$

(5.5)

where the diffuse radiance vector $D_{P}^{-1} \triangleq L_{P}^{-1} \cdot S_{P}$ is the component-wise multiplication of the illuminant and surface reflectance spectra, and $\|\|$ denotes the $L_2$-norm of the argument vectors.

Note that the minimisation above involves $2|P| + n$ unknowns, where $|P|$ is the number of pixels in patch $P$. Hence, it becomes computationally intractable when the surface area is large. In practice, the selected patches need only be large enough so as to gather useful statistics from the radiance information. Moreover, as mentioned earlier, we can further reduce the degrees of freedom of the unknowns by noting that the spectral radiance vectors at all pixels in the same surface lie in a $2$-dimensional subspace $Q \subset \mathbb{R}^n$, spanned by the diffuse radiance vector $D_{P}^{-1}$ and the light vector $L_{P}^{-1}$. This is a characteristic of the dichromatic model that has been widely utilised by prior work on colour constancy [49, 50, 164, 165].

Having all the pixel radiance vectors $I(u)$ at hand, one can obtain the subspace $Q$ via Singular Value Decomposition (SVD). Denote the two basis vectors resulting from this SVD operation $z_1$ and $z_2$ and, accordingly, let the subspace be $Q = \text{span}(z_1, z_2)$. Since $D_{P}^{-1} \in Q$, we can parameterise $D_{P}^{-1}$ up to scale as $D_{P}^{-1} = vz_1 + z_2$.

Likewise, the light vector $L_{P}^{-1} \in Q$ can also be decomposed as $L_{P}^{-1} = w_1z_1 + w_2z_2$, where the values of $w_1$ and $w_2$ are two known scalars. Furthermore, the dichromatic plane hypothesis also implies that, given the light vector $L_{P}^{-1}$ and the surface diffuse radiance vector $D_{P}^{-1}$, one can decompose any pixel radiance $I(u)$ into a linear combination of the former two vectors. In other words,

$$I(u) = g(u)D_{P}^{-1} + k(u)L_{P}^{-1} = (g(u)v + k(u)w_1)z_1 + (g(u) + k(u)w_2)z_2$$

(5.6)

Having obtained the basis vectors $z_1, z_2$, we can compute the mapping of the pixel radiance $I(u)$ onto the subspace $Q$. This is done with respect to this basis by means of
Equation 5.5, we have

\[ I(u) = \tau_1(u)z_1 + \tau_2(u)z_2 \]  

(5.7)

Further, by equating the right hand sides of Equations 5.6 and 5.7, we obtain

\[ g(u) = \frac{w_2\tau_1(u) - w_1\tau_2(u)}{w_2v_w - w_1} \]  

(5.8)

\[ k(u) = \frac{\tau_2(u)v - \tau_1(u)}{w_2v_w - w_1} \]  

(5.9)

From Equations 5.8 and 5.9, we note that \( g(u) \) and \( k(u) \) are univariate rational functions of \( v \). Moreover, \( D_{\tau}^{-1} \) is a linear function with respect to \( v \). We also observe that the term \( R(u) \) is only dependent on \( g(u) \). Therefore, the objective function in Equation 5.5 can be reduced to a univariate rational function of \( v \). Thus, substituting the Equations 5.8 and 5.9 into the first and second term on the right hand side of Equation 5.5, we have

\[
F(P)_{L^1} = \sum_{u \in P} \|I(u) - \frac{w_2\tau_1(u) - w_1\tau_2(u)}{w_2v_w - w_1}(vz_1 + z_2) - \frac{\tau_2(u)v - \tau_1(u)}{w_2v_w - w_1}L_{L^1}\|^2 \\
+ \sum_{u \in P} \frac{\alpha}{(w_2v_w - w_1)^2} \left[ \left( \frac{\partial m(u)}{\partial x(u)} \right)^2 + \left( \frac{\partial m(u)}{\partial y(u)} \right)^2 \right] \\
= \sum_{u \in P} \frac{1}{(w_2v_w - w_1)^2} \left[ \|I(u)w_2 - (w_2\tau_1(u) - w_1\tau_2(u))z_1 - \tau_2(u)L_{L^1}\|v \\
- \left( I(u)w_1 - (w_2\tau_1(u) - w_1\tau_2(u))z_2 - \tau_1(u)L_{L^1}\right) \|^2 \\
+ \frac{\alpha}{(w_2v_w - w_1)^2} \sum_{u \in P} \left[ \left( \frac{\partial m(u)}{\partial x(u)} \right)^2 + \left( \frac{\partial m(u)}{\partial y(u)} \right)^2 \right] \\
= \sum_{u \in P} \|p(u) - q(u)\|^2 + \frac{\alpha N}{(w_2v_w - w_1)^2} \\
= \sum_{u \in P} \left( \frac{w_1}{w_2}p(u) - q(u) \right)^2 + \frac{\alpha N}{(w_2v_w - w_1)^2} \\
= \sum_{u \in P} \|p(u)\|^2 + \frac{2}{w_2v_w - w_1} \sum_{u \in P} \left( \frac{w_1}{w_2}p(u) - q(u) \right) \\
+ \frac{1}{(w_2v_w - w_1)^2} \left( \sum_{u \in P} \left( \frac{w_1}{w_2}p(u) - q(u) \right)^2 + \alpha N \right) 

(5.10)
5.1. RECOVERY OF THE REFLECTION MODEL PARAMETERS

where $\langle \cdot, \cdot \rangle$ denotes the inner-product of two vectors, and

$$
m(u) = w_2 \tau_1(u) - w_1 \tau_2(u)$$

$$
p(u) = I(u)w_2 - (w_2 \tau_1(u) - w_1 \tau_2(u))z_1 - \tau_2(u)L_{t-1}$$

$$
q(u) = I(u)w_1 - (w_2 \tau_1(u) - w_1 \tau_2(u))z_2 - \tau_1(u)L_{t-1}$$

$$
N = \sum_{u \in P} \left[ \left( \frac{\partial m(u)}{\partial x(u)} \right)^2 + \left( \frac{\partial m(u)}{\partial y(u)} \right)^2 \right]$$

Note that $p(u), q(u), w_1$ and $w_2$ are known given the vector $L_{t-1}$. With the change of variable $r = \frac{1}{w_2 - w_1}$ we can write the right hand side of Equation 5.10 as a quadratic function of $r$ whose minimum is attained at

$$r^* = -\frac{\sum_{u \in P} \left( \frac{p(u)}{w_2} \right) w_2 (p(u) - q(u))}{\sum_{u \in P} \left( \frac{w_1}{w_2} \right) (p(u) - q(u))^2 + \alpha N} \quad (5.11)$$

This gives the corresponding minimiser $v^* = \frac{1}{w_2} (\frac{1}{r^*} + w_1)$. Hence, given the illuminant spectrum $L_{t-1}$, one can recover $g_p, k_p$ by substituting the optimal value of $v$ into Equations 5.8 and 5.9. The diffuse radiance component is computed as $D_P = v^* z_1 + z_2$, and the spectral reflectance at wavelength $\lambda$ is given by $S'_P(\lambda) = \frac{D'_P(\lambda)}{L_{t+1}(\lambda)}$.

Recovery of the Illuminant Spectrum

In the second step of each iteration $t$, we solve for $L'$ and $S'_{P_1}, \ldots, S'_{P_r}$ given $g'_P$ and $k'_P$. Since the second term $R(u)$ in Equation 5.4 is wavelength-independent, the optimisation problem in line 6 of Algorithm 4 can be reduced to minimising

$$F^*(I)_{|g', k'} = \sum_{P \in P} \sum_{u \in P} \left( I(u) - g'_P(u)D_P - k'_P(u)L \right)^2$$

$$= \sum_{P \in P} \sum_{u \in P} \sum_{i=1}^n (I(u, \lambda_i) - g'_P(u)D_P(\lambda_i) - k'_P(u)L(\lambda_i))^2 \quad (5.12)$$

where $D_P = L \cdot S_P$.

Since the objective function 5.12 is quadratic, and, therefore convex with respect to $L$ and $D_P$, the optimal values of these variables can be obtained by equating the respective partial derivatives of $F^*(I)_{|g', k'}$ to zero. These partial derivatives are given
CHAPTER 5. REFLECTION PARAMETER RECOVERY

by

\[
\frac{\partial F^*(I)|_{g^*,k^*}}{\partial L(\lambda_i)} = -2 \sum_{P \in P} \sum_{u \in P} (I(u, \lambda_i) - g'(u)D_p(\lambda_i) - k'(u)L(\lambda_i)) k'(u)
\]

\[
\frac{\partial F^*(I)|_{g^*,k^*}}{\partial D_p(\lambda_i)} = -2 \sum_{u \in P} (I(u, \lambda_i) - g'(u)D_p(\lambda_i) - k'(u)L(\lambda_i)) g'(u)
\]

Equating the above equations to zero, we obtain

\[
L(\lambda_i) = \frac{\sum_{P \in P} \sum_{u \in P} [k'(u)I(u, \lambda_i) - g'(u)k'(u)D_p(\lambda_i)]}{\sum_{P \in P} \sum_{u \in P} (k'(u))^2} (5.13)
\]

\[
D_p(\lambda_i) = \frac{\sum_{u \in P} [g'(u)I(u, \lambda_i) - g'(u)k'(u)L(\lambda_i)]}{\sum_{u \in P} (g'(u))^2} (5.14)
\]

From Equations 5.13 and 5.14, the illuminant spectrum can be solved in closed form as

\[
L^*(\lambda_i) = \frac{\sum_{P \in P} \sum_{u \in P} k'(u)I(u, \lambda_i) - \sum_{P \in P} \frac{(\sum_{P \in P} g'(u)k'(u))(\sum_{P \in P} g'(u)I(u, \lambda_i))}{\sum_{u \in P} (g'(u))^2}}{\sum_{P \in P} \sum_{u \in P} (k'(u))^2 - \sum_{P \in P} \frac{(\sum_{P \in P} g'(u)k'(u))^2}{\sum_{u \in P} (g'(u))^2}} (5.15)
\]

**Recovery of Image Shading, Reflectance and Specularity**

Note that, in the optimisation scheme above, we recover the reflectance, shading and specularity factors for pixels in each patch \( P \in \mathcal{P} \) used for the recovery of the illuminant spectrum. This implies that, although we have only computed the variables \( g(u), k(u) \) and \( S(u) \) for pixel-sites \( u \in \mathcal{P} \), we have been able to recover the illuminant spectrum \( L \). Since \( L \) is a global photometric variable in the scene, we can recover the remaining dichromatic variables making use of \( L \) in a straightforward manner. These include shading, reflectance and specularity factors for all image pixels.

For this purpose, we assume the input scene is composed of smooth surfaces with slowly varying reflectance. In other words, the neighbourhood of each pixel can be regarded as a locally smooth patch made of the same material, \( i.e. \) all the pixels in the neighbourhood share the same spectral reflectance. Given the illuminant spectrum, we can obtain the shading, specularity and reflectance of the neighbourhood at the pixel of interest by applying the procedure corresponding to line 4 in Algorithm 4. This corresponds to the application of the first of the two steps used in the optimisation method presented in the section above.
5.1. RECOVERY OF THE REFLECTION MODEL PARAMETERS

The pseudocode of this algorithm is summarised in Algorithm 5. Note that the assumption of smooth surfaces with slowly varying reflectance is applicable to a large category of scenes where surfaces have a low degree of texture, edges and occlusion. Following this assumption, the reflectance at each pixel is recovered as the shared reflectance of its surrounding patch. To estimate the shading and specularity, one can apply the closed-form formulae of these, as shown in Equations 5.8 and 5.9. These formulae yield exact solutions in the ideal condition, which requires that all the pixel radiance vectors lie in the same dichromatic hyperplane spanned by the illuminant spectrum and the diffuse radiance vector.

However, in practice, it is common for multi-spectral images to contain noise which breaks down this assumption and renders the above quotient expressions numerically unstable. Therefore, to enforce a smooth variation of the shading factor across pixels, we recompute the shading and specularity coefficients after obtaining the spectral reflectance. This is due to the observation that the reflectance spectrum is often more stable than the other two variables, i.e. shading and specularity factors. Specifically, one can compute the shading and specular coefficients as those resulting from the projection of pixel radiance onto the subspace spanned by the illuminant spectrum and the diffuse radiance spectrum vectors.

Algorithm 5 Estimate the shading, specularity and reflectance of an image knowing the illuminant spectrum

Require: Image $I$ with radiance $I(u, \lambda)$ for each band $\lambda \in \{\lambda_1, \ldots, \lambda_n\}$ and the illuminant spectrum $L$

Ensure: $g(u), k(u), S(u, \lambda)$ where $g(u), k(u)$: the shading and specularity at pixel location $u$. $S(u, \lambda)$: the diffuse reflectance of at pixel $u$ and wavelength $\lambda$.

1: for all $u \in I$ do
2:  $\mathcal{N} \leftarrow$ Neighbourhood of $u$
3:  $[g_\mathcal{N}, k_\mathcal{N}, S_\mathcal{N}] \leftarrow \arg\min_{g_\mathcal{N}, k_\mathcal{N}, S_\mathcal{N}} F(P)|_L$
4:  $S(u) \leftarrow S_\mathcal{N}$
5: end for
6: return $g(u), k(u), S(u)$
Similar to other photometric methods based on the dichromatic model, this framework breaks down when dichromatic hyper-plane assumption is violated, i.e., the illuminant spectrum is co-linear to the diffuse radiance spectrum of the material. This renders the subspace spanned by the radiance spectra of the patch pixels to collapse to a 1-dimensional space. As a consequence, a Singular Value Decomposition of these radiance spectra does not succeed in finding two basis vectors of the subspace. Since the diffuse component is a product of the illuminant power spectrum and the material reflectance, this condition implies that the material has a uniform spectral reflectance. In other words, the failure case only happens when the input scene contains a single material with a uniform reflectance, i.e., one that resembles a shade of gray.

This failure case is very rare in practice. In fact, when the scene contains more than one material, as more uniform albedo patches are sampled from the scene, there are more opportunities to introduce the non-collinearity between the illuminant spectrum and surface diffuse radiance spectrum. In short, our method guarantees the recovery of dichromatic model parameters on scenes with more than one distinct albedo.

5.2 Imposing Smoothness Constraints

In Section 5.1.2, we addressed the need of enforcing the smoothness constraint on the shading field \( g = \{g(u)\}_{u \in I} \) using the regularisation term \( R(u) \) in Equation 5.2. In Equation 5.3, we present a regulariser that encourages the slow spatial variation of the shading field. There are two reasons for using this regulariser in the optimisation framework introduced in the previous sections. Firstly, it yields a closed-form solution for the surface shading and reflectance, given the illuminant spectrum. Secondly, it is reminiscent of smoothness constraints imposed upon shape from shading approaches and, hence, it provides a link between other methods in the literature, such as that in [187] and the optimisation method in the previous sections. However, we need to emphasise that the optimisation procedure above by no means implies that the framework is not applicable to alternative regularisers. In fact, our target function is flexible in the sense that other regularisation functions can be formulated dependent on the surface at hand.
5.2. IMPOSING SMOOTHNESS CONSTRAINTS

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Robust kernel $\rho_\sigma(\eta)$</th>
<th>Influence function $\Gamma_\sigma(\eta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tukey</td>
<td>$\rho_\sigma(\eta) = \begin{cases} \sigma \left(1 - \left(1 - \left(\frac{\eta}{\sigma}\right)^2\right)^3\right) &amp; \text{if }</td>
<td>\eta</td>
</tr>
<tr>
<td>Li</td>
<td>$\rho_\sigma(\eta) = \sigma \left(1 - \exp\left(-\frac{\eta^2}{\sigma^2}\right)\right)$</td>
<td>$\Gamma_\sigma(\eta) = \exp\left(-\frac{\eta^2}{\sigma^2}\right)$</td>
</tr>
<tr>
<td>Huber</td>
<td>$\rho_\sigma(\eta) = \begin{cases} \eta^2 &amp; \text{if }</td>
<td>\eta</td>
</tr>
</tbody>
</table>

Table 5.1: Robust kernels and influence functions.

In this section, we introduce a number of alternative regularisers on the shading field that are robust to noise and outliers and adaptive to the surface shading variation. To this end, we commence by introducing robust regularisers. We then present extensions based upon the surface curvature and the shape index.

To quantify the smoothness of shading, an option is to treat the gradient of the shading field as the smoothness error. In Equation 5.3, we have introduced a quadratic error function of the smoothness. However, in certain circumstances, enforcing the quadratic regulariser as introduced in Equation 5.2 causes the undesired effect of over-smoothing the surface. This well-known phenomenon has been experienced in a number of developments [27, 83] in the field of Shape from Shading. It is worth noting in passing that ample work exists in the literature addressing the over-smoothing tendency of quadratic regularisers used for enforcing smoothness constraints on gradients [44, 187, 190].

As an alternative, we utilise kernel functions stemming from the field of robust statistics. Formally speaking, a robust kernel function $\rho_\sigma(\eta)$ quantifies an energy associated with both the residual $\eta$ and its influence function, i.e. measures sensitivity to changes in the shading field. Each residual is, in turn, assigned a weight as defined by an influence function $\Gamma_\sigma(\eta)$. Thus the energy is related to the first-moment of the influence function as $\frac{\partial \rho_\sigma(\eta)}{\partial \eta} = \eta \Gamma_\sigma(\eta)$. Table 5.1 shows the formulae for Tukey’s bi-weight [75], Li’s Adaptive Potential Functions [104] and Huber’s M-estimators [80].
5.2.1 Robust Shading Smoothness

Having introduced the above robust estimators, we proceed to employ them as regularisers for the target function. Here, several possibilities exist. One of them is to directly minimise the shading variation by defining robust regularisers with respect to the shading gradient. In this case, the regulariser $R(u)$ is given by the following formula

$$R(u) = \rho_\sigma \left( \left\| \frac{\partial g}{\partial x} \right\| + \rho_\sigma \left\| \frac{\partial g}{\partial y} \right\| \right)$$  \hspace{1cm} (5.16)

Despite effective, the formula above still employs the gradient of the shading field as a measure of smoothness. In the next section, we explore the use of curvature as a measure of consistency.

5.2.2 Curvature Consistency

Alternatively, one can instead consider the intrinsic characteristics of the surface at hand given by its curvature. Specifically, Ferrie and Lagarde [44] have used the global consistency of principal curvatures to refine surface estimates in Shape from Shading. Moreover, ensuring the consistency of curvature directions does not necessarily imply a large penalty for discontinuities of orientation and depth. Therefore, this measure can avoid oversmoothing, which is a drawback of the quadratic smoothness error.

The curvature consistency can be defined on the shading field by treating it as a manifold. To commence, we define the structure of the shading field using its Hessian matrix

$$H = \begin{pmatrix}
\frac{\partial^2 g}{\partial x^2} & \frac{\partial^2 g}{\partial x \partial y} \\
\frac{\partial^2 g}{\partial x \partial y} & \frac{\partial^2 g}{\partial y^2}
\end{pmatrix}$$

The principal curvatures of the manifold are hence defined as the eigenvalues of the Hessian matrix. Let these eigenvalues be denoted by $\kappa_1$ and $\kappa_2$, where $\kappa_1 \geq \kappa_2$. Moreover, we can use the principal curvatures to describe local topology using the Shape Index [94] defined as follows

$$\phi = \frac{2}{\pi} \arctan \left( \frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2} \right)$$  \hspace{1cm} (5.17)

The observation above is important because it permits casting the smoothing process of the shading field as a weighted mean process, where the weight assigned to a
5.2. IMPOSING SMOOTHNESS CONSTRAINTS

pixel is determined by the similarity in local topology, i.e. the shape index, about a local neighbourhood. Effectively, the idea is to favour pixels in the neighbourhood that belong to the same or similar shape class as the pixel of interest. This is an improvement over the quadratic smoothness term defined in Equation 5.3 because it avoids the indiscriminate averaging of shading factors across discontinuities. That is, it is by definition edge preserving.

For each pixel $u$, we consider a local neighbourhood $N$ around $u$ and assign a weight to each pixel $u^*$ in the neighbourhood as

$$w(u^*) = \exp\left(\frac{-(\phi(u^*) - \mu_\phi(N))^2}{2\sigma_\phi^2(N)}\right),$$

where $\mu_\phi(N)$ and $\sigma_\phi(N)$ are the mean and standard deviation of shape index over the neighbourhood $N$. Using this weighting process, we obtain an adaptive weighted mean regulariser as follows

$$R(u) = \left(\frac{g(u) - \sum_{u^* \in N} w(u^*) g(u^*)}{\sum_{u^* \in N} w(u^*)}\right)^2$$  \hspace{1cm} (5.18)

This approach can be viewed as an extension of the robust regulariser function with a fixed kernel, presented in Equation 5.16. To regulate the level of smoothing applied to a neighbourhood, we consider the shape index statistics [94] so as to adaptively change the width of the robust kernel. The rationale behind adaptive kernel widths is that a neighbourhood with a great variation of shape index requires stronger smoothing than one with a smoother variation. The regulariser function is exactly the same as Equation 5.16, except for the kernel width which is defined pixel-wise as

$$\sigma(u) = \exp\left(-\left(\frac{1}{K_\phi |N|} \sum_{u^* \in N} (\phi(u^*) - \phi(u))^2\right)^{1/2}\right)$$  \hspace{1cm} (5.19)

where $N$ is a neighbourhood around the pixel $u$, $|N|$ is the cardinality of $N$ and $K_\phi$ is a normalisation term.

With the above formulation of the kernel width, it can be observed that a significant variation of the shape index within the neighbourhood corresponds to a small kernel width, causing the robust regulariser to produce heavy smoothing. In contrast, when the shape index variation is small, a lower level of smoothing occurs due to a wider kernel width.

Note that the use of the robust regularisers introduced earlier in this section as an alternative to the quadratic regulariser does not preclude the applicability of the optimisation framework described in Section 5.1.4. In fact, the change of regulariser only
affects the formulation of the target function in Equation 5.10, in which the shading factor $g(u)$ can be expressed as a univariate function as given in Equation 5.8. Since all the above robust regularisers are only dependent on the shading factor, the resulting target function is still a function of the variable $r \doteq \frac{1}{w_2 + \cdots + w_l}$. Further, by linearisation of the robust regularisers, one can still numerically express the regulariser as a quadratic function of the variable $r$. Subsequently, the closed-form solution presented earlier stands as originally described.

5.3 Adaptation to Trichromatic Imagery

In this section, we show how to utilise the optimisation method above to recover the dichromatic parameters from trichromatic images. To this end, we transform the dichromatic model for multispectral images into one for trichromatic imagery. Let us denote the spectral sensitivity function of the trichromatic sensor $c$ (where $c \in \{R, G, B\}$) by $C_c(\lambda)$. The response of the sensor $c$ to the spectral irradiance arriving at the location $u$ is given by $I_c(u) = \int_{\Omega} E(u, \lambda) C_c(\lambda) d\lambda$, where $E(u, \lambda)$ is the image irradiance and $\Omega$ is the spectrum of the incoming light. Furthermore, it is well-known that the image irradiance is proportional to the scene radiance $I(u, \lambda)$, i.e. $E(u, \lambda) = K_{opt} \cos^4 \beta(u) I(u, \lambda)$, where $\beta(u)$ is the angle of incidence of the incoming light ray on the lens and $K_{opt}$ is a constant only dependent on the optics of the lens [78]. Hence, we have

$$I_c(u) = K_{opt} \cos^4 \beta(u) \int_{\Omega} I(u, \lambda) C_c(\lambda) d\lambda$$

$$= K_{opt} \cos^4 \beta(u) \int_{\Omega} (g(u) L(\lambda) S(u, \lambda) + k(u) L(\lambda)) C_c(\lambda) d\lambda$$

$$= K_{opt} \cos^4 \beta(u) \int_{\Omega} L(\lambda) S(u, \lambda) C_c(\lambda) d\lambda + K_{opt} \cos^4 \beta(u) k(u) \int_{\Omega} L(\lambda) C_c(\lambda) d\lambda$$

$$= g^*(u) D_c(u) + k^*(u) L_c$$

where $g^*(u) = K_{opt} \cos^4 \beta(u) g(u)$ and $k^*(u) = K_{opt} \cos^4 \beta(u) k(u)$.

Here we notice that $D_c(u) = \int_{\Omega} L(\lambda) S(u, \lambda) C_c(\lambda) d\lambda$ and $L_c(u) = \int_{\Omega} L(\lambda) C_c(\lambda) d\lambda$ are the $c$ component of the surface diffuse colour corresponding to the location $u$ and of the illuminant colour, respectively.
The dichromatic cost function for the trichromatic image $I$ of a scene is formulated as

$$F(I) \triangleq \sum_{u \in I} \left[ \sum_{c \in \{R,G,B\}} \left[ I_c(u) - (g^*(u)D_c(u) + k^*(u)L_c) \right]^2 + \alpha R(u) \right]$$

(5.20)

where $R(u)$ is a spatially varying regularisation term, as described in Equation 5.2.

It is worth noticing that the cost function in Equation 5.20 is a special case of Equation 5.2, where $n = 3$. Hence, the method of recovering the dichromatic parameters, as elaborated upon in Sections 5.1.3 and 5.1.4 can be applied to this case in order to recover the trichromatic diffuse colour $D(u) = [D_R(u), D_G(u), D_B(u)]^T$ and illuminant colour $L = [L_R, L_G, L_B]^T$, as well as the shading and specular factors $g(u)$ and $k(u)$ up to a multiplier.

### 5.4 Experiments

In this section, we perform experiments on a number of image databases so as to verify the accuracy of the recovered dichromatic parameters. Our datasets include indoor and outdoor multispectral and RGB images with uniform and cluttered backgrounds, under natural and artificial lighting conditions. For this purpose, we acquired in-house two multi-spectral image databases captured in the visible and near-infrared ranges. These consist of indoor images captured under artificial light sources and outdoor images under natural sunlight and skylight. From these two databases, two trichromatic image databases are synthesized for the spectral sensitivity functions of a Canon 10D and a Nikon D70 camera sensor and the CIE standard RGB colour matching functions [34]. Apart from these databases, we also compared the performance of our algorithm with the alternatives on the benchmark dataset reported by Barnard et al. in [9].

The indoor database includes images of 51 human subjects, each captured under one of 10 directional light sources with varying directions and spectral power. The light sources are divided into two rows. The first of these is placed above the camera system and the second one at the same height as the cameras. The main direction of the lights is adjusted so as to point towards the centre of the scene. The imagery was acquired using a pair of OKSI Turnkey Hyperspectral Cameras. These cameras are equipped with Liquid Crystal Tunable Filters which allow multi-spectral images to
be resolved up to 10 nm in both the visible (430–720 nm) and the near infrared (650–990 nm) wavelength ranges. To obtain the ground truth illuminant spectrum for each image, we measured the average radiance reflected from a white calibration target, i.e. a LabSphere Spectralon, illuminated by the light sources under consideration. Using the same camera system and calibration target, we captured the outdoor images of a paddock from four different viewpoints, each from seven different viewing angles at different times of the day.

In the following experiments, we explore the utility of the recovered parameters of the dichromatic model for multiple applications. Therefore, we present the main bulk of the experiments in Section 5.4.1, where we demonstrate the effectiveness of our method for illumination spectrum recovery. In Section 5.4.2, we present results for skin recognition and material clustering tasks. The purpose of the section is two-fold, one of which is to assess the robustness of the recovered reflectance for material recognition, the other is to reaffirm the accuracy of the illumination spectrum recovery results presented in Section 5.4.1. Lastly, we explore the use of the recovered shading and specularity coefficients for specularity removal in Section 5.4.3.

5.4.1 Illumination Spectrum Recovery

For our experiments on illumination spectra recovery, we compare the results yielded by our method to those delivered by the colour constancy method proposed by Finlayson and Schaefer [49]. In their work [49], they estimated illuminant colours based on the dichromatic model without prior assumptions on the illuminant statistics. Although their experiments were performed on trichromatic imagery, this method can be adapted to multispectral data in a straightforward manner. Their approach relies on the dichromatic plane hypothesis that, the colour vectors of pixels with the same reflectance span a two-dimensional space. Utilising this idea, illumination estimation is cast as an optimisation problem so as to maximise the total projection length of the light colour vector on all the dichromatic planes. Geometrically, this approach predicts the illuminant colour as the intersection of dichromatic planes, which may lead to a numerically unstable solution when the angle between dichromatic planes are small.

Finlayson and Schaefer’s method can be adapted to multispectral images as fol-
lows. First, we employ our automatic patch selection method to provide homogeneous patches as input for their colour constancy algorithm. Secondly, we solve the eigen-system of the sum of projection matrices on the dichromatic planes. The light colour vector is the eigenvector corresponding to the largest eigenvalue.

The other alternative used here is akin to the spectrum deconvolution approach proposed by Sunshine et al. [160] to recover the absorption bands characteristic of the surface material chemistry. This method makes use of the upperbound envelope of a reflectance spectrum, also known as its continuum, which can be regarded as a reflectance spectrum without any absorption feature. For illuminant recovery, we view the estimated illuminant spectrum as the continuum of the radiance spectra at all the pixels. The work in [160] assumes that the continuum is a linear function of the wave number, i.e. the reciprocal of wavelength, on the log reflectance scale. Making use of this assumption, it then fits this parametric form to the continuum of the radiance spectra to recover the illuminant. Note that the resulting illuminant does not rely on patch selection and is therefore independent of the number of patches.

The section is organised as follows. We commence by providing results on hyperspectral imagery. We then turn our attention to light colour recovery in trichromatic imagery. We conclude the section by providing a noise perturbation analysis.

**Multispectral Light Spectrum Recovery**

As mentioned above, we first focus our attention on the use of our dichromatic parameter recovery algorithm for illuminant spectrum estimation in hyperspectral imagery. To this end, we have performed experiments using 1, 5, 10, 20, 30, 40 and 50 automatically selected patches of uniform albedo. Each patch has a size of $20 \times 20$ pixels. The accuracy of light spectrum recovery is measured as the Euclidean deviation angle between the estimated and ground truth spectrum in $n$ dimensions, where $n$ is the number of sampled wavelengths. These results are then compared against those obtained by the method of Finlayson and Schaefer [49] and Sunshine et al.’s [160] on the same number of patches.

Table 5.2 shows the means and standard deviations of the angular errors, in degrees, over all images in the indoor face database versus of the number of selected
Table 5.2: Accuracy versus the number of patches used for our illuminant estimation method on the multispectral face image database captured under both the visible and near-infrared spectra, in degrees, compared to Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s method.

<table>
<thead>
<tr>
<th>No. patches</th>
<th>Visible spectrum</th>
<th>Near-infrared spectrum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Our method</td>
<td>F &amp; S</td>
</tr>
<tr>
<td>1</td>
<td>17.25 ± 6.07</td>
<td>20.55 ± 10.53</td>
</tr>
<tr>
<td>5</td>
<td>5.62 ± 5.54</td>
<td>7.52 ± 4.90</td>
</tr>
<tr>
<td>10</td>
<td>5.81 ± 5.63</td>
<td>7.42 ± 4.19</td>
</tr>
<tr>
<td>20</td>
<td>6.21 ± 5.22</td>
<td>7.37 ± 3.88</td>
</tr>
<tr>
<td>30</td>
<td>6.49 ± 5.53</td>
<td>7.32 ± 3.65</td>
</tr>
<tr>
<td>40</td>
<td>6.66 ± 5.86</td>
<td>7.29 ± 3.54</td>
</tr>
<tr>
<td>50</td>
<td>6.82 ± 5.84</td>
<td>7.26 ± 3.50</td>
</tr>
</tbody>
</table>

The results are reported with a weight $\alpha = 100000$ assigned to the regularisation term in Equation 5.2. In this experiment, the regularisation term is defined to be the smoothness of shading variation, as shown in Equation 5.3. To obtain an optimal value of $\alpha$, we perform a procedure similar to the grid search employed in cross validation. The procedure involves applying our algorithm on a randomly sampled portion of the database several times for different parameter values and then selecting the value that yields the highest overall performance.

As shown in Table 5.2, our algorithm achieves a higher accuracy than the alternative methods when using no more than 20 homogeneous patches for both spectral ranges. It is noticeable that even with a single patch, our method still significantly outperforms Finlayson and Schaefer’s method [49]. This observation is consistent with the well-known fact that the dichromatic plane method and its variants require at least two homogeneous surfaces to compute the intersection between the dichromatic planes.
5.4. EXPERIMENTS

Figure 5.3: Accuracy versus number of patches used of our illuminant estimation method on the multispectral face image database, in degrees, compared to Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s method. The results for both the visible (left) and near-infrared ranges (right) are shown.

In addition, the angular error of the estimated illuminant decreases as the number of patches increases from 1 to 5. Similar to Finlayson and Schaefer’s method, our algorithm performs better when employing more than one patch. The explanation for this rests on the following geometric interpretation. That is, the intersection of two or more dichromatic hyperplanes corresponding to the patches is more well-constrained than a vector lying on a single dichromatic hyperplane.

However, as the number of patches grows beyond 5, the angular error of our method tends to increase. The reason for this trend is that the reliability of the illuminant estimate rests on the degree of specularity of the selected patches. Our algorithm selects the most specular patches from the scene which are ranked in the decreasing order of specularity. As more patches are selected, the objective function involves a larger number of patches with weak or no specularity. The dichromatic hyperplane obtained by a Singular Value Decomposition on weakly specular or diffuse patches as described in Section 5.1.3 may not be reliable as these patches provide little information of the illuminant spectrum. As a results, the constraints introduced by these dichromatic hyperplanes may degrade the accuracy of the illuminant estimate.

Although our methods remains more accurate than Finlayson and Schaefer’s method
Figure 5.4: Ground truth illuminant spectra and those estimated by our method and the alternatives from the image of a human subject in the dataset illuminated by a high-oblique light direction (from above and to the left of the camera center). Here we show the estimated illuminant spectra in both the visible (left column) and near-infrared (right column). From top to bottom: the spectra estimated using 1, 5, 10, 40 and 50 patches from the image.
5.4. EXPERIMENTS

<table>
<thead>
<tr>
<th>No. patches</th>
<th>Visible spectrum</th>
<th>Near-infrared spectrum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Our method</td>
<td>F &amp; S</td>
</tr>
<tr>
<td>1</td>
<td>13.93 ± 2.83</td>
<td>22.17 ± 5.84</td>
</tr>
<tr>
<td>5</td>
<td>13.89 ± 1.25</td>
<td>14.06 ± 1.47</td>
</tr>
<tr>
<td>10</td>
<td>13.80 ± 1.17</td>
<td>14.09 ± 1.35</td>
</tr>
<tr>
<td>20</td>
<td>14.00 ± 1.43</td>
<td>14.03 ± 1.33</td>
</tr>
<tr>
<td>30</td>
<td>13.69 ± 1.53</td>
<td>14.00 ± 1.23</td>
</tr>
<tr>
<td>40</td>
<td>13.44 ± 2.09</td>
<td>13.99 ± 1.22</td>
</tr>
<tr>
<td>50</td>
<td>13.85 ± 1.48</td>
<td>13.97 ± 1.24</td>
</tr>
</tbody>
</table>

Table 5.3: Accuracy, in degrees, versus the number of patches used for our illuminant estimation method on the multispectral outdoor image database captured under both the visible and near-infrared spectra compared to Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s approach.

in the visible range, its accuracy is slightly lower than the latter with more than 20 patches in the near-infrared range. Nonetheless, our method is able to achieve a reasonable estimate with a small number of homogeneous patches. Lastly, we can conclude that Sunshine et al.’s method [160] is, in general, inferior to the other two.

To illustrate the statistics in Table 5.2, we show, in Figure 5.4, the plots of the estimated spectra of a light source illuminating an indoor scene. The plots show spectra in both the visible (left column) and infra-red (right column) ranges. In each row, from top to bottom, we show the results yielded by our method and the alternatives using a different number of homogeneous patches for illuminant estimation. As before, we show the plots for 1, 5, 10, 40 and 50 patches. The ground truth spectra, the spectra estimated by our method, Finlayson and Schaefer’s [49] and Sunshine et al.’s method [160] are drawn in red, blue, green and magenta, respectively. Note that the highest value of each spectra is normalised to unity.

This visual illustration is consistent with a common trend in Table 5.2, that increasing the number of patches from 1 to 5 yields a significant improvement of accuracy for illumination estimation. Noticeably, our method outperforms the alternatives in recovering illumination spectra in the visible range. Meanwhile, its performance for the
near-infrared range is comparable to Finlayson and Schaefer’s [49] and is better than Sunshine et al.’s method [160]. Moreover, our method is more robust than the others even when it uses a single homogeneous patch for light estimation.

In Table 5.3 and Figure 5.5, we show the accuracy of the recovered spectrum of natural sunlight illuminating the outdoor scene in our dataset. Here our algorithm is applied with a regularisation weight $\alpha = 10000000$. With this setting, our method always outperforms the alternatives in the visible range. Using 20 or more randomly selected patches is sufficient for our method to improve performance over Finlayson and Schaefer’s [49] on the near-infrared images. As before, our algorithm significantly outperforms the alternative methods in the case of a single uniform albedo patch. Figure 5.5 also illustrates the stability of our method with respect to the increase in the number of selected patches. It is also noticeable that the accuracy of all the algorithms for the outdoor image database is lower than that for the face database probably due to the low availability of uniform reflectance patches across the scene. These trends are visually demonstrated using a number of sample plots of the estimated spectra of natural sunlight, as shown in Figure 5.6.

Finally, we examine the robustness of our algorithm to noise in spectral imagery. To
Figure 5.6: Ground truth illuminant spectra and those estimated by our method and the alternatives from the image of an outdoor scene. Here we show the estimated illuminant spectra in both the visible (left column) and the near-infrared (right column) ranges. From top to bottom: the spectra estimated using 1, 10, 20, 30 and 40 patches from the image.
Figure 5.7: Accuracy of the estimated illuminant spectra versus the standard deviation of Gaussian noise. The vertical axis shows the angular deviation of the estimated spectra from the corresponding ground-truth, while the horizontal axis shows the standard deviation of Gaussian noise as the percentage of maximum brightness of the imagery. The performance for the visible spectrum is shown in the left-hand image, while that corresponding to the infra-red spectrum is shown on the right-hand panel.

Do this, we perturb the multispectral face image database with various levels of additive Gaussian noise. The noise has an increasing standard deviation between 0.5 and 2% of the maximum image brightness, with increments of 0.5%. In Figure 5.7, we plot the performance of our algorithm and the alternatives across various levels of noise in the visible (left-hand panel) and infra-red (right-hand panel) ranges. For our algorithm and Finlayson and Schaefer’s one [49], we employ all the homogenous patches recovered from the images. The regularisation weight for our method is $\alpha = 100000$. As shown in Figure 5.7, our method achieves a lower mean and standard deviation of the angular error than the other two in the visible spectrum. Moreover, in the infra-red spectrum, our method greatly outperforms Sunshine et al.’s [160] by more than two standard deviations of the angular error in the recovered illuminant spectra.

On the degradation in performance with an increasing level of noise, the Sunshine et al.’s method is most stable because it considers the upper bound of all the radiance spectra in an image, which is least affected by the level of Gaussian noise. However, it is the least accurate method of the three alternatives. Ours and Finlayson & Schaefer’s
5.4. EXPERIMENTS

<table>
<thead>
<tr>
<th>No. of patches</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CMF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our method</td>
<td>3.84 ± 2.86</td>
<td>4.18 ± 2.89</td>
<td>4.16 ± 2.79</td>
<td>3.99 ± 2.61</td>
<td>3.94 ± 2.58</td>
<td>3.84 ± 2.54</td>
<td>3.83 ± 2.55</td>
</tr>
<tr>
<td>Sunshine</td>
<td>6.09 ± 2.60</td>
<td>6.09 ± 2.60</td>
<td>6.09 ± 2.60</td>
<td>6.09 ± 2.60</td>
<td>6.09 ± 2.60</td>
<td>6.09 ± 2.60</td>
<td>6.09 ± 2.60</td>
</tr>
<tr>
<td><strong>Canon 10D</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our method</td>
<td>3.91 ± 2.87</td>
<td>4.34 ± 3.01</td>
<td>4.15 ± 2.72</td>
<td>4.09 ± 2.52</td>
<td>3.98 ± 2.45</td>
<td>3.90 ± 2.36</td>
<td>3.86 ± 2.34</td>
</tr>
<tr>
<td>F &amp; S</td>
<td>25.37 ± 10.81</td>
<td>3.72 ± 3.00</td>
<td>3.22 ± 2.17</td>
<td>2.82 ± 1.70</td>
<td>2.62 ± 1.46</td>
<td>2.54 ± 1.40</td>
<td>2.46 ± 1.35</td>
</tr>
<tr>
<td>Sunshine</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
</tr>
<tr>
<td><strong>Nikon D70</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our method</td>
<td>3.88 ± 2.82</td>
<td>4.26 ± 2.88</td>
<td>4.17 ± 2.88</td>
<td>3.91 ± 2.49</td>
<td>3.93 ± 2.46</td>
<td>3.85 ± 2.41</td>
<td>3.86 ± 2.40</td>
</tr>
<tr>
<td>F &amp; S</td>
<td>25.67 ± 10.83</td>
<td>4.26 ± 3.94</td>
<td>3.22 ± 2.30</td>
<td>2.74 ± 1.67</td>
<td>2.59 ± 1.50</td>
<td>2.48 ± 1.37</td>
<td>2.43 ± 1.31</td>
</tr>
<tr>
<td>Sunshine</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
<td>5.84 ± 2.06</td>
</tr>
</tbody>
</table>

Table 5.4: Accuracy of our illuminant estimation method with $\alpha = 10000$ on the synthesized RGB face image database, in degrees, compared to Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s approach. We show the results on RGB face images synthesized from the multispectral face imagery for the Stiles and Burch’s colour matching functions (rows 2–4), the spectral sensitivity response functions of a Canon 10D (rows 5–7) and a Nikon D70 camera (rows 8–10).

appear to degrade linearly with the level of noise, although the latter one degrades at a slower rate than our method. This can be explained by the fact that Finlayson & Schaefer’s method relies on an eigenvalue decomposition, which is equivalent to our method with a zero-regularisation term. Albeit obtaining a more robust solution for the illuminant, their method does not take surface shading and highlights into account. Not only can our approach estimate the illuminant spectrum, it is also capable of computing all the dichromatic parameters, while maintaining a reasonable estimation using regularisation.

Trichromatic Light Recovery

Next, we turn our attention to the utility of our parameter recovery method for the purpose of illuminant colour estimation from trichromatic images. To this end, we generate RGB imagery from the multispectral face and outdoor databases mentioned
previously. These are synthesized by simulating a number of trichromatic spectral responses, including the CIE-1932 colour matching functions [34] and the camera sensors for a Nikon D70 and a Canon 10D. Furthermore, we apply our method and the alternatives to the Mondrian and specular image datasets as described by Barnard et al [10]. We also compare the performance of our algorithm with several colour constancy approaches described in [9] by the same authors.

To illustrate the effect of varying the value of $\alpha$, we perform experiments with $\alpha = 10000$ and $\alpha = 100$. In Table 5.4 and Figure 5.8 we show results for the light estimation accuracy on the RGB face images with $\alpha = 10000$ and a patch size of $20 \times 20$ pixels. Our method outperforms the alternatives in terms of estimation accuracy and stability when the number of patches (or the number of available intrinsic colours in the scene) is 5 or less. Another general trend is that our method and Finlayson and Schaefer’s [49] improve their accuracy as the number of selected patches increases. However, this improvement is marginal for our method when we use 20 or more patches. Meanwhile, the method of Finlayson and Schaefer [49] tends to achieve a closer estimate to the groundtruth when it is applied on a sufficiently large number of patches from the images synthesized for the Canon 10D and Nikon D70 sensors. Interestingly, for images simulated using the colour matching functions, which emulate the human visual perception, our method achieves a similar accuracy to Finlayson and Schaefer’s [49] across all the number of selected patches, while being more stable (with lower variance of angular error). In all our experiments, the approach of Sunshine et. al [160] is the one that delivers the worst performance.

Table 5.5 and Figure 5.9 show the accuracy of the illuminant colour estimation on the outdoor RGB image database, with $\alpha = 100$ and a patch size of $20 \times 20$ pixels. The major trend of these statistics is that our method achieves an accuracy significantly higher than those achieved by the others. The difference in performance is in the order of several standard deviations of the angular error. While the performance of our method is slightly degraded as the number of patch increases (above 20), the stability of the estimate remains constant for Canon 10D and Nikon D70 images, and even improves for the images simulated for the CIE 1932 standard colour matching functions. It is also important to notice that our method performs better on the imagery
5.4. EXPERIMENTS

(a) CMF
(b) Canon 10D
(c) Nikon D70

Figure 5.8: Accuracy versus number of image patches for our illuminant estimation method with $\alpha = 10000$ on the RGB images synthesized from the multi-spectral face imagery, in degrees, compared to Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s approach. The figures include results for simulated RGB images as captured with (a) Stiles and Burch’s colour matching functions (b) a Canon 10D camera sensor (c) a Nikon D70 camera sensor.
the multispectral outdoor imagery for the Stiles and Burch’s colour matching functions (rows 10 and Sunshine RGB outdoor image database, in degrees, compared to Finlayson & Schaefer’s method (F & S)).

- When applied to cluttered scenes with a high variation in colour and texture.
- synthetised using the CIE 1932 standard. This is consistent with the results reported for the RGB face images. Overall, our method appears to outperform the alternatives when applied to cluttered scenes with a high variation in colour and texture.

Next, we turn our attention to the illuminant estimation accuracy on the Mondrian and specular image datasets reported in [10]. To account for the level of texture density in some of the images, we choose a patch size of 10 × 10 pixels, which is small enough so that the assumption of uniform albedo across each patch still holds. In this experiment, a patch is regarded to be of homogeneous reflectance if 75% or more of the patch pixels deviate by less than 1 degree from their projection on the dichromatic plane of the patch. We also enforce a criterion that precludes the selection of highly contrasting patches which contain more than one material or saturated highlight pixels. Specifically, we rank patches in an image by their contrast levels and select the most contrasting ones, excluding those in the top 10% in each image of the Mondrian dataset. For the specular dataset, we exclude the top 30% percent of the patches in

<table>
<thead>
<tr>
<th>No. of patches</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIE 1932</td>
<td>0.46 ± 0.38</td>
<td>0.50 ± 0.36</td>
<td>0.97 ± 1.60</td>
<td>1.32 ± 2.02</td>
<td>1.40 ± 1.72</td>
<td>2.18 ± 2.23</td>
<td>2.34 ± 2.27</td>
</tr>
<tr>
<td>Our method</td>
<td>23.93 ± 6.66</td>
<td>10.74 ± 2.27</td>
<td>11.23 ± 2.03</td>
<td>11.28 ± 0.98</td>
<td>11.04 ± 0.91</td>
<td>10.95 ± 1.00</td>
<td>10.99 ± 0.89</td>
</tr>
<tr>
<td>Sunshine</td>
<td>11.35 ± 1.46</td>
<td>11.35 ± 1.46</td>
<td>11.35 ± 1.46</td>
<td>11.35 ± 1.46</td>
<td>11.35 ± 1.46</td>
<td>11.35 ± 1.46</td>
<td>11.35 ± 1.46</td>
</tr>
<tr>
<td>F &amp; S</td>
<td>0.94 ± 1.32</td>
<td>0.82 ± 1.22</td>
<td>1.08 ± 1.56</td>
<td>1.08 ± 1.63</td>
<td>1.67 ± 2.06</td>
<td>1.76 ± 2.08</td>
<td>2.22 ± 2.37</td>
</tr>
<tr>
<td>Our method</td>
<td>22.41 ± 9.68</td>
<td>9.33 ± 2.22</td>
<td>9.42 ± 2.06</td>
<td>9.68 ± 1.45</td>
<td>9.66 ± 1.14</td>
<td>9.73 ± 1.03</td>
<td>9.63 ± 0.99</td>
</tr>
<tr>
<td>Sunshine</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
</tr>
<tr>
<td>Canon 10D</td>
<td>0.78 ± 1.42</td>
<td>0.66 ± 0.66</td>
<td>0.90 ± 1.02</td>
<td>1.60 ± 2.17</td>
<td>1.99 ± 2.63</td>
<td>2.24 ± 2.57</td>
<td>2.70 ± 2.82</td>
</tr>
<tr>
<td>Sunshine</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
</tr>
<tr>
<td>Nikon D70</td>
<td>0.78 ± 1.42</td>
<td>0.66 ± 0.66</td>
<td>0.90 ± 1.02</td>
<td>1.60 ± 2.17</td>
<td>1.99 ± 2.63</td>
<td>2.24 ± 2.57</td>
<td>2.70 ± 2.82</td>
</tr>
<tr>
<td>Sunshine</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
<td>10.33 ± 1.40</td>
</tr>
</tbody>
</table>

Table 5.5: Accuracy of our illuminant estimation method with α = 100 on the synthesized RGB outdoor image database, in degrees, compared to Finlayson & Schaefer’s method (F & S) and Sunshine et al.’s approach. We show the results on RGB outdoor images synthesized from the multispectral outdoor imagery for the Stiles and Burch’s colour matching functions (rows 2–4), the spectral sensitivity response functions of a Canon 10D (rows 5–7) and a Nikon D70 camera (rows 8–10).
5.4. EXPERIMENTS

Figure 5.9: The accuracy of our illuminant estimation method with $\alpha = 100$ on the RGB images synthesized from the multi-spectral outdoor image database, in degrees, compared to Finlayson & Schaefer’s method (F & S) and Sunshine et al.’s method, versus the number of patches used. The figures include results for simulated RGB images as captured with (a) Stiles and Burch’s colour matching functions [34] (b) a Canon 10D camera sensor and (c) a Nikon D70 camera sensor.
Table 5.6: Accuracy of our illuminant estimation method with $\alpha = 1000$ on the Mondrian and specular datasets reported in [10], in degrees, compared to Finlayson and Schaefer’s method (F & S) and Sunshine et al.’s approach. The accuracy is measured in degrees.

<table>
<thead>
<tr>
<th>No. patches</th>
<th>Standard dynamic range (8 bits)</th>
<th>Extended dynamic range (16 bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Our method</td>
<td>F &amp; S</td>
</tr>
<tr>
<td>5</td>
<td>7.78 ± 6.23</td>
<td>9.88 ± 10.71</td>
</tr>
<tr>
<td>10</td>
<td>7.78 ± 6.46</td>
<td>10.06 ± 10.89</td>
</tr>
<tr>
<td>20</td>
<td>7.77 ± 6.69</td>
<td>9.41 ± 9.70</td>
</tr>
<tr>
<td>30</td>
<td>7.75 ± 6.90</td>
<td>9.29 ± 9.76</td>
</tr>
<tr>
<td>40</td>
<td>7.82 ± 6.81</td>
<td>9.12 ± 9.41</td>
</tr>
<tr>
<td>50</td>
<td>7.88 ± 6.77</td>
<td>9.25 ± 9.54</td>
</tr>
</tbody>
</table>

Figure 5.10: Accuracy of our illuminant estimation method with $\alpha = 1000$ for the Mondrian and specular datasets with 8-bit and 16-bit dynamic ranges described in [10]. Our method is compared to that of Finlayson and Schaefer (F & S) and Sunshine et al.

Table 5.6 and Figure 5.10 show the accuracy when 1, 5, 10, 20, 30, 40 and 50 patches are used. Our results are consistent with previous experiments, which shows that our method outperforms the alternatives on both the 8-bit and 16-bit datasets. This
is reflected not only by a lower mean of angular error yielded by our method, but also a lower standard deviation of its performance. In addition, our method delivers a variance of angular error which is almost constant with 5 or more selected patches. Further, the performance of our method improves slightly by increasing the number of selected patches up to 30 for the 8-bit dynamic range and up to 50 for the 16-bit dynamic range. To some extent, our estimator appears to be insensitive to the dynamic range of the input image. This shows that our method is more stable and robust to variations in the scene reflectance.

In comparison to the benchmark methods reported by Barnard et al. [9], our method ranks second, just below the gamut mapping methods presented in [52]. The methods in [52] deliver an accuracy between 5.6 and 7.1 degrees for the 16-bit images and between 6.3 and 8.3 degrees for the 8-bit images, as shown in Table II in [9]. However, these results were reported for imagery that had already undergone several processing steps including segmentation, scaling and clipping operations. On the other hand, our method does not require any preprocessing and, moreover, it is capable of recovering all the dichromatic-model parameters and is applicable to hyperspectral imagery with trichromatic imagery as a particular case.

5.4.2 Skin Recognition and Material Clustering

We now turn our attention to the illumination invariance of the spectral reflectance recovered by our algorithm and its applications to recognition tasks. Firstly, we focus on using the spectral image reflectance extracted according to Section 5.1.4 for skin recognition. This task can be viewed as a classification problem where the skin and non skin spectra comprise positive and negative classes, respectively. In this manner, we can assert the robustness and consistency of both the illuminant spectrum and surface reflectance recovered by our algorithm at training time, and those yielded by the method for skin recognition at the testing phase.

In this experiment, we compare the skin recognition performance yielded using the reflectance spectra recovered by our method as the feature for classification to those results yielded by the classifier using a number of alternative features. To this end, we present the results for two variants of our recovered reflectance, both estimated by the
procedure described in Section 5.1.4. For the first variant, the ground-truth illuminant spectrum is supplied as input. For the second one, we use the estimated illuminant spectra obtained by the experiments in Section 5.4.1. By comparing the performance in these two cases, we can assess the robustness of the recovered reflectance when the estimated illuminant spectra is used as compared to the ground-truth. In addition, we also compare these variants with a number of alternatives. The first of these is the spectral reflectance obtained by normalising the raw image radiance spectra by the measured ground-truth illuminant. The second case is where the classifier is applied to the raw radiance spectra. Lastly, we use the principal components resulting from performing subspace projection via Linear Discriminant Analysis (LDA) on the original radiance spectra.

This experiment was performed on the face image database captured in the visible range described earlier. To obtain a training data-set, we selected skin and non skin regions from an image captured under a light source placed in a high-oblique position in front of the subject. On average, there were 856 skin pixels and 7796 non-skin pixels selected from several regions in each image as training data. Subsequently, each of the features described above was extracted from each training set and used as input to a Support Vector Machine (SVM) classifier [30] with a Radial Basis Function (RBF) kernel. In addition, the parameters were selected using 5-fold cross validation at training time. To classify skin versus non skin pixels, the resulting SVM model was applied to the test images of the same subject. The test images, each with a size of $340 \times 400$ pixels, were acquired under other illuminant conditions.

In Figure 5.11 we present the skin segmentation maps obtained using the input features described above. The top row shows the training images of a number of sample subjects, with skin training regions enclosed in red rectangular boundaries and non-skin training areas enclosed in blue rectangles. The second row, from top to bottom, shows the test images for the subjects in the top row. Note that the illuminant directions and power spectra in the two rows differ, as can be observed in the shading and shadows. In fact, the training images are illuminated by the light source placed in a high-oblique position in front of the subjects whereas the test images are illuminated by a frontal light source with the same direction as the viewing direction. The bottom
5.4. EXPERIMENTS

Figure 5.11: Skin segmentation results. Top row: sample training images for skin recognition, with labelled skin regions (with red borders) and non-skin regions (in blue borders). Second row: the test images of the corresponding subjects, captured under a different illumination direction. The third to last row are the skin probability map obtained using different features. Third row: obtained using the reflectance estimated given the ground-truth illuminant spectrum. Fourth row: yielded using the reflectance estimated given the estimated illuminant spectrum. Fifth row: yielded using the reflectance obtained by normalising radiance spectra by the ground-truth illuminant spectrum. Sixth row: yielded using the raw radiance of the input images. Seventh row: recovered making use of the top 20 LDA components of the raw radiance spectra.
five rows are the skin probability maps yielded by the SMVs trained using the features described above. In the figure, lighter pixels are classified as being more likely to be skin. The third and fourth rows correspond to the variants of our recovered reflectance, with ground-truth illuminant and estimated illuminant spectra supplied as input, respectively. The fifth, sixth and seventh rows correspond to the reflectance obtained by normalising the image radiance by the ground-truth illuminant spectrum, the raw radiance and the top 20 LDA components of the image radiance spectra, respectively.

From Figure 5.11, we can conclude that the skin reflectance spectra recovered by our method are, in fact, invariant to illuminant power and direction. This stems from the fact that the reflectance features delivered by our method yield the most visually accurate skin maps. In many cases, non-skin face details such as eyebrows and mouth are correctly distinguished from skin. Furthermore, the results of the two reflectance variants are highly consistent. This is due to the low difference between the estimated illuminant and the ground truth, which varies typically between 1 and 3 degrees.

On the other hand, the reflectance features used for the results on the fifth row, although being illuminant invariant, still yield falsely classified skin pixels. The poor classification results obtained by these features can be explained by the variation induced by the illuminant spectrum and the surface shading. This is evident at pixels near the face boundary and the highlight positions. This is expected in the fifth row since normalising radiance by illuminant power does not achieve surface shading-independence and does not discount the specular components inherent in the dichromatic model. In contrast, our method achieves the recovery of the reflectance free of geometry and specularity artifacts. Thus it is able to recognise skin pixels at grazing angles and specular spikes. In addition, normalised raw radiance spectra and their LDA components, as employed for the classification on the sixth and seventh rows are not illuminant invariant. Therefore these cannot cope with the change in illumination between the training and test images. As shown in the last two rows, this results in much more false negatives in skin areas and false positives in other materials as compared to the reflectance features yielded by our method.

In order to provide a quantitative analysis, in Table 5.7 we show the performance of the above skin segmentation schemes in terms of the classification rate (CR), the
5.4. EXPERIMENTS

Correct detection rate (CDR) and false detection rate (FDR). Since we aim to compare different reflectance features, rather than different parameter settings of the classifier, we favour these metrics over the Receiver Operating Characteristic (ROC) curve. The correct detection rate is the percentage of skin pixels correctly classified. The false detection rate is the percentage of non-skin pixels incorrectly classified. The classification rate is the overall percentage of skin and non-skin pixels classified accurately. The table shows the segmentation accuracy measures over all the visible face images of all the subjects in the dataset illuminated by the frontal light source. The rows of the table correspond to the different features described earlier for skin classification. As expected, the reflectance recovered by our method achieves the highest skin recognition rates. This is consistent with the qualitative results above. Furthermore, the overall performance difference between the two reflectance variants based upon our method, i.e. when the estimated and the ground-truth light spectrum are used, is less than 4%. This demonstrates the robustness of our reflectance estimation method to errors in the input illuminant spectrum. As before, the reflectance obtained by normalising radiance by illuminant power performs better than the raw radiance spectra and their LDA components. Again, the radiance feature and its LDA components yield the most false positives and negatives.

Next, we examine the utility of the spectral reflectance recovered by our algorithm for unsupervised material clustering on multispectral images. This experiment can be regarded as an extension of the skin segmentation application. In addition, it is complementary to skin clustering algorithms using trichromatic features, which has been described elsewhere in the literature [131]. It also compares the clustering accuracy on the estimated reflectance to that on the measured (ground-truth) reflectance. Here, we perform a clustering algorithm based on a deterministic annealing approach [76] using the three reflectance features mentioned in the previous experiment. These features include the reflectance estimated given the ground-truth illuminant spectrum, the reflectance yielded given an estimated illuminant spectrum, and the reflectance obtained by normalising radiance spectra by the ground-truth illuminant spectrum. The clustering algorithm is initialised with a single cluster for all the materials. As the algorithm proceeds, new material clusters are introduced by splitting the existing ones. Thus,
<table>
<thead>
<tr>
<th>Feature</th>
<th>CDR(%)</th>
<th>FDR(%)</th>
<th>CR(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflectance recovered with ground-truth light</td>
<td>85.12 ± 13.36</td>
<td>5.10 ± 6.30</td>
<td>90.94 ± 6.12</td>
</tr>
<tr>
<td>Reflectance recovered with the estimated light</td>
<td>79.79 ± 22.58</td>
<td>8.32 ± 15.78</td>
<td>87.00 ± 13.63</td>
</tr>
<tr>
<td>Reflectance computed by illuminant normalisation</td>
<td>70.63 ± 16.95</td>
<td>5.50 ± 5.69</td>
<td>84.75 ± 8.04</td>
</tr>
<tr>
<td>Raw radiance</td>
<td>47.27 ± 20.58</td>
<td>12.48 ± 11.64</td>
<td>71.23 ± 9.17</td>
</tr>
<tr>
<td>Top 20 LDA components of raw radiance</td>
<td>44.83 ± 31.47</td>
<td>7.14 ± 11.42</td>
<td>73.62 ± 11.54</td>
</tr>
</tbody>
</table>

Table 5.7: Accuracy of several skin pixel recognition methods, each using a different reflectance-based feature as input for classification.

...the resulting number of clusters is data-dependent and does not need to be specified as input.

In Figure 5.12, we show the clustering maps of the images of several human subjects, with each row corresponding to a reflectance feature. The resulting material clusters are marked with different shades of gray. In fact, there is a high level of similarity between the clustering results yielded by the reflectance features estimated with the estimated illuminant spectrum and with the ground truth illuminant spectrum provided as input, as shown in rows 2 and 3. This demonstrates, again, the robustness of our reflectance estimation method to errors in the input illuminant spectrum. In these clustering maps, all the materials are well-separated from each other. Moreover, there are very few misclassified pixels within each cluster. On the faces, the skin pixels are clearly distinguished from the neighbouring regions. Notably, the background regions displaying printed faces, such as that in the third column, are correctly clustered as paper. This result demonstrates that the spectral variation of material reflectance is a better feature for classification than trichromatic colour. Note that using trichromatic imagery, it would have been virtually impossible to set apart materials with the same apparent color, such as real faces from printed ones. In the last row, we use the ground...
5.4. EXPERIMENTS

Figure 5.12: Material clusters, with each material marked by a different shade of gray. Top row: A band of the input images, shown at 670 nm. Second row: material clustering maps, obtained with the clustering feature being the reflectance estimated given the ground-truth illuminant. Third row: material clustering maps resulting from the use of the reflectance feature recovered with an estimated illuminant spectrum as input. Fourth row: material clustering maps, using the reflectance obtained by normalising the input radiance image by the ground-truth illuminant spectrum.

truth (measured) reflectance as feature for the clustering algorithm. Compared to our estimated reflectance, the measured reflectance produces noisier clustering maps, with a substantial number of pixels made of the same material assigned to various clusters. In other words, our reflectance recovery method improves the clustering performance by reducing measurement noise in the raw reflectance spectra.

5.4.3 Specularity Removal

Having estimated the illuminant spectra and surface reflectance, and as mentioned previously, we can employ the procedure in Section 5.1.4 to separate the diffuse from the
specular component in multi-spectral imagery. This is feasible in situations where the spectral reflectance varies slowly within a small spatial neighbourhood of the scene. Note that this assumption is a valid one for many real-world surfaces. Thus, each local neighbourhood can be considered as a smooth homogeneous surface. As a result, the diffuse component at a location $u$ in patch $P$ is estimated as $\hat{D}(u) = g(u)(L \cdot S_P)$. The specularity component is given by $k(u)L$.

Here, we perform highlight removal on the indoor human face image database
presented earlier. As mentioned in previous sections, we commence by estimating the illuminant spectra. We consider a neighbourhood of size $11 \times 11$ around each image pixel and assume that the neighbourhood has a common reflectance. As discussed in Section 5.1.4, a practical enforcement of smooth variation of the shading factor entails reprojecting the pixel radiance onto the subspace spanned by the illuminant spectrum and the diffuse radiance spectrum vectors. In this experiment, we employ a projection that minimises the $L_2$-norm of the distance between pixel radiance and this two-dimensional subspace.

Figure 5.13 shows the resulting shading and specularity coefficients estimated for a number of sample face images in the multispectral dataset. The top row shows the input images illuminated from a high oblique light direction. The second and third rows show the shading and specular coefficients, respectively, as yielded from our method. The last two rows show the results produced by the alternative highlight removal method in [134]. The alternative uses a probabilistic framework based upon the statistics arising from Lambertian reflectance in diffuse illumination. Note that the alternative is only applicable to a single-band greyscale image compliant with the Lambertian model. Thus, to compare the two methods, we apply the alternative to the brightest band in each of the input images. Also, we applied the alternative method to face images under the assumption in [134] regarding the collinearity of the light and viewer directions.

As observed from Figure 5.13, our method is successful at detecting and separating the specular from the diffuse component at typical highlight locations, such as noses, eyelids and foreheads. In addition, our method produces smooth matte diffuse images that capture the variation in the geometry of faces. Note that our method does not require the illumination direction a priori. On the other hand, Ragheb and Hancock’s method [134] assumes the collinearity of the illumination and viewing directions. Therefore, it cannot cope with the application setting shown for our method, where the light source is placed at a high-oblique position with respect to the camera. As expected, the alternative tends to miss important highlight points and generates false specular spikes.

Since our method makes the uniform albedo assumption on input surface patches,
CHAPTER 5. REFLECTION PARAMETER RECOVERY

Figure 5.14: A comparison between the highlight removal results of our method and that by Mallick et al. [108]. First column: the input image. Second and third columns: the diffuse and specular components resulting from our method. Fourth and fifth columns: the diffuse and specular components yielded by Mallick et al.’s method.

it tends to generate highlights in highly textured areas and along material boundaries. However, note that fine-scale relief texture of rough and highly textured areas may cause specularity that is only detected by elaborate measurements, as discussed in the work of Wang and Dana [175]. Since the background of the multispectral images in Figure 5.13 may be viewed as textured regions in the images, it may give rise to the highlights detected by our method, as shown in the third row of Figure 5.13.

Now we turn our attention to the application of our method to specularity detection and removal on trichromatic images. We compare the performance of our method with the highlight removal method by Mallick et al. [108], which employs a partial differential equation to erode the specularity at each pixel. In Figure 5.14, we show the highlight removal results for both our method and the alternative proposed in [108]. As before, our method performs better on smooth, homogeneous surfaces than on textured
areas, such as those in the second and third rows. On the smooth surfaces, most of the specular spikes are detected by our method, although there may be false specularities along the occlusion boundaries and material boundaries. On the other hand, the alternative produces smoother diffuse components, which are more intuitively correct. In addition, it detects more specular pixels on smooth surfaces.

As can be seen in Figure 5.14, our method is able to detect the expected specular spikes, as shown in the third column. Note that our method may introduce undesirable specularities along edges. This can be observed in the specularity maps in the third row. This is because patches at these locations are not of uniform albedo. Notably, the specularity map in the second row shows the underlying texture variation of the pear, which may be the cause of specularity being scattered over the fruit skin. In the second column, we show the diffuse component recovered by our method, where the diffuse colours at specular points are approximated from the neighbouring non-specular pixels.

5.5 Concluding Remarks

In this chapter, we have presented an optimisation approach to recovering reflection parameters from a single multispectral image as a solution to the dichromatic model. The recovery process is based upon optimising a cost function that imposes a smoothness constraint on dichromatic surfaces with uniform reflectance. The method serves several purposes, including scene illuminant estimation, reflectance-based recognition and clustering, and specularity removal. Departing from the dichromatic model, we have presented a cost function which can be optimised using a coordinate descent approach applied to automatically selected surface patches. We have also elaborated upon the use of a number of regularisation strategies so as to enforce smoothness constraints upon the cost function. Hence, the method is quite general in nature. Moreover, it is not only applicable to multispectral and hyperspectral imagery, but also extensible to trichromatic imagery in a straightforward manner. We have performed experiments on real-world imagery where, firstly, the illuminant spectra resulting from our method are shown to outperform those delivered by the alternative methods, especially with a small number of selected homogeneous surface patches. Secondly, we have shown
how the reflectance spectra recovered by our method can be used for skin recognition and material clustering purposes. Finally, we have shown how specularities can be removed from images so as to facilitate further computer vision tasks.
Chapter 6

Shape and Refractive Index Recovery from Spectro-Polarimetric Imagery

I can see it particularly clearly in the twilight when I stare at the zenith; the whole sky seems to be covered by a network, as it were, and everywhere I look I see this characteristic pattern. It is very pleasing to be able to determine the direction of polarization without an instrument in this way, and even obtain an estimate of its degree.

Marcel Gilles Jozef Minnaert
Belgian astronomer, 1893-1970
Light and Color in the Outdoors [117]

Chapter 5 focused on a method for recovering reflection parameters based on the dichromatic model. In this chapter, we focus on the recovery of both surface normals and refractive index from a single-view through the use of spectro-polarimetric imagery. As mentioned in Section 1.1.3, there is little work on the simultaneous recovery of normals and refractive index in this setting. The problem still remains unexplored despite a variety of techniques using polarisation for shape recovery and material recognition, as we have reviewed in Section 2.5. With polarimetric imaging, we can exploit polarisation as an additional source of information to recover the surface orientation in conjunction with the material refractive index.
Our recovery method confers a number of advantages over prior techniques. Firstly, it does not make prior assumptions on the illumination condition and the index of refraction. This is a major advantage over existing approaches which require either accurate measurements of shape or refractive index, or complicated instrument setups. Secondly, the setting is simple, requiring only a single view of the scene. This setting allows the application of the method to a variety of scenes captured in unconstrained environments. On the other hand, our method requires a specialised spectro-polarimetric camera for image acquisition.

Here, we focus on dielectric surfaces that undergo diffuse polarisation due to subsurface scattering and transmission from the object surface into the air. The diffuse polarisation of the reflection process is modelled by the Fresnel transmission theory. We present a method for estimating the azimuth angle of surface normals from the spectral variation of the phase of polarisation. Moreover, we estimate the zenith angle of surface normals and index of refraction simultaneously in a well-posed optimisation framework. We achieve this by introducing two additional constraints to the problem, \textit{i.e.} the surface integrability and the material dispersion equations. This yields an iterative solution which is computationally efficient due to the use of closed-form solutions for both the zenith angle and the refractive index in each iteration. To demonstrate the effectiveness of our approach, we show results on shape recovery and rendering for both real-world and synthetic imagery.

The remainder of the chapter is organised as follows. In Section 6.1, we commence by providing the background on the polarisation of electromagnetic waves. This consists of an introduction to the polarisation theory for dielectric surfaces and the properties of diffuse polarisation due to subsurface scattering. The section also introduces the Fresnel transmission ratio and the phase of polarisation concepts which appear throughout the chapter. In Section 6.2, we present our method for the recovery of surface normals and refractive index from a set of spectro-polarimetric images captured in a single view. The section begins with a formulation of the polarisation image radiance as a sinusoidal function. This permits the estimation of the azimuth angle of the surface normals from the polarisation phase across the wavelength domain. Subsequently, we formulate the recovery of the zenith angle and the refractive index as an optimisation
problem with an integrability constraint over the surface and a dispersion equation for the material refractive index. In Section 6.3, we show how this optimisation problem can be tackled in an elegant and efficient manner. In Section 6.4 we illustrate the utility of the method for the purposes of recovering the shape and refractive index from synthetic and real-world imagery. Finally, in Section 6.5 we draw conclusions and suggest directions for future investigation.

6.1 Preliminaries

6.1.1 Polarisation of Light

In this section, we provide a brief physics background on the theory of the polarisation of electromagnetic waves. Based on the wave theory, light propagation results in harmonic vibrations in its transmission medium. As an electromagnetic wave, light is associated with a magnetic field and an electric field that are mutually orthogonal and vibrate perpendicular to the propagation direction. Such electric and magnetic fields can be represented as field vectors parallel to a plane perpendicular to the propagation direction. According to [18, 73], the polarisation of light characterises the pattern of vibration of the electric and magnetic field vectors as light propagates across space and time.

Throughout the section we introduce a number of concepts relating polarisation, diffuse reflection and Fresnel transmission theory. For mnemonic purposes, in Figure 6.1, we provide a list of notation used in the chapter. In addition, we consider a right-handed reference coordinate system where the origin is located at the viewpoint, the positive $z$–axis coincides with the propagation direction of the reflected light as observed in the line of sight. The positive $x$–axis points towards the right-hand side of the field of view. Within this right-handed reference system, surface normals can be specified by their azimuth (tilt) and zenith (slant) angles with respect to the positive $x$ and $z$ axes.

The electric field of an electromagnetic wave can be viewed as a superposition of two harmonic plane waves with orthogonal planes of vibration [18]. Figure 6.2(a) depicts a circularly polarised illumination flux with its electric field vector $\vec{E}$ rotating in...
a helicoidal pattern about the propagation direction of the reflected light as observed
from the viewpoint. In the figure, the propagation direction is along the $z$-axis. As
shown in Figure 6.2(a), the electric field vector can be decomposed into two sinusoidal
components $\vec{E}_x$ and $\vec{E}_y$ vibrating in the $x-z$ and $y-z$ planes. In the figure, the two
components have equal magnitudes and $\vec{E}_y$ is shifted by $1/4$-wavelength ($\pi/2$) with re-
spect to $\vec{E}_x$. As a result, the front of the electric field follows a trajectory with a circular
projection in the $x-y$ plane as it varies with respect to time and space. Figure 6.2(b)
illustrates the circular rotation of the electric field, while its $x$ and $y$ components vibrate

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_0(u, \lambda)$</td>
<td>Transmitted irradiance at a pixel $u$ and a wavelength $\lambda$ through a polariser oriented at an angle $\theta$ from a reference axis.</td>
</tr>
<tr>
<td>$\phi(u, \lambda)$</td>
<td>Phase of polarisation.</td>
</tr>
<tr>
<td>$I_{max}(u, \lambda)$</td>
<td>Maximum and minimum transmitted radiance on the Transmitted Radiance Sinusoid (TRS).</td>
</tr>
<tr>
<td>$I_{min}(u, \lambda)$</td>
<td>Surface normal at pixel $u$.</td>
</tr>
<tr>
<td>$\alpha(u)$ and $\theta(u)$</td>
<td>Azimuth and zenith angle at pixel $u$.</td>
</tr>
<tr>
<td>$\eta(u, \lambda)$</td>
<td>Index of refraction at pixel $u$ and wavelength $\lambda$.</td>
</tr>
<tr>
<td>$\vec{E}$</td>
<td>Electric field vector with components $\vec{E}_x$ and $\vec{E}_y$, wavenumber $k$ and phase difference $\varepsilon$.</td>
</tr>
<tr>
<td>$\vec{E}<em>{un}$ and $\vec{E}</em>{pol}$</td>
<td>Unpolarised and completely polarised components.</td>
</tr>
<tr>
<td>$I_{\parallel}$ and $I_{\perp}$</td>
<td>Parallel and perpendicular components of the incident irradiance.</td>
</tr>
<tr>
<td>$I_{T \parallel}$ and $I_{T \perp}$</td>
<td>Parallel and perpendicular components of the transmitted radiance.</td>
</tr>
<tr>
<td>$F_{\parallel}$ and $F_{\perp}$</td>
<td>Parallel and perpendicular Fresnel reflection coefficients.</td>
</tr>
<tr>
<td>$R(u, \lambda)$</td>
<td>Fresnel transmission ratio.</td>
</tr>
<tr>
<td>$C_m(u)$</td>
<td>Dispersion coefficients of Sellmeier’s equations.</td>
</tr>
<tr>
<td>$B_m(u)$ and $D_m(u)$</td>
<td>Dispersion coefficients of Sellmeier’s equations.</td>
</tr>
</tbody>
</table>

Figure 6.1: Notation used throughout Chapter 6.
In general, the polarisation of a light wave can range from being unpolarised, to being completely polarised. When the phase difference between its orthogonal components varies in a random manner with respect to time, the resultant electric field is incoherent or unpolarised, producing an electric field vector rotating isotropically and randomly in the plane perpendicular to the light propagation direction. On the other hand, completely polarised light is the result of the coherent superposition of two orthogonal harmonic components with a fixed-phase difference with respect to time. In this case, the light wave is elliptically polarised since the front of the electric field traces out an ellipse whose major and minor axes are determined by the phase difference of the harmonic components.

A general polarisation state can be represented as a mixture of an unpolarised state with a completely polarised state. Consider a light wave with wavelength $\lambda$ propagating...
ing in the $z$ direction. Let the unpolarised and completely polarised components be $\vec{E}_{un}$ and $\vec{E}_{pol}$, respectively. The fronts of $\vec{E}_{un}$ and $\vec{E}_{pol}$ can be projected onto the $x-y$ plane so as to trace a circle and an ellipse as shown in Figure 6.3(a). The total polarisation state $\vec{E}_{pol}$ can be further decomposed into the $x$ and $y$ components $\vec{E}_x$ and $\vec{E}_y$. The latter polarisation components are sinusoidal in nature, oscillating in the $x-z$ and $y-z$ planes with magnitudes $E_{0x}$ and $E_{0y}$, respectively. Mathematically, the vectorial form of $\vec{E}_x$ and $\vec{E}_y$ at a point with coordinates $[0,0,z_0]^T$ at time $t$ is expressed as

$$\vec{E}_x = E_{0x} \cos(kz_0 - \omega t)$$

(6.1)

$$\vec{E}_y = E_{0y} \cos(kz_0 - \omega t + \varepsilon)$$

(6.2)

where the wavenumber $k$ is given by $k = \frac{2\pi}{\lambda}$, $\omega$ is the frequency of the wave, and $\varepsilon$ is the phase difference between the two components. The vectors $\vec{E}_x$ and $\vec{E}_y$ along the $x$ and $y$ axes have magnitudes $E_{0x}$ and $E_{0y}$, respectively.

Thus, the total electric-field vector $\vec{E}$ can be obtained as the linear combination of the unpolarised and the completely polarised field vectors [73], which is given by

$$\vec{E} = \vec{E}_{un} + [E_{0x} \cos \psi, E_{0y} \cos(\psi + \varepsilon)]^T$$

(6.3)

where $[\cdot]^T$ denotes the transpose of a vector and $\psi = kz_0 - \omega t$.

---

Figure 6.3: (a) The superposition of a unpolarised wave $\vec{E}_{un}$ with an elliptically polarised one $\vec{E}_{pol}$. The elliptically polarised wave can be further decomposed into two component vectors $\vec{E}_x$ and $\vec{E}_y$, whose lengths oscillate sinusoidally with magnitudes $E_{0x}$ and $E_{0y}$. The shape and orientation of the major and minor axes of the ellipse are determined by the ratio $E_{0x}/E_{0y}$ and the phase difference between $\vec{E}_x$ and $\vec{E}_y$. (b) The transmitted wave through a polariser with the transmission axis direction $\vec{P}$ oriented at an angle $\theta$ with respect to the $x$-axis.
6.1. PRELIMINARIES

Equation 6.3 provides a general analytical expression of the electric field with respect to time and space. Now we consider the decomposition of the above light wave into polarisation components oriented at different directions in the plane perpendicular to the propagation direction (i.e. the \(x-y\) plane). The power of these polarisation components can be measured by placing a linear polariser in front of the camera, as shown in Figure 6.3(b). In effect, the polariser only transmits the polarisation component of the incoming light that vibrates along its transmission axis. Let the transmission axis \(\vec{P}\) lie on the \(x-y\) plane at an angle \(\theta\) with respect to the \(x\)-axis, i.e. \(\vec{P} = [\cos \theta, \sin \theta]^T\).

For the randomly polarised component \(E^*_{\text{un}}\), the irradiance transmitted along any direction \(\vec{P}\) is half of its original energy \(I_{\text{un}}\), which is given by the time-averaged square of its amplitude \(I_{\text{un}} = \langle E^*_{\text{un}} \rangle_t\), where \(\langle . \rangle_t\) denotes average over time. Similarly, the transmitted irradiance for the completely polarised component \(E^*_{\text{pol}}\) is the time-averaged energy along the orientation \(\vec{P}\) per unit area. Therefore, the total transmitted irradiance at the pixel \(u\) and wavelength \(\lambda\) of the incoming light through the polariser is related to the electric field as

\[
I_\theta(u, \lambda) = \langle |\vec{E} \cdot \vec{P}|^2 \rangle_t
\]

\[
= \frac{1}{2} I_{\text{un}} + \left( \langle E_{0x} \cos \psi \cos \theta + E_{0y} \cos(\psi + \varepsilon) \sin \theta \rangle \right)^2_t
\]

\[
= \frac{1}{2} I_{\text{un}} + E_{0x}^2 \cos^2 \theta \langle \cos^2 \psi \rangle_t + E_{0y}^2 \sin^2 \theta \langle \cos^2(\psi + \varepsilon) \rangle_t
\]

\[
+ E_{0x} E_{0y} \sin \theta \cos \theta \langle \langle \cos(2\psi + \varepsilon) \rangle_t + \cos \varepsilon \rangle
\]

\[
= \frac{1}{2} \left( I_{\text{un}} + E_{0x}^2 \cos^2 \theta + E_{0y}^2 \sin^2 \theta + E_{0x} E_{0y} \cos \varepsilon \sin 2\theta \right)
\]

(6.4)

where, as before, \(\langle . \rangle_t\) denotes the average value over time.

In Equation 6.4, we use the equalities \(\langle \cos^2 \psi \rangle_t = \langle \cos^2(\varepsilon + \omega t) \rangle_t = \frac{1}{2}\), \(\langle \cos^2(\psi + \varepsilon) \rangle_t = \langle \cos^2(\varepsilon + \omega t + \varepsilon) \rangle_t = \frac{1}{2}\) and \(\langle \cos(2\psi + \varepsilon) \rangle_t = 0\). The right-hand side of this Equation can be rewritten as a sinusoid function

\[
I_\theta(u, \lambda) = \frac{1}{2} \left( I_{\text{un}} + \frac{1}{2} (E_{0x}^2 + E_{0y}^2) + \frac{1}{2} (E_{0x}^2 - E_{0y}^2) \cos 2\theta + E_{0x} E_{0y} \cos \varepsilon \sin 2\theta \right)
\]

\[
= \frac{1}{2} \left( I_{\text{un}} + \frac{1}{2} (E_{0x}^2 + E_{0y}^2) + I_s \cos 2(\theta - \phi) \right)
\]

(6.5)

where \(I_s = \left( \frac{1}{4} (E_{0x}^2 + E_{0y}^2)^2 - E_{0x}^2 E_{0y}^2 \sin^2 \varepsilon \right)^{1/2}\) and \(\phi = \frac{1}{2} \arctan \left( \frac{2E_{0x} E_{0y} \cos \varepsilon}{E_{0x}^2 - E_{0y}^2} \right)\).
Thus, from Equation 6.5, we can conclude that, in general, the transmitted radiance of a polarisation component is a sinusoid function of the transmission angle \( \vartheta \) with an amplitude \( \frac{1}{2} I \), and a phase of polarisation \( \phi \).

## 6.1.2 Polarisation Upon Diffuse Reflection

In Section 6.1.1, we arrived at a formulation of the transmitted irradiance of light through a linear polariser. This formulation is concerned with the polarisation of light in general, whether it is caused by reflection upon a surface, refraction through the boundary between media, or scattering through the atmosphere, haze or fog. In this section, we focus on the polarisation caused by diffuse reflection upon a dielectric surface.

In several cases, diffuse reflection is the result of multiple scattering from micro-facets [128]. On the other hand, for translucent or multi-layered materials, diffuse reflection is attributed to incident light penetrating the surface, scattering within the dielectric body and refracting back into the transmission medium. Here, we may assume that after penetration, light is largely depolarised by the internal scattering process within the dielectric body. Later, the refraction of scattered light from the dielectric body through the material-air boundary induces polarisation. Therefore, the following theory of polarisation, which is subject to Snell’s law of refraction and Fresnel reflection theory, applies to translucent or multi-layered materials with subsurface scattering.

Figure 6.4(a) shows a diagram of the diffuse reflection process described above. Here, refraction is the result of the change of velocity as the internally scattered ray \( \vec{k}_i \) travels from the material body with refractive index \( \eta \) to the air whose index is unity. Assume that this ray is incident on the surface boundary at an angle \( \theta_i \). After emerging from the surface, the transmitted ray \( \vec{k}_T \) is bent at an emittance (reflection) angle \( \theta \) with respect to the surface normal \( \vec{N} \). The propagation direction of the transmitted ray and the surface normal both lie in the plane of reflection, as shown in Figure 6.4(a). Note that \( \theta \) is also the zenith angle of the surface normal with respect to the viewing direction.

Note that the electric field vectors associated with the rays \( \vec{k}_i \) and \( \vec{k}_T \) are always perpendicular to their propagation directions. Furthermore, these electric field vectors
can be decomposed into two orthogonal harmonic components, i.e. one in the plane of reflection while the other one perpendicular to it. In Figure 6.4(a), we have denoted $\vec{E}_i^\parallel$ and $\vec{E}_i^\perp$ as the parallel and perpendicular components of the electric field associated with $\vec{k}_i$. Similarly, $\vec{E}_T^\parallel$ and $\vec{E}_T^\perp$ are the parallel and perpendicular components of the transmitted waves travelling in the direction $\vec{k}_T$.

Furthermore, Figure 6.4(b) presents a cross-section of the emitted electric field as

Figure 6.4: (a) Polarisation upon diffuse reflection from a dielectric surface with a normal vector $\vec{N}$ and refractive index $\eta$. The incident light waves penetrate the surface, scatter inside the dielectric body and finally refract through the material-air boundary. The electric fields of the waves incident on and transmitted through the material-air interface are represented by vectors rotating in a plane perpendicular to the propagation direction. These electric field vectors, in turn, can be decomposed into two orthogonal components which are parallel and perpendicular to the plane of emittance, which contains both the surface normal and the emittance direction. The vectors $\vec{E}_i^\parallel$ and $\vec{E}_i^\perp$ are the components of the internal incident field on the material-air surface boundary before refraction. The vectors $\vec{E}_T^\parallel$ and $\vec{E}_T^\perp$ are the components of the wave emerging from the surface after refraction. (b) The emitted electric field, where the surface normal $\vec{N}$ is aligned with the $y$-axis. $\vec{E}_T^\parallel$ vibrates along the $y$-axis because it is in the plane of reflection. The perpendicular component $\vec{E}_T^\perp$ lies in the $x$-axis and $\vec{P}$ is the orientation of the polariser’s transmission axis.
viewed from the camera’s position, where the surface normal $\vec{N}$ is aligned with the $y$ axis. In this view, $E_T^\parallel$ vibrates along the $y$ axis because it is coplanar with $\vec{N}$. The perpendicular component $E_T^\perp$ lies in the $x$-axis as it is orthogonal to $E_T^\parallel$. In the Figure, $\hat{P}$ is the orientation of the polariser’s transmission axis.

6.1.3 Fresnel Transmission Ratio and Phase of Polarisation

We now peruse two important concepts arising from the polarisation of light upon diffuse reflection, the Fresnel transmission ratio and the phase of polarisation. To commence, let us denote the radiance of the parallel and perpendicular components emitted at the material surface as $I_T^\parallel$ and $I_T^\perp$ respectively. With these notations, the transmission ratios of these components through the material-air interface are related to the Fresnel reflection coefficients as follows

$$\frac{I_T^\perp}{I_{i\perp}} = 1 - F_\perp$$  \hspace{2cm} (6.6)

$$\frac{I_T^\parallel}{I_{i\parallel}} = 1 - F_\parallel$$  \hspace{2cm} (6.7)

where $I_{i\parallel}$ and $I_{i\perp}$ are the parallel and perpendicular components of the incident irradiance, respectively and $F_\perp$ and $F_\parallel$ are the Fresnel reflection coefficients for the parallel and perpendicular polarisation components.

Making use of the assumption that the light scattered inside the dielectric body is unpolarised, we can conclude that its electric field vector oscillates in all the directions perpendicular to the propagation direction with equal energy and probability. This means that the magnitudes $E_{i\parallel}$ and $E_{i\perp}$ of the perpendicular and parallel components of the internally scattered ray $\mathbf{k}_i$ are equal, i.e. $E_{i\parallel} = E_{i\perp}$. This, in turns, leads to $I_{i\parallel} = I_{i\perp}$ as the radiance is proportional to the square of the electric field amplitude. With this in mind, we take the ratio of the left and right-hand sides of Equation 6.6 and 6.7 to obtain

$$\frac{I_{T\perp}}{I_{T\parallel}} = \frac{1 - F_\perp}{1 - F_\parallel}$$  \hspace{2cm} (6.8)

We refer to the right hand side of Equation 6.8 as the Fresnel transmission ratio, which is a function of the normal zenith angle and the refractive index.

We can further explore the relationship between the phase of polarisation $\phi$ resulting from subsurface scattering and the surface geometry. To do this, we consider the
6.2 SHAPE AND REFRACTIVE INDEX RECOVERY

Departing from the theory on polarisation developed in Section 6.1, in this section we present a method for the recovery of the surface orientation and index of refraction from polarisation imagery. The technique discussed here is applicable to convex and continuously twice-differentiable surfaces with material refractive index following the
dispersion equation. In Section 6.2.1, we commence by extracting the phase angle and
the maximal and minimal radiance of the Transmitted Radiance Sinusoid (TRS) from
the polarimetric imagery. In Section 6.2.2, we present a method for disambiguating
and estimating the azimuth angle of the surface normals from the wavelength-indexed
spectrum of phase angles. In addition, we compute the Fresnel transmission ratio from
the maximal and minimal radiance, from which to estimate the zenith angle of the
surface normal and the refractive index. In Section 6.2.3, we formulate the estima-
tion as an optimisation problem that takes into account the data error for the Fresnel
transmission ratio, the surface integrability and the dispersion of material refractive
index across the spectrum. Having formulated the objective function, we present the
optimisation process entirely in Section 6.3.

The process of shape and refractive index recovery can be summarised as follows.

1. Decomposition of the spectral polarimetric imagery into polarisation compo-
nents, including the phase, maximal and minimal radiance of the Transmitted
Radiance Sinusoid (TRS), as described in Section 6.2.1.

2. Rough estimation of the azimuth angle from the phase angle for each pixel and
wavelength, allowing a 180-degree ambiguity. The relationship between these
two angles is described in Section 6.2.2.

3. Joint estimation of the zenith angle and refractive index from the Fresnel trans-
mision ratio, as presented in Section 6.2.3.

4. Disambiguation between the two candidates of the azimuth angle using shading
information indicated by the estimated zenith angle, as presented in Section 6.2.2

5. Re-estimation of the azimuth angle as presented in Section 6.2.2.

6. Integration of the surface normal field to reconstruct the surface depth.

It is noted that the 180-degree ambiguity between the candidates of the azimuth
angle resulting from the Step 2 does not affect the joint estimation of the zenith angle
and refractive index in Step 3. This is because the objective functions involved in this
estimation are invariant to the 180-degree shift in the azimuth angle. As we shall show
in Section 6.2.3, the relevant objective function consists of the square of the cosine and sine of the azimuth angle. Once the zenith angle has been obtained, we perform disambiguation of the azimuth angle based on the shading information suggested by the estimated zenith angle.

### 6.2.1 Decomposing Polarisation Images

As illustrated in Figure 6.4(a), the polarisation of light reflected from a surface can be measured by mounting a linear polariser in front of the camera’s optics. By rotating the polariser, one can capture polarisation components that are oriented at different angles on the plane orthogonal to the light propagation direction. Here, we note that the analysis in Section 6.1.1 can be applied directly to this polarisation imaging system. As a result, the radiance at each point in the scene varies sinusoidally with respect to the angle corresponding to the rotation of the polariser. This variation is depicted in Figure 6.5, where the radiance $I$ fluctuates along a sinusoid curve bounded between $I_{\text{min}}$ and $I_{\text{max}}$. This sinusoid is termed the Transmitted Radiance Sinusoid (TRS).

As established in Section 6.1.1, we maintain the notation $\theta$ for the polariser orientation angle and $\phi$ for the phase angle of the sinusoid. Since we aim at employing polarimetric imaging spectroscopy data, we employ $\lambda$ to denote the wavelength index for the spectra at each of the image pixels $u$ in the image. Moreover, assume that the po-

![Figure 6.5: The transmitted radiance through a linear polariser varies sinusoidally with respect to the polariser orientation angle.](image-url)
larimetric spectral imagery comprises \( N \) multispectral image cubes \( I_{\theta_1}, I_{\theta_2}, \ldots, I_{\theta_N} \). Each of these image cubes is captured at \( N \) discrete values of the polariser orientation angle \( \theta \), which we denote as \( \theta_j, j = \{1, 2, \ldots, N\} \). The variation of the image intensity at the pixel \( u \), wavelength \( \lambda \) and polariser orientation \( \theta_j \) is then given by

\[
I_{\theta_j}(u,\lambda) = \frac{I_{\max} + I_{\min}}{2} + \frac{I_{\max} - I_{\min}}{2} \cos(2\theta_j - 2\phi)
\] (6.10)

where \( I_{\min} \) and \( I_{\max} \) are the minimum and maximum radiance, respectively, at the pixel \( u \) and \( \phi \) is the corresponding polarisation angle. Note that here, for the sake of brevity, we have dropped the wavelength, pixel and polarisation angle variables from the notations for \( I_{\min} \) and \( I_{\max} \).

The recovery of \( I_{\min} \), \( I_{\max} \) and \( \phi \) in Equation 6.10 may be effected in a number of ways. Note that, by capturing the same scene with three or more polariser orientations, one can fit a sinusoidal curve through the intensity-polariser angle pairs using a numerical nonlinear least square fitting algorithm. This method is, however, not efficient since the optimisation has to be performed per pixel.

Alternatively, these parameters can be obtained making use of the method in [182], where three images are acquired so as to compute the phase, intensity and degree of polarisation. However, the method in [182] is susceptible to noise corruption since it employs only three images. Here, we employ an alternative that yields a stable estimation of the intensity, phase and degree of polarization by solving an over-determined linear system of equations. The method is akin to that described by Nayar et al. in [124]. We commence by rewriting Equation 6.10 for each pixel-site \( u \) and wavelength \( \lambda \) in the following vector form

\[
I_{\theta_j}(u,\lambda) = \begin{bmatrix} 1 \\ \cos(2\theta_j) \\ \sin(2\theta_j) \end{bmatrix}^T \begin{bmatrix} \frac{I_{\max} + I_{\min}}{2} \\ \frac{I_{\max} - I_{\min}}{2} \cos(2\phi) \\ \frac{I_{\max} - I_{\min}}{2} \sin(2\phi) \end{bmatrix} = a_j^T x \] (6.11)

where \( a_j = \begin{bmatrix} 1 \cos(2\theta_j) \sin(2\theta_j) \end{bmatrix}^T \) and \( x = \begin{bmatrix} \frac{I_{\max} + I_{\min}}{2} \\ \frac{I_{\max} - I_{\min}}{2} \cos(2\phi) \\ \frac{I_{\max} - I_{\min}}{2} \sin(2\phi) \end{bmatrix} \).
6.2. SHAPE AND REFRACTIVE INDEX RECOVERY

After collecting $N \geq 3$ measurements at three or more polariser orientations, we arrive at the following over-determined linear system

$$I = Ax$$  \hspace{1cm} (6.12)$$

where $I = \begin{bmatrix} I_{\theta_1}(u, \lambda) \\ I_{\theta_2}(u, \lambda) \\ \vdots \\ I_{\theta_N}(u, \lambda) \end{bmatrix}$ and $A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_N^T \end{bmatrix}$

Equation 6.12 is well-constrained when the number of polariser angles is $N \geq 3$. Moreover, the coefficient matrix $A$ depends solely on the polariser angles and, therefore, allows for an efficient solution of Equation 6.12 over all the image pixels and wavelengths simultaneously.

Having obtained the solution $x^* = [x_1, x_2, x_3]^T$, the maximal and minimal radiance on the sinusoid and the phase of polarisation at each pixel $u$ and wavelength $\lambda$ can be recovered in a straightforward manner making use of the following relations

$$I_{\text{max}} = x_1 + \sqrt{x_2^2 + x_3^2}$$
$$I_{\text{min}} = x_1 - \sqrt{x_2^2 + x_3^2}$$
$$\phi = \frac{1}{2} \arctan \frac{x_3}{x_2}$$

6.2.2 Azimuth Angle and Phase of Diffuse Polarisation

Now, we turn our attention to the disambiguation and the estimation of the azimuth angle based on the phase angle $\phi$. Recall that, in Section 6.1.3, we observed that the maximum transmitted radiance, i.e. $I_{\text{max}}$ occurs when the polariser’s angle is parallel to the plane of reflection. Moreover, note that the TRS in Figure 6.5 reaches its maximum at the phase angle. Consequently, the phase angle $\phi$ must either coincide with the azimuth (tilt) angle $\alpha$ of the surface normal or differ from it by $\pi$ radians. In other words, at each pixel $u$, either of the two cases given by $\alpha(u) = \phi$ and $\alpha(u) = \phi \pm \pi$ must hold. This reduces the problem of recovering the azimuth of the surface normal to that of disambiguating the two possible solutions for $\alpha(u)$. 
CHAPTER 6. SHAPE AND REFRACTIVE INDEX RECOVERY

Disambiguation of the Azimuth Angle

To this end, we assume that the surface under study is convex, i.e. the surface normals point in the opposite direction to the gradient of surface shading. This assumption allows us to select the candidate azimuth angle with the closer orientation to the negative gradient direction. The surface shading at a pixel $u$ is given by $\cos \theta(u)$, where $\theta(u)$ is the zenith angle of the surface normal. Suppose that the zenith angle has been obtained through the process described in Section 6.3. By sorting the zenith angles available, we are able to detect both the boundary and singular points on the surface and use these to disambiguate the azimuth angles.

The disambiguation process commences by first rotating the azimuth angles by $180^\circ$ where necessary to satisfy the convexity requirement. In addition, for convex objects, the surface normals at the occluding boundary always point outward from the body of the object. We make use of this observation to initialise the azimuth angles at the occlusion boundary. Also, since diffuse polarisation occurs most strongly at the occlusion boundary, the disambiguation of azimuth angle based on polarisation performs most reliably at these positions. On the other hand, weak polarisation is often observed at singular points, i.e. where the surface normal is perpendicular to the image plane. Therefore, abrupt changes in azimuth angle are permitted near these locations. As a result, we start smoothing the azimuth angle at the boundary pixels and propagate the operation toward the singular points of the surface. The smoothing operation proceeds iteratively, aiming to preserve the directional similarity of the azimuth angle within a neighbourhood.

Note that the disambiguation method above only copes with convex surfaces. Although the reconstruction of shapes consisting of a mixture of convex and concave parts from normal fields is an important problem, it is outside the scope of this paper. Rather, here we focus on the recovery of the surface normals themselves from polarisation information. Thus, we limit our attention to only convex surfaces. It is worth noting in passing that to fully recover the height of a complex surface with both convex and concave parts, the method proposed by Zhu and Shi [191] could be used. The algorithm in [191] flips surface patches and adjusts their heights so as to satisfy a global integrability constraint. It resolves this global disambiguation problem by computing
the optimal height of singular points. To compute the height at the singular points, the authors construct a graph with singular points as vertices, on which the global integrability constraint can be stated as a Max-cut problem. The solution to this problem can then be found by making use of semidefinite programming. Having obtained the height at singular points, the height of the whole surface is propagated from these points via the application of fast marching and patch stitching.

Estimation of the Azimuth Angle

In general, the disambiguation process above provides an estimation of the azimuth angle $\phi^*(u, \lambda)$ at each pixel $u$ and wavelength $\lambda$ of the spectral imagery. However, the estimated phase angle for each wavelength may vary widely due to weak polarisation and noise corruption. Hence, we make use of the weighted mean of the spectral phase of polarisation across the spectrum as an alternative to the phase angle at any given wavelength. The weights are chosen to reflect the fitting error of the TRS curve to the intensity-polariser angle pairs at each wavelength. Here, the fitting error $\epsilon(u, \lambda)$ is quantified using the $\ell^2$-norm of the residual $\epsilon(u, \lambda)$ for the Equation 6.12, i.e. $\epsilon(u, \lambda) = \|I - Ax\|_2$, where $I, A$ and $x$ have been computed per pixel and wavelength as per the previous section. The weight $w(u, \lambda)$ associated with the phase of polarisation at each wavelength is defined via a kernel weighting function. Here we employ the Epanechnikov kernel, which is a popular choice in the parameter estimation literature. The kernel is given by

$$w(u, \lambda) = \begin{cases} 1 - \frac{\epsilon^2(u, \lambda)}{h} & \text{if } \frac{\epsilon^2(u, \lambda)}{h} < 1 \\ 0 & \text{otherwise} \end{cases}$$

(6.13)

where $h$ is a bandwidth parameter.

Since the azimuth angle is a directional quantity, instead of averaging the disambiguated azimuth angles $\phi^*(u, \lambda)$ directly, we estimate the mean of the sines and cosines of these angles for each pixel-site as follows

$$\langle \sin(\phi^*(u)) \rangle_\lambda = \frac{\sum_\lambda \sin(\phi^*(u, \lambda))w(u, \lambda)}{\sum_\lambda w(u, \lambda)}$$

$$\langle \cos(\phi^*(u)) \rangle_\lambda = \frac{\sum_\lambda \cos(\phi^*(u, \lambda))w(u, \lambda)}{\sum_\lambda w(u, \lambda)}$$
where $\langle \cdot \rangle_\lambda$ denotes the mean value across wavelengths. Thus, the estimated azimuth angle at pixel $u$ becomes

$$
\alpha^*(u) = \begin{cases} 
\arctan \left( \frac{\langle \sin(\phi^*(u)) \rangle_\lambda}{\langle \cos(\phi^*(u)) \rangle_\lambda} \right) & \text{if } \langle \cos(\phi^*(u)) \rangle_\lambda > 0 \\
\arctan \left( \frac{\langle \sin(\phi^*(u)) \rangle_\lambda}{\langle \cos(\phi^*(u)) \rangle_\lambda} \right) + \pi & \text{if } \langle \cos(\phi^*(u)) \rangle_\lambda < 0 \\
\frac{\pi}{2} & \text{if } \langle \cos(\phi^*(u)) \rangle_\lambda = 0
\end{cases}
$$

(6.14)

### 6.2.3 Zenith Angle and Refractive Index

To fully determine the surface shape, one needs to estimate the zenith angle of the surface normals with respect to the viewing direction. Following the previous section, where we showed how the azimuth angle of surface normals can be estimated from the phase of polarisation, we now provide an account of the joint estimation of the zenith angle of surface normals and material refractive index from the polarimetric spectral imagery under consideration.

Following Section 6.1.3, we can conclude that the TRS curve reaches its minimal value when the transmission axis is perpendicular to the plane of reflection. This is, in fact, the complement of the observation made in the Section 6.1.3 regarding $I_{\max}$. As a result, we have the equalities $I_{\max} = I_{T\parallel}$ and $I_{\min} = I_{T\perp}$ and obtain the ratio of the recovered minimal and maximal radiance on the TRS curve as

$$
\frac{I_{\min}}{I_{\max}} = \frac{I_{T\perp}}{I_{T\parallel}} = \frac{1 - F_{\perp}}{1 - F_{\parallel}}
$$

(6.15)

Following the expressions of the Fresnel reflection coefficients in [181] and Snell’s law of refraction, we can relate the ratio on the right-hand side of Equation 6.15 to the zenith angle and refractive index through the equation

$$
\frac{I_{\min}}{I_{\max}} = \frac{\cos \theta(u) \sqrt{\eta^2(u, \lambda) - \sin^2 \theta(u)} + \sin^2 \theta(u)}{\eta(u, \lambda)}
$$

(6.16)

The right-hand side of Equation 6.16 is the Fresnel transmission ratio. To further simplify computation, we take its square root to obtain the following function with respect to the zenith angle and the refractive index

$$
R(u, \lambda) \triangleq \frac{\cos \theta(u) \sqrt{\eta^2(u, \lambda) - \sin^2 \theta(u)} + \sin^2 \theta(u)}{\eta(u, \lambda)}
$$

(6.17)
6.2. SHAPE AND REFRACTIVE INDEX RECOVERY

Furthermore, from Equation 6.16 we note that the above function can be related to the radiance $I_{\text{max}}$ and $I_{\text{min}}$ recovered in Section 6.2.1 as follows

$$R(u, \lambda) = \sqrt{\frac{I_{\text{min}}}{I_{\text{max}}}} \quad (6.18)$$

It is worth noting in passing that prior literature has mainly focused on the recovery of the zenith angle of the surface normals and the index of refraction from the degree of polarisation [4, 5, 8, 118, 119, 163]. Indeed, all of these methods depart from the Fresnel reflection theory to arrive at an equation similar to Equation 6.16. However, the main limitation to their practical application resides on their reliance upon either known index of refraction [4, 5, 8, 118, 119], imagery captured from multiple viewpoints [118] or under several known light source directions [163]. The need for multiple measurements and instrumental setups makes them impractical for shape and material analysis on real-world data.

Moreover, the estimation of the zenith angle and index of refraction cannot be performed without further constraints. This is because the Fresnel transmission ratio only provides a single equation per wavelength relating the zenith angle and index of refraction, i.e. the number of variables to be estimated is one more than the number of equations. Therefore, the problem is ill-posed in general.

To deal with these limitations, we propose two additional constraints in order to recover the zenith angle and the refractive index in a well-formed manner. While the integrability constraint enforces spatial consistency between neighbouring surface locations, the dispersion constraint aims to resolve the ill-posedness of the joint estimation of the zenith angle and refractive index.

**Integrability Constraint**

We commence by formulating the integrability constraint with respect to the zenith and azimuth angles of the surface normal. Note that the surface under study can be represented by a two-dimensional twice-differentiable function with a continuous second derivative [54]. As a result, its cross derivatives are the same irrespective of the order of the differentiated variable.

For the reference coordinate system presented before, let the surface height function at the pixel $u$ be $Z(u)$. By definition, the normalised surface normal at $u$ is given
CHAPTER 6. SHAPE AND REFRACTIVE INDEX RECOVERY

by

\[ \vec{N} = \frac{1}{\sqrt{Z_x^2 + Z_y^2 + 1}} [-Z_x, -Z_y, 1]^T \]  \hspace{1cm} (6.19)

where \( Z_x \) and \( Z_y \) are the surface gradients on the \( x \) and \( y \) axes on our coordinate system.

The normalised surface normal direction can also be represented with respect to the azimuth and zenith angles as follows

\[ \vec{N} = \begin{bmatrix} \cos \alpha(u) \sin \theta(u) \\ \sin \alpha(u) \sin \theta(u) \\ \cos \theta(u) \end{bmatrix} \]  \hspace{1cm} (6.20)

From Equations 6.19 and 6.20, we have

\[ Z_x = -\cos \alpha(u) \tan \theta(u) \]  \hspace{1cm} (6.21)
\[ Z_y = -\sin \alpha(u) \tan \theta(u) \]  \hspace{1cm} (6.22)

Recall that, in Section 6.2.2, we obtained the normal azimuth angles \( \alpha(u) \) and treat them as constants in Equations 6.21 and 6.22. As a consequence, the cross derivatives can be rewritten as \( Z_{xy} = -\cos \alpha(u) \frac{\partial \tan \theta(u)}{\partial y} \) and \( Z_{yx} = -\sin \alpha(u) \frac{\partial \tan \theta(u)}{\partial x} \). Since the integrability constraint states that \( Z_{xy} = Z_{yx} \), it can be parameterised with respect to the azimuth and zenith angles as

\[ \cos \alpha(u) \frac{\partial \tan \theta(u)}{\partial y} = \sin \alpha(u) \frac{\partial \tan \theta(u)}{\partial x} \]  \hspace{1cm} (6.23)

According to the chain rule, \( \partial \tan \theta(u) = \frac{\partial \theta(u)}{\cos^2 \theta(u)} \). Therefore, the integrability constraint in Equation 6.23 can be rewritten as

\[ \cos \alpha(u) \theta_y(u) = \sin \alpha(u) \theta_x(u) \]  \hspace{1cm} (6.24)

with \( \theta_x(u) \) and \( \theta_y(u) \) being the derivatives of \( \theta(u) \) with respect to \( x \) and \( y \).

Material Dispersion

In order to impose further constraints upon the wavelength domain, we note that the index of refraction for a wide range of materials is governed by the material dispersion equation [18]. Specifically, the dispersion equation constrains the refractive index
spectrum to a low-dimensional space with dimension equal to the number of dispersion coefficients. The number of these coefficients is generally lower than the number of spectral bands. The implication is hence that the number of variables to be estimated in the spectral domain can be significantly reduced, especially for spectral images with a high spectral resolution.

Thus, we introduce dispersion equations as a constraint on the refractive index for our optimisation scheme. Among several approximating functions of material dispersion in the physics literature, perhaps Cauchy’s and Sellmeier’s [18] are the most popular. In the former, Cauchy modelled the empirical relationship between the refractive index of a material and the wavelength of light as follows

\[
\eta(u, \lambda) = \sum_{m=1}^{M} C_m(u) \lambda^{-2(m-1)}
\]  

where the index of refraction depends solely on the wavelength and the material-specific dispersion coefficients \(C_m(u), m \in \{1, \ldots, M\}\).

In addition, Sellmeier’s dispersion equation [146] can handle anomalous dispersive regions by including additional coefficients to represent vacuum wavelengths, \(i.e\). where the wave front moves across vacuum, and holds for a wide range of wavelengths, including the ultraviolet, visible and infrared spectrum. Sellmeier’s dispersion equation is given by

\[
\eta^2(u, \lambda) = 1 + \sum_{m=1}^{M} \frac{B_m(u) \lambda^2}{\lambda^2 - D_m(u)}
\]  

where \(B_m(u)\) and \(D_m(u)\) are the material-specific dispersion coefficients.

The dispersion equations above allow a representation of the index of refraction as a linear combination of \(M\) rational functions of wavelength. With these representations, the estimation of refractive index can be treated as that of computing the dispersion coefficients. In practice, an expansion containing up to the sixth term is sufficient to represent a wide range of materials including crystals, liquids, glasses, gases and plastics [89]. For spectral imagery comprising more than seven wavelength-indexed bands, the number of equations relating the Fresnel transmission ratio to the zenith angle and refractive index exceeds the number of dispersion coefficients, rendering the problem solvable. As a result, the problem becomes a well-constrained one that can be formulated in a minimisation setting.
CHAPTER 6. SHAPE AND REFRACTIVE INDEX RECOVERY

Objective Function

Having introduced the integrability and material dispersion constraints in Sections 6.2.3 and 6.2.3, we now focus on the formulation of an objective function for the estimation of the zenith angle and refractive index. The rationale behind our cost function lies in the use of the additional constraints, including integrability and dispersion equations so as to allow the recovery of the shape and index of refraction to be performed without prior knowledge or predetermined illumination conditions.

The cost function aims at satisfying Equation 6.18, which equates the square root of the Fresnel transmission ratio defined in Equation 6.17 to the quantity \( \sqrt{\frac{I_{\min}}{I_{\max}}} \) while taking into account the integrability of the surface and the material dispersion equation. Our objective function is given by two terms. The first of these accounts for the error of the Fresnel transmission ratio \( R(u, \lambda) \) in Equation 6.17 with respect to the ratio \( r(u, \lambda) = \sqrt{\frac{I_{\min}}{I_{\max}}} \) as computed from the image radiance. The second term measures the error of the integrability constraint described in Equation 6.24. Thus, the cost function is given by

\[
E = \int_S \int_W \left( R(u, \lambda) - r(u, \lambda) \right)^2 d\lambda du + \beta(u) \int_S \left( \cos \alpha(u) \frac{\partial \tan \theta(u)}{\partial y} - \sin \alpha(u) \frac{\partial \tan \theta(u)}{\partial x} \right)^2 du
\]

subject to the chosen dispersion equation, i.e. Equation 6.25 or 6.26, where \( S \) is the image spatial domain and \( W \) is the wavelength range.

In Equation 6.27, we assume to have obtained an estimation of the azimuth angle \( \alpha(u) \) up to a 180-degree ambiguity, as described in Section 6.2.2, and treat it as a constant in the cost function. We note that this cost function is invariant to the 180-degree shift in the azimuth angle, i.e.

\[
\left( \cos \alpha(u) \theta_x(u) - \sin \alpha(u) \theta_y(u) \right)^2 = \left( \cos(\alpha(u) + \pi) \theta_x(u) - \sin(\alpha(u) + \pi) \theta_y(u) \right)^2
\]

Therefore, we can utilise the rough estimate of the azimuth angle obtained in Section 6.2.2 without an adverse effect in the estimation of the zenith angle and the refractive index.

In addition, the Fresnel transmission ratio \( R(u, \lambda) \) is related to the zenith angle and refractive through Equation 6.17. The regularisation parameter \( \beta(u) \) is spatially varying.
and weighs the relative importance between the data closeness and surface smoothness imposed by the integrability constraint. Here, we use the spatial dependence of $\beta(u)$ on the surface location so as to reflect the reliability of the azimuth angle $\alpha(u)$ estimated from polarisation information. To quantify the reliability of the estimate $\alpha(u)$, we employ the degree of polarisation as follows

$$
\beta(u) = \gamma \langle \frac{I_{\text{max}}(u, \lambda) - I_{\text{min}}(u, \lambda)}{I_{\text{max}}(u, \lambda) + I_{\text{min}}(u, \lambda)} \rangle_{\lambda}
$$

(6.28)

where $\gamma$ is a scaling constant for the whole image and $\langle \cdot \rangle_{\lambda}$ denotes the mean value across wavelengths.

### 6.3 Optimal Zenith Angle and Refractive Index

In this section, we adopt an iterative approach to the recovery of both the zenith angle and the index of refraction. The algorithm proceeds in a coordinate descent manner [20] to minimise the cost functional in Equation 6.27. The step sequence of the minimisation strategy is summarised in Algorithm 6. At each iteration, the first step of the coordinate descent approach minimises the objective function with respect to the zenith angle $\{\theta(u) | u \in S\}$ while fixing the index of refraction. The second step recovers the index of refraction $\{\eta(u, \lambda) | u \in S, \lambda \in W\}$ for the current estimate of the zenith. We iterate between these two steps until convergence is reached.

The algorithm is initialised with a uniform material refractive index $\eta_0$ across both the spatial and spectral domains, as indicated in Line 1. It terminates once the estimation of the parameters stabilises, i.e. the change between the estimates in successive iterations falls below a pre-determined threshold $\tau_\theta$ for the zenith angle and a threshold $\tau_\eta$ for the refractive index. For both, the zenith angle and refractive index, the change is measured as the $\ell^1$-norm of the difference between successive estimates. In the following two subsections, we elaborate further on the details of the optimisation steps above.
Algorithm 6 Estimation of the zenith angle and refractive index from a polarimetric spectral image with $K$ wavelength-indexed bands.

Require: Fresnel transmission ratio $r(u, \lambda)$ for each pixel $u$ and band $\lambda \in \{\lambda_1, \ldots, \lambda_K\}$.

Ensure: $(\theta(u), \eta(u, \lambda)) | u \in S, \lambda \in W$, where

1. $\theta(u):$ the zenith angle at the pixel $u$.
2. $\eta(u, \lambda):$ the refractive index at pixel $u$ and wavelength $\lambda$.

1: $\eta(u, \lambda) \leftarrow \eta_0 \forall u \in S, \lambda \in W$
2: while true do
3:   $\theta_{old}(u, \lambda) \leftarrow \theta(u, \lambda)$
4:   $\eta_{old}(u, \lambda) \leftarrow \eta(u, \lambda)$
5:   $\theta(u) \leftarrow \arg\min_{\theta(u)} \mathcal{E}$
6:   $\eta(u, \lambda) \leftarrow \arg\min_{\eta(u, \lambda)} \mathcal{E}$
7:   if $|\theta_{old}(u) - \theta(u)| < \tau_{\theta}$ and $|\eta_{old}(u, \lambda) - \eta(u, \lambda)| < \tau_{\eta}, \forall u$ and $\lambda$ then
8:      break
9:   end if
10: end while
11: return $\theta(u), \eta(u, \lambda)$

6.3.1 Recovery of the Zenith Angle

Here we derive a solution for the zenith angle while keeping the refractive index fixed in each iteration. We note that the original cost functional in Equation 6.27 is complex, involving a non-linear Fresnel transmission ratio function $R(u, \lambda)$ with respect to the zenith angle $\theta(u)$. To minimise this cost functional, we need to solve a highly non-linear Euler-Lagrange equation whose analytical solution cannot be derived in a straightforward manner. To this end, we opt for an equivalent, yet simpler formulation of the cost functional, which expresses the data error term using the inverse function of the Fresnel transmission ratio in Equation 6.17. Eventually, we reformulate the cost functional with a quadratic function of the zenith angle $\theta(u)$ in the data error term, whose minimum can be derived analytically.

With the refractive index $\eta(u, \lambda)$ fixed, we can invert the Fresnel transmission ratio
function in Equation 6.17 with respect to the zenith angle \( \theta(u) \) as follows

\[
\sin \theta(u) = \frac{\eta(u, \lambda) \sqrt{1 - R^2(u, \lambda)}}{\sqrt{\eta^2(u, \lambda) - 2R(u, \lambda)\eta(u, \lambda) + 1}}
\]  

(6.29)

Note that the zenith angle is a geometric variable independent of the wavelength. However, the above equation provides wavelength-dependent estimates of the zenith angle. In practice, these estimates may not be the same across the spectrum due to measurement error and noise corruption. If the index of refraction is at hand, the value of \( r(u, \lambda) \) computed from the ratio of maximal and minimal image irradiance can be used as an estimate for the function \( R(u, \lambda) \). This yields a wavelength-dependent estimate of the zenith angle \( \theta(u) \) at the current iteration, which is given by

\[
\varphi(u, \lambda) = \arcsin \left( \frac{\eta(u, \lambda) \sqrt{1 - r^2(u, \lambda)}}{\sqrt{\eta^2(u, \lambda) - 2r(u, \lambda)\eta(u, \lambda) + 1}} \right)
\]  

(6.30)

Note that in Equation 6.30, we use the notation \( \varphi(u, \lambda) \) for the wavelength-dependent estimate and distinguish it from the wavelength-independent zenith angle \( \theta(u) \). We take profit of this wavelength dependency and instead of directly minimising the original cost functional in Equation 6.27, we seek to recover a zenith angle close to the wavelength-dependent estimates in Equation 6.30 while satisfying the integrability constraint. Thus, we employ the alternative cost functional

\[
\mathcal{E}_1 = \int_S \int_W (\tan \theta(u) - \tan \varphi(u, \lambda))^2 \, d\lambda \, du
\]

\[
+ \beta(u) \int_S \left( \cos \alpha(u) \frac{\partial \tan \theta(u)}{\partial y} - \sin \alpha(u) \frac{\partial \tan \theta(u)}{\partial x} \right)^2 \, du
\]

\[
= \int_S \int_W (g(u) - \tan \varphi(u, \lambda))^2 \, d\lambda \, du
\]

\[
+ \beta(u) \int_S \left( \cos \alpha(u)g_y(u) - \sin \alpha(u)g_x(u) \right)^2 \, du
\]  

(6.31)

as an alternative to that in Equation 6.27.

Equation 6.31 poses the minimisation problem in a simpler setting. The merit of the alternative cost functional is that the Fresnel ratio error is quantified as a quadratic form of the function \( g(u) \). This is important since this quadratic form is more tractable than the original error term, which contains a rational function in the expression for \( R(u, \lambda) \).
Moreover, we can rewrite Equation 6.31 as follows

\[ E_1 = \int_S f(u, \theta(u), \theta_x(u), \theta_y(u)) du \]  

(6.32)

by letting \( f(\cdot) \) be given by

\[ f(u, \theta(u), \theta_x(u), \theta_y(u)) \triangleq \int_W (\theta(u) - \varphi(u, \lambda))^2 d\lambda + \beta(u) \left( \cos \alpha(u) \theta_x(u) - \sin \alpha(u) \theta_y(u) \right)^2 \]

(6.33)

The function \( f(\cdot) \) is important since it permits the use of calculus of variations to recover the minimiser of the functional in Equation 6.32. We do this by noting that the minima must satisfy the following Euler-Lagrange equation

\[ \frac{\partial f}{\partial \theta} = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial \theta_x} \right) + \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial \theta_y} \right) \]  

(6.34)

By computing the derivatives of \( f \) so as to satisfy the Euler-Lagrange equation above, we arrive at the following differential equation

\[ \theta(u) \int_W d\lambda - \int_W \varphi(u, \lambda) d\lambda = \beta(u) \times \left( \sin^2 \alpha(u) \theta_{xx}(u) - \sin 2\alpha(u) \theta_{xy}(u) + \cos^2 \alpha(u) \theta_{yy}(u) \right) \]

(6.35)

where \( \theta_{xx}(u), \theta_{xy}(u) \) and \( \theta_{yy}(u) \) are the second order and covariant derivatives of \( \theta(u) \) with respect to the \( x \) and \( y \) axes of the coordinate system.

In the discrete case, where the imagery is acquired at \( K \) wavelength-indexed bands \( \lambda \in \{ \lambda_1, \ldots, \lambda_K \} \), we have \( \int_W d\lambda = K \). Therefore, \( \theta(u) \) satisfies the differential equation

\[ \theta(u) = \frac{1}{K} \int_W \varphi(u, \lambda) d\lambda + \frac{\beta(u)}{K} \left( \sin^2 \alpha(u) \theta_{xx}(u) - \sin 2\alpha(u) \theta_{xy}(u) + \cos^2 \alpha(u) \theta_{yy}(u) \right) \]

(6.36)

We note that Equation 6.36 is a second-order partial differential equation with respect to \( \theta(u) \). We further enforce the continuity and differentiability of the spatial domain by assuming that the function \( \theta(u) \) is continuously twice-differentiable, \( i.e. \theta_{xy}(u) = \theta_{yx}(u) \). This assumption permits the decomposition of \( \theta(u) \) into an orthogonal basis of integrable two-dimensional functions, in a similar manner to that in [54]. Since digital images have a limited band of spatial frequencies, the surface shading can be expressed as a finite linear combinations of the real part of Fourier basis functions, which are cosine functions. This representation allows an analytical solution to the functional minimisation problem above. Moreover, we will show later that this representation also leads to a computationally efficient solution to Equation 6.36.
6.3. OPTIMAL ZENITH ANGLE AND REFRACTIVE INDEX

Note that the function \( \theta(u) \) can be viewed as a discrete function on a two-dimensional lattice. Let the size of the lattice be \( X \times Y \), where \( X \) and \( Y \) are the image width and height, respectively. Based on the Nyquist–Shannon sampling theorem [149], the zenith angle \( \theta(u) \) can be reconstructed using frequency components of up to one-half of the sampling frequency of the image. In image processing, these sampling frequencies \( \nu \), where \( \nu = (\nu_x, \nu_y) \), are often chosen such that \( \nu_x = \frac{2\pi i}{X} \), where \( i = 0, 1, \ldots, X - 1 \), and \( \nu_y = \frac{2\pi j}{Y} \), where \( j = 0, 1, \ldots, Y - 1 \) [59]. With these two-dimensional frequency components, the function \( \theta(u) \) can be reconstructed with an orthogonal set of Fourier basis functions \( e^{i(\nu^T u)} = e^{i(\nu_x u_x + \nu_y u_y)} \), where \( i \) is the imaginary number and \( u = (x_u, y_u) \) is the pixel location. Formally, this is given by

$$\theta(u) = \sum_{\nu} \kappa_{\nu} e^{i(\nu^T u)} \quad (6.37)$$

Intuitively, Equation 6.37 means that the shading of the surface can be decomposed into a linear combination of Fourier components with a range of frequencies matching that of the input imagery. In the equation, \( \kappa_{\nu} \) is the coefficient (weight) of the Fourier basis function \( e^{i(\nu^T u)} \), which can be computed making use of the equation

$$\kappa_{\nu} = \frac{1}{|S|} \sum_{u} \theta(u)e^{-i(\nu^T u)}$$

where \( |S| \) represents the number of image pixels.

Similarly, the partial derivatives of \( \theta(u) \) can also be expressed in terms of the Fourier basis, as follows

$$\theta_{xx}(u) = -\sum_{\nu} \kappa_{\nu} \nu_x^2 e^{i(\nu^T u)} \quad (6.38)$$

$$\theta_{xy}(u) = -\sum_{\nu} \kappa_{\nu} \nu_x \nu_y e^{i(\nu^T u)} \quad (6.39)$$

$$\theta_{yy}(u) = -\sum_{\nu} \kappa_{\nu} \nu_y^2 e^{i(\nu^T u)} \quad (6.40)$$

Let \( h(u) = \frac{1}{K} \int_{\Phi} \varphi(u, \lambda) d\lambda \). By substituting Equations 6.37, 6.38, 6.39 and 6.40 into Equation 6.36, we obtain

$$h(u) = \sum_{\nu} \kappa_{\nu} e^{i(\nu^T u)} \left( 1 + \frac{\beta(u)}{K} \left( \sin^2 \alpha(u) \nu_x^2 - \sin 2\alpha(u) \nu_x \nu_y + \cos^2 \alpha(u) \nu_y^2 \right) \right) \quad (6.41)$$
Note that Equation 6.41 applies to every image location $u$ and every spatial frequency $\nu$. By making use of the expressions for $h(u)$ at every image location and frequency, we can construct a linear system with respect to the unknown vector $\mathbf{U} = [\kappa_u]^T$, which is, effectively, the concatenation of all the Fourier coefficients. The recovery of the coefficients $\kappa_u$ can be then effected by solving the linear system $LU = H$, with $H = [h(u)]^T$, which is the vectorial concatenation of the known function values $h(u)$ for all the image locations $u$ and $L$ is a matrix with rows and columns indexed to the image pixels and spatial frequencies, respectively. In other words, the matrix element $L_{u,\nu}$ corresponding to a given pixel $u$ and a given frequency $\nu$ is $L_{u,\nu} = e^{i(u^Tv)} \left( 1 + \frac{\beta(u)}{K} \left( \sin^2 \alpha(u)\nu_x^2 - \sin 2\alpha(u)\nu_x\nu_y + \cos^2 \alpha(u)\nu_y^2 \right) \right)$. With a chosen Fourier basis and the azimuth angle $\alpha(u)$ obtained as described in Section 6.2.2, the matrix $L$ can be computed in a straightforward manner. With the coefficients $\kappa_u$ at hand, the zenith angle $\theta(u)$ can be recovered through the application of Equation 6.37.

### 6.3.2 Recovery of the Refractive Index

With the zenith angle obtained in the previous section, we now turn our attention at the estimation of the refractive index at each image location making use of the Fresnel transmission ratio. To derive the refractive index directly from the Fresnel transmission ratio in Equation 6.17, we are required to solve a quadratic equation involving the index of refraction $\eta(u, \lambda)$. This expression is given by

$$
\left( \cos^2 \theta(u) - r^2(u, \lambda) \right) \times \eta^2(u, \lambda) + 2r(u, \lambda) \sin^2 \theta(u) \times \eta(u, \lambda) - \sin^2 \theta(u) = 0 \quad (6.42)
$$

In general, the quadratic equation above yields no more than two real-valued roots. The choice of refractive index value depends on the physically plausibility of these roots, i.e. the refractive index for dielectrics must be a real value greater than one. This choice is straightforward if only one of the roots is physical plausible.

In the case where the two roots are plausible, we adopt an iterative approach which iteratively selects the root closer to the refractive index average at the same wavelength within the local spatial neighbourhood. This approach works under the assumption that there is a single solution to the refractive index at a number of pixels in the image. Initially, we label these pixels as having their refractive index uniquely determined. At
6.3. OPTIMAL ZENITH ANGLE AND REFRACTIVE INDEX

Each iteration, we assign the refractive index of those pixels with two plausible solutions making use of the regions whose refractive index is already determined. We do this by selecting the root which is in better accordance with the average of the refractive indices within its spatial neighbourhood. A pseudocode of this iterative procedure is illustrated in Algorithm 7.

Algorithm 7 Refractive Index Selection

Require: \( \eta(u, \lambda) \): Solutions to the refractive index at pixel \( u \) and wavelength \( \lambda \)

Ensure: \( \eta^*(u, \lambda) \): The uniquely determined refractive index at pixel \( u \) and wavelength \( \lambda \)

1: for all wavelength \( \lambda \) do
2: for all pixel \( u \) with a single physically plausible refractive index at wavelength \( \lambda \) do
3: \( \text{determined}(u, \lambda) \leftarrow \text{true} \)
4: end for
5: while there are pixels with two plausible refractive indices do
6: for all pixel \( u \) with two solutions \( \eta_1(u, \lambda) \) and \( \eta_2(u, \lambda) \) do
7: \( N(u) \leftarrow \text{the spatial neighbourhood of } u \)
8: \( \bar{\eta}(u, \lambda) \leftarrow \text{Average}_{v \in N(u), \text{determined}(v, \lambda) = \text{true}} \eta(v, \lambda) \)
9: \( \eta^*(u, \lambda) \leftarrow \eta_i(u, \lambda) \) which is closer to \( \bar{\eta}(u, \lambda) \)
10: \( \text{determined}(u, \lambda) \leftarrow \text{true} \)
11: end for
12: end while
13: end for
14: return \( \eta^*(u, \lambda) \forall \) pixel \( u \) and wavelength \( \lambda \)

When neither root is a physically plausible solution to Equation 6.42 exists, we opt for an approximation of the Fresnel transmission ratio that provides a single solution to the refractive index. To this end, we consider an approximating function that is a product of two separable factors, one of which contains solely the refractive index while the other depends on the zenith angle. By adopting a formulation similar to Schlick’s approximation of the Fresnel reflection coefficient [145], we employ the expression

\[
R^*(u, \lambda) = d + c \ (1 - \cos \theta(u))^b
\]  

(6.43)
where \(b, c\) and \(d\) are constants.

Figure 6.6(a) shows this graph for values of \(\eta\) ranging between 1.2 and 3, with an increment of 0.3. Empirical observations of this graph give rise to a power function with respect to the zenith angle. At the end-points, \(R = 1\) as \(\theta(u) = 0\) and \(R = \frac{1}{\eta(u, \lambda)}\) as \(\theta(u) = \frac{\pi}{2}\). To satisfy these conditions, it is necessary that \(c = \frac{1}{\eta(u, \lambda)} - 1\) and \(d = 1\). Thus, we have

\[
R'(u, \lambda) = 1 + \left(\frac{1}{\eta(u, \lambda)} - 1\right)(1 - \cos \theta(u))^{b}
\] (6.44)

In Equation 6.44, \(b\) is a pre-determined power that provides the best fit with respect to the true Fresnel transmission ratio over a range of parameter values. In this chapter, we consider material refractive indices between 1 and 3 and \(\theta(u) \in [0, \frac{\pi}{2}]\). Using a one-dimensional search for the power \(b\), we find that \(b = 1.4\) minimises the \(\ell^1\)-error between the approximating and the true Fresnel transmission ratio. In Figure 6.6(b), we have plotted the approximating Fresnel ratio for \(b = 1.4\) and noted its similarity to the Fresnel ratio function.

To verify the approximation accuracy for the Fresnel transmission ratio, in Figure 6.7, we plot the error function for \(b = 1.4\), where it is represented as a surface with respect to both the zenith angle and refractive index. From the Figure, we observe that the absolute error is below 0.04 for all the combinations of zenith angle and refractive index.
tive index within the considered range. The error is largest near grazing zenith angles ($70^\circ \leq \theta(u) < 90^\circ$) for small refractive indices or at around $40^\circ$ for refractive indices larger than 2.3. However, these are extreme cases and, generally, the material and the geometry under study do not fall into these cases.

Using the approximated Fresnel ratio function, we arrive at a single approximate solution for the refractive index at each pixel and wavelength. Given the zenith angle $\theta(u)$ at the current iteration, the index of refraction is estimated to be

$$
\eta(u, \lambda) = \frac{(1 - \cos \theta(u))^b}{(1 - \cos \theta(u))^b - 1 + r(u, \lambda)}
$$

(6.45)

Next, we apply the following strategy to ensure the physical plausibility of the approximating solution in Equation 6.45. We note that the approximating refractive index is physically plausible, i.e. $\eta(u, \lambda) > 1$ if $(1 - \cos \theta(u))^b - 1 + r(u, \lambda) > 0$ and $r(u, \lambda) < 1$. In the case where $r(u, \lambda) = 1$, we can conclude that $\theta(u) = 0$ as can be observed in Figure 6.6(a). However, in this case, the refractive index value can be arbitrary. Otherwise, when $r(u, \lambda) < 1$, i.e. $\theta(u) \neq 0$, we can guarantee the physical plausibility of the solution in Equation 6.45 by scaling the zenith angles at all the image pixels such that $\cos \theta(u) < 1 - (1 - r(u, \lambda))^\frac{1}{2}$ for all $u$ and $\lambda$.

Finally, having obtained a refractive index spectrum per surface location, we fit

![Figure 6.7: The approximation error for the Fresnel transmission ratio plotted for refractive indices between 1 and 3 and zenith angles in the range of $[0, \frac{\pi}{2}]$.](image)
the spectrum to one of the dispersion equations in Equation 6.25 or 6.26. As a result, we recover the dispersion coefficients in these equations and compute the closest approximation on the dispersion curve for the refractive index obtained earlier.

The Cauchy’s dispersion equation, i.e. Equation 6.25, can be rewritten in the following matrix form

\[ n(u) = \Lambda C(u) \] (6.46)

where \( n(u) = \begin{bmatrix} \eta(u, \lambda_1) \\ \vdots \\ \eta(u, \lambda_K) \end{bmatrix}, C(u) = \begin{bmatrix} C_1(u) \\ \vdots \\ C_M(u) \end{bmatrix}, \) and \( \Lambda = \begin{bmatrix} 1 & \lambda_1^{-2} & \ldots & \lambda_1^{-(M-1)} \\ 1 & \lambda_2^{-2} & \ldots & \lambda_2^{-(M-1)} \\ \vdots \\ 1 & \lambda_K^{-2} & \ldots & \lambda_K^{-(M-1)} \end{bmatrix}, \)

where \( M \) is the number of dispersion coefficients used.

Since \( n(u) \) and \( \Lambda \) are known, the system in Equation 6.46 is over-determined if the number of dispersion coefficients is chosen such that \( M \leq K \), where \( K \) is the number of wavelength indexed bands in the imagery. In this case, we can solve for the coefficients \( C(u) \) and compute the refractive index spectrum according to the dispersion equation.

For Sellmeier’s equation, i.e. Equation 6.26, the fitting task can be posed as a non-linear least-squares optimisation problem. If the number of dispersion coefficients \( M \) is chosen such that \( M \leq K \) then the optimisation problem above becomes well-constrained and the coefficients \( C(u) \) can be solved numerically by standard line-search or trust-region techniques.

6.4 Experiments

In this section, we perform experiments to illustrate the utility of our method for the purpose of recovering the shape and refractive index of objects in the scene. To this end, we report results on synthetic and real-world spectro-polarimetric images. The use of synthetic data permits us to effect a quantitative analysis where the accuracy of the estimated parameters can be compared to the ground-truth. On the other hand, the use of real-world images allows us to illustrate the utility of the recovered surface orientation and index of refraction for the purpose of generating views under novel lighting conditions and view points.
6.4. EXPERIMENTS

For all our experiments, we employed the Cauchy's dispersion equation as a constraint on the refractive index to estimate the optimal parameters of the cost functional in Equation 6.27. Here, we consider the eighth-order Cauchy's dispersion equation consisting of five terms. The fitting of refractive index spectra to this dispersion equation was performed via constrained linear least-squares [35].

6.4.1 Synthetic Data

We commence by performing experiments on a synthetic dataset. Our dataset comprises spectro-polarimetric images of eight 3D surfaces rendered with the Wolff diffuse reflectance model for dielectrics [181]. To generate our dataset, we render these synthetic surfaces with the refractive index of the 19 plastic and liquid materials reported in [89], and the spectral reflectance of Polyvinyl Chloride (PVC) under five illumination directions. These amount to a total of 1520 combinations of shape and photometric parameters. For each combination, five polarisation images were generated corresponding to five polariser orientations at 0°, 30°, 45°, 60° and 90° in the anti-clockwise direction with respect to the horizontal axis of the rendering context, \( i.e. \) the image coordinate system. All the multispectral images in this dataset are 30 bands in length, spanning the 430 – 720\( \text{nm} \) range, with a 10\( \text{nm} \) step between successive bands. In this experiment, we have set the bandwidth \( h \) of the Epanechnikov kernel in Section 6.2.2 to unity and the scaling constant \( \gamma \) of the integrability term in Equation 6.28 to 0.05.

The different illumination conditions simulated in our dataset are described as follows. Each of the illuminants assumes the power spectrum of sunlight, which we acquired using a StellarNet spectrometer. Each illuminant direction is in the horizontal ground plane that contain the viewing direction and points towards the surface under study. We have used five light source directions, which we denoted \( L_1, L_2, L_3, L_4 \) and \( L_5 \). These form angles of \(-26.5°, -14°, 0°, 14° \) and \(26.5° \) with respect to the viewing direction, where a negative angle means the light direction is on the left-hand side with respect to the view point and a positive angle means otherwise.

In Figure 6.8, we show the variation of shading with respect to the illumination direction. In the Figure, we show the shapes in our datasets. The first three rows include
synthetic images under three instances of single illumination directions, which are $L_3$, $L_4$ and $L_5$, and the last two are instances of two simultaneous illuminant directions, which are $L_2 + L_4$ and $L_1 + L_5$. The images are rendered in pseudo trichromatic colours synthesized from the multispectral radiance of the $45^\circ$ polarisation component, using the Stiles and Burch colour matching functions [157]. Note that we already presented the method of simulating the trichromatic responses of spectral images in Chapter 3.

Similarly, Figure 6.9 shows the variation of shading with respect to the angle of polarisation. Each row, from top to bottom, respectively shows the polarisation component oriented at $0^\circ$, $30^\circ$, $45^\circ$, $60^\circ$, $90^\circ$. The images of the above components have been synthesized with the refractive index of Polystyrene under the frontal illumination direction. As before, we have employed the Stiles and Burch colour matching functions [157] in order to generate the pseudo trichromatic colours.

In Figure 6.10, we present the needle maps estimated by our method for the synthetic shapes under the frontal light direction $L_3$, and the oblique light directions originating from the right-hand side of the viewing position, at angles of $14^\circ$ ($L_4$) and $26.5^\circ$ ($L_5$). The top row shows the ground-truth needle maps corresponding to the surface normals of the synthetic shapes. The second, fourth and sixth rows respectively show the needle-maps recovered from the images rendered under the illuminant directions $L_3$, $L_4$ and $L_5$. In addition, we show the corresponding angular error of the estimated normal fields for the above illumination conditions in the third, fifth and seventh rows. The error maps in these rows are depicted in grayscale values indicating the dot products between the normalised estimated surface normals and the ground truth.

We note that, for the Dome, Ridge, Torus, Two-Domes, Vase and Tea pot, the surface normal orientation has been successfully recovered almost everywhere except for regions with strong shadowing. This observation applies to the shapes above regardless of the illumination direction and shading variations. The low estimation errors for the surface normals of these shapes can be verified visually by inspecting the error maps in the third, fifth and seventh rows. The strategy of smoothing the azimuth angles we described in Section 6.2.2 tends to be effective on convex surfaces, but breaks down at non-convex parts for shapes consisting of a mixture of concave and convex surfaces, such as the Mozart and the Duck. Here, the horizontal symmetry of the recovered sur-
6.4. EXPERIMENTS

Illumination direction \( L_3 \) (frontal)

Illumination direction \( L_4 \) (14° to the right)

Illumination direction \( L_5 \) (26.5° to the right)

Illumination direction \( L_2 + L_4 \)

Illumination direction \( L_1 + L_5 \)

Dome  Ridge  Torus  Two domes  Mozart  Vase  Duck  Tea pot

Figure 6.8: Synthetic images of our test shapes rendered with the refractive index of Polystyrene under five different illumination conditions. The rows, from top to bottom, correspond to the combinations of illumination directions \( L_3, L_4, L_5, L_2 + L_4 \) and \( L_1 + L_5 \). The images are rendered in pseudo trichromatic colours synthesized from the multispectral radiance of the polarisation component oriented at 45° with respect to the horizontal axis of the image coordinate system.
Figure 6.9: Pseudo trichromatic rendering of synthetic spectral-polarimetric images of eight shapes synthesized with the refractive index of Polystyrene under the frontal light source direction. Each row, from top to bottom, shows the polarisation component oriented at 0°, 30°, 45°, 60°, 90°.
face normals for a number of shapes, such as the Dome, Ridge, Torus, Two-Domes and Vase, indicates that our estimation method relies on polarisation rather than shading to reveal the surface geometry. The results also imply that our method is insensitive to changes in illumination power spectrum and direction.

Now, we turn our attention to the accuracy of the estimated shading maps. The first row of Figure 6.11 shows the ground-truth shading maps. These are followed by the estimated shading maps in the subsequent rows. Note that the recovered shading is in strong agreement with the ground truth, with the exception of strongly shadowed regions, as exhibited earlier in the recovered needle-maps in Figure 6.10. Once again, these results are invariant to changes in illumination power spectrum and direction.

With the surface normal estimation in hand, we reconstruct the surface depth by means of normal field integration. To this end, we make use of the surface integration method introduced by Frankot and Chellappa [54] for the purpose of surface reconstruction. In the first row of Figure 6.12, we present the surface rendering for all the shapes in the dataset. The following rows show rendering of the recovered surfaces when the input imagery is illuminated from several illumination directions, including $L_3$, $L_4$ and $L_5$. All these surfaces are rendered under the same novel lighting direction. As shown, the reconstruction of most shapes that adhere to our convexity assumption, i.e. except for Mozart and Duck, is almost perfect with respect to the ground truth. This result is consistent with the qualitative ones in the Figures 6.10 and Figures 6.11. Here, we also note that strong shadows caused by the oblique illumination directions $L_4$ and $L_5$ do not convey any information, and therefore, do not benefit surface reconstruction within these regions. The effect of dark regions on the reconstruction results can be observed in the rendering of the Ridge, Two-Domed shape and Duck in the third and fourth rows. However, the reconstruction of non-shadowed parts still exhibits the curvature of the original shapes.

Now, we turn to a more quantitative set of results. In Table 6.1, we show the accuracy of the recovered needle map, which is quantified as the Euclidean angle, in degrees, between the estimated surface normal direction and the corresponding ground truth, averaged over the pixels of every image. In columns 2–6, we report the mean and standard deviation of the angular error for each combination of shape and illuminant
Figure 6.10: The estimated needle maps of eight different shapes as compared to the corresponding groundtruth. First row: ground truth needle maps. Second, fourth and sixth rows: The needle maps recovered from images illuminated under the illumination directions $L_3$, $L_4$ and $L_5$, respectively. Third, fifth and seventh rows: The error of the estimated needle maps for the illumination directions $L_3$, $L_4$ and $L_5$, respectively.
Figure 6.11: The estimated shading maps of eight different shapes as compared to the corresponding groundtruth. First row: The ground truth shading maps. From the second to the fourth row: The shading maps recovered from images illuminated under the illumination directions $L_3$, $L_4$ and $L_5$, respectively.
CHAPTER 6. SHAPE AND REFRACTIVE INDEX RECOVERY

Figure 6.12: The estimated surface depths of eight different shapes as compared to the corresponding groundtruth. First row: The ground truth depth. From the second to the fourth row: The depth maps recovered from images illuminated under the illumination directions $L_3$, $L_4$ and $L_5$, respectively.
Table 6.1: The accuracy of the recovered surface normals for several illumination directions computed as the average absolute angular difference (in degrees) between the estimated surface normal directions and the ground truth.

<table>
<thead>
<tr>
<th>Shape</th>
<th>Surface normal error (degrees)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$L_3$</td>
<td>$L_4$</td>
<td>$L_5$</td>
<td>$L_2 + L_4$</td>
<td>$L_1 + L_5$</td>
</tr>
<tr>
<td>Dome</td>
<td>2.0170 ± 1.6373</td>
<td>2.2202 ± 1.8689</td>
<td>1.9844 ± 1.5550</td>
<td>2.0170 ± 1.6373</td>
<td>2.0171 ± 1.6374</td>
</tr>
<tr>
<td>Ridge</td>
<td>2.1008 ± 0.3520</td>
<td>2.1008 ± 0.3520</td>
<td>3.3999 ± 0.3343</td>
<td>2.1008 ± 0.3520</td>
<td>2.1008 ± 0.3520</td>
</tr>
<tr>
<td>Torus</td>
<td>1.7772 ± 1.0043</td>
<td>1.7817 ± 1.0138</td>
<td>1.8052 ± 0.9981</td>
<td>1.7772 ± 1.0037</td>
<td>1.7776 ± 1.0040</td>
</tr>
<tr>
<td>Test shape</td>
<td>2.0476 ± 1.4064</td>
<td>2.0690 ± 1.404</td>
<td>2.1141 ± 1.3934</td>
<td>2.0477 ± 1.4067</td>
<td>2.0477 ± 1.4064</td>
</tr>
<tr>
<td>Mozart</td>
<td>14.4993 ± 0.8102</td>
<td>14.5186 ± 1.0945</td>
<td>14.6955 ± 1.0344</td>
<td>14.5253 ± 0.8032</td>
<td>14.5694 ± 0.7750</td>
</tr>
<tr>
<td>Vase</td>
<td>1.7351 ± 1.2868</td>
<td>1.7421 ± 1.3019</td>
<td>1.7230 ± 1.2740</td>
<td>1.7351 ± 1.2868</td>
<td>1.7351 ± 1.2868</td>
</tr>
<tr>
<td>Duck</td>
<td>5.6798 ± 1.3044</td>
<td>6.5201 ± 1.1333</td>
<td>5.6263 ± 1.1393</td>
<td>5.6743 ± 1.3233</td>
<td>5.2331 ± 1.2778</td>
</tr>
<tr>
<td>Tea pot</td>
<td>2.0005 ± 1.0508</td>
<td>2.0005 ± 1.0508</td>
<td>2.0004 ± 1.0479</td>
<td>2.0005 ± 1.0508</td>
<td>2.0005 ± 1.0508</td>
</tr>
</tbody>
</table>

The numerical results in Table 6.1 show that the surface normals were recovered at a high level of accuracy, being lower or equal to 3.3999 degrees for most of the shapes, except for Mozart and Duck. The error outliers for the Mozart and Duck shapes are understandable since they are more complex than the other shapes, being composed of both convex and concave regions. Overall, the quantitative results are largely consistent with the needle maps and error maps in Figure 6.10.

Likewise, Table 6.2 shows negligible surface reconstruction error for most of the shapes except for the Mozart and the Duck. Among these shapes, the highest error is observed for the Ridge, at i.e. approximately 3%–6% of the maximal depth value. As before, the higher depth reconstruction error for the Mozart and Duck is due to the complexity introduced by the combination of convex and concave regions. Further, the depth error for the Duck can be partly attributed to dark areas in the input images when synthesized under the oblique light source directions ($L_3$ and $L_4$).
CHAPTER 6. SHAPE AND REFRACTIVE INDEX RECOVERY

<table>
<thead>
<tr>
<th>Test shape</th>
<th>Normalised depth error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$L_3$</td>
</tr>
<tr>
<td>Dome</td>
<td>0.0040 ± 0.0062</td>
</tr>
<tr>
<td>Ridge</td>
<td>0.0313 ± 0.0034</td>
</tr>
<tr>
<td>Torus</td>
<td>0.0049 ± 0.0028</td>
</tr>
<tr>
<td>Test shape</td>
<td>0.0059 ± 0.0049</td>
</tr>
<tr>
<td>Mozart</td>
<td>0.0932 ± 0.0441</td>
</tr>
<tr>
<td>Vase</td>
<td>0.0095 ± 0.0085</td>
</tr>
<tr>
<td>Duck</td>
<td>0.0958 ± 0.0383</td>
</tr>
<tr>
<td>Tea pot</td>
<td>0.0116 ± 0.0025</td>
</tr>
</tbody>
</table>

Table 6.2: The accuracy of surface depth reconstruction for several illumination directions. This is computed as the mean absolute difference between the estimated and the ground-truth depth, with both the depth maps being normalised to the range between 0 and 1.

We note that the errors in Tables 6.1 and 6.2 often vary only slightly with respect to illumination direction. In some cases, the errors for images under the frontal and the combined illuminant directions ($L_2 + L_4$ and $L_1 + L_5$) are lower than those for images under the oblique illumination directions. The increase in error for the oblique lighting directions is partly due to the dark image areas in the input imagery. In fact, the variation of the surface normal error is within one degree and that of the depth error is within 0.03 when the illuminant shifts from the frontal direction to the most oblique direction (at an angle of 26.5° from the camera axis). Moreover, the standard deviations of the surface normal error and depth error in Tables 6.1 and 6.2 are negligible, peaking at 1.8689 and 0.0442 degrees, respectively. This means that the recovered shape is substantially insensitive to the material refractive indices used in our experiment. These observations, again, support the claim that polarisation is a good cue to surface orientation because it is robust to changes in illumination direction and material refractive index. As such, this differentiates Shape from Polarisation from Shape from Shading and Photometric Stereo methods, which attribute geometric cues to image shading.

Finally, we turn our attention to the accuracy of the recovered refractive index. In Table 6.3, we report the Euclidean angular difference, in degrees, between the average
6.4. EXPERIMENTS

<table>
<thead>
<tr>
<th>Test shape</th>
<th>(L_3)</th>
<th>(L_4)</th>
<th>(L_5)</th>
<th>(L_2 + L_4)</th>
<th>(L_1 + L_5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dome</td>
<td>0.0285 ± 0.0345</td>
<td>0.0284 ± 0.0344</td>
<td>0.0285 ± 0.0345</td>
<td>0.0285 ± 0.0345</td>
<td>0.0285 ± 0.0345</td>
</tr>
<tr>
<td>Ridge</td>
<td>0.0315 ± 0.0385</td>
<td>0.0315 ± 0.0385</td>
<td>0.0315 ± 0.0385</td>
<td>0.0315 ± 0.0385</td>
<td>0.0315 ± 0.0385</td>
</tr>
<tr>
<td>Torus</td>
<td>0.0303 ± 0.037</td>
<td>0.0303 ± 0.037</td>
<td>0.0303 ± 0.037</td>
<td>0.0303 ± 0.037</td>
<td>0.0303 ± 0.037</td>
</tr>
<tr>
<td>Test shape</td>
<td>0.0373 ± 0.0474</td>
<td>0.0367 ± 0.0464</td>
<td>0.0347 ± 0.0432</td>
<td>0.0373 ± 0.0474</td>
<td>0.0373 ± 0.0474</td>
</tr>
<tr>
<td>Mozart</td>
<td>0.0302 ± 0.0367</td>
<td>0.0305 ± 0.0371</td>
<td>0.0302 ± 0.0368</td>
<td>0.0302 ± 0.0367</td>
<td>0.0302 ± 0.0367</td>
</tr>
<tr>
<td>Vase</td>
<td>0.0321 ± 0.0389</td>
<td>0.0320 ± 0.0389</td>
<td>0.0310 ± 0.0378</td>
<td>0.0321 ± 0.0389</td>
<td>0.0321 ± 0.0389</td>
</tr>
<tr>
<td>Duck</td>
<td>0.0540 ± 0.0487</td>
<td>0.0551 ± 0.0509</td>
<td>0.0542 ± 0.0498</td>
<td>0.0540 ± 0.0487</td>
<td>0.0540 ± 0.0487</td>
</tr>
<tr>
<td>Tea pot</td>
<td>0.0317 ± 0.0387</td>
<td>0.0317 ± 0.0387</td>
<td>0.0318 ± 0.0387</td>
<td>0.0317 ± 0.0387</td>
<td>0.0317 ± 0.0387</td>
</tr>
</tbody>
</table>

Table 6.3: Means and standard deviations of the Euclidean angular difference (in degrees) between the estimated refractive indices and the ground truth, across all the reported materials.

<table>
<thead>
<tr>
<th>Test shape</th>
<th>(L_3)</th>
<th>(L_4)</th>
<th>(L_5)</th>
<th>(L_2 + L_4)</th>
<th>(L_1 + L_5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dome</td>
<td>0.0459 ± 0.0362</td>
<td>0.0458 ± 0.0360</td>
<td>0.0460 ± 0.0362</td>
<td>0.0459 ± 0.0362</td>
<td>0.0459 ± 0.0362</td>
</tr>
<tr>
<td>Ridge</td>
<td>0.0462 ± 0.0366</td>
<td>0.0462 ± 0.0366</td>
<td>0.0462 ± 0.0366</td>
<td>0.0462 ± 0.0366</td>
<td>0.0462 ± 0.0366</td>
</tr>
<tr>
<td>Torus</td>
<td>0.0462 ± 0.0366</td>
<td>0.0462 ± 0.0366</td>
<td>0.0462 ± 0.0366</td>
<td>0.0462 ± 0.0366</td>
<td>0.0462 ± 0.0366</td>
</tr>
<tr>
<td>Test shape</td>
<td>0.0770 ± 0.0563</td>
<td>0.0729 ± 0.0555</td>
<td>0.0660 ± 0.0504</td>
<td>0.0770 ± 0.0563</td>
<td>0.0770 ± 0.0563</td>
</tr>
<tr>
<td>Mozart</td>
<td>0.0468 ± 0.0371</td>
<td>0.0469 ± 0.0374</td>
<td>0.0461 ± 0.0367</td>
<td>0.0468 ± 0.0371</td>
<td>0.0468 ± 0.0371</td>
</tr>
<tr>
<td>Vase</td>
<td>0.0542 ± 0.0411</td>
<td>0.0535 ± 0.0409</td>
<td>0.0499 ± 0.0390</td>
<td>0.0542 ± 0.0411</td>
<td>0.0542 ± 0.0411</td>
</tr>
<tr>
<td>Duck</td>
<td>0.1334 ± 0.0555</td>
<td>0.1330 ± 0.0606</td>
<td>0.1314 ± 0.0585</td>
<td>0.1334 ± 0.0555</td>
<td>0.1334 ± 0.0555</td>
</tr>
<tr>
<td>Tea pot</td>
<td>0.0497 ± 0.0392</td>
<td>0.0497 ± 0.0392</td>
<td>0.0497 ± 0.0391</td>
<td>0.0497 ± 0.0392</td>
<td>0.0497 ± 0.0392</td>
</tr>
</tbody>
</table>

Table 6.4: Means and standard deviations of the mean squared error between the estimated absolute values of the refractive indices and the ground truth, across all the reported materials.
estimated refractive index spectra over all the pixels in an image and the ground truth refractive index used for synthesizing that image. Similarly, in Table 6.4, we report the mean-squared error of the estimated refractive index as compared to the ground-truth. The angular error measure indicates the level of similarity between the recovered refractive index spectrum and the ground truth in terms of the shape of the spectrum. The low angular errors reported in Table 6.3, peaking at 0.0551 degrees, show that we can recover the refractive index with a high level of accuracy. In addition, the mean-squared error of the absolute refractive index peaks at 0.1334 for the Duck shape and does not exceed 0.077 for the other shapes. Moreover, the refractive index estimate is largely consistent across all the illumination directions for each shape. In fact, the change in lighting direction hardly affects the resulting refractive index. This observation is consistent with the shape recovery results, demonstrating that the recovered refractive index is robust against changes in illumination conditions.

6.4.2 Real-world Imagery

In this section, we report experimental results on real-world multispectral images acquired in-house. To acquire the imagery, we have used a benchtop hyperspectral camera equipped with an Acousto-Optical Tunable Filter (AOTF) which selects wavelengths in the range of 450 – 650 nm at a spectral resolution of 10 nm. By applying radio-frequency acoustic waves to compress and relax an optically anisotropic crystal [38, 61, 66], the incoming light is diffracted into an ordinary and an extraordinary component polarised in different directions. In our camera system, the filter only transmits the ordinary ray while blocking the extraordinary ray using total internal reflection. By rotating the camera box about its optical axis, we effectively selected the plane of polarisation of the polarised light at each wavelength.

We acquired multispectral images of five different objects made of matte plastic and porcelain. To measure polarisation, we captured the image of these objects when the polariser’s transmission axis is oriented at each of seven different angles given by 45, 60, 75, 90, 105, 120 and 135 degrees in the clock-wise direction with respect to the vertical axis of the camera. The imagery was captured using two unpolarised artificial sunlights as illuminants. These simultaneously illuminated the left and right hand side
of the objects under study. Note that the algorithm does not require prior knowledge of
the illuminant power spectrum and direction. Here, we used two illuminants to ensure
that there were no dark shadows in the captured images. To present rendering results,
as presented later in the section, we estimated the material reflectance from unpolarised
multispectral images of the same objects captured under an illuminant direction aligned
to the viewing direction.

Recall that our method delivers the surface orientation and material index of refraction
from input polarisation images. For the recovery of the reflectance, we assume that
the light penetrating the object’s sub-surface does not scatter very far and therefore be-
ing emitted at almost the same place as the incident point. Under this assumption,
the incidence and reflection angles can be assumed to be the same. This is impor-
tant since it allows the use of the illuminant power spectrum and the Wolff reflectance
model [181] to estimate the material reflectance (albedo), after the surface orientation
and index of refraction have been recovered by our algorithm.

Figure 6.13 shows sample input images of the five objects under study, which we
name hereafter Bear, Statue, Pig, Dinosaur and Pine Tree. The left-hand panels show
the trichromatic pseudocolour of the input multispectral images, simulated with the
Stiles and Burch colour matching functions [157]. These correspond to the images of
the objects with the polariser’s transmission axis forming an angle of 45° with respect
to the vertical axis of the camera. The second and third columns show the recovered
needle-maps and shading maps of the objects in the first column. Note that the nee-
dle maps show clear overall surface contours, capturing the surface normal orientation
along occlusion boundaries. The shading maps are intuitively correct, with minor dis-
tortions at the object material boundaries. This is due to the fact that our method is
based solely on polarisation information. Therefore, the changes at material bound-
daries have been interpreted by our algorithm as a variation in object geometry.

Having obtained the surface orientation and material properties of the objects un-
der study, we generate novel views of the original objects under novel illuminants.
For such a purpose, we reconstructed the surface height from surface orientation us-
ing the integration method introduced by Frankot and Chellappa [54]. Subsequently,
we rendered the trichromatic texture from the estimated material refractive index and
Figure 6.13: Shape estimation and rendering for real-world images. First column: The 45° polarisation component images rendered in trichromatic pseudo-colours. Second column: Needle maps. Third column: Shading maps. Fourth and fifth columns: Two-view stereo rendering of the input shapes under $L_3$. In the fourth row, the view direction is $10^\circ$ to the left of the original view. In the fifth row, the view direction forms an angle of $10^\circ$ to the right of the original view.
6.4. EXPERIMENTS

Figure 6.14: A novel view of the objects shown in Figure 6.13, rendered under the light directions $L_3$ (first row) and $L_2 + L_4$ (second row).

reflectance, using the rendering equation of the Wolff reflectance model [181]. This texture was then mapped onto the recovered three-dimensional surfaces. For validation, we captured the ground truth images of these objects under incandescent lights with directions coplanar to the horizontal axis of the image plane. Specifically, from left to right, these light directions point towards the illuminated object, forming angles of $-45$, $-30$, $0$, $30$ and $45$ degrees with the viewing direction, where, as before, a negative angle means the light is located on the left-hand side of the camera and vice versa. We denote these light directions $L_1$, $L_2$, $L_3$, $L_4$ and $L_5$, respectively, so as to be consistent with the synthetic dataset.

In the fourth and fifth columns of Figure 6.13, we show a pair of stereo images of the original objects rendered under the frontal illumination direction $L_3$. The left and right images were generated for two view directions forming $10^\circ$ to the left and right of the original view direction. Indeed, the novel views reveal the object shapes near the left and right occlusion boundaries, with applications to three-dimensional visualisation of objects using single-view images. Figure 6.14 shows the rendering of the original objects from another viewing direction, under the frontal illuminant $L_3$ and the two simultaneous illuminants $L_2 + L_4$. The three-dimensional structure of the shapes captured by the views in Figure 6.14 are perceptually consistent with the stereo pairs in Figure 6.13.
Figure 6.15: First and third columns: Pseudo trichromatic colour rendering of real-world objects from the frontal view under $L_3$ and $L_2 + L_4$, respectively. Second and fourth columns: The trichromatic colour rendering of ground-truth unpolarised multispectral images corresponding to the lighting conditions in the first and third columns.
In Figure 6.15, we present the rendering results for the above objects illuminated by novel incandescent illuminants with different illumination directions and viewed from the original position. In the first and third columns, we show the images rendered under the illuminants $L_3$ and $L_2 + L_4$. The second and fourth columns show the ground truth images captured under the same lighting condition as those in the first and third columns. The rendering under the frontal light $L_3$ shows consistency in shapes and shading with the ground truth. For the two-illuminant setting $L_2 + L_4$, the rendering results show irregularities at material boundaries as material changes are interpreted as geometric changes by polarisation information. Apart from these exceptions, the surface shading is consistent with the ground truth elsewhere.

We proceed to provide a quantitative analysis for the rendering results shown in Figure 6.15. In Table 6.5, we show the rendering accuracy under the novel lighting directions. The error is measured as the angular difference between the rendered image reflectance spectra and their ground-truth on a per-pixel basis. It is worth stressing that, our error measures are obtained from the multi-spectral images rather than the RGB values yielded by the color matching functions. Recall that we estimate the image reflectance spectra from the rendering equation of the Wolff reflectance model [181]. The numerical results reported in Table 6.5 are the mean and standard deviation across pixels in each image. The results here are consistent with the qualitative results presented previously in the sense that the rendering quality is better for the cases of the frontal illuminant and the two simultaneous illuminants. This is due to the fact that in such conditions, the objects are fully illuminated and the rendered image has a smooth shading variation. On the other hand, the lower rendering accuracy for the oblique illuminant directions is due to non-smooth shading where shadows occur across material boundaries. Nonetheless, our method can produce rendering results that are in good accordance with the ground truth. In addition, it delivers shading maps which conform well to the geometry of the object under study, based solely on polarisation information. This is an important characteristic of our method since it does not employ shading or chrominance as a cue.

Finally, in Figure 6.16, we plot the refractive index spectra of a number of regions in the real-world images. The first column shows the $45^\circ$ polarisation component ren-
Figure 6.16: Refractive index spectra of a number of selected regions in the input images. First column: The 45° polarisation component images rendered in pseudocolour, with the selected regions indicated by rectangular bounding boxes. Second to fifth columns: The mean and standard deviation of the estimated refractive index spectra over the pixels in each selected region, plotted in colours matching those of the region boundaries depicted in the first column.
6.5. CONCLUDING REMARKS

<table>
<thead>
<tr>
<th></th>
<th>$L_3$</th>
<th>$L_4$</th>
<th>$L_5$</th>
<th>$L_1 + L_5$</th>
<th>$L_2 + L_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bear</td>
<td>11.63 ± 2.90</td>
<td>12.22 ± 3.67</td>
<td>12.94 ± 5.95</td>
<td>12.65 ± 4.71</td>
<td>11.48 ± 3.17</td>
</tr>
<tr>
<td>Statue</td>
<td>12.32 ± 3.43</td>
<td>14.11 ± 3.24</td>
<td>14.18 ± 4.02</td>
<td>15.75 ± 4.03</td>
<td>13.46 ± 2.90</td>
</tr>
<tr>
<td>Pig</td>
<td>10.70 ± 3.40</td>
<td>11.78 ± 3.47</td>
<td>12.94 ± 4.14</td>
<td>12.87 ± 4.43</td>
<td>10.53 ± 2.91</td>
</tr>
<tr>
<td>Dinosaur</td>
<td>10.67 ± 3.76</td>
<td>12.19 ± 6.94</td>
<td>14.01 ± 8.15</td>
<td>9.02 ± 3.94</td>
<td>8.27 ± 3.60</td>
</tr>
<tr>
<td>Pine Tree</td>
<td>10.82 ± 2.69</td>
<td>11.33 ± 3.51</td>
<td>14.92 ± 4.81</td>
<td>13.05 ± 3.94</td>
<td>10.35 ± 3.99</td>
</tr>
</tbody>
</table>

Table 6.5: The angular deviation (in degrees) between the spectral reflectance images rendered for the frontal viewing direction and the ground truth images. The mean and standard deviation of these errors across pixels are reported for each image.

Table 6.5 shows the angular deviation (in degrees) between the spectral reflectance images rendered in pseudocolour, with the selected regions indicated by rectangular bounding boxes. The remaining columns show the mean and standard deviation of the estimated refractive index for each wavelength, computed across the pixels in each selected region. Note that the error plots are shown in colours that match those of the region bounding boxes overlaid on the images in first column. Indeed, in Figure 6.16, we notice that Region 2 of the Bear, Regions 1 and 4 of the Dinosaur and Region 4 of the Pine Tree, which are all made of white-painted plastic materials, correspond to refractive index spectra with very similar shapes. This observation demonstrates the consistency of the refractive index recovered by our method for similar materials across a variety of object shapes and surface orientations.

6.5 Concluding Remarks

This chapter has focused on shape and refractive index recovery based on the polarisation information in a multispectral or hyperspectral image acquired from a single-view. Departing from the theory of polarisation in the electromagnetic spectrum, the method hinges on the analysis of polarisation upon diffuse reflection from dielectric surfaces. By capturing the polarisation of light emitted from these surfaces, we observe the variation of the spectral radiance with respect to the polariser angle along a Transmitted Radiance Sinusoidal (TRS) curve. The recovery process commences with the decomposition of the input imagery into the components of the TRS, i.e. its phase, minimal
and maximal radiance values. We have provided a link between the azimuth angle of surface normal and the phase of polarisation and have shown how to disambiguate the azimuth angle of surface normals from the wavelength-indexed spectrum of the phase angles. We have also drawn upon the Fresnel transmission ratio between the minimal and maximal radiance on the TRS to jointly recover the zenith angle of the surface normal and the refractive index. To make the recovery problem well-posed, we have enforced integrability on the surface and employed the material dispersion as a constraint on the refractive index. We solve the estimation problem using an iterative optimisation approach and derive a closed-form solution for the zenith angle and the refractive index in each iteration. Lastly, we have demonstrated the merit of our method for the purpose of shape and refractive index recovery from synthetic and real-world imagery. The experimental results demonstrate the utility of the method for applications involving non-contact measurement of refractive indices of dielectrics.
Chapter 7

Conclusion

Spectral imaging is an emerging and promising trend in Computer Vision. In this thesis, we have focussed on the issues of spectrum representation, photometric invariants and shape recovery in spectral imaging. The work presented here offers potential applications in a variety of fields including, but not limited to, surface inspection, material analysis, biometrics, bio-security and agriculture such as early pathological diagnosis of skin and plant leaves, visualisation and Computational Photography, e.g. colour reproduction, photorealistic rendering, illumination and material editing.

In Section 7.1, we summarise the core contributions of the work presented in this thesis. Subsequently, we discuss future research directions and suggest improvements to the current work in Section 7.2.

7.1 Summary of Contributions

The theoretical foundation of the thesis commenced with the preliminaries on spectral imagery in Chapter 3. We introduced spectral imagery as a means of capturing physical properties of the scene, including illuminant power spectra, the reflection geometry, material reflectance and microscopic properties of the surface. We noted that these properties exist independently of the colour captured by trichromatic sensors and perceived by the human vision system. As such, we presented an image formation model of spectral imagery independent of the colour responses recorded by trichromatic sensors. Through this image formation model, we have shown how the spectral
distribution of the irradiance image of the scene accounts for the colours interpreted by trichromatic sensors. The theoretical framework presented in Chapter 3 has been applied throughout the thesis to generate trichromatic images as both experimental data and colour visualisation of spectral images.

With this fundamental understanding of spectral imagery, we proceeded to make the following contributions to the literature

- In Chapter 4, we presented a compact spectrum representation for efficient recognition and storage and accurate reconstruction of spectral imagery.

- In Chapter 5, we presented an approach to the recovery of the illuminant power spectrum, spectral reflectance, shading factor and specular coefficient from a single spectral image.

- In Chapter 6, we provided a framework for the simultaneous recovery of shape and refractive index from spectro-polarimetric images captured from a single view.

The specific details of these contributions are addressed as follows.

7.1.1 Spectrum Representation

In Chapter 4, we formulated a novel representation of reflectance spectra based on the control points of B-Spline curves interpolated to the discrete wavelength-indexed samples of the input spectra. Departing from the initial interpolating curve, we employed a knot removal scheme in the parameter domain to retain important characteristic knots and control points while removing the others. This procedure arrives at a stable representation robust to local perturbation and noise in the spectra and effective for reconstruction. With this representation, we performed segmentation and recognition using the control points of the interpolating B-Spline as discriminative features.

This representation confers a number of advantages. Firstly, it is compact, yet sufficient for reconstruction based on a cost function that balances the reconstruction error with the representation length. Secondly, due to the local support property of splines, it allows accurate reconstruction of continuous free-form spectra and is readily
applicable to any spectral photometric invariants. Thirdly, the representation presented here provides a set of common basis functions to compare discrete spectral samples with different spectral resolutions and lengths, which may have been acquired using heterogeneous devices. Finally, since the representation permits the reconstruction of a continuous spectrum from the discrete samples, a number of closed-form analytical operations such as derivative analysis, which were impossible to perform on discrete spectral data, are now readily applicable to the resulting continuous spectra.

More importantly, the representation does not only apply to individual spectra but can be extended to multispectral and hyperspectral image cubes. This is effected through the enforcement of a common knot vectors and common B-Spline basis functions for the spectra at all the pixels in the input image. As a result of this step, the discrimination between spectra are attributed to the control points of the resulting B-Spline curves, which are used by recognition tasks as the sole discriminative features. Finally, obtaining a common knot vector and a common set of basis functions for all the pixels in a spectral image is the key to a compact representation of spectral images with a large number of pixels.

Our representation also offers advantages over other low-dimensional representations of spectral reflectance. Most of the alternatives employ a linear mixture of basis functions, some of which are principal components obtained directly from Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), or Gaussian or wavelet basis functions. Despite being effective for recognition, these basis functions may not be suitable for accurately representing and reconstructing continuous spectra from discrete spectral values. On the other hand, since B-Spline basis functions possess the local support property, they allow a more accurate reconstruction that conforms to the spectral variation of the original discrete spectral samples.

7.1.2 Reflection Parameter Recovery

In Chapter 5, we addressed the problem of recovering reflection parameters from a single spectral image. To this end, we modelled the image irradiance making use of the dichromatic reflection model [147]. The method aims to simultaneously recover the illuminant power spectrum, the material reflectance spectrum, the shading factor
and the specular coefficient at every surface location.

The recovery problem is cast as an optimisation framework based on an objective function which consists of a data error term evaluated with respect to the dichromatic model and a regulariser to quantify the spatial smoothness of the scene geometry. The objective function can incorporate alternative spatial smoothness regularisers, which are based on robust kernels of the shading factor, surface curvatures or shape index.

To solve the recovery problem, we first selected uniform-reflectance image patches that meet the dichromatic criteria, i.e. exhibit both diffuse and specular reflection. Subsequently, we described an iterative optimisation procedure to optimise the above objective function on the selected patches. In each iteration, we alternated the target subsets of variables to optimised between the illuminant power spectrum and the other parameters of the dichromatic model. Following this procedure, we expressed the objective function as a convex (quadratic) function with respect to each variable subset so as to derive optimal closed-form solutions to the target variables in each optimisation step. Once the illumination power spectrum is obtained as a result of this iterative process, we recovered the reflectance, shading factor and specularity at each pixel, assuming that its local spatial neighbourhood is a dichromatic image patch with uniform-reflectance.

To our knowledge, the work presented in this chapter is one of the first on spectral photometric invariants and is readily extensible to trichromatic images. The optimisation framework generalises the idea of intersecting the dichromatic planes representing different materials in the scene, which was proposed by Finlayson and Schaefer for trichromatic images [49], by taking the scene geometry into account. Due to its ability to operate on a single image, the method has implications in real-time image editing such as re-illumination, material editing and colour balancing. Furthermore, the recovered spectral reflectance, in conjunction with the spectral representation we devised in Chapter 4, may benefit material recognition and scene analysis tasks.

### 7.1.3 Shape and Refractive Index from Polarisation

Following the work in Chapter 5, we further extended our study to the recovery of photometric invariants and shape from spectro-polarimetric images acquired from a
7.1. SUMMARY OF CONTRIBUTIONS

single-view in Chapter 6. We proposed a recovery framework that hinges on a number of physics laws including the Fresnel reflection and refraction theory, the polarisation of light upon reflection and refraction and the dispersion of material refractive index. Our study focuses on the analysis of polarisation upon diffuse reflection from dielectric surfaces. To the best of our knowledge, the work in this Chapter is the first one in Computer Vision addressing the estimation of both shape and refractive index from spectro-polarimetric imagery. Our experimental results are encouraging and can have potential applications in visualisation, three-dimensional surface rendering and non-contact measurement of material refractive index.

Moreover, our recovery method offers a number of advantages. Firstly, it permits the recovery of material refractive index without the need for accurate calibration of optical bench measurements. In fact, this information is an important physical property which can be used to quantify subsurface structure. Furthermore, our method only requires the input imagery to be captured from a single view, thus avoiding the calibration of multiple acquisition devices. More importantly, the estimation results are independent of the illumination power spectrum and direction since these factors are cancelled in the Fresnel transmission ratio. Finally, our method does not assume any knowledge of the material under study other than the dispersion equation. As a result, real-time applications may favour our approach over alternatives that require either accurate measurements of shape or refractive index or complicated instrument setups.

In Chapter 6, we presented the recovery of shape as that of the azimuth angle and zenith angle of surface normals. To estimate the azimuth angle, we commenced with the observation that the variation of the scene radiance with respect to the polariser angle along a Transmitted Radiance Sinusoidal (TRS) curve. Next, we presented a linear system to decompose the polarimetric imagery into the phase, minimal and maximal radiance values of the TRS. Subsequently, we related the azimuth angle to the phase of polarisation and showed how to recover and disambiguate the azimuth angle from the wavelength-indexed spectrum of the phase angles.

In addition, we proposed an optimisation framework to simultaneously recover the zenith angle and refractive index. The optimisation method aims to minimise the residuals of the Fresnel transmission ratio, i.e. the ratio of the minimal to the maximal
radiance on the TRS, with respect to a function relating the zenith angle and refractive index. To render the optimisation problem well-posed, we introduced additional constraints based on surface integrability and the dispersion of material refractive index. We then approached the problem with an iterative procedure which yields a closed-form solution to the zenith angle and the refractive index in each iteration.

7.2 Future Work and Discussion

In the future, it is desirable to consider a representation of both the spatial and spectral domains of multispectral and hyperspectral images. This is a step forward from the spectrum representation presented in Chapter 4. In fact, the spatial variation of spectral signatures conveys information of surface regularity, material distribution and texture, that would be useful for scene description and recognition purposes.

Since the main aim of our spectral representation is to reconstruct the original shape of the sampled spectrum, our optimisation approach employed an objective function based on a reconstruction error. However, for the classification of spectra, one could consider the discriminative power of a spectral image descriptor for different kinds of materials. The descriptor can then be derived as an optimal one that maximises both the inter-class discrimination and the intra-class similarity.

In addition, we note that the photometric invariance method in Chapter 5 assumes uniform illumination over the scene. In the future, we intend to extend the applicability of this work to scenes under complex lighting conditions, i.e. spatially-varying illumination, where the sources may consist of multiple lights with different power spectra, which illuminate the different locations of the scene with different mixture proportions of their power. In addition, our method assumes uniform reflectance over the observed image regions. This aims to reduce the number of variables and to render the optimisation problem well-posed. As a result, the method does not cope well with highly texture surfaces. In summary, a practical extension of the presented method would be designed to handle this kind of surfaces and spatially-varying illumination.

In Chapter 6, we focussed our study on diffuse reflection from dielectric surfaces. While this phenomenon is often due to the penetration, subsurface scattering and back-
refraction of light at the surface boundary, polarisation may occur as light is directly reflected at the air-material interface. The latter form of polarisation is called specular polarisation. At the current state of our work, we only deal with materials that diffusely polarise light. As a future development, we intend to develop a mixture model of these two different polarisation phenomena.

It is worth noting that there are limitations to the use of polarisation for shape and refractive index estimation. Polarisation information is a useful cue to object shape near the grazing angle between the viewing direction and the surface normal, where the surface orientation induces a strong degree of polarisation. However, toward the singular points of the surface, where weak polarisation occurs, polarisation is not a reliable source of information for determining the azimuth angle of surface normals and the refractive index. This is because the Transmitted Radiance Sinusoid (TRS) becomes nearly flat, resulting in unstable estimation of the phase angle due to phase shift. Therefore, the azimuth angle at these locations usually experiences abrupt changes in its orientation. Nonetheless, given a zero degree of polarisation, it is certain that the surface location under observation is a singular point, i.e. the surface normal points toward the view position. At these singular points, the refractive index becomes indeterminate since any arbitrary refractive index would satisfy the Fresnel transmission ratio equation, i.e. Equation 6.16. Since shape reconstruction is most reliable at the object boundary, our method can potentially combine occluding contours recovered from multiple views to enhance the quality of shape reconstruction.

Although the degree of diffuse polarisation is commonly weaker and more difficult to measure than that of the specular polarisation, our polarisation model is relevant and useful for a variety of multi-layered and translucent materials such as skin, glass and polymers. Our method has an advantage over shape-from-shading or photometric stereo approaches that may not be able to operate on these surface materials due to the lack of texture and shading information. For metals and highly reflective materials, specular polarisation becomes dominant, thus providing a valuable cue to shape and refractive index recovery.
Appendix A

List of Publications

The following publications have been published by the author during his Ph.D. candidature. A number of these are available on the author’s website at http://users.cecs.anu.edu.au/~huynh/


Bibliography


