Illumination Invariant Outdoor Perception

Rishi Ramakrishnan

A thesis submitted in fulfilment
of the requirements of the degree of
Doctor of Philosophy

Australian Centre for Field Robotics
School of Aerospace, Mechanical and Mechatronic Engineering
The University of Sydney

Submitted September 2015; revised March 2016

Declaration

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the University or other institute of higher learning, except where due acknowledgement has been made in the text.

Rishi Ramakrishnan

September 2015

Abstract

Rishi Ramakrishnan
The University of Sydney

Doctor of Philosophy
September 2015

Illumination Invariant Outdoor Perception

This thesis proposes the use of a multi-modal sensor approach to achieve illumination invariance in images taken in outdoor environments. The approach is automatic in that it does not require user input for initialisation, and is not reliant on the input of atmospheric radiative transfer models.

While it is common to use pixel colour and intensity as features in high level vision algorithms, their performance is severely limited by the uncontrolled lighting and complex geometric structure of outdoor scenes. The appearance of a material is dependent on the incident illumination, which can vary due to spatial and temporal factors. This variability causes identical materials to appear differently depending on their location.

Illumination invariant representations of the scene can potentially improve the performance of high level vision algorithms as they allow discrimination between pixels to occur based on the underlying material characteristics. The proposed approach to obtaining illumination invariance utilises fused image and geometric data. An approximation of the physical processes involved in outdoor illumination is used to derive scaling factors for each pixel. This has the effect of relighting the entire scene using a single illuminant that is common in terms of colour and intensity for all pixels. The approach is extended to the multi-image scenario through the use of an overlapping region, meaning that the resultant dataset is both spatially and temporally illumination invariant.

This thesis also proposes illumination invariant radiometric normalisation methods, where the aim is to obtain the reflectance spectra of materials in the scene. The

incident illumination is measured using in situ measurements and this is used in conjunction with an illumination model to estimate the per-pixel lighting conditions in order to normalise outdoor imagery. The advantage of this technique is that the reflectance estimates are less susceptible to variability due to the geometry of the scene. This can lead to performance gains for classification and clustering within shadowed regions.

The proposed illumination invariance approach is evaluated on several datasets and shows that spatial and temporal invariance can be achieved without loss of spectral dimensionality. The system requires very few tuning parameters, meaning that expert knowledge is not required in order for its operation. This has potential implications for robotics and remote sensing applications where perception systems play an integral role in developing a rich understanding of the scene.

Acknowledgements

I would like to thank my supervisors Juan Nieto and Steve Scheding for their guidance and advice throughout the past few years. Their support has been invaluable and has allowed me to complete this thesis.

Thank you to my parents and brother for your support and encouragement which has led me to where I am today.

I would also like to thank my friends in ACFR and at USYD; Alex, Andrew, Jeremy, Jono, Lloyd, Victor, Suchet, Tatsumi and Zach. I am indebted to Zach for registering countless datasets which has made my work much easier. Special thanks to Victor for the daily coffees, biscuits and conversations which have kept me going. I am also grateful to Lloyd, Sarah and Victor for their support during the final few weeks which is greatly appreciated. A big thanks is also extended to my undergraduate friends who motivated me to start and end (by consistently asking me when I was finished) this work.

Special acknowledgement must also be made to Richard, Javier, Bang, Ben and Stewart, whose work with the hyperspectral cameras and sensor vehicle have allowed me to gather all the datasets required. This work has been made possible by the Rio Tinto Centre for Mining Automation.

Contents

List of Figures

List of Tables

List of Algorithms

Nomenclature

Notations and Symbols

t	time
$\mathbf{x}$	point in $3D$ space
i	index of a planar patch in $3D$ space
$\mathbf{N}_x$	surface normal at point x
$\mathbf{N}_i$	surface normal at region i
A_i	area of region i
α_i	angle between surface normal and vector towards sun position
Γ_i	sky factor at region i
$\mathbf{L}$	radiance (function of wavelength)
$\mathbf{B}$	radiosity (function of wavelength)
$\mathbf{f}_r$	bidirectional reflectance distribution function (function of wavelength)
ρ	albedo (function of wavelength)
Ω_x	hemisphere positioned at x
$\mathbf{x} \rightarrow \boldsymbol{\Theta}$	vector leaving point $\mathbf{x}$ in the direction of $\boldsymbol{\Theta}$
$\mathbf{x} \leftarrow \boldsymbol{\Psi}$	vector incident on point $\mathbf{x}$ from the direction of $\boldsymbol{\Psi}$
$\boldsymbol{\Theta} \leftrightarrow \boldsymbol{\Psi}$	the direction vectors Θ and Ψ can be interchanged
$\mathbf{E}_{sun}$	extraterrestrial sunlight irradiance spectrum (function of wavelength)
$\boldsymbol{\tau}$	transmittance function (function of wavelength)
$\mathbf{E}_{sun}\boldsymbol{\tau}$	terrestrial sunlight irradiance spectrum (function of wavelength)
$\mathbf{E}_{sky}$	diffuse skylight irradiance spectrum (function of wavelength)
F_{ij}	form factor between regions i and j
V	line of sight visibility
$\cap$	intersection of two sets

Abbreviations

ACFR	Australian Centre for Field Robotics
BRDF	Bidirectional Reflectance Distribution Function
DIS	Downwelling Irradiance Sensor
EMD	Earth Mover's Distance

CG	Computer Graphics
CVIP	Computer Vision and Image Processing
GPS	Global Positioning System
LIDAR	Light Detection and Ranging *(also commonly known as 'ladar' or a 'laser range scanner')*
LBP	Loopy Belief Propagation
LI	Linear Interpolation
MRF	Markov Random Field
PSNR	Peak Signal-to-Noise Ratio
RANSAC	Random Sample Consensus
RGB	Red, Green and Blue
SAM	Spectral Angle Mapper
SNR	Signal-to-Noise Ratio
SSIM	Structural Similarity
SWIR	Short Wave Infra-Red
VNIR	Visible Near Infra-Red

Glossary of Terms

Calibration Board: Surface with known reflectance characteristics.

ColorChecker Board: Surface containing 24 coloured patches used for ground truth evaluation.

Channel:

Diffuse: Material that reflects light uniformly in all directions.

Diffuse Skylight: Extraterrestrial sunlight that has been scattered by the atmosphere.

Downwelling Irradiance Sensor: Device for measuring the incident illumination.

Extraterrestrial Sunlight: Light emitted by the sun, prior to passing through the atmosphere of the earth.

Irradiance: Amount of power incident upon a surface, per unit projected area.

Radiance: Amount of power incident upon (or emitted by) a surface, per unit solid angle and per unit projected area.

Radiometric Normalisation: The process of converting pixel intensity values to reflectance.

Sky Dome: Hemispherical light source through which diffuse skylight is modelled.

Sky Factor: Amount of the sky dome visible from a region.

Sun Angle: Angle between the surface normal of a region and the vector toward the sun position.

Terrestrial Sunlight: Extraterrestrial sunlight that is not scattered by the atmosphere and arrives at the surface of the earth.

Datasets

Table 1 – Summary of the datasets used for experimental results. Datasets were captured under different weather conditions, ranging from overcast to sparse clouds, and with both hyperspectral and standard RGB cameras. The LIDAR resolution is equal in both the azimuth and elevation directions.

					Invariance	
ID	**Location**	**Camera**	**LIDAR Res. (deg)**	**Weather**	**Spatial**	**Temporal**
1	ColorChecker	VNIR	-	sunny	-	-
2	Shale	VNIR	0.040	partly cloudy	✓	×
3	USYD Hall	SWIR	0.050	partly cloudy	✓	×
4	Seymour Centre	RGB	0.041	sparse clouds	✓	×
5	Seymour Centre	RGB	0.041	sparse clouds	✓	✓
6	Seymour Centre	RGB	0.041	overcast	✓	✓
7	USYD Hall	RGB	0.030	sparse clouds	✓	×
8	ACFR Lawns	RGB	-	sunny/overcast	✓	✓

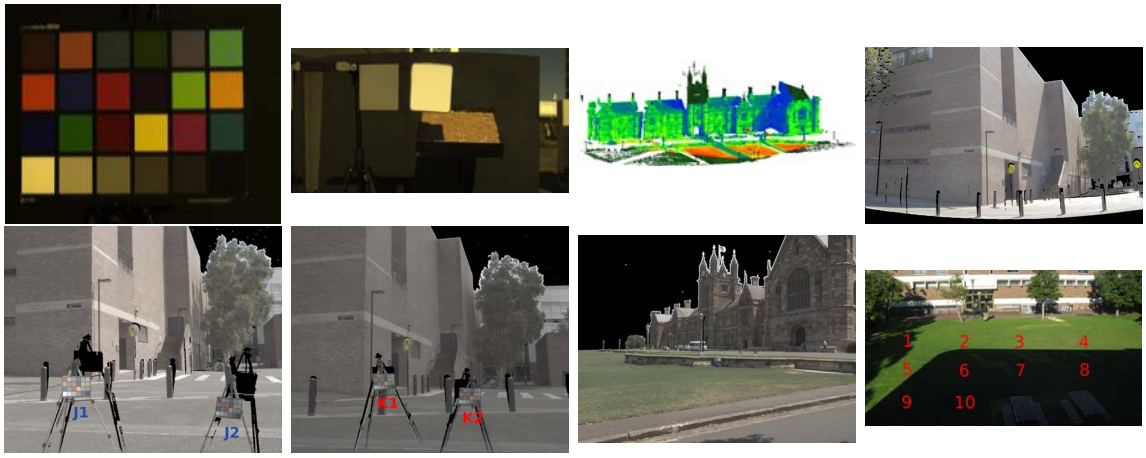


Figure 1 – Thumbnails of Datasets. *Top row, from left to right:* Datasets 1 to 4. *Bottom row, from left to right:* Datasets 5 to 8.

Chapter 1

Introduction

The aim of this thesis is to develop an automatic technique for achieving illumination invariance in images taken in the outdoor environment. In the uncontrolled lighting situations that exist in outdoor scenarios, the incident illumination is not constant at all regions in the scene. Illumination invariance aims to minimise these variations and this thesis presents a principled approach to the problem via the modelling of the physical processes involved, focusing on outdoor perception scenarios during daylight hours.

1.1 Motivation

The appearance of a scene is dictated by the process in which light interacts with the environment and is subsequently measured by a sensing device. Light is emitted by an illumination source with a characteristic colour and intensity depending on the properties of the source. As it travels through a scene, it is absorbed and re-emitted by the materials it interacts with. For example, it may be affected by the gaseous molecules that are in its path, or it may strike the surface of an object in the scene. In cases where the light reflects off an object, the colour and intensity of the light is altered by the characteristics of the material it encountered and the angle at which it struck. The reflected light may travel in the direction of the camera, or it may act

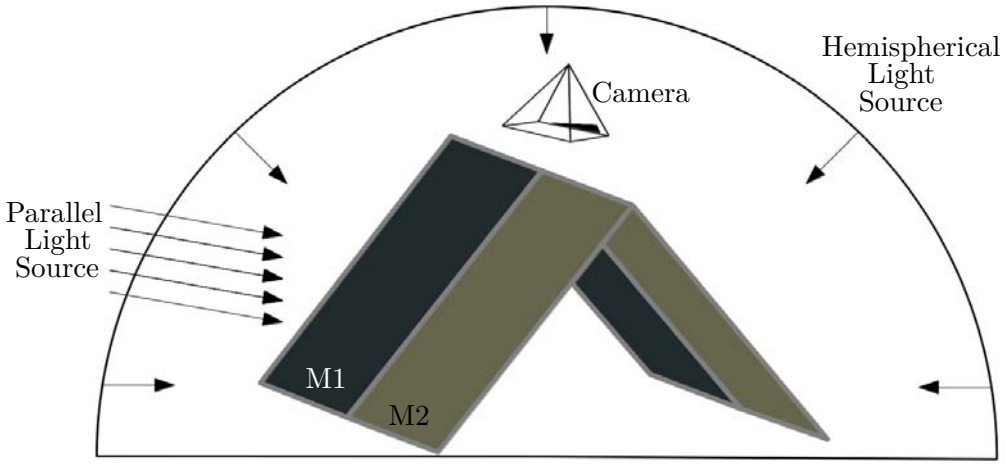


Figure 1.1 – In this example, two materials M1 (dark blue - 33,42,44) and M2 (dark green - 102,102,77) are illuminated by a parallel (orange - 255,170,85) and hemispherical (grey - 85,85,85) light source. RGB values indicating the colour of the illumination sources and material are shown in brackets.

as a secondary illumination source for other regions in the scene. The three main components affecting the appearance are therefore the incident illumination at each point in the scene, underlying material characteristics and the scene structure [57]. Altering each component has a significant impact on the measurement.

In outdoor scenarios, the complex geometry of the scene has the potential to occlude light sources and change the relative orientation of a region with respect to the light source. Coupled with the fact that there are multiple illumination sources [70] means that a material in the scene will be illuminated differently depending on its location in the scene.

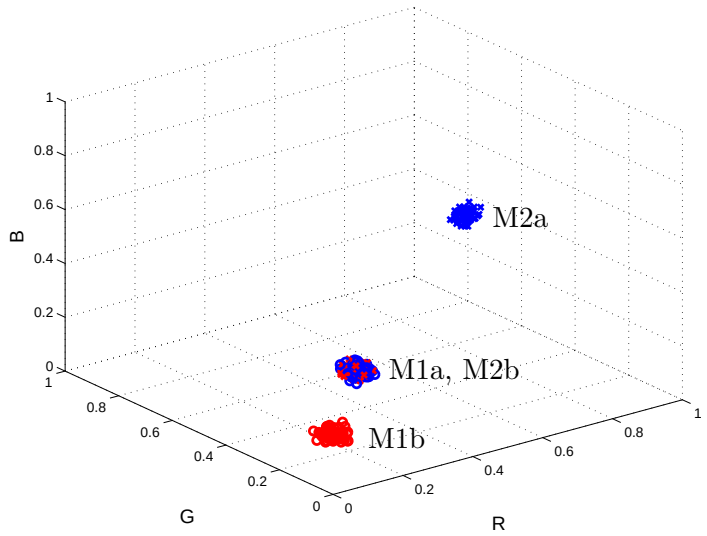
As a contrived example of the worst case scenario, consider a scene consisting of two uniform materials (M1 and M2) separated by a boundary as shown in Figure 1.1. This scene is illuminated by a parallel and hemispherical light source, with one side of the structure being occluded from the parallel source. A camera is positioned above the scene and the captured image is combined with additive Gaussian noise (mean: 0, standard deviation: 0.005) for visualisation purposes. The resultant image is shown in Figure 1.2. A number of interesting features are noticeable. Firstly, the appearance of M1 and M2 when exposed to different amounts of incident illumination (regions

Figure 1.2 – Resultant image when captured by a camera situated above the scene. The region $M2a$ changes colour as it is exposed to the orange parallel light source.

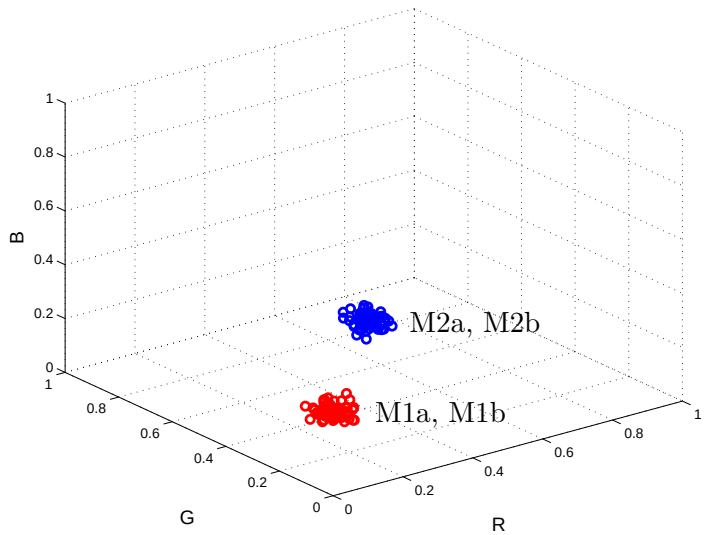
M1a and M2b), appear similar to one another. This is due to the fact that the measured intensity is dependent on both the incident illumination and underlying material reflectance [22]. Secondly, despite the fact that the scene consists of two materials, analysis of the colour intensities (Figure 1.3a) shows that three distinct clusters are present. This means that it is unlikely that classification and/or clustering algorithms will be able to uniquely identify the materials in the scene. The purpose of attaining illumination invariance in an image is to remove the dependence on incident illumination, leading to a much more discriminative feature space (Figure 1.3b). This allows algorithms such as support vector machines, linear classifiers and clustering methods to more easily separate the data in a manner that is more characteristic of underlying material properties.

Perception systems form an integral part of robotics and remote sensing platforms as they provide valuable input for high level algorithms such as localisation, mapping, object recognition and classification. In particular, visual perception has received widespread interest due to the rich information it provides. The decreasing cost of sensors and the desire for high resolution scene understanding means that modern sensor suites typically contain multiple sensor modalities. Examples of this include cameras, LIDARs and GPS, each of which provide a unique perspective of the environment.

These sensors are attractive due to their non-destructive nature of perceiving the



(a) Original image.



(b) Illumination invariant image.

Figure 1.3 – A plot of the RGB intensities shows three distinct clusters, with M1a and M2b being grouped together. An illumination invariant representation increases the separability of the data since the appearance is indicative of the underlying material reflectance characteristics.

environment. They allow perception to occur from afar, which is important as often the terrain could be inaccessible or hazardous. The recent development of multispectral and hyperspectral field based sensors has meant that remote sensing applications such as vegetation analysis and classification can utilise the discriminative nature of high spectral resolution data. Using a multiple sensor modality approach, such as

(a) 9:00am **(b)** 10:30am **(c)** 12:00pm **(d)** 2:00pm **(e)** 3:30pm

Figure 1.4 – Temporal illumination variations occur due to the change in position of the sun throughout the day, affecting the shadows and incident illumination angle at each pixel in the image. The varying atmospheric conditions also impact the appearance, with clouds attenuating the sunlight in the final image.

fusing imagery with geometric information from LIDARs, RADARs and stereo-vision cameras, can also allow 3D geological maps to be developed that assist in performing volumetric semantic classification [55] and mineralogy distribution analysis [41].

In order for high level algorithms to operate reliably, they must be robust to illumination variations in the scene. An example of this is in localisation algorithms, where ideally the same scene would appear identically at different times of day [13]. Spatially, the geometry of the scene induces visible artefacts such as shadows, where a region is occluded from a light source. A second spatial variation occurs due to the relative orientation between a region's surface normal and the light source direction. The angle affects the intensity of the incident illumination through a cosine weighting (Lambert's cosine law [59]). Temporally, two main factors affect the appearance of a scene. When the illumination sources are dynamic in nature, the shadows and relative angle to the source will change over time, significantly altering the appearance of a scene as seen in Figure 1.4. Also, the colour and intensity of an illumination source may change over time.

This is problematic as identical materials under different illumination conditions vary in appearance. An object in the scene will be perceived differently based on its location and the time of day under which it was observed. Consequently, the performance of algorithms such as clustering, segmentation and classification experience a degradation in performance when operating in uncontrolled lighting scenarios such as those in the outdoor environment [26]. While the human perception system is able to cater for illumination variations, it remains challenging for computer vision systems to do

the same.

1.2 Thesis Contributions

This thesis proposes an approach to reduce the influence of variable illumination in images taken in the outdoor environment. The complementary nature of geometry and imagery is exploited in order to model the influence of each illumination source independently of one another. An illumination invariant image is obtained by calculating scaling factors that allow the entire scene to be relit using a single, common illuminant.

With the addition of a direct measurement of the scene illumination, this thesis also proposes an illumination invariant radiometric normalisation technique, typically required for remote sensing purposes. This allows measurements in the field to be compared against laboratory obtained data [71] and is useful for supervised high level algorithms such as classification and change detection between scans taken on different days.

The specific contributions are:

- An approach to attaining illumination invariance in single image scenarios that can be initialised automatically. This is critical as its lack of dependency on user input means it is less susceptible to the introduction of user induced errors. The method retains the full spectral resolution of the data, which provides high level algorithms with extremely discriminative information.

- An approach to obtain temporally illumination invariant images for multi-image scenarios through the use of overlapping regions. This is useful for sensing platforms that may re-visit a location at different times, or may move throughout the scene during data collection. In this scenario, generating datasets that are consistent with one another is vital when comparing materials or objects between images taken at different times.

- An approximation method to the calculation of sky factors for all regions in the scene which targets applications where computational speed is a priority.

- An iterative sampling and smoothing technique for both sky factors and indirect illumination approximation. This method allows the outdoor illumination model to be accurately estimated, with the inclusion of indirect illumination potentially increasing the performance of applications such as un-mixing and classification.

- The final contribution is an approach for performing illumination invariant radiometric normalisation. This converts the pixel values captured by the camera and converts it to reflectance. By accounting for geometric variations in the scene, robust reflectance estimates can be provided for applications such as material classification.

1.3 Publications

The work in this thesis has led to the following publications:

- Rishi Ramakrishnan, Juan Nieto and Steve Scheding. Shadow Compensation for Outdoor Perception. In *International Conference on Robotics and Automation*, pages 4835-4842, IEEE, 2015. - This paper presents the spatial illumination invariance system proposed in Chapter 3 as well as a fast approximation to sky factors for robotics applications.

- Rishi Ramakrishnan, Juan Nieto and Steve Scheding. Illumination Invariance in Outdoor Environments. *International Journal of Computer Vision.* 2015. Under review - This paper included additional experiments for achieving spatial illumination invariance, the approach for temporal illumination invariance, automatic initialisation and Markov Random Field approach to generating a shadow map presented in Chapter 3.

- Rishi Ramakrishnan, Juan Nieto and Steve Scheding. Verification of Sky Models for Image Calibration. *Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on*, pages 907-914, IEEE, 2013. - This paper investigates modelling the sky colour through a number of parametric approaches typically used in the computer graphics community for rendering purposes. These models are then used in a preliminary application of inverse reflectometry for simple, diffusely reflecting objects in an outdoor environment.

1.4 Thesis Structure

This thesis is structured as follows:

Chapter 2 introduces the outdoor illumination model and presents an overview of the relevant literature. Illumination invariance can be reformulated in a number of ways, depending on which research community it is approached from. The three fields focused on are (i) Computer Vision and Image Processing (CVIP), which utilises purely image based data and attempts to find an image space that removes the influence of illumination; (ii) Computer Graphics (CG), which utilises the scene structure, final image and illumination information to estimate the material reflectance in a process known as inverse reflectometry; and (iii) Remote Sensing, which can use either single or multiple sensor modalities to estimate the underlying reflectance of the scene (radiometric normalisation).

Chapter 3 proposes an illumination invariant perception system. This method uses a multi-modal sensor suite to perceive the environment and takes into account variations due to both spatial and temporal factors. The automatic initialisation and sky factor approximation methods are also presented here.

Chapter 4 extends this technique to illumination invariant radiometric normalisation with a focus on remote sensing applications. Approximation methods for both

sky factors and indirect illumination in large scale scenes are presented. The normalisation procedure is derived with and without the inclusion of indirect illumination, which for some applications may be necessary.

Chapter 5 presents the results of extensive evaluation of the illumination invariant perception systems developed in this thesis. The results and their implications are analysed through the use of ground truth data.

Chapter 6 discusses the conclusion and future work involved in this thesis, as well as the open problems that are left to be solved.

Chapter 2

Background

This chapter presents the outdoor illumination model in Section 2.1, which describes the measurement process through which the appearance of a scene is captured. In doing so, the causes of illumination variation in a scene can be understood. This is followed by a review of the literature involved with the development of illumination invariant representations of images.

2.1 Outdoor Illumination Model

The relationship between material, illumination and the imaging sensor must first be described in order to understand the root causes of illumination variation, and identify which factors are accounted for in the literature. In this section, an illumination model for the outdoor environment is derived from the rendering equation commonly used in the computer graphics community. This technique allows each illumination source to be treated independently.

2.1.1 Derivation

The fields of computer graphics, robotics and remote sensing are linked through their generation/use of images. While computational graphical modelling develops an

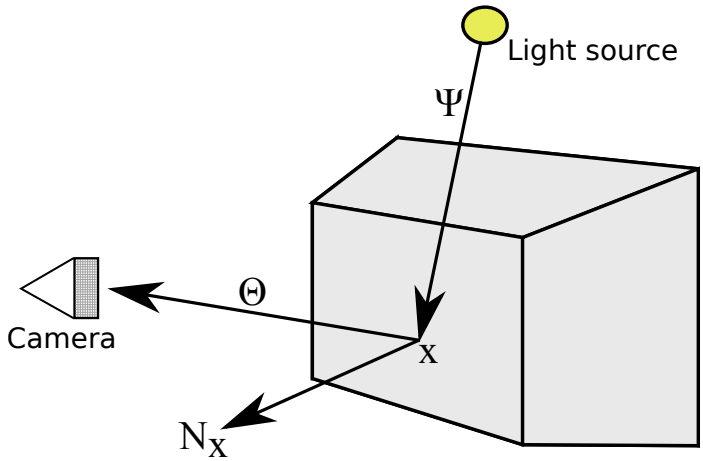


Figure 2.1 – A simple scene layout consists of an object on which the point x lies. Incident illumination arrives from the light source in the direction Ψ and is reflected depending on the material characteristics of the surface. The reflected light leaves x in the direction Θ, which is towards the camera.

image using the scene and illumination structure, both robotics and remote sensing applications use the image to feed information into low and high level algorithms. The common factor between these fields is the concept of radiance, which is the amount of power incident upon (or emitted by) a surface, per unit solid angle and per unit projected area (W/sr/m^2/nm). It is of primary importance as image intensities prior to post-processing are proportional to radiance [17].

The radiance L of a point x in the scene, at wavelength λ, in the direction of the camera Θ is a function of the Bidirectional Reflectance Distribution Function (BRDF) f_r, emitted radiance L_e and incident irradiance [17]:

$$L(x \to \Theta, \lambda) = L_e(x \to \Theta, \lambda) + \int_{\Omega_x} f_r(x, \Psi \leftrightarrow \Theta, \lambda) L(x \leftarrow \Psi, \lambda) \cos(N_x, \Psi) d\omega_\Psi, \tag{2.1}$$

where Ω_x is the hemisphere above point x oriented in accordance with the normal N_x and Ψ is the incident illumination direction. The notation $x \to \Theta$ indicates a vector leaving in direction Θ from point x, while $x \leftarrow \Psi$ indicates a vector incident upon x with direction Ψ. The symbol $\leftrightarrow$ indicates that the direction vectors can be interchanged and is due to the reciprocity principle of the BRDF [17]. Emitted radiance occur due to the conversion of energy into light and can commonly occur when

objects are heated and begin to glow a certain colour depending on the temperature.

This formulation is the rendering equation used in the computer graphics community and can be approximated in a number of ways in order to generate realistic looking images. Radiosity based approaches were first introduced in [33] and originally modelled heat transfer between surfaces. This method assumes diffuse reflections from all surfaces within the scene and accounts for inter-reflections, meaning it is classed as a global illumination technique. The derivation of the radiosity approximation is reproduced from [17] and is used to form the outdoor illumination model. The primary quantity of concern when using this method is radiosity, which is the power arriving or leaving per unit area per nanometre (W/m^2/nm).

Diffuse Reflection Assumption of the World

The BRDF is a multi-dimensional function describing the reflectance characteristics of a material. It is dependent on the incoming and outgoing light directions, both of which are parameterised by their azimuth and zenith angles with respect to the local coordinate system where the z-axis is aligned with the surface normal. Assuming that all regions in the scene are composed of a diffusely reflecting material (uniformly reflects light in all directions), the BRDF becomes independent of the incoming and outgoing light directions and can be simplified using a Lambertian shading model:

$$f_r(x, \Theta \leftrightarrow \Psi, \lambda) = \frac{\rho(x, \lambda)}{\pi}, \tag{2.2}$$

where $\rho(x, \lambda)$ is the albedo of the material (proportion of light reflected by the surface) at point x ranging from 0 to 1, and the division by π is used for normalisation of the function. Essentially, the diffuse assumption means that the surface reflects light uniformly on the hemisphere and will appear identical, regardless of the viewing angle.

The radiance at point x is also independent of the outgoing light direction, so Equation (2.1) becomes:

$$L(x, \lambda) = L_e(x, \lambda) + \rho(x, \lambda) \int_{\Omega_x} \frac{L(x \leftarrow \Psi, \lambda) \cos(N_x, \Psi)}{\pi} d\omega_\Psi. \tag{2.3}$$

The domain of the integral is transformed from being over the hemisphere, to an integral over all surfaces S in the scene:

$$L(x,\lambda) = L_e(x,\lambda) + \frac{\rho(x,\lambda)}{A_i}\int_{S_i}\int_{S_j}\frac{L(y,\lambda)\cos(\Psi,N_x)\cos(-\Psi,N_y)}{\pi r_{xy}^2}V(x,y)dA_y dA_x, \tag{2.4}$$

where A_i is the area of surface i, r_{xy} is the distance between point x and y on surfaces i and j respectively, and $V(x,y)$ is the binary visibility function.

Scene Discretisation

Discretising the scene into small homogeneous patches and multiplying by π to convert radiance into radiosity B, means Equation (2.4) can be expressed in a discrete form:

$$B_i(\lambda) = B_{e,i}(\lambda) + \rho_i(\lambda)\sum_{j=1}^{N} F_{ij}E_j(\lambda), \tag{2.5}$$

where E_j is the incident irradiance from surface j. The terms B and E both have the same units, but B will be used to indicate radiosity leaving the region of interest and E will indicate illumination sources from other regions. Here, the double integral is incorporated into a purely geometrical term known as the form factor F_{ij}:

$$F_{ij} = \frac{1}{A_i}\int_{S_i}\int_{S_j}\frac{\cos(\Psi,N_x)\cos(-\Psi,N_y)}{\pi r_{xy}^2}V(x,y)dA_y dA_x. \tag{2.6}$$

Intuitively, the form factor expresses the energy transfer between two surfaces [11].

Illumination Source Modelling

In the outdoor environment, there are several sources of illumination that need to be considered when analysing visual data. The most dominant sources are terrestrial sunlight and diffuse skylight [32, 70]. These sources vary in their colour and intensity, and must be considered separately from one another in order to develop the outdoor illumination model. Figure 2.2 presents the model used in this thesis and shows the

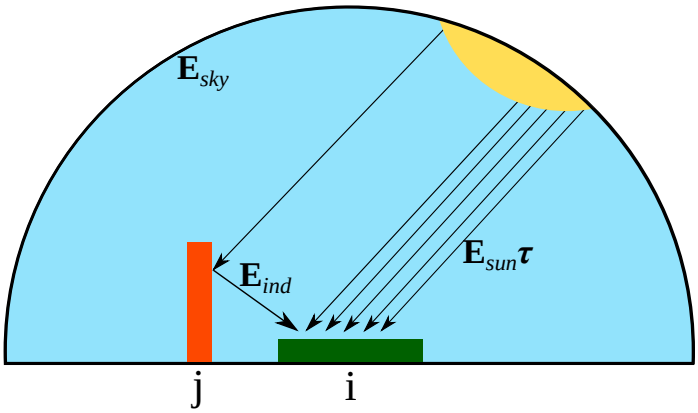


Figure 2.2 – The outdoor illumination model consists of terrestrial sunlight $\mathbf{E}_{sun}\boldsymbol{\tau}$, diffuse skylight $\mathbf{E}_{sky}$ and indirect illumination $\mathbf{E}_{ind}$.

three sources being considered; terrestrial sunlight $\mathbf{E}_{sun}\boldsymbol{\tau}$, diffuse skylight $\mathbf{E}_{sky}$ and indirect illumination $\mathbf{E}_{ind}$, all of which are parameterised by wavelength λ. Splitting Equation (2.5) into the sum of the individual light source components gives:

$$B_i(\lambda) = B_{e,i}(\lambda) + \rho_i(\lambda)\left[F_{i,sun}E_{sun}(\lambda)\tau(\lambda) + F_{i,sky}E_{sky}(\lambda) + \sum_{j=1}^{N} F_{ij}E_j(\lambda)\right], \tag{2.7}$$

which allows the influence of each source to be analysed. The derivation of the form factors for each illumination component is reproduced from [12].

Extraterrestrial sunlight $\mathbf{E}_{sun}$ consists of light emitted from the sun in a spherical pattern. Due to the large distances involved, by the time this light reaches the earth it can be treated as a parallel light source [12, 71]. As the light travels from the top of the atmosphere to the ground, it is absorbed at particular wavelengths based on the atmospheric conditions. The fraction of light that reaches the surface of the earth is known as the solar path transmittance $\boldsymbol{\tau}$ [71] and an example of this is shown in Figure 2.3. The light striking the earth is known as terrestrial sunlight and is found by multiplying the $\mathbf{E}_{sun}$ with $\boldsymbol{\tau}$ for each wavelength.

Since parallel light sources have no area, several modifications are required when calculating the form factor $F_{sun,i}$ between the sun and a patch i. Firstly, the light

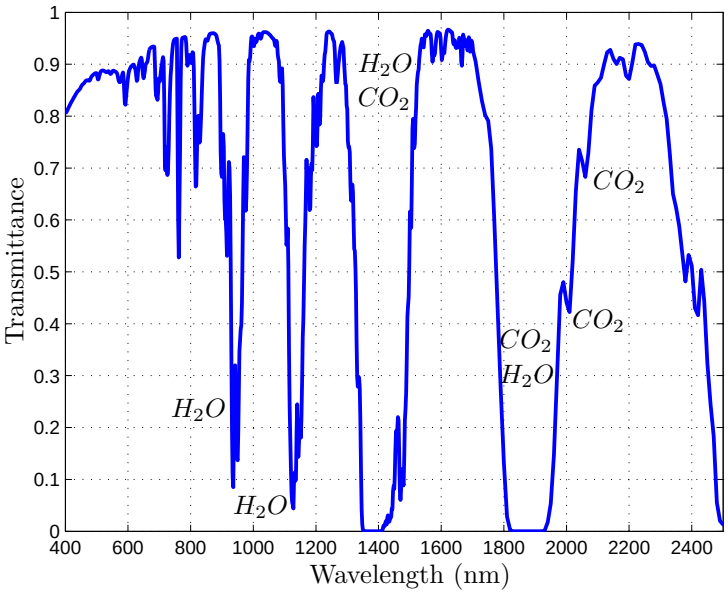


Figure 2.3 – Example transmittance τ of the atmosphere generated using a computational atmospheric radiative transfer model. The different molecules absorb light at characteristic wavelengths and this function will change depending on the weather conditions.

source $E_j(\lambda)$ in Equation (2.5), is expressed in terms of power instead of radiosity. This is achieved by symbolically multiplying the irradiance $E_{sun}(\lambda)\tau(\lambda)$ by the area of the light source A_{sun}. The form factor between the direct sunlight source and patch i is given by the projected visible area of patch i:

$$F_{sun,i} = \int_{A_{sun}} \cos\theta_i V(-\Theta_{sun}, x_i) dA_i, \tag{2.8}$$

where Θ_{sun} is the directional vector towards the position of the sun. If visibility is assumed to be uniform over the entire surface of i, the form factor can be simplified to:

$$F_{sun,i} = V_{sun,i} A_i \cos\theta_i. \tag{2.9}$$

The reciprocity principle for form factors is:

$$A_1 F_{1,2} = A_2 F_{2,1},$$

which intuitively means that once one form factor is known, the other is as well. This

means that $F_{i,sun}$ can be expressed as:

$$F_{i,sun} = A_{sun} V_{sun,i} \cos\theta_i. \tag{2.10}$$

To account for the fact that the light source is expressed in terms of power, the form factor is divided by A_{sun} and so the radiosity leaving patch i due to direct sunlight is given by:

$$\begin{aligned} B_{i,sun}(\lambda) &= \rho_i(\lambda) E_{sun}(\lambda) \tau(\lambda) A_{sun} \frac{F_{i,sun}}{A_{sun}} \\ &= \rho_i(\lambda) E_{sun}(\lambda) \tau(\lambda) V_{sun,i} \cos\theta_i \end{aligned} \tag{2.11}$$

Once extraterrestrial sunlight enters the earth's atmosphere, Rayleigh scattering due to particles smaller than the wavelength leads to the blue sky visible during daylight hours. At times near dusk and dawn, the distance travelled through the atmosphere increases, leading to a red tinge on the horizon.

The form factor for skylight can be modelled by assuming the light source is a hemispherical light source [12, 70], referred to as the sky dome. Therefore, the form factor is the solid angle subtended to the sky dome, weighted by the angle of incidence and normalised to sum to unity:

$$F_{i,sky} = \int_\Omega \frac{H(\omega)\cos\theta_i}{2\pi} V(\omega, dA_i) d\omega, \tag{2.12}$$

where $H(\omega)$ is the ratio between the skylight radiance in direction ω and the radiance in the zenith direction. The radiosity of patch i due to skylight is therefore:

$$B_{i,sky}(\lambda) = \rho_i(\lambda) E_{sky}(\lambda) F_{i,sky}, \tag{2.13}$$

$$= \rho_i(\lambda) E_{sky}(\lambda) \int_\Omega \frac{H(\omega)\cos\theta_i}{2\pi} V(\omega, dA_i) d\omega. \tag{2.14}$$

Substituting Equations (2.11) and (2.14) back into the discrete rendering equation, allows a radiation model to be generated that clearly isolates the influence of each

illumination source:

$$B_i(\lambda) = B_{e,i}(\lambda) + \rho_i(\lambda)[V_{i,sun}E_{sun}(\lambda)\tau(\lambda)\cos\theta_i + F_{i,sky}E_{sky}(\lambda) + \sum_{j=1}^{N} F_{ij}B_j(\lambda)], \tag{2.15}$$

$$L_i(\lambda) = L_{e,i}(\lambda) + \frac{\rho_i(\lambda)}{\pi}[V_{i,sun}E_{sun}(\lambda)\tau(\lambda)\cos\theta_i + \Gamma_i E_{sky}(\lambda) + \sum_{j=1}^{N} F_{ij}L_j(\lambda)\pi], \tag{2.16}$$

where Γ_i is referred to as the sky factor and is equal to the form factor $F_{i,sky}$.

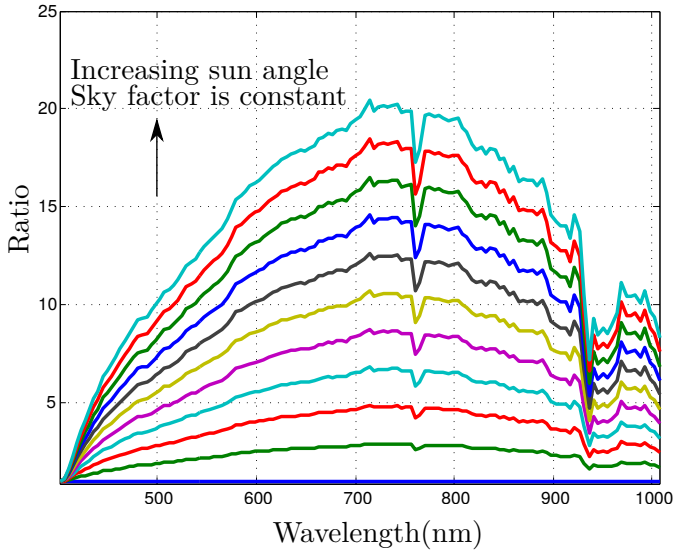
The causes of illumination variation in a scene and its influence on the appearance of a material can be seen through Equation (2.16). The visibility, sun angle and sky factor terms dictate the intensity of the direct sunlight and diffuse skylight illumination sources. If the reflectance of the material is constant and these terms vary, the appearance of the material captured by the camera will also change as shown in Figure 2.4.

2.1.2 Component Evaluation

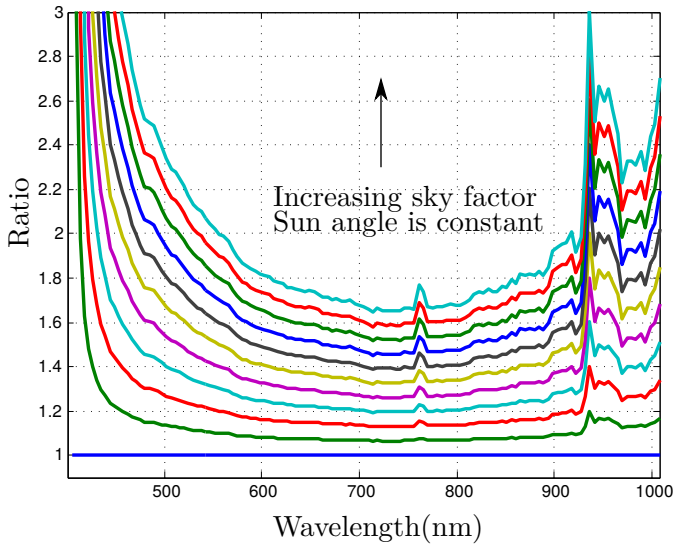
Radiance

One of the key components for generating an illumination invariant representation of the scene is to obtain observations of the scene radiance. This is obtained through the use of imaging sensors such as consumer grade RGB and multispectral cameras, as well as hyperspectral line scanners. These sensors differ in the portion of the electromagnetic spectrum that they are able to sense, as well as their respective sensor response functions.

Consumer grade cameras typically record measurements in three colour bands; red, green and blue, with the spectral response of each channel being wideband in nature. Following the measurement of the scene radiance, the raw sensor data is typically processed by removing sensor noise using dark current subtraction, debayering, white



(a) Decreasing the sun angle not only changes the intensity, but also the shape of the spectra. Each illumination spectra is normalised by the 90° sun angle spectra.



(b) In this example, increasing the sky factor predominantly affects the magnitude of the spectra as the sun angle is small. Each illumination spectra is normalised by the 0 sky factor spectra.

Figure 2.4 – Influence of the sun angle and sky factor on the appearance of a material of green reflectance.

balancing, gamma correction and conversion to a colour space. The latter three steps represent non-linear forms of post-processing.

Hyperspectral line scanners offer greater spectral resolution compared to consumer grade cameras through the use of narrow band sensor responses, but this comes at the cost of data storage as well as price. Typically, these scanners measure the radiation at many spectral bands at each pixel and can sense various regions of the electromagnetic spectrum including visible, near, short and long wave infra-red regions. This makes hyperspectral scanners extremely useful for characterising the chemical constituents of scenes, as different materials possess unique spectral signatures [71]. During processing, sensor noise is removed using dark current subtraction, the image is corrected for smear and the data may be converted to radiance units if required using precise calibration data.

While Equation (2.16) expresses the radiance as a function of the material and illumination conditions, it is not required that the data be radiometrically calibrated. As long as the pixel intensity of the camera is linearly related to radiance, the rendering equation still holds. Therefore, for consumer grade cameras, the image prior to the gamma correction and colour space conversion can be used. White balancing algorithms that operate on each channel independently can be used to improve the range of the image. For hyperspectral scanners, the data can be used both before and after calibration.

Visibility and Sun Angle

Modern robotics and remote sensing vehicles perceive the environment using a multitude of sensors. The navigational sensors typically found on board these vehicles include Global Positioning System (GPS) and Inertial Measurement Units (IMUs). These allow the pose of the vehicle to be determined with respect to the earth and the point cloud generated by the LIDAR sensor can therefore be geo-registered. Given the location of the vehicle and the time at which the scans are taken, the position of the sun can be accurately determined [63]. Therefore, ray tracing from each position in the scene to the sun location and determining whether it is occluded or not allows the visibility term $V_{sun,i}$ to be calculated.

Once the position of the sun is determined, the sun angle θ_i is trivial to compute using the dot product between the sun position vector and the normal at each position in the point cloud.

Sky Factor Calculation

Determining the sky factor for each region is vital in determining the influence of skylight on appearance. There are several methods used to calculate the sky factors that differ based on their accuracy and computational time. In this section the common approaches to sky factor estimation are detailed with the *full* calculation of sky factors referring to the computationally expensive, but accurate methods used to determine sky factors.

Assuming that the sky dome is of uniform colour and intensity, the sky factor integral (Equation 2.12) can be simplified:

$$F_{i,sky} = \int_{\Omega} \frac{\cos\theta_i}{2\pi} V(\omega, dA_i) d\omega. \tag{2.17}$$

This assumption allows the influence of skylight to be determined without needing to calculate the skylight strength at each position on the sky dome. The uniform assumption is commonly used in the remote sensing community and for some computer vision applications [50]. It is valid for most daylight hours, with the exception of times near dawn and dusk when the horizon can appear red.

In order to evaluate Equation (2.17), a sampling procedure can be used. A number of rays are sent from region i in the direction of points on the sky dome chosen from a uniform distribution. Rays that are not occluded before they reach the sky dome are weighted by their angle of incidence and accumulated. After a large number of rays are traced, the result is normalised by 2π and this determines the sky factor. As is standard with Monte Carlo based approaches, the larger the number of samples, the more accurate the estimation becomes. Cosine distributed sampling of the sky dome can also be used, in which case the number of samples reaching the sky dome is just normalised to estimate the sky factor.

In remote sensing contexts, the sky factor is typically found by selecting a larger number of points on the sky dome, and determining the visibility to each point. The sky factor is then estimated as:

$$\Gamma_i = \frac{N_i}{N_{total}}, \tag{2.18}$$

where N_i is the number of non-occluded points and N_{total} is the total number of rays traced to the sky dome. This approach ignores the weighting of the incident angle but has been used successfully [10, 62].

2.1.3 Summary

In this section, the outdoor illumination model was presented and is shown to consist of three main light sources. Terrestrial sunlight illuminates the scene as a parallel light source, diffuse skylight acts as a uniform hemispherical light source and indirect illumination is caused by inter-reflection of light within the scene.

The outdoor illumination model shows that illumination variation in the scene is caused by a number of parameters. The visibility and sun angle term dictate the intensity of the terrestrial sunlight source, while the sky factor term describes the amount of diffuse skylight visible from a region. Using the radiosity rendering formulation, indirect illumination is governed by the form factor geometrical term. Each source is treated independently and the combination of them is used to calculate the total incident illumination at a point in the scene.

This model is used throughout this thesis in order to derive illumination invariance in a principled manner. With some modifications (such as the addition of path radiance and the adjacency effect), the illumination model has been used in the literature, particularly in remote sensing applications for radiometric normalisation of hyperspectral images [10, 29, 78].

2.2 Literature Review

The problem of illumination invariance and its various formulations has received widespread attention, with the literature being split in terms of the different research communities. The computer vision and image processing field tends to use image based techniques to solve what is known as colour constancy. This problem is motivated by human perception, which allows us to perceive materials under different illumination conditions to be constant. The remote sensing community remodels illumination invariance as a radiometric normalisation issue, where pixel intensity is converted into material reflectance. This is useful for supervised learning techniques which require laboratory obtained spectral libraries. Finally, the field of inverse reflectometry from the computer graphics community is vital when attempting to reconstruct material characteristics. This thesis combines elements from each of these three communities to develop an illumination invariant imaging system for the outdoor environment.

2.2.1 Computer Vision and Image Processing

Computer Vision and Image Processing (CVIP) techniques for illumination invariance and shadow detection tend to focus on approaches using information from the image only. Early works recognised that images are a combination of intrinsic properties such as illumination and material properties. Once the illumination effects are removed from the image, the remaining details form an illumination invariant representation of the scene. CVIP has also focused on colour constancy, motivated by the perception system of humans.

Colour constancy techniques aim to develop an illumination invariant representation of an image. The dependency of the pixel intensity on the scene illuminant and underlying material characteristics mean that it is an ill-posed problem [1]. A number of approaches have been developed, with one of the simplest being the Gray-World approach, which operates on the principle that the average colour in the scene is neutral. This can lead to poor results when the assumption does not hold.

One of the seminal works in the field of colour constancy involved the development of the Retinex approach in [44, 45] that uses the maximum response values in each camera channel to estimate the illuminant. This method assumes that image gradients due to a change in material are greater than those due to variations in illumination [74]. The reliance on a white reflective surface is sometimes not valid and so several modification to the theory were made to take into account local pixel values. An alternative approach to colour constancy include gamut based approaches [21, 27] which exploit the fact that the range of illuminants is constrained by the image colour. Subsequently, an estimate of the illumination can be obtained.

While the concept of colour constancy is promising, it must be noted that these algorithms do not typically account for geometric variability in the scene. Most algorithms assume a single illuminant is present at each pixel [46], however this is not the case for outdoor scenes where skylight is the predominant source within areas of shadow. Without factoring in multiple illuminants, the shadows will remain in the processed image and this is not desirable for a number of applications.

The use of depth data from a stereo vision camera in combination with the Retinex approach was proposed in [83]. In this work, edges in the depth map indicate sharp changes in orientation and are used to determine where spatial comparisons are made. The multiple illumination colour constancy scenario was approached in [32] through a sampling and illuminant estimation scheme, while [46] analyses the shadow edges in the image. These approaches outperform colour constancy methods that assume single illumination sources, but do not account for geometric variability in the scene.

Intrinsic images were proposed in [4] as a method of separating out the illumination, material and range components that constitute an image. Using a diffuse assumption of the scene, the pixel intensity is a combination of the illumination source I, incident illumination angle α and material reflectance ρ:

$$L(\lambda) = I(\lambda)\rho(\lambda)\cos\alpha, \tag{2.19}$$

where indirect sources of illumination are neglected. The decomposition of the pixel

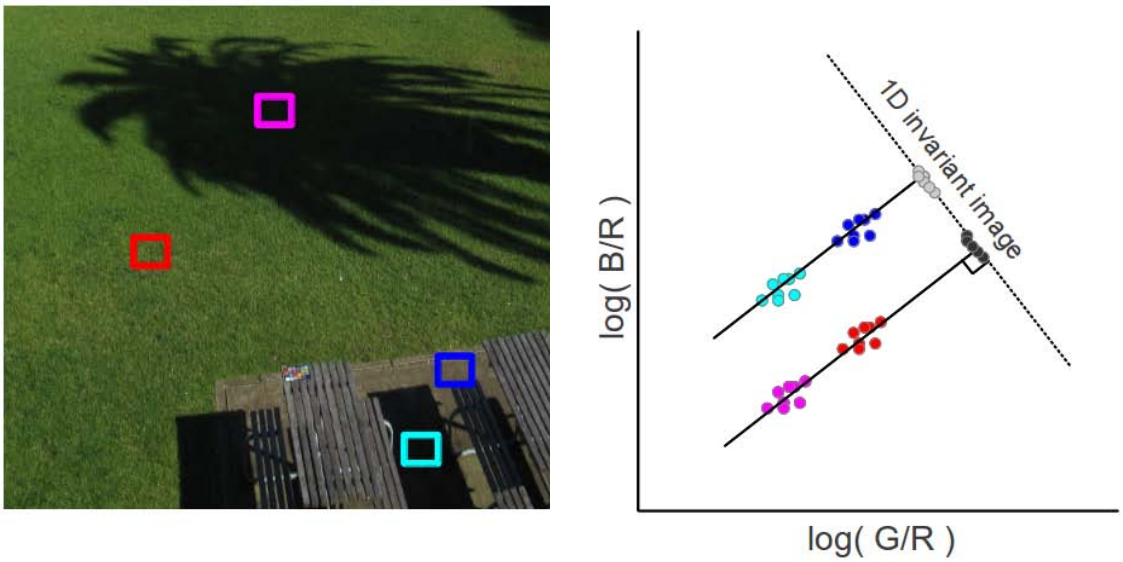
intensity is ill posed, with many combinations of illumination, reflectance and orientation providing identical pixel intensities.

In order to reduce the search space, several assumptions about the nature of the scene must be made. Namely, assuming surfaces are continuous and have uniform reflectance, while illumination varies smoothly with the exception of shadow boundaries.

Due to the ill-posed nature of the intrinsic image problem, [82] uses image sequences obtained from a fixed camera and a maximum likelihood approach. In the modified setup, the scene reflectance remains constant while the illumination varies throughout the sequence. The problem remains under-determined and so [82] uses the fact that derivative filters (both horizontal and vertical directions) applied to natural images tend to yield sparse outputs whose histograms can be approximated by Laplacian distributions. A median filter is then used in order to calculate the maximum likelihood estimate of the filtered reflectance image. While this method allows the image to be decomposed into intrinsic components, its reliance on a fixed camera and image sequences makes it undesirable for robotics and remote sensing applications, where the sensing platform is typically moving and observing different aspects of the environment.

One of the most recent and popular methods for shadow removal that recovers $1D$, $2D$ and $3D$ invariant images was developed in a series of papers in [19, 20, 22–26]. Here, the dimensionality is an indication of how many spectral channels are present in the image. The method stems from the premise that after taking logarithms of chromaticity band-ratios, a change in illumination will cause each point to move on a line as seen in Figure 2.5. These lines will be approximately parallel to each other and the direction is independent of the material. Projecting onto a line orthogonal to the illumination direction allows a $1D$ illumination invariant image to be generated. Originally, the illumination invariant line direction was calculated using the spectral sensitivity functions of the camera.

In order to automate the calculation of the required orthogonal line, [24, 26] utilise an entropy minimisation procedure. For all possible directions of the line, the log



(a) Identical materials appear differently when illuminated by different sources.

(b) Identical materials under different illumination conditions lie on a straight line in the log-chromaticity plot.

Figure 2.5 – Varying the illumination on a material will cause the log chromaticity band ratio point to move along a straight line. The direction of the line is independent of the material and by projecting onto an orthogonal line, a grayscale image that is invariant to illumination variations can be obtained.

chromaticity points of the image are projected onto the candidate line and the entropy is measured. Shannon's entropy is used in [24], while a quadratic entropy measure is used in [26] which results in a smoother function that is simpler to minimise. At the correct angle for illumination invariance, the entropy measure will be minimised as there will be maximum separation between the different materials.

To obtain a $3D$ illumination invariant representation of the scene, [23] solve a Poisson equation which may be computationally expensive and results in artefacts [20]. Using a number of non-random paths in the image and performing $1D$ integration, [19] was able to recover the $3D$ invariant image rapidly, but it still contains artefacts due to the path directions. This work was extended in [20] to produce improved results by closing shadowed regions and controlling the paths along which integration takes place. While the illumination invariant approach can provide visually pleasing results, the presence of high dynamic ranges and sensor noise can bring about poor results [42]. It is also not clear how the algorithm will perform at long wavelengths, such

as those captured when using hyperspectral cameras in the infra-red region of the electromagnetic spectrum. In the derivation of the illumination invariant approach, the illumination is modelled via Wien's approximation, which decreases in accuracy at long wavelengths.

The log-chromaticity approach developed in [22] was used for robotic localisation in [13]. Through the utilisation of an intrinsic image instead of the standard RGB colourspace, the localisation accuracy was able to increase as illumination variance decreased significantly. However, the same benefits of using a $1D$ invariant representation were not found when performing matching between images [80]. While the invariant image allows discrimination between classes within an image, its performance degrades when comparing multiple images. Through a hybrid system using both invariant and RGB representations, [80] showed that improved robustness could be achieved when facing illumination variability in the scene.

Model based approaches provide an alternative means to obtaining illumination invariant representations of scenes. In [42, 43], a set of weak cues is combined to form a strong cue to estimate the illumination conditions. The features used include the sky colour variations, cast shadows and variations in shading of objects such as buildings and pedestrians. Each edge pixel in the image is classified using training data to determine whether it is a shadow or not and colour information is re-integrated using the method developed in [23].

A random field technique was developed in [31] to isolate material and lighting influences in the image. The reflectance is modelled as being composed of a sparse set of basis functions, with the total energy function of the system containing three components. The first component encourages smooth variations in lighting, while the second uses information from Retinex to distinguish material and shading edges. The final term promotes a small number of material in the scene, with optimisation being performed via coordinate descent.

Shadow detection and compensation is performed using a graphical model based method in [35]. Regions containing similar materials are identified and edge weights are assigned based on whether they have similar illumination conditions. Using train-

ing data, a classifier is developed allowing shadow detection to be performed, with the benefit that soft shadows can also be detected through the use of non-adjacent pairwise connections in the graph. Shadow removal is then performed by using a simple model of the incident illumination that accounts for direct and indirect sources, and relighting the entire scene so each pixel is illuminated by all the sources.

One of the state of the art model based approaches was developed in [2, 3], which uses a single image to estimate shape, illumination and reflectance. By using statistical models of the depth and reflectance, and assuming that material variation is sparse, optimisation can be performed to extract the intrinsic images. Through the incorporation of noisy depth data, [3] was able to improve on the results.

The use of near-infrared (NIR) data for shadow detection was investigated in [28]. Potential regions of shadow are identified by their low intensity in the RGB and NIR channels, and the ratio map of NIR to RGB is used to confirm whether or not the pixels are occluded from the sun. While the method is able to accurately identify shadows, several parameters must be selected and a specialised camera setup is required. To compensate for the effects of shadows, [67] used this approach for shadow removal.

The benefit of using image based approaches is that the illumination conditions are typically derived from the scene measurements, meaning there is no dependence on external systems. However, while these approaches have shown promising results, they are limited by the fact that they are unable to account for topographic variability. This is due to the lack of explicit modelling of the orientations of the surfaces in the scene, which can be obtained through a second sensor modality. Also, obtaining a $3D$ invariant image is still a difficult problem, with several approaches requiring additional parameters to be tuned.

2.2.2 Computer Graphics

Generating photo-realistic images using a computer has been a long term goal for the computer graphics community, with technological progress allowing advanced tech-

niques to be developed. Rendering consists of combining knowledge of the scene structure, material characteristics and illumination sources within a scene, and generating an image from the viewpoint of a virtual camera [17].

From a computer graphics perspective, illumination invariance is equivalent to the problem of inverse reflectometry. Given the scene structure, illumination sources and final image, inverse reflectometry aims to estimate the underlying reflectance characteristics of all objects in the scene. The discrimination between variations in the image due to surface geometry, reflectance and illumination is not a trivial task [15] and a large amount of research has been performed in this area.

The first group of inverse reflectometry techniques either have knowledge about the location of the illumination sources, or directly take measurements of them. This reduces the number of parameters in the rendering model that are required to be estimated.

A simple method for measuring the illumination in outdoor scenarios is to take images of the sky dome. In order to estimate the reflectance of a large scale outdoor scene, [84] takes a series of photographs under sunny conditions. A parameterised sky model [58] is then fitted to these measurements in order to estimate the illumination conditions. This allows two different reflectance functions to be obtained; one based on sunlight and the other based on skylight. The main drawback is that all surfaces in the scene must be illuminated by sunlight in at least one of the captured images.

Surface reflection estimation on a small scale scene was performed in [47–49] through the use of a highly structured scene. The geometry was captured using a laser scanner, before the optimum position for illumination sources in the scene were chosen in order to observe both diffuse and specular reflectance. This uses an iterative updating procedure to estimate the parameters of the reflectance model, with inverse radiosity being used in [47] and photon mapping being used in [49]. In outdoor scenarios, the illumination sources are uncontrolled making it difficult to implement such a method. The use of inverse radiosity for reflectance estimation in the presence of inter-reflections was shown to be quite effective.

The requirement in [84] for imaging under clear conditions was relaxed in [14] through the use of external hardware being placed in the scene. Three spheres (light probes) are placed in the scene and a set of images are taken under different illumination conditions. A shiny, black sphere allows the direction of the sun to be estimated as it appears as a specular highlight, while its intensity is measured using a diffusely reflecting gray sphere.

Through the use of a laser scanner to obtain the geometry, iterative inverse global illumination is used to estimate the reflectance. For each iteration, the radiance of each point is estimated and compared to original image values. This allows the diffuse and specular reflectance components for each material to be updated for the next iteration until convergence is reached.

A sampling based method was developed in [68] for inferring both illumination and reflectance from a scene. The scene illumination was modelled as a number of light sources on a hemisphere and by using the region within a shadow and ray tracing to the sources, a system of linear equations could be developed. If the surface had a known Lambertian reflectance, the illumination source could be estimated, while a second image without the occluding object is required if the reflectance is unknown. This method was extended upon in [69] through the development of an iterative method to estimate the illumination and the reflectance of the scene. However, this requires that the material of the surface being occluded is uniform.

2.2.3 Remote Sensing

Through technological advancements in recent times, remote sensing has expanded from aerial and satellite platforms, to also include those that are based in the field. This unique perspective of the environment brings about new challenges and opens the door to new applications. Remote sensing observes the environment from afar, with application including vegetation analysis, surveillance and thematic mapping. The obtained data can come from sensors including laser range finders and cameras, and is typically characterised by its high spatial resolution.

The aim of achieving illumination invariance in an image can be reformulated into the remote sensing problem of radiometric normalisation. Once imagery is collected, for example hyperspectral data from field or aerial based platforms, pre-processing steps of dark current subtraction and smear correction are applied. The resultant data is either proportional, or if calibration data is used, equal to the observed radiance in the scene. Reflectance estimates are obtained by dividing by the incident illumination and are characteristic of the underlying material. Essentially, this representation of the image is independent of illumination in the scene.

Applications such as geophysical parameter estimation, vegetation analysis and sub-pixel detection use reflectance data and the algorithms operate accurately when the estimates are plausible and the relative reflectance between materials is maintained [56]. The various methods for normalisation can be split based on the type of data they use.

Empirical normalisation methods are often used for aerial and satellite based imagery when direct or simulated measurements of the illumination sources are unavailable. These include the use of residual images, continuum removal, Internal Average Relative Reflectance (IARR) and the empirical line method [71].

Residual images are obtained by initially selecting a wavelength and scaling each pixel spectra so that the value in this wavelength is equal to the maximum value in the entire image. The mean of this normalised radiance image is subtracted from all pixels in order to produce a reflectance estimate that compensates for both geometric and solar irradiance. One of the simplest methods for normalisation is the use of IARR, where each pixel is divided by the mean spectra of the entire image. Compared to residual images, this method is unable to account for variable topography in the scene.

The empirical line method [72] utilises multiple reflectance measurement in the scene, which can be obtained by the user with instruments such as spectrometers. The reflectance is plotted against the captured radiance measurements in order to calculate offset and gain values for each wavelength. These are applied to the image to obtain a reflectance estimate, though it does not account for scene geometry [73]. The Variable Empirical Coefficient Algorithm (VECA) [30] uses fused digital elevation maps and

imagery in order to perform topographic correction. The algorithm operates based on the relationship between observed radiance and the sun angle (angle between the surface normal and the sun position). The sun angle and radiance are highly correlated and can be approximated using a linear regression equation. A multiplicative, wavelength dependent scaling factor is calculated which increases and decreases shaded and sunlit intensities respectively. While this method is simple to operate, it does not take into account geometric influences such as the sky factor, as well as indirect illumination sources.

While the reflectance spectra obtained using these empirical methods often retains the characteristic absorption bands required for discrimination between materials, they do have limitations. The main issue is that they only produce relative reflectance estimates. This may be useful for some applications, but when comparing the magnitude of spectra to those obtained from a laboratory, discrepancies arise. This is significant when dealing with spectra of materials that may be similar in shape but varying in magnitude.

Computational atmospheric radiative transfer models and skylight simulation software range in complexity from those that are relatively simple [37, 58, 60], which are popular for computer graphics purposes, to the more advanced SMARTS [34] and MODTRAN [5] which has been evaluated and used for many remote sensing applications. The models simulate the propagation of light through the atmosphere, with parameters such as atmospheric composition and visibility being obtained from meteorological station measurements, or seasonal and location dependent estimates. An iterative based approach using vegetation within the scene was developed in [65] in order to estimate the visibility and reflectance parameters. The ATCOR system [64] builds on these approaches and provides a method for the radiometric correction of airborne imagery. The MODTRAN model is used for atmospheric correction, while a digital elevation map is used to account for topographic variability in the scene through the calculation of slope, aspect and elevation.

A drawback of using an atmospheric radiative transfer model is that they are highly parameterised. This means that each variable needs to be either known from direct

measurements, or estimated by the user. While for some areas, the measurements of oxygen, water and ozone in the atmosphere are known, this data is not available everywhere. For example, if the location of the sensor is unknown and only the relative position of sun is available, then computational models cannot be used. Another issue is that the light sensing device must be calibrated to the same units as that of the model. It is therefore ideal to obtain a more accurate estimation.

The radiosity method was used in a remote sensing application in order to perform plant canopy modelling [7]. Radiosity approaches focus on discrete elements in the scene and can retain local reflectance characteristics. This is opposed to radiative transfer techniques that utilise volumetric elements and smooth location, shape and orientation information over local neighbourhoods to provide a continuous scene representation. The discretised nature of the radiosity approach allows spatially dependent BRDF information to be retained, but this comes at a significant computation cost in terms of memory and the number of equations required to account for the emission, transmission and reflection of light by each element.

Several hardware based methods exist which are used to directly measure the incident illumination in a scene. Diffuse calibration boards whose reflectance characteristics are precisely known are placed within the scene being imaged. During a post-processing step, the average pixel intensity of the calibration board is calculated and this is used to normalise the entire image (flat-field correction) [54]. Similarly, Downwelling Irradiance Sensor (DIS) (or Cosine Integrators) provide a similar direct measurement of the spectral shape and magnitude of the incident illumination [8].

The main limitation of these techniques is that the resultant image retains the dependence on the scene geometry. The measurement of incident illumination obtained from a calibration board is only valid at the location of the board. Elsewhere in the scene, the geometry varies with occlusions and indirect illumination influencing the appearance of the underlying materials. Ideally, the calibration board would be placed at all points in the scene, however this is not a practical option. A similar limitation is seen when using Downwelling Irradiance Sensors, especially for field based data, where the sensor is often placed on-board a sensor platform with a fixed orientation.

The drawbacks of a purely image based system for perceiving the environment is that it does not account for the variable illumination in the scene. In the outdoor environment, the captured pixel intensity is dependent on the geometry between the illumination source, scene and the camera. Active illumination reduces these dependencies by illuminating the scene with a source from a known position.

An early form of active sensing was to use laser to not only obtain the depth from the scene, but also to measure the intensity of the return signal. This is dependent on the underlying reflectance characteristics of the scene and it is therefore possible to perform basic materials classification after post-processing to remove the effects of surface orientation.

Using an aerial based platform, [39] combined the use of range and reflectance data to identify objects on the ground. This approach was used as it is robust to the weather and insensitive to shadowing, which is a significant challenge for outdoor perception.

Active sensing was also utilised for underground mining [38], where a laser was used for mapping. Reflectance from materials such as coal and colour panels was shown to vary, indicating that it is indeed possible to use an active sensing approach. Similarly, [18] used a LIDAR based approach for underground mapping and was able to highlight areas of water leakage through simple analysis of the intensity returns. Regions that are moist are distinguished from those that are dry by their low intensity returns and are this is useful for observing discontinuities.

To increase the spectral resolution of active sensing approaches, [56] developed a broadband white laser illumination source in conjunction with a hyperspectral camera. The spectral resolution of such a system allows greater material discrimination as characteristic absorption bands in the reflectance data can be found. A calibration board placed in the scene was required in order to measure the illumination source strength at a specific position.

Active illumination reduces the influence of variable illumination, especially in shadows where the performance of high level algorithms degrades. It is robust to illumination conditions and allows remote sensing to occur at both day and night time.

The major limitation though, is the power required for illumination. When using mobile platforms in the outdoor environment, where power requirements must be carefully constrained, it is often infeasible to illuminate a scene with a light source. The distance between the scene and the source, as well as the overpowering nature of sunlight, mean that high power is required in order obtain a return signal. While lasers provide reasonable active sensing observations, they typically only operate for one wavelength. In order to classify materials for example, more discriminative data is preferred.

Purely image based radiometric normalisation fail to account for geometric influences on illumination. Therefore, the next logical step is to combine the geometric data with imagery. Fused LIDAR and hyperspectral data was used in [55] for geological mapping. This provides valuable information for volumetric estimation of materials but illumination variations still pose a significant problem.

A digital surface model was used in [29] with hyperspectral data. Ray tracing is used in conjunction with a support vector machine method developed in [78] to detect regions of shadow. The radiance L captured by the camera of location x in the scene is modelled using a modified version of Equation (2.16):

$$L(x,\lambda) = [V(x)\frac{E_{sun}(\lambda)\tau(\lambda)}{E_{sky}(\lambda)}\cos\alpha(x) + \Gamma(x)]\rho(x,\lambda)E_{sky}(\lambda). \tag{2.20}$$

The aim of the system is to relight all points so that they are fully illuminated by the sun and sky, and in order to do this the user selects multiple regions in the image that represent the same material under different illumination conditions. This is done for several materials and non-linear optimisation is used to estimate $\frac{E_{sun}(\lambda)\tau(\lambda)}{E_{sky}(\lambda)}$ and $\rho(\lambda)E_{sky}(\lambda)$ for each input material.

This allows correction of original image to occur such that each pixel is illuminated by the same source. The new relit radiance L^* is calculated as:

$$L^*(x,\lambda) = L(x,\lambda)\frac{\cos\alpha(x)\frac{E_{sun}(\lambda)\tau(\lambda)}{E_{sky}(\lambda)} + 1}{V(x)\cos\alpha(x)\frac{E_{sun}(\lambda)\tau(\lambda)}{E_{sky}(\lambda)} + \Gamma(x)}. \tag{2.21}$$

While the technique visibly improves the perception of shadowed areas, it requires user input and amplifies noise at longer wavelengths. This is because the signal-to-noise ratio at these bands is low in the outdoor environment.

Through the integration of atmospheric radiative transfer models into the system, [40] and [10] manage to bypass the requirement for user selection of regions in the scene. Forward modelling is used in [40] for target detection with fused LIDAR and hyperspectral data. Such a technique is used because the user knows what material is being searched for in the scene. The purpose of a forward modelling system is to estimate what the radiance captured by the camera should be, if the material in the scene is a target.

The outdoor illumination model utilised excludes self-emission, but includes the adjacency effect E_{adj} and up-welling radiance L_u:

$$L(x,\lambda) = [V(x)E_{sun}(\lambda)\tau(\lambda)\cos\alpha(x) + \Gamma(x)E_{sky}(\lambda) + E_{adj}(\lambda)]\frac{\rho(x,\lambda)}{\pi} + L_u(\lambda). \quad (2.22)$$

Up-welling radiance occurs when light is scattered by the atmosphere in the direction of the sensor. The illumination parameters are estimated using an atmospheric radiative transfer model (MODTRAN [5]). While forward modelling is ideal for applications such as target detection, in applications where the materials in the scene are unknown to the user, it is not operable.

Illumination invariance is achieved in [29] through relighting, but it is unable to estimate the underlying reflectance characteristics of the materials in the scene. This is because the illumination spectra cannot be explicitly determined without more information.

This issue is alleviated in [10] by integrating a computational atmospheric radiative transfer model into the system. A similar outdoor illumination model is used, with the addition of up-welling radiance L_{up}:

$$L(x,\lambda) = [V(x)E_{sun}(\lambda)\tau(\lambda)\cos\alpha(x) + \Gamma(x)Esky(\lambda)]\frac{\rho(x,\lambda)}{\pi}\tau_2(\lambda) + L_{up}(\lambda), \quad (2.23)$$

where $\tau_2(\lambda)$ is the transmittance from x to the sensor.

The additional terms of τ_2 and L_{up} can be incorporated into the model as they can be estimated using atmospheric radiative transfer models. Using this integrated system allows the illumination at each pixel to be estimated individually and this was found to be useful for target detection. The limitation though, along with the method of [40], is that it is dependent on the atmospheric radiative transfer model for which the parameters as mentioned previously, may not be available in all areas.

Through the combination of geometric and image data, improved illumination invariance has been achieved. The spectra of materials that were originally occluded from the sun can be either relit or normalised to have a similar shape and magnitude to an identical material exposed to sunlight. While the addition of a second sensor modality increases the performance of high level algorithms, it does have some drawbacks. Firstly, there is the cost of the ranging sensor that must be taken into account, as well as the multi-modal sensor calibration that must be performed. However, the benefits of using such a system are clear and these are not major limitations.

2.3 Summary

The concept of illumination invariance can be reformulated into several different problems depending on the final application. For computer vision and image processing, achieving illumination invariance has usually focused on projecting the data into a lower dimensional form and compensating for the variations. This is contrasted to computer graphics and remote sensing, where the task of attaining reflectance estimates for objects in the scene is of critical importance.

Typically, the inverse reflectometry approach to illumination invariance has attempted to estimate both specular and diffuse reflectance parameters. However, keeping in mind that the final application for the illumination invariant system being developed in this thesis will be for robotics and remote sensing platforms, this may be of unnecessary resolution. When utilising supervised classification and clustering methods,

there is typically an underlying diffuse reflection estimation of the material used for supervision. This comes from the fact that the dimensionality of the models becomes too large when incorporating all parameters such as the incoming and outgoing light directions, and spectral wavelength.

From the literature, it is clear that fused point cloud and image data complements one another and is useful for deriving the underlying reflectance properties of the scene. This is because appearance is dependent on geometry, illumination and material, and it is non-trivial to isolate each of these components. Few illumination invariant techniques are available that allow the original number of spectral channels to be retained, while at the same time being independent of both the user and atmospheric radiative transfer models.

Chapter 3

Physically Based Illumination Invariance

This chapter proposes an automatic method for achieving illumination invariance in outdoor environments through the utilisation of a physically based approach. The term *'physically based'* will be used to describe the method and indicates that it is derived from an approximation of the physical processes involved. Chapter 2 showed that current methods for doing so are dependent on highly parameterised external atmospheric models [10], or require selecting materials under differing illumination conditions within the image [29]. Purely image based solutions such as those presented in [25] project the data into lower dimensional spaces, thereby losing valuable colour information, and require several tuning parameters in order to recover the full colour spectrum.

The proposed method draws on the outdoor illumination model presented in Chapter 2 to estimate the influence of each illumination component on a per-pixel basis. An overview of the approach is shown in Figure 3.1. Through fused point cloud and imagery data, an initialisation process is used to approximate the ratio of the two dominant illumination sources; terrestrial sunlight and diffuse skylight. Scaling factors are then calculated for each pixel in order to relight all regions with respect to a common illuminant. This method is extended to the multiple image scenario through

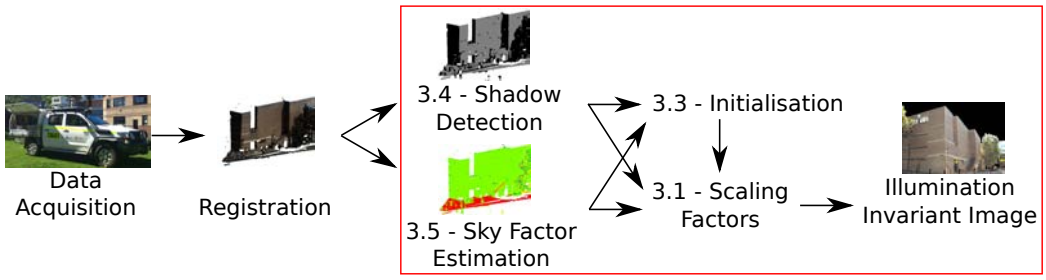


Figure 3.1 – Generating an illumination invariant representation of the scene is performed by utilising an image fused with a geo-registered point cloud. Following initialisation through an automatic process, scaling factors are found to relight each pixel with a common illuminant.

the use of an overlapping region, which allows images in a sequence to be relit with respect to a single illuminant, ensuring the consistency of a material's appearance throughout the entire dataset.

Section 3.2 details the proposed relighting method for a single image scenario, and this is extended to multiple images in Section 3.3. The method for automatically initialising the system is outlined in Section 3.4, along with a shadow smoothing technique in Section 3.5. Section 3.6 presents an efficient method for approximating sky factors in the scene. This method reduces the computational time required for illumination invariance and is suitable when real time performance is desired.

3.1 Data Collection and Pre-Processing

In order to obtain an illumination invariant image in the outdoor environment using the proposed method, both visual and geometric data must first be obtained. This thesis focus's on the use of consumer grade RGB and hyperspectral cameras, combined with LIDAR sensors to measure the range. These two modalities can be registered together using reflective markers, or markerless based approaches [76].

The fused point cloud is geo-registered using GPS measurements and the normals are calculated [75]. To assist in generating a shadow map (binary image showing which areas are occluded from the sun), surface reconstruction techniques can be

used. However, these often need to be refined by hand to remove spurious faces and so in this thesis, each return from the LIDAR is represented by a circular disc whose radius is dependent on the range and angle to the sensor (see Appendix A). Ray tracing is then used to calculate the line-of-sight visibility from each disk to the sun position which is accurately known [63].

3.2 Spatial Illumination Invariance

When sensing the environment from a single position, the incident illumination at each pixel captured by the camera varies due to geometric factors such as orientation and occlusions, as seen in Equation (2.16). Therefore, the appearance of a material under different illumination conditions can change dramatically (see Figure 2.4) and negatively affect the performance of high level algorithms [26]. In this section, the single image scenario is analysed and scaling factors for each pixel are derived that allow an image to be developed where the incident illumination at each pixel is an (unknown) constant source.

The impact of having a constant incident illumination source for each pixel is that identical materials appear similar irrespective of their location within the image. In other words, an illumination invariant image has been generated that retains the original number of spectral channels as that of the original data. Subsequently, the discriminative power of colour, which is characteristic of the underlying material, can be used to inform high level algorithms that may have previously relied on either $1D$ illumination invariant intensity levels or illumination dependent colour values.

Several relaxations to the illumination model and the evaluation of its components must be made in order to achieve illumination invariance. Ideally, illumination invariance would be obtained by generating an image that is purely dependent on the material characteristics. However, this is ill posed since the appearance is dependent on two unknowns (material and illumination). As a consequence, the resultant illumination invariant image will be a function of both material reflectance and an unknown illuminant that is common for all pixels. This is in contrast to radiometric

Figure 3.2 – Illumination invariance in the single image scenario requires initialisation points across a shadow boundary $\{A, A'\}$ obtained from the same material.

normalisation which will be discussed in Chapter 4, which removes any dependence on the illumination source, but requires an additional direct measurement of the incident illumination.

3.2.1 Initialisation

Given two points $\{A, A'\}$ obtained from the same material with one occluded from the sun as shown in Figure 3.2, it is possible to find the ratio between terrestrial sunlight and skylight irradiance. Assuming no emission in the scene and that indirect illumination is negligible, the radiance can be approximated as:

$$L_A(\lambda) = \frac{\rho_A(\lambda)}{\pi}[E_{sun}(\lambda)\tau(\lambda)\cos\alpha_A + \Gamma_A E_{sky}(\lambda)], \tag{3.1}$$

$$L_{A'}(\lambda) = \frac{\rho_{A'}(\lambda)}{\pi}[\Gamma_{A'} E_{sky}(\lambda)]. \tag{3.2}$$

Since the points are obtained from the same material source, $\rho_A = \rho_{A'}$, a terrestrial sunlight-skylight ratio can be derived:

$$\begin{aligned} \frac{L_A(\lambda)}{E_{sun}(\lambda)\tau(\lambda)\cos\alpha_A + \Gamma_A E_{sky}(\lambda)} &= \frac{L_{A'}(\lambda)}{\Gamma_{A'} E_{sky}(\lambda)}, \\ \frac{E_{sun}(\lambda)\tau(\lambda)}{E_{sky}(\lambda)} &= \frac{\Gamma_{A'} L_A(\lambda) - \Gamma_A L_{A'}(\lambda)}{L_{A'}(\lambda)\cos\alpha_A}. \end{aligned} \tag{3.3}$$

Therefore, given a pair of points obtained from the same material with one occluded from the sun, if indirect illumination is negligible and there is no emitted radiance, Equation (3.3) shows that the ratio between the two main illumination sources; terrestrial sunlight and skylight, is independent of the material. This ratio plays a crucial role in deriving scaling factors for each pixel as it allows the main illuminants to be expressed in terms of each other.

The determination of which pixels are occluded from the sun can be performed in a number of ways. The use of a multi-modal sensor system means that the obtained point cloud can be geo-registered

3.2.2 Scaling Factors

The illumination invariant image obtained is a function of both material reflectance and an unknown illuminant that is common for all pixels. In order to develop such a representation, each pixel needs to be multiplied by an appropriate scaling factor based on its original incident illumination properties.

For a pixel B' occluded from the sun, assuming indirect illumination is negligible and there is no emission, the radiance can be described by:

$$L_{B'}(\lambda) = \frac{\rho_{B'}(\lambda)}{\pi}[\Gamma_{B'} E_{sky}(\lambda)]. \tag{3.4}$$

In order to relight the point with respect to skylight (denoted by $L_{B'-relit}(\lambda)$), the radiance $L_{B'}(\lambda)$ needs to be multiplied by a scaling factor:

$$\begin{aligned} L_{B'-relit}(\lambda) &= \frac{\rho_{B'}(\lambda)}{\pi}[E_{sky}(\lambda)], && (3.5)\\ &= \frac{\rho_{B'}(\lambda)}{\pi}[E_{sky}(\lambda)]\left[\frac{\Gamma_{B'}}{\Gamma_{B'}}\right], \\ &= L_{B'}(\lambda)\left[\frac{1}{\Gamma_{B'}}\right]. && (3.6) \end{aligned}$$

The scaling factor in this case is purely geometric and independent of wavelength. Therefore, multiplication by this scaling factor will not influence spectral shape, only

its intensity.

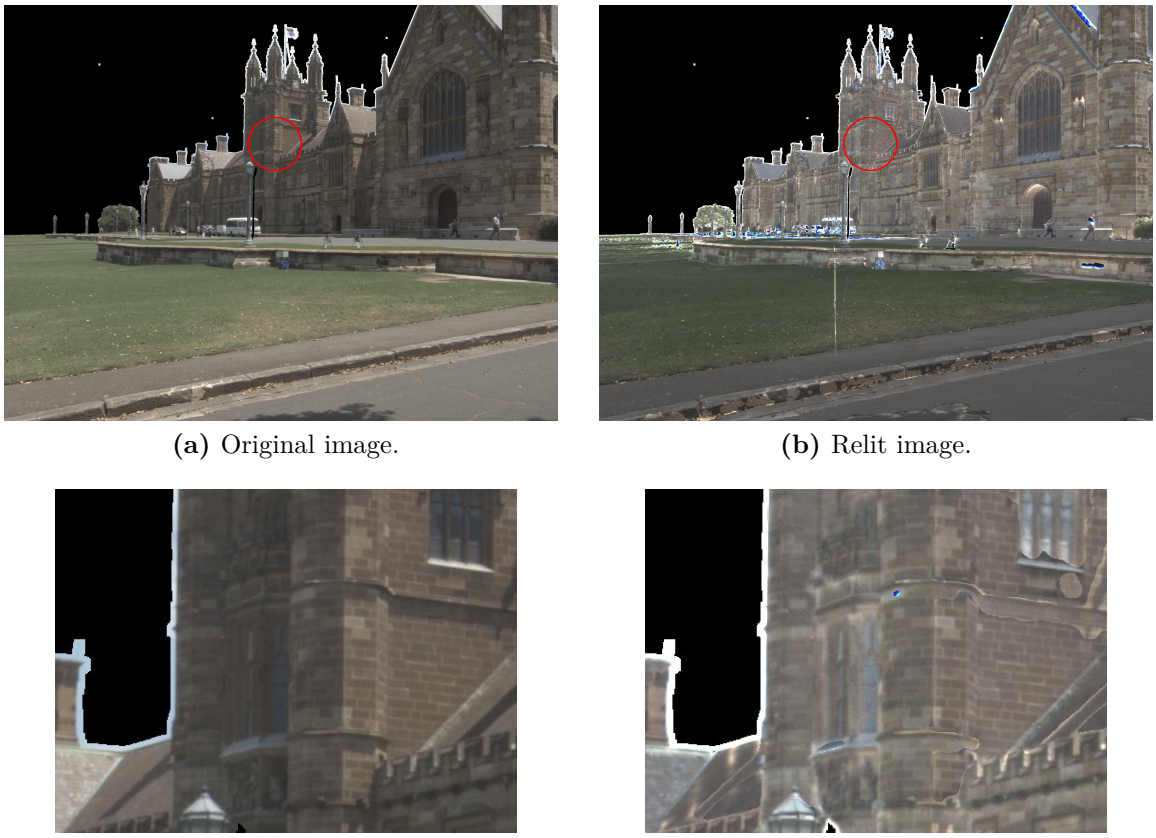
For a pixel C that is exposed to terrestrial sunlight, the relit radiance can also be found by multiplying the original observed radiance by a scaling factor:

$$\begin{aligned} L_{C-relit}(\lambda) &= \frac{\rho_C(\lambda)}{\pi}[E_{sky}(\lambda)], \qquad (3.7) \\ &= \frac{\rho_C(\lambda)}{\pi}[E_{sky}(\lambda)]\left[\frac{E_{sun}(\lambda)\tau(\lambda)\cos\alpha_C + \Gamma_C E_{sky}(\lambda)}{E_{sun}\tau(\lambda)\cos\alpha_C + \Gamma_C E_{sky}(\lambda)}\right], \\ &= L_C(\lambda)\left[\frac{E_{sky}(\lambda)}{E_{sun}(\lambda)\tau(\lambda)\cos\alpha_C + \Gamma_C E_{sky}(\lambda)}\right], \\ &= L_C(\lambda)\left[\frac{1}{\frac{E_{sun}(\lambda)\tau(\lambda)}{E_{sky}(\lambda)}\cos\alpha_C + \Gamma_C}\right], \qquad (3.8) \end{aligned}$$

where the terrestrial sunlight-skylight ratio is known from the selection of a pair of points along a shadow boundary and Equation (3.3). The scaling factor in this case is wavelength dependent and must be applied on each spectral channel independently. This means that errors in the calculation of the terrestrial sunlight-skylight ratio will negatively affect the relighting process. These errors may arise due to an incorrect assumption of diffusivity, negligible indirect illumination or no emitted radiance.

Therefore, Equations (3.6) and (3.8) show that all pixels can be relit to using a common illumination source (skylight), through the calculation of scaling factors per pixel. As an example of the effectiveness of the algorithm, Figure 3.3 presents the results obtained by relighting Dataset 7 using the relighting algorithm. Dataset 7 presents a complex scene structure, whose image is captured by a consumer grade RGB camera under clear sky conditions. There is no ground truth data available in this scene, so only qualitative evaluation may be performed. In order to apply the relighting equations to a consumer grade RGB camera, each channel is treated as having a Dirac delta sensor response and linear images are used.

The relit image is obtained using the *full* sky factor calculation and demonstrates that a high degree of spatial illumination invariance is achieved when compared to the original image. This is seen by comparing Figures 3.3c and 3.3d, which are regions of uniform material in the original and relit image highlighted by the red



(a) Original image.

(b) Relit image.

(c) Cropped region of uniform material from the original image.

(d) Cropped region of uniform material from the relit image.

Figure 3.3 – Relighting Dataset 7, captured using a consumer grade RGB camera under clear sky conditions. The regions highlighted by the red circles indicate regions of approximately uniform material exposed to different amounts of illumination. Relighting the image is seen to increase the spatial illumination invariance properties of the image.

circles in Figures 3.3a and 3.3b respectively. The intensity and colour of the brown brick varies in the original image as one side of the building is occluded from the sun. Through relighting, all sides of the building appear similarly coloured and of equal intensity, thereby indicating higher spatial illumination invariance. Some artefacts are noticeable in the relit image, such as bright blue colours on the edge of the roof and front of the building. This is due to errors in the calculation of the surface normal, leading to the incorrect scaling factor being applied during relighting.

Skylight, instead of terrestrial sunlight, is chosen as the common illuminant source

because only the non-occluded scaling factor is wavelength dependent. The scaling factor for the occluded region is a constant, which is advantageous as these regions typically have a low SNR. If they were to be multiplied by a wavelength dependent scaling factor (as is the case when relighting with respect to full exposure to sunlight and skylight as shown in Appendix B) it would amplify the noise, especially in the case of hyperspectral imagery. This is compared to consumer grade cameras which integrate spectral noise through the use of a wideband sensor response for each colour channel.

The drawback of using skylight as the common illuminant is that the resultant image will have a lower dynamic range than the original image. However, the relit image can be linearly scaled without loss of invariance, so this does not pose a problem.

Compared to the shadow compensation method of [29], the proposed approach does not require the selection of a known number of materials under different illumination conditions. This selection is required as optimisation is used to estimate both the terrestrial sunlight-skylight ratio, and the function $\rho(\lambda)E_{sky}(\lambda)$ for each selected material. The resultant relit image amplifies noise at longer wavelengths, while the proposed method does not have this property since relighting is performed with respect to the weaker illumination source.

A second advantage of the method presented in this thesis is that it does not require the use of highly parametrised atmospheric models such as MODTRAN [5] whose parameters are not available at all locations, and only needs a single image compared to the multiple image requirement of [79]. The state of the art illumination invariant image generation method of [25] has recently been applied to robotics applications [13], however this suffers from the fact that it loses the discriminative information of colour once it projects down to a single dimension [80] and also amplifies noise [51]. The proposed method retains the full spectral dimensionality of the original data and keeps noise amplification to a minimum.

3.3 Temporal Illumination Invariance

Spatial illumination invariance was achieved in Section 3.2 by relighting the scene with respect to diffuse skylight. However, when subsequent images of the scene are taken at different times and the scene is relit, identical materials may appear different if the skylight spectra has changed in colour or intensity. This is due to the relit image being dependent on both the material and illumination source as shown in Equations (3.5) and (3.7). The illumination spectra can vary due to changes in either the weather conditions or sun position over time. Different weather conditions cause variations in the transmittance and diffuse skylight spectra, while the movement of the sun affects the visibility and sun angle terms as shown in the outdoor illumination model (Equation 2.16). An example of the temporal factors affecting the colour and intensity of the illuminant is two images taken in clear and overcast conditions. In the first image, the diffuse skylight illumination source will be skewed to short wavelengths and have a high intensity. This is compared to the gray, low intensity skylight spectra present in the second image.

In this section, a method to compensate for temporal shifts in illumination is proposed. This technique builds upon the single image approach by relighting each image independently and then determining the scaling factors between the different illuminants through the use of an overlapping region. There is no constraint on the temporal scales involved in this process, with the physically based approach catering for variations on the order of hours, days and years. The purpose of temporal illumination invariance is to achieve consistency between images, allowing for identical materials to maintain their appearance throughout the dataset.

For this scenario, it is assumed that each image can be made illumination invariant using the method described in Section 3.2. This means that shadows are present in a way that initialisation points can be selected. A small overlapping region is required in order to calculate the scaling factors between each image as shown in Figure 3.4, where the point D is common in images taken at times j and k.

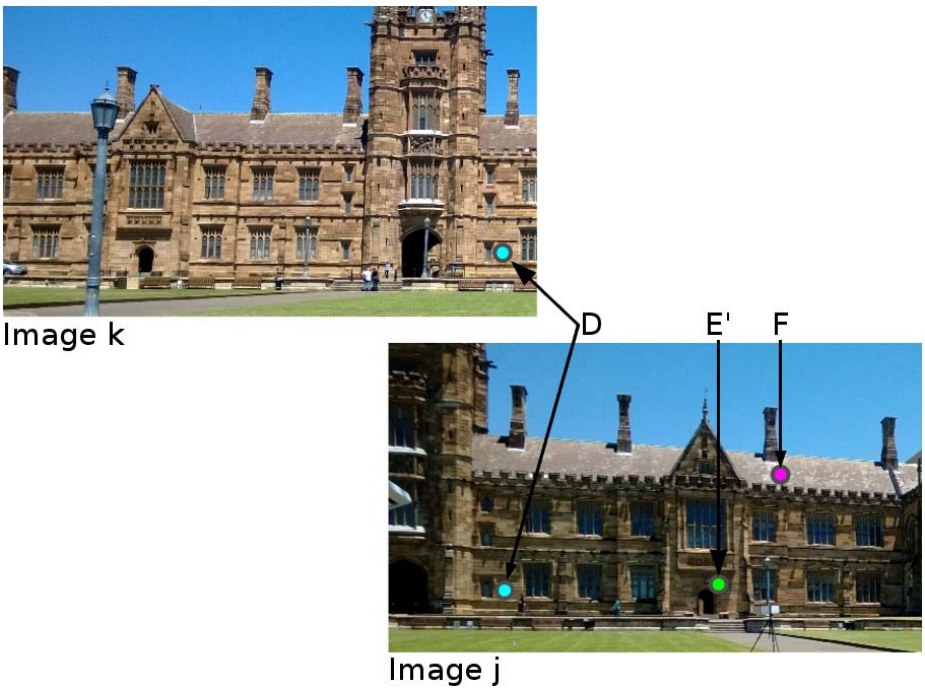


Figure 3.4 – The multi-image scenario consists of images acquired at two different times and places (optionally). Changes in weather conditions cause the colour and intensity of the illumination sources to vary, so performing illumination invariance on each image individually will not yield consistent results.

3.3.1 Derivation

The multi-image scenario is shown in Figure 3.4, where two overlapping images taken at times j and k are captured, and the aim is to generate a large scale map. In this scenario, the sun angle and visibility, as well as the transmittance, skylight and radiance spectra are temporally dependent. This dependence is included in the outdoor illumination model through the inclusion of subscripts denoting the time i at which the image was captured:

$$L_{M,i}(\lambda) = \frac{\rho_M(\lambda)}{\pi}[V_{M,i}E_{sun}(\lambda)\tau_i(\lambda)\cos\alpha_{M,i} + \Gamma_M E_{sky,i}(\lambda)], \tag{3.9}$$

where M is an arbitrary point in the scene, and the sky factor and reflectance is temporally independent as the scene structure and composition is static.

In the following derivation of the required scaling factors, points in image j will be

relit with the skylight spectra present in image k. The radiance of a point E' in image j that is occluded from the sun is given by:

$$L_{E',j}(\lambda) = \frac{\rho_{E'}(\lambda)}{\pi}[\Gamma_{E'} E_{sky,j}(\lambda)]. \tag{3.10}$$

To relight an occluded point with respect to $E_{sky,k}$, the radiance must be multiplied by a scaling factor:

$$\begin{aligned}
L_{E'-relit,k}(\lambda) &= \frac{\rho_{E'}(\lambda)}{\pi}[E_{sky,k}(\lambda)], \\
&= \frac{\rho_{E'}(\lambda)}{\pi}[E_{sky,k}(\lambda)] \left[\frac{\Gamma_{E'} E_{sky,j}(\lambda)}{\Gamma_{E'} E_{sky,j}(\lambda)}\right], \\
&= \frac{\rho_{E'}(\lambda)}{\pi}[\Gamma_{E'} E_{sky,j}(\lambda)] \left[\frac{1}{\Gamma_{E'}}\right] \left[\frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)}\right], \\
&= L_{E',j}(\lambda) \left[\frac{1}{\Gamma_{E'}}\right] \left[\frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)}\right], \\
&= L_{E'-relit,j}(\lambda) \left[\frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)}\right].
\end{aligned} \tag{3.11}$$

Therefore, the scaling factor consists of a geometric term (the inverse sky factor) and the ratio between the skylight spectra present in both images (the temporal skylight scaling factor).

The radiance of a point F in image j that is exposed to terrestrial sunlight is given by:

$$L_{F,j}(\lambda) = \frac{\rho_F(\lambda)}{\pi}[E_{sun}(\lambda)\tau_j(\lambda)\cos\alpha_F + \Gamma_F E_{sky,j}(\lambda)]. \tag{3.12}$$

To relight this point with respect to $E_{sky,k}(\lambda)$, the radiance must be multiplied by a scaling factor:

$$\begin{aligned}
L_{F-relit,k}(\lambda) &= \frac{\rho_F(\lambda)}{\pi}[E_{sky,k}(\lambda)], \\
&= \frac{\rho_F(\lambda)}{\pi}[E_{sky,k}(\lambda)] \left[\frac{E_{sun}(\lambda)\tau_j(\lambda)\cos\alpha_F + \Gamma_F E_{sky,j}(\lambda)}{E_{sun}(\lambda)\tau_j(\lambda)\cos\alpha_F + \Gamma_F E_{sky,j}(\lambda)}\right], \\
&= \frac{\rho_F(\lambda)}{\pi}[E_{sun}(\lambda)\tau_j(\lambda)\cos\alpha_F + \Gamma_F E_{sky,j}(\lambda)]
\end{aligned}$$

$$\left[\frac{E_{sky,k}(\lambda)}{E_{sun}(\lambda)\tau_j(\lambda)\cos\alpha_F + \Gamma_F E_{sky,j}(\lambda)}\right],$$
$$= L_{F,j}(\lambda)\left[\frac{1}{\frac{E_{sun}(\lambda)\tau_j(\lambda)}{E_{sky,j}(\lambda)}\cos\alpha_F + \Gamma_F}\right]\left[\frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)}\right],$$
$$= L_{F-relit,j}(\lambda)\left[\frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)}\right]. \tag{3.13}$$

Here, the scaling factor consists of the terrestrial sunlight-skylight ratio present in image j, which is known due to the assumption that initialisation points are available in both images, and the temporal skylight scaling factor. Equations (3.11) and (3.13) show that in order to relight occluded and non-occluded points in image j with respect to $E_{sky,k}(\lambda)$, the temporal skylight scaling factor must be known.

3.3.2 Temporal Skylight Scaling Factor

The temporal skylight scaling factor is obtained using a region of overlap between the two images. In the case where the scene geometry is obtained via two separate scans, the two point clouds captured at times j and k are registered together. This can be performed using algorithms such as Iterative Closest Point (ICP)[6], or through the use of accurate navigation sensors to generate a large scale geo-registered map. The use of multiple point clouds can aid in achieving a more complete geometric understanding of the scene, especially if the scans are taken from different positions. For example, parts of an object may be visible in one scan, but occluded in the other.

Overlapping points between the images are then found by projecting the point cloud of image j, into the image space of image k and determining which points are within the boundaries of the image. These points will therefore be assigned observations from both images and can be used to determine the temporal skylight scaling factor. Four situations arise when using a point within the overlapping region to calculate the scaling factor and this is based on whether it is occluded or exposed to terrestrial sunlight in both images.

In the first case to be considered, the point D within the overlap is occluded from

the sun in both images. Therefore, the radiance can be described by:

$$L_{D,j}(\lambda) = \frac{\rho_D(\lambda)}{\pi}[\Gamma_D E_{sky,j}(\lambda)], \tag{3.14}$$

$$L_{D,k}(\lambda) = \frac{\rho_D(\lambda)}{\pi}[\Gamma_D E_{sky,k}(\lambda)], . \tag{3.15}$$

Dividing Equation (3.15) by Equation (3.14) means that the temporal skylight scaling factor can be calculated as:

$$\frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)} = \frac{L_{D,k}(\lambda)}{L_{D,j}(\lambda)}. \tag{3.16}$$

When the point D is occluded in image j and exposed to terrestrial sunlight in image k, the radiance can be described as:

$$L_{D,j}(\lambda) = \frac{\rho_D(\lambda)}{\pi}[\Gamma_D E_{sky,j}(\lambda)], \tag{3.17}$$

$$L_{D,k}(\lambda) = \frac{\rho_D(\lambda)}{\pi}[E_{sun}(\lambda)\tau_k(\lambda)\cos\alpha_{D,k}(\lambda) + \Gamma_D E_{sky,k}(\lambda)]. \tag{3.18}$$

Dividing Equation (3.18) by Equation (3.17) gives:

$$\begin{aligned}
\frac{L_{D,k}(\lambda)}{L_{D,j}(\lambda)} &= \frac{E_{sun}(\lambda)\tau_k(\lambda)\cos\alpha_{D,k} + \Gamma_D E_{sky,k}(\lambda)}{\Gamma_D E_{sky,j}(\lambda)}, \\
&= \frac{E_{sun}(\lambda)\tau_k(\lambda)}{E_{sky,j}(\lambda)}\left[\frac{\cos\alpha_{D,k}}{\Gamma_D}\right] + \frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)}, \\
&= \frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)}\left[\frac{E_{sun}(\lambda)\tau_k(\lambda)}{E_{sky,k}(\lambda)}\left(\frac{\cos\alpha_{D,k}}{\Gamma_D}\right) + 1\right].
\end{aligned}$$

Solving for the temporal skylight scaling factor yields:

$$\frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)} = \frac{\frac{L_{D,k}(\lambda)}{L_{D,j}(\lambda)}}{\frac{E_{sun}(\lambda)\tau_k(\lambda)}{E_{sky,k}(\lambda)}\left[\frac{\cos\alpha_{D,k}}{\Gamma_D}\right] + 1}, \tag{3.19}$$

where the terrestrial sunlight-skylight ratio in image k is already known since it is assumed that the single image system can be initialised.

It is trivial to show that for the case when D is exposed to terrestrial sunlight in

image j and occluded in image k, that the temporal skylight scaling factor is:

$$\frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)} = \frac{\frac{E_{sun}(\lambda)\tau_j(\lambda)}{E_{sky,j}(\lambda)}\left[\frac{\cos\alpha_{D,j}}{\Gamma_D}\right]+1}{\frac{L_{D,j}(\lambda)}{L_{D,k}(\lambda)}}. \tag{3.20}$$

In the final case when D is exposed to terrestrial sunlight in both images, the radiance is described by:

$$L_{D,j}(\lambda) = \frac{\rho_D(\lambda)}{\pi}[E_{sun}(\lambda)\tau_j(\lambda)\cos\alpha_{D,j} + \Gamma_D E_{sky,j}(\lambda)], \tag{3.21}$$

$$L_{D,k}(\lambda) = \frac{\rho_D(\lambda)}{\pi}[E_{sun}(\lambda)\tau_k(\lambda)\cos\alpha_{D,k} + \Gamma_D E_{sky,k}(\lambda)], . \tag{3.22}$$

Dividing Equation (3.22) by Equation (3.21) gives:

$$\begin{aligned}\frac{L_{D,k}(\lambda)}{L_{D,j}(\lambda)} &= \frac{E_{sun}(\lambda)\tau_k(\lambda)\cos\alpha_{D,k} + \Gamma_D E_{sky,k}(\lambda)}{E_{sun}(\lambda)\tau_j(\lambda)\cos\alpha_{D,j} + \Gamma_D E_{sky,j}(\lambda)}, \\ &= \frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)}\left[\frac{\frac{E_{sun}(\lambda)\tau_k(\lambda)}{E_{sky,k}(\lambda)}\cos\alpha_{D,k} + \Gamma_D}{\frac{E_{sun}(\lambda)\tau_j(\lambda)}{E_{sky,j}(\lambda)}\cos\alpha_{D,j} + \Gamma_D}\right].\end{aligned} \tag{3.23}$$

Solving for the temporal skylight scaling factor gives:

$$\frac{E_{sky,k}(\lambda)}{E_{sky,j}(\lambda)} = \frac{L_{D,k}(\lambda)}{L_{D,j}(\lambda)}\left[\frac{\frac{E_{sun}(\lambda)\tau_j(\lambda)}{E_{sky,j}(\lambda)}\cos\alpha_{D,j} + \Gamma_D}{\frac{E_{sun}(\lambda)\tau_k(\lambda)}{E_{sky,k}(\lambda)}\cos\alpha_{D,k} + \Gamma_D}\right], \tag{3.24}$$

where the terrestrial sunlight-skylight ratios for both images are known.

Therefore, given a point in the region of overlap, Equations (3.16), (3.19), (3.20) and (3.24) can be used to estimate the temporal skylight scaling factor required to perform multi-image illumination invariance. Given multiple points in the overlapping region, a number of estimates can be generated and the average scaling factor can be used to reduce the influence of noise.

However, it is noted that the best estimate of the temporal skylight scaling factor will most likely come from the case where the overlapping point is in shadow in both images. This is because the calculation of the scaling factor is independent of the

terrestrial sunlight-skylight ratio, sun angle and sky factor terms. Any errors in the calculation of these terms would subsequently produce noisy estimates of the temporal skylight scaling factor.

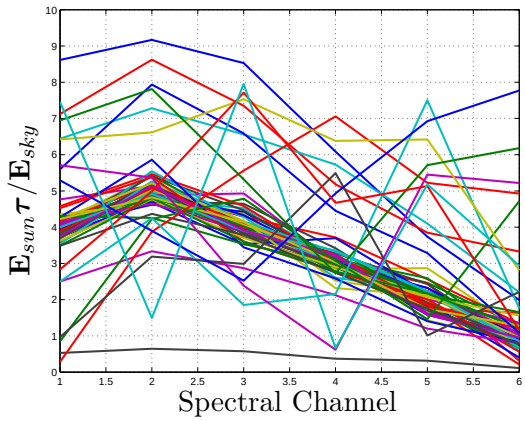
Compared to the IRMs used in [79], the proposed multi-image relighting method is more robust to the scene complexity. The IRMs are calculated for the different orientations within the overlapping region, and if a new orientation is found outside that region, then their method is unable to perform relighting. Since orientation is explicitly accounted for through the relative sun angle term, the proposed method does not face this restriction.

In this section, a system to achieve illumination invariance via accurate modelling of the physical processes involved was developed that takes into account temporal changes in illumination colour and intensity. This is performed through the use of an overlapping region, allowing each scene to be relit with respect to the diffuse skylight illuminant present in one of the images. In order to generate both spatially and temporally illumination invariant images, the terrestrial sunlight-skylight ratio must be known. While this could be estimated from user selected points, an automatic process is desired as it has the potential to reduce errors and allows the system to be easily integrated into online applications.

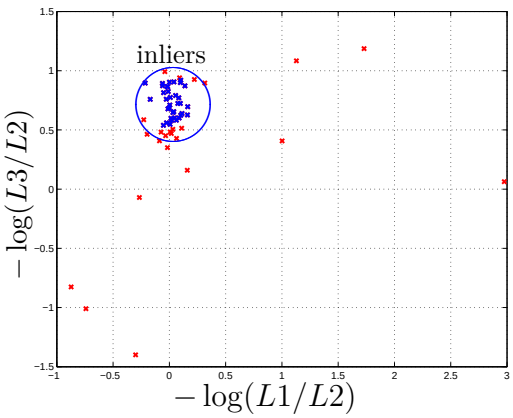
3.4 Automatic Initialisation

A requirement of the illumination invariant system proposed in this thesis is that it should be independent of user input. The reliance on a user to input initialisation points in [13] and [29] is undesirable as it may lead to the selection of points that are not from the same material, and is also impractical for large scale mapping applications. This leads to incorrect estimation of the illumination sources and therefore, negatively impact upon the performance of these illumination invariance algorithms. This section proposes an automatic initialisation method for estimating the terrestrial sunlight-skylight ratio in the scene.

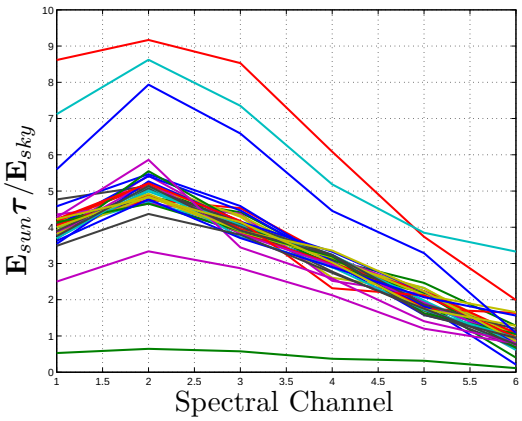
The terrestrial sunlight-skylight ratio (Equation 3.3) is seen to be material independent, which allows multiple point pairs to be used to generate the estimate, as long as the points in each pair are obtained from the same material. The proposed method consists of three stages that filter out the estimates from the point pair candidates based on spectral shape and intensity. A toy example is shown in Figure 3.5 for a hyperspectral camera with six spectral channels.



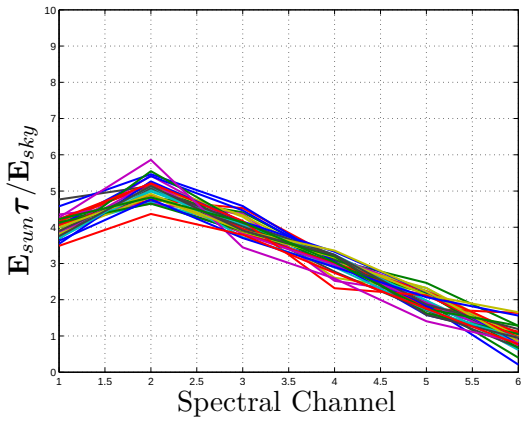
(a) First stage - Point pair candidates from along the shadow boundary generate a number of estimates.



(b) Log-chromaticity plot of the first stage (red) and the second stage (blue). (L_1, L_2, L_3) are chosen as spectral channels 1, 3 and 5 respectively.



(c) Second stage - Refining the estimates from the first stage based on the spectral shape.



(d) Third stage - Refining the estimates from the second stage based on the spectral intensity.

Figure 3.5 – At each stage, the terrestrial sunlight-skylight ratio estimate is refined based on spectral shape and intensity.

3.4.1 Stage 1 - Candidate Selection

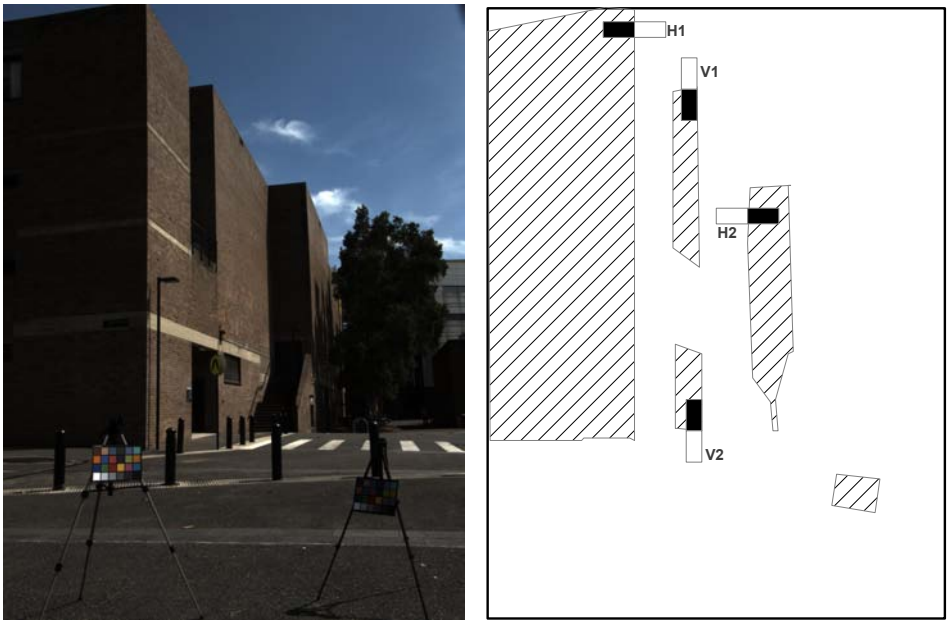
The initial stage of automatic initialisation is candidate selection. The requirement to calculate the terrestrial sunlight-skylight ratio using two points is that they be obtained from the same material. This motivates the use of a spatial smoothness assumption of materials in the scene, which means that adjacent pixels are highly likely to be from the same material. This is true for the majority of pixels in the scene [2]. The shadow boundaries are therefore analysed in order to find two adjacent pixels with one occluded from terrestrial sunlight.

In the first stage of automatic initialisation, a large number of potential candidates are obtained by analysing the shadow map as shown in Figure 3.6. Four binary filters are convolved over the shadow map in order to extract the shadow boundaries in the scene. The candidate pairs selected will include pairs from different materials, and this will produce incorrect estimates of the terrestrial sunlight-skylight ratio. This is shown in Figure 3.5a, where a large number of candidates are selected and the terrestrial sunlight-skylight ratio is estimated. Point pairs that are selected from uniform materials will produce ratio estimates that are close together, while those that are selected from different materials will produce inconsistent estimates.

3.4.2 Stage 2 - Spectral Shape Refinement

The second stage rejects outliers from the previously selected candidates based on spectral shape. In order to enforce consistency, three spectral channels (L_1, L_2, L_3) are initially selected with an example of this being shown in Table 3.1 for different camera types. The selection of bands is fairly robust as long as the bands are not within the destructive water absorption bands in which the SNR is extremely small.

The ratio between the different spectral channels is used to generate a $2D$ representation of each potential candidate $(\frac{L_1}{L_2}, \frac{L_3}{L_2})$. Rejecting outliers is performed by taking the negative logarithm of each component and applying Random Sample Consensus (RANSAC) independently to each of the two dimensions, to fit a zero-order



(a) Original scene. **(b)** Shadow map and binary filters used to extract candidates.

Figure 3.6 – Automatic initialisation stage 1. The shadow map is filtered using four binary filters (H1, H2, V1, V2) to extract the edges and generate point candidates either side of the shadow boundary.

Table 3.1 – Suggested wavelengths/bands for various camera types for automatic initialisation that lie outside destructive water absorption wavelengths.

	L_1	L_2	L_3
RGB	R	G	B
VNIR	400nm	550nm	600nm
SWIR	1073nm	1243nm	1319nm

polynomial model. Those points that are inliers in both dimensions are selected for further refinement in the third stage. The advantage of using the logarithm space is that it allows a greater degree of separation between the candidates. Figure 3.5b shows that the outlier rejection step successfully removes inconsistent candidates in the log-chromaticity space, with the remaining estimates having similar shape to one another as seen in Figure 3.5c.

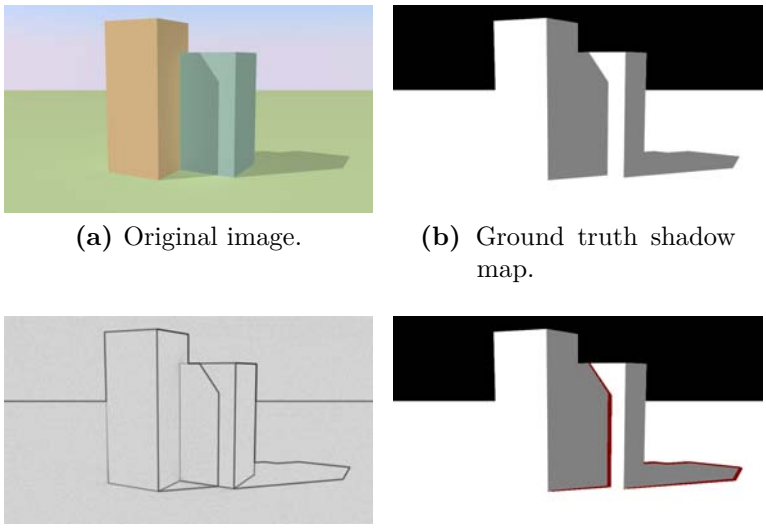
3.4.3 Stage 3 - Spectral Intensity Refinement

While the candidates that reach this stage possess similar spectral ratios, they will occur at different intensities. Therefore, another outlier rejection stage is required that refines the candidates based on intensity. Similarly to the second stage, the negative logarithms of the terrestrial sunlight-skylight ratio are analysed using the RANSAC algorithm operating independently on each band to fit a zero-order polynomial model. The common inliers are selected in order to finalise the selection of initialisation points as seen in Figure 3.5d.

While Equation (3.3) implies material independence, in practice this is not always the case. Factors such as the low reflectance of a material affect the signal-to-noise ratio of the image data, while large angles between the sun and the surface normal mean that the denominator becomes close to zero. Therefore, there is a reliance on the algorithm identifying multiple initialisation points of various orientations and material types. The mean of the estimates is used to develop the final estimation of the terrestrial sunlight-skylight ratio.

In this section, a three stage method for automatic initialisation of the proposed for achieving spatial illumination invariance was developed. The technique generates a large number of candidates to estimate the terrestrial sunlight-skylight ratio. This is followed by a two stage filtering process based on their spectral shape and intensity.

The first stage for generating candidates requires an accurate shadow map of the scene in order to determine the shadow boundaries. While the utilisation of a physical model based method means ray tracing to the known sun position can give reasonable shadow detection results, there are scenarios where this does not suffice. For instance, situations may arise where the complete geometry of the scene is unknown (e.g. when only the front surface of the scene is obtained via a LIDAR scan), or the point cloud is geo-registered incorrectly due to sensor noise. Therefore, further refinement of the shadow map is required and a graphical model approach to this is presented in Section 3.5. Developing a more accurate shadow map has the effect of increasing the chance of correctly selecting pixel either side of the shadow boundary in stage 1 of



(a) Original image.

(b) Ground truth shadow map.

(c) Pairwise potential weights.

(d) Initial shadow map estimate. Misalignment is shown in red.

Figure 3.7 – Shadow smoothing process.

the proposed automatic initialisation technique, and reduces shadow boundary edge artefacts from arising during relighting.

3.5 Shadow Detection

Identifying whether a pixel is exposed or occluded from the sun is required for the first stage of the automatic initialisation process described in Section 3.4.1 and when determining which scaling factors to use when relighting an image. A simple, ray tracing based approach can be inaccurate due to phenomena such as navigational sensor noise leading to incorrect geo-registration, or poor surface reconstruction. This means the predicted shadows may not be aligned with the image. Therefore, a graphical model initialised using ray tracing is used to develop a shadow map.

Shadow detection using purely image based data faces a number of challenges. From the outdoor illumination model (Equation 2.16), the pixel intensity captured by a camera is dependent on the material and incident illumination within the scene. The influence of terrestrial sunlight on pixel intensity is governed by (i) whether or not

the corresponding point in the scene has line-of-sight visibility to the sun, and (ii) the cosine of the angle between the surface normal and the vector in the direction of the sun position. Diffuse skylight also influences pixel intensity, especially on overcast days when there is a large amount of attenuation of terrestrial sunlight. The combination of these illumination factors with the material reflectance means that applying a simple threshold on the pixel intensities is inadequate to determine regions of shadows. In the scenario where the sensor platform contains a multi-modal sensor suite, combining the information from various sensors enables shadow estimation to operate in a principled manner. The main assumption that is required for the proposed technique to work well is that the initial estimate for the shadow map is reasonably accurate and only requires slight tweaking. In the ideal case, the LIDAR sensor and the camera are close to each other so that all points observed by the camera are also associated with a range measurement. In practice, there are typically some parts of the scene that are visible in the image but occluded from the LIDAR sensor, but as long as there are no new material classes introduced in these regions, then the smoothing technique can still operate as expected.

3.5.1 Shadow Map Smoothing

Surface reconstruction methods for dense point clouds often require manual tuning in order to fill holes in the mesh or remove erroneous faces. Methods such as the disc approximation used in [62] contain holes in the surface representation due to the estimation of the size of each disc. Therefore, ray tracing to determine shadowed regions will be inaccurate and further post-processing of the shadow map is required. A second source of error in the shadow map is induced by incomplete sensing of the geometry of the scene. This can occur when the LIDAR is occluded from certain regions in the scene. Finally, noise in the navigational sensors may cause incorrect geo-registration of the point cloud, meaning that shadow boundaries in the image are not aligned with the shadow map.

Image processing techniques such as morphological operators, diffusion and smooth-

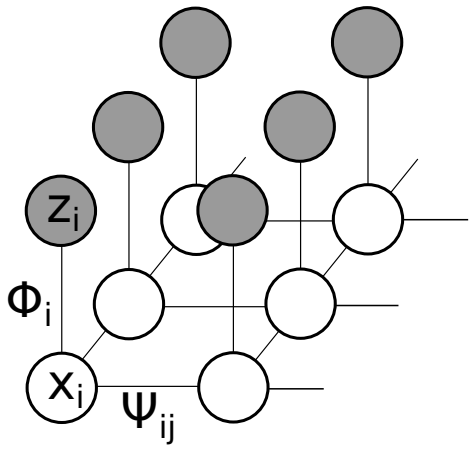


Figure 3.8 – A grid structured graphical model is used to smooth the shadow map.

ing can be used to reduce the noise in the shadow map, but these require tuning parameters that change for each scene. In this thesis, the use of a Conditional Random Field (CRF) is proposed, in order to smooth the shadow map for use in the illumination invariant system.

The model has an undirected grid structure as shown in Figure 3.8. In this model, each pixel in the image is represented by a node in the graph. The underlying state $\mathbf{x} = (x_1, ..., x_k, ...x_n)$ represents the label of the pixel, where $x_k \in \{\text{occluded}, \text{exposed}, \text{no return}\}$. The '*occluded*' and '*exposed*' labels indicate whether the pixel is in shadow or in sunlight respectively. The '*no return*' state indicates areas in the image where no LIDAR return was observed and typically represents areas of sky and far-field objects that do not require processing. The observations in the model $\mathbf{z} = (z_1, ..., z_i, ...z_n)$ are obtained by ray tracing from the $3D$ point corresponding to each pixel to the sun position and checking for occlusions $z_i \in (\text{occluded}, \text{exposed})$.

Unary Potential The observation likelihoods $P(z_i|x_i)$ are represented by unary potentials for each node. As opposed to other methods that require user input in the form of bounding boxes or strokes to select foreground/background areas, the proposed method utilises information from the ray traced shadow map. Once again, the complementary nature of the different sensors provides valuable information. Intensity histograms h are developed for each label $\omega = \{h_{occluded}(z_i), h_{exposed}(z_i), h_{no\ return}(z_i)\}$

as per [9]. The histograms are normalised to sum to one and the unary potential is calculated as:

$$\phi_i(x_i, z_i) = -\log(h_{x_i}(z_i)). \tag{3.25}$$

This entire process is independent of the user and the initial trimap does not place any constraints on the final label of a node, allowing them to switch labels if they fit the model.

Pairwise Potential Spatial consistency is enforced in the graphical model through the inclusion of pairwise potentials. This is used so that regions of similar intensity in the image have similar assigned labels. Near regions of high contrast such as edges where illumination changes are present, the spatial consistency constraint is relaxed allowing the model to utilise the information from the unary potential instead.

To calculate the edges in the scene, a standard deviation filter *sdf* is used using horizontal and vertical regions depending on the direction of the link in the graphical model. This edge detection procedure tends to be more robust than Canny or Sobel edge detectors as it does not rely on thresholds. The pairwise potential is given by:

$$\psi_{i,j}(x_i, x_j, z_i, z_j) = \begin{cases} (1 - \sqrt{sdf})^3 & \text{if} \quad x_i \neq x_j, \\ 0 & \text{otherwise}, \end{cases} \tag{3.26}$$

with the square root and cubic functions accentuating the boundaries.

Inference The total energy of the system is given by:

$$E(\mathbf{x}, \mathbf{z}) = \sum_i \phi_i(\mathbf{x}, \mathbf{z}) + \theta \sum_{(i,j) \in N} \psi_{ij}(\mathbf{x}, \mathbf{z}), \tag{3.27}$$

where $\mathcal{N}$ is the set of neighbouring pixels in the image and θ is the weighting value dictating the degree of smoothness. The solution to the labelling problem can be solved by minimising the energy function using many different inference techniques, with graph cuts [9] empirically shown to be an effective method.

The key parameter to tune is θ and this will vary based on the complexity and nature of the scene. Therefore, optimisation of the parameters occurs by minimising a cost function. This has the benefit of not requiring labelled training data to find the best value, while also allowing for flexibility between different datasets.

The edge map $\mathcal{E}_0$ acquired by applying edge detection techniques such as Canny and Sobel on the original image will identify edges that occur due to sharp changes in material or incident illumination, and is treated as a ground truth map. For a specific θ value, the edge map $\mathcal{E}_{\hat{x},\theta}$ is determined by running edge detection on the labelling $\hat{\mathbf{x}}$. Ideally, all edges that occur in $\mathcal{E}_{\hat{x},\theta}$ should match those present in $\mathcal{E}_0$, so the cost function is developed to encourage true positives, while penalising false positives:

$$c(\hat{\mathbf{x}}, \theta) = |\mathcal{E}_0 \cap \mathcal{E}_{\hat{\mathbf{x}},\theta}| - |\mathcal{E}_{\hat{\mathbf{x}},\theta} \setminus \mathcal{E}_0|. \tag{3.28}$$

Penalising false positives is required as without this term, the best cost value that could be achieved would be to label all pixels as edges.

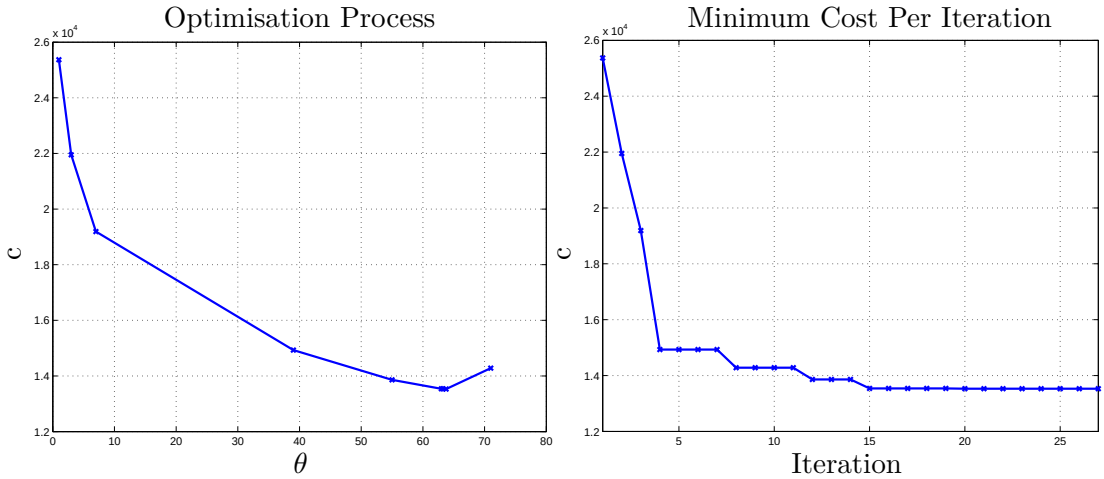
To determine the optimal weighting value $\hat{\theta}$, the *patternsearch* optimisation algorithm implemented in MATLAB with a lower bound of zero and no upper bound is used, in an attempt to find:

$$\hat{\theta} = \arg\min_{\theta}[-c(\hat{\mathbf{x}}, \theta)]. \tag{3.29}$$

An example of this process is shown in Figure 3.9. This optimisation technique is appealing as it does not require gradients of the cost function to be calculated. The cost function is evaluated for a number of different parameter values within a mesh around the current optimal value and at each iteration, the mesh decreases in size. The use of an optimisation based approach to determine the tuning parameters allows the shadow detection technique to be robust to the complexity of the scene.

3.6 Fast Sky Factor Approximation

The computational complexity in the *full* method (see Section 2.1.2) for calculation of sky factors lies in determining the occlusions in the scene, which depends on scene



(a) In this example, the best minimum cost value is found at $\theta = 63.75$.

(b) The improvement in the cost metric becomes smaller as the process approaches the parameter value that minimises the metric.

Figure 3.9 – Example of the optimisation process for calculating $\hat{\theta}$.

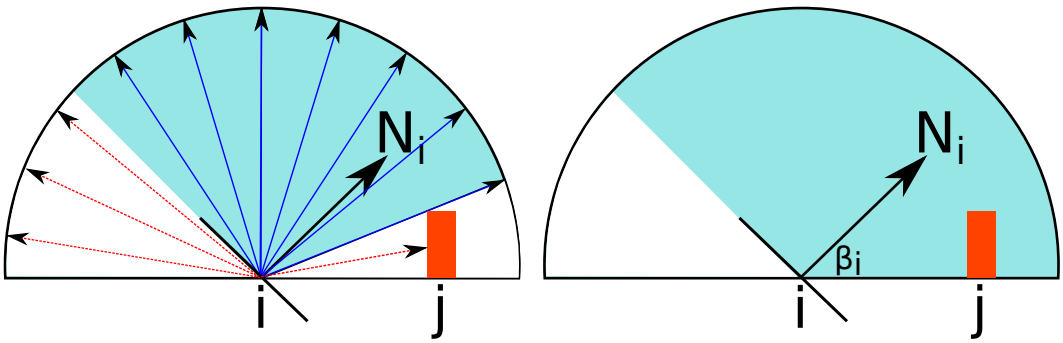
complexity. For remote sensing contexts this is not an issue, but for robotics scenarios this is a major bottleneck that prevents the *full* calculation from being used. This thesis proposes a *fast* approximation of sky factors for use in robotics applications. The method neglects occlusions in the scene and uses the orientation of the region in order to generate the sky factor estimate:

$$\Gamma_i = \begin{cases} \dfrac{\frac{\pi}{2} + \beta_i}{\pi} & N_i.z > 0, \\ \dfrac{\frac{\pi}{2} - \beta_i}{\pi} & N_i.z < 0, \end{cases} \tag{3.30}$$

where $N_i.z$ is the z component of the normal at region i and β_i (in radians) is the angle between the normal in and the ground plane.

In terms of computational saving, the *fast* method runs in $O(1)$ time, compared to $O(kn)$ in the worst case scenario for the *full* method, where k is the number of rays traced and n is the number of regions in the scene. This means that as more data is incorporated into the scene model, the *fast* method can quickly generate an upper bound value for the sky factor value.

The limitation of the *fast* approximation to sky factors is shown in Figure 3.11, which



(a) The *full* calculation requires ray tracing to points on the sky dome and determining whether the ray is occluded or not.

(b) The *fast* method ignores occlusions and uses the surface normal to calculate an upper bound of the sky factor.

Figure 3.10 – Calculation of sky factors using the *full* and *fast* methods.

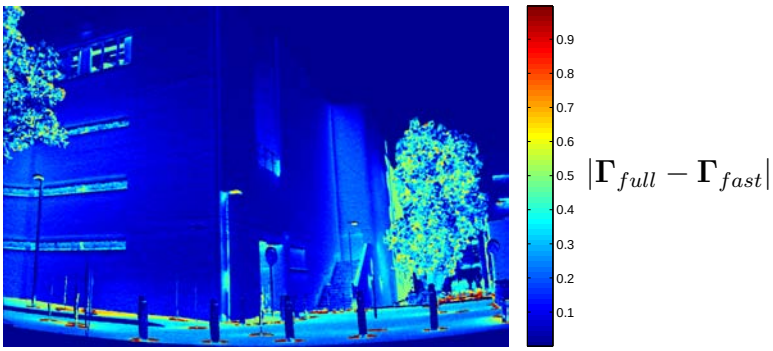


Figure 3.11 – Example illustrating the main regions of error between the *full* and *fast* approximation of sky factors. The accuracy is seen to decrease close to the intersection of orthogonal walls and in foliage.

shows the absolute error of the method (magnitude difference between *full* and *fast* approximations). Close to the intersection of orthogonal structures, the approximation decreases in accuracy as occlusions become more prevalent. An example of this is the ground plane and the vertical wall of the building. Another area of high error is in foliage, where leaves and branches are under sampled by the LIDAR sensor leading to inaccurate representation of the tree and thus, inaccurate normal calculation.

3.7 Summary

Perception systems robust to illumination variation are a vital component of both robotic and remote sensing platforms. They must be robust to illumination variation in the scene which can occur both spatially and temporally.

In this chapter a method for relighting a scene using fused LIDAR and camera data has been presented. An automatic, spatial illumination invariance system using a single image was proposed and this was extended to the multi-image scenario. This allows temporal changes to be accounted for, which can occur due to a change in sun position over time, or a change in weather conditions at the time of data acquisition. For robotics applications a computationally efficient method for approximating sky factors was also developed.

The key advantages of the proposed approach in the single image scenario is that it does not require multiple images, highly parametrised atmospheric models or user selection of initialisation points. The method retains the full dimensionality of the original data, thereby allowing high level algorithms to utilise the full discriminative nature of colour, which is characteristic of the underlying material.

In this chapter, in order to solve the illumination invariance problem the entire scene is relit with respect to a common illuminant. However this illuminant remains unknown as there is insufficient information available to solve for it. In Chapter 4, this method is extended to radiometric normalisation. Through the use of external measurements of the scene illumination, an explicit approximation of the illumination sources can be calculated and the material reflectance can be estimated.

Chapter 4

Illumination Invariant Radiometric Normalisation

This chapter presents a method for illumination invariant radiometric normalisation of data, which converts the pixel intensities to reflectance. Chapter 3 presented a method to generate illumination invariant images by relighting the entire scene with respect to a common, but unknown illuminant. Completely removing the dependence on illumination is ill posed as the pixel intensity equation described in Equation (2.16) has two unknowns; the material and illumination properties of the scene. However, the problem of recovering reflectance estimates becomes well posed if a direct measurement of the illumination is obtained. Reflectance is important for remote sensing as it allows high level algorithms to operate in a supervised manner. For example, the radiometrically normalised data may be compared against laboratory obtained spectral measurements for the purposes of classification.

Common methods for normalisation range from the use of highly parameterised atmospheric radiative transfer models such as MODTRAN [5], to the use of hardware based methods such as calibration boards and downwelling irradiance sensors. In the ideal scenario, a measurement of the incident illumination at each point in the scene would be recorded. Normalisation would then proceed by dividing each pixel by its corresponding measurement. This approach however, is not practical due to the large,

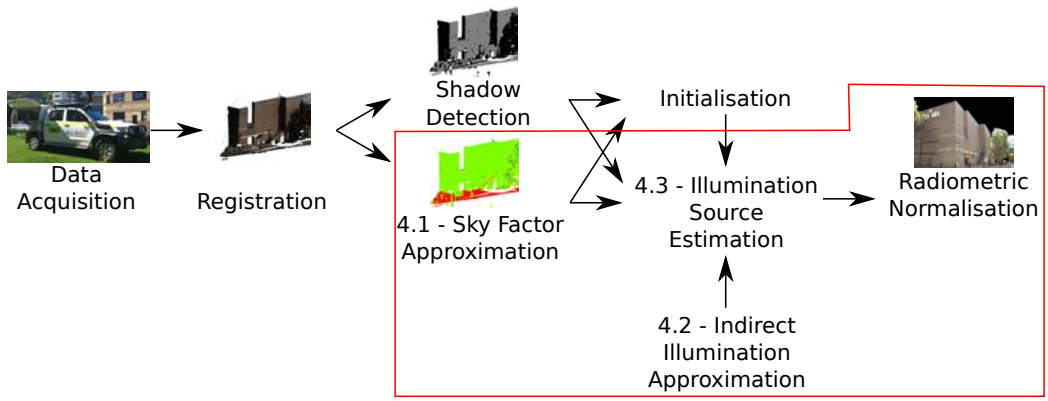


Figure 4.1 – The proposed radiometric normalisation method utilises a physically based model in order to achieve illumination invariance. The terrestrial sunlight, diffuse skylight and optionally, indirect illumination sources are explicitly calculated through the use of direct measurements.

complex and potentially hazardous environments encountered during remote sensing. While computational models such as MODTRAN are useful, the atmospheric conditions present at the time of data acquisition are required to be known, which is not always possible. This is the main motivation behind using fused geometrical and appearance data, and inferring the illumination sources from measurements.

In the remote sensing scenario, computation time is generally not a priority due to the large amounts of data involved. This allows for more precise and higher resolution models to be used, for example, indirect illumination sources can be included. This can be potentially valuable in situations where indirect illumination acts as a strong illumination source. An example of this is within regions of the scene that occluded from the sun. As a proportion of the total incident illumination, indirect sources play a more significant role compared to regions that are exposed to sunlight, which due to its intensity, can dominate the other sources.

The proposed method for radiometric normalisation first estimates the terrestrial sunlight-skylight ratio using the automatic method described in Section 3.4. An in situ measurement of the illumination present in the scene is then taken and this allows the spectra of the illumination sources to be known. Through the utilisation of fused geometry and imagery data, indirect illumination sources can be approximated and accounted for during radiometric calibration.

For large scale, complex scenes this thesis proposes a method for approximating sky factors using a small number of samples in Section 4.1. A summary of the radiosity method for determining indirect illumination and the proposed approximation method is provided in Section 4.2. Finally, the illumination invariant radiometric normalisation method is developed in Section 4.3.

4.1 Large Scale Sky Factor Estimation

Obtaining an estimate for sky factors for each pixel in the image can be performed in a number of ways. A ray tracing method to selected points on the sky dome was used in both [10] and [62], however this is computationally expensive. The need to determine whether each ray is occluded makes the process difficult to use in large scale complex scenes. A *fast* estimate for sky factors was proposed in Section 3.6 for robotics applications, however this has the limitation of ignoring occlusion in the scene. For robotics applications, such an approximation is acceptable as low latency, high accuracy on low cost computers is a high priority requirement, but for remote sensing purposes this requirement can be relaxed. It is still computationally expensive to calculate per-pixel sky factors, so this section proposes a sampling based method that uses loopy belief propagation [53] to approximate sky factors for the entire image from a smaller selection of observations. The proposed method for large scale sky factor approximation is split into an iterative sampling stage, that generates new *full* sky factor estimates at a sparse number of locations, and a smoothing stage that factors in the geometric properties of the scene.

An overview of the approximation approach for an image with N pixels in each spectral channel is presented in Algorithm 1, where:

- i is the iteration number,
- $\mathbf{\Gamma} = \{\Gamma(1), \Gamma(2) \ldots \Gamma(N)\}$ is the true, underlying, but unknown sky factor value,
- $\hat{\mathbf{\Gamma}}_i = \{\hat{\Gamma}_i(1), \hat{\Gamma}_i(2) \ldots \hat{\Gamma}_i(N)\}$ is the estimate of the sky factors at iteration i,

Algorithm 1 The large scale sky factor approximation is split into a sampling and smoothing stage.

```
1. Iterative Sampling
training set T_0 ← ∅
validation set V_0 ← ∅
error ← 1
for i:=1 to maxIterations do
    V_i ← best candidate sampling of image
    T_i ← T_{i-1} ∪ importance sampling of squared error map
    Γ̂_i ← interpolate training set
    calculate error
    T_{i+1} ← T_i ∪ V_i
    if SSIM''(Γ̂_i, Γ_fast) < κ then
        break
    end if
end for
2. Smoothing
Loopy Belief Propagation
```

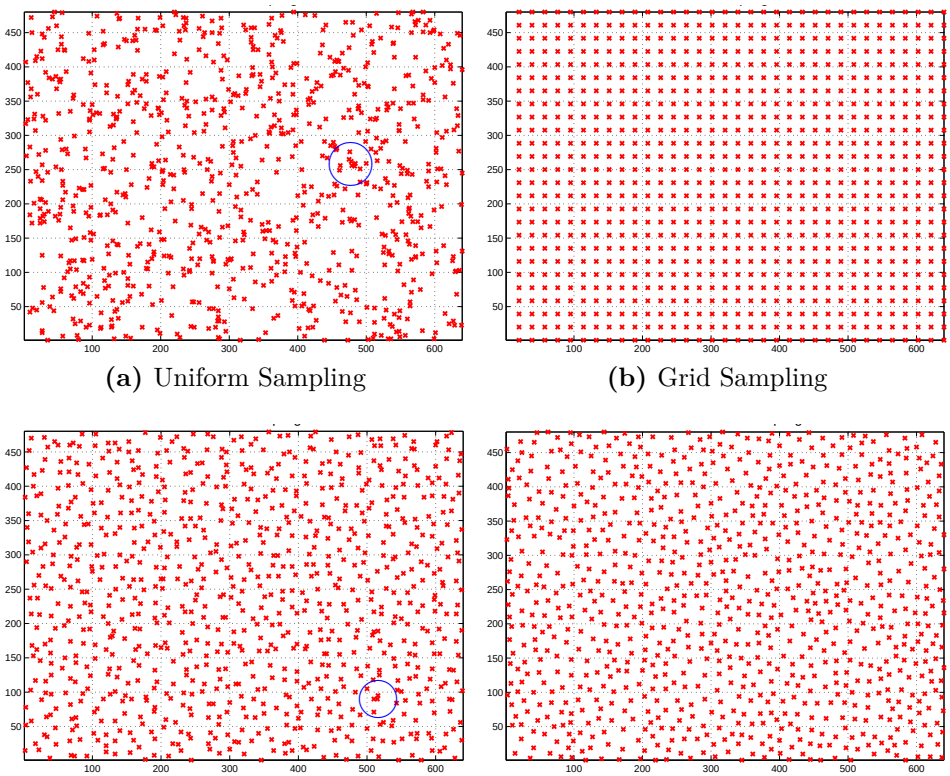
- $\mathbf{\Gamma}_{fast}$ is the *fast* approximation of the sky factors,
- $\mathbf{T}_i$ is the training set at iteration i and is used to generate $\hat{\mathbf{\Gamma}}_i$,
- $\mathbf{V}_i$ is the validation set at iteration i and is used to evaluate $\hat{\mathbf{\Gamma}}_i$ against $\mathbf{\Gamma}$.

4.1.1 Iterative Sampling

The proposed algorithm consists of two sampling procedures, one to generate a validation set, and the other to generate a training set.

Validation Set

The purpose of the validation set $\mathbf{V}_i$ at iteration i, is to evaluate how well the sky factor estimate $\hat{\mathbf{\Gamma}}_i$ fits the underlying function $\mathbf{\Gamma}$. The validation set must therefore provide sufficient spatial sampling of the image and there are a number of different strategies that can be employed. Uniform random sampling of the image space can be implemented by generating a $2D$ random variable, indicating the row and column



(a) Uniform Sampling **(b)** Grid Sampling

(c) Stratified Sampling **(d)** Best Candidate Sampling

Figure 4.2 – Potential sampling methods include uniform, grid, stratified and best candidate sampling. The blue circles in the uniform and stratified sampling cases indicate areas of high localised clustering of samples. Best candidate sampling is seen to have greater spatial sampling properties as it has minimal local clustering and aliasing.

in the image space chosen using a uniform distribution. Uniform random sampling does not guarantee sufficient spatial sampling of the image space, so grid sampling can be used instead. The image is divided into cells of equal width and height, and the sample is positioned at the centre of each cell.

A drawback of grid sampling is that it has a tendency to lead to aliasing effects because the required sampling resolution and frequency content of the scene are dependent on each other [59]. A common method in the computer graphics community for avoiding aliasing is to use stratified sampling. Here, the image is divided into regions (strata), and a sample is obtained using a uniform distribution from each region.

Stratified sampling still has the potential for samples to cluster together near the

edges of adjacent strata as shown in Figure 4.2c. A solution to this is a method known as Mitchell's best candidate sampling [59], which at each creation of a new sample, generates a set of candidates. The distance between each candidate and the closest samples already selected is calculated, and the one with the largest distance is chosen. This leads to a reduced clustering of sample points, while providing sufficient spatial coverage.

Using the best candidate sampling method for generating a validation set allows an error function to be calculated to quantify how well the linearly interpolated estimate $\hat{\mathbf{\Gamma}}_i$ fits the underlying sky factor values. The error metric for each iteration is calculated as:

$$error_i = \frac{100}{N_v} \sum_{j \in V_i} G(j) = \frac{100}{N_v} \sum_{j \in V_i} \frac{|\hat{\Gamma}_i(j) - \Gamma(j)|}{\Gamma(j)}, \tag{4.1}$$

where N_v is the number of validation points used at each iteration, and $\Gamma(j)$ is the sky factor value obtained through the $full$ ray tracing approach.

Training Set

The training set $\mathbf{T}_i$ is used to generate the linearly interpolated estimate $\hat{\mathbf{\Gamma}}_i$ of the sky factor function at each iteration. The sampling procedure targets areas in the image space where the difference between the estimation and the underlying function is large. This is achieved by evaluating $G(j)^2$ at each validation point and linearly interpolating to generate a grayscale image. The training set is generated by sampling N_T points from this image, which will be focussed on regions of high error.

In addition to sampling from the squared error function, the validation points $\mathbf{V}_{i-1}$ used in the previous iteration, are added to the training set $\mathbf{T}_i$. A new validation set $\mathbf{V}_i$ is then generated at each iteration. The intuition behind this is that sampling from the squared error function will yield points close to the validation set and if the set were not changed, new training points would just cluster around the same positions.

Termination Criteria

The Structural Similarity metric (SSIM)[81] is used as a criteria to determine whether the iteration stage should terminate. At each iteration the structural similarity between the current and *fast* approximation of the sky factors is calculated. The *fast* approximation is used as it closely aligns with the geometry of the scene.

A second order differential of the SSIM metric (with respect to the iteration number) is calculated at each iteration, as this allows the algorithm to determine whether the rate of change has plateaued. When this occurs, using additional samples does not improve the similarity between the current and *fast* approximation.

The termination criteria is therefore determined by:

$$SSIM''(\hat{\mathbf{\Gamma}}_i, \mathbf{\Gamma}_{fast}) < \kappa, \tag{4.2}$$

where κ is a user defined threshold and $SSIM''$ is the second derivative of the SSIM metric with respect to the iteration number. The selection of κ dictates how much change is acceptable in order to terminate the sampling procedure. Values close to zero tend to yield reasonable results. Due to the generation of new random samples at each iteration, the metric exhibits variations between iterations, so the mean over three iterations is used to to filter out the fluctuations.

4.1.2 Markov Random Field Smoothing

Calculating *full* sky factors for all pixels in the image is computationally expensive and if the surface representation is incomplete or noisy, then these errors will introduce artefacts in the sky factor estimate. This method also does not take into account the spatial correlations in the scene that exist due to the structure. For most pixels in the scene (with some exceptions), adjacent sky factors will be approximately equal. This spatial correlation between pixels means the approximation problem is well suited for inference on a Markov Random Field (MRF). Other options include the use of anisotropic diffusion and guided image filtering [36], which are edge preserving image

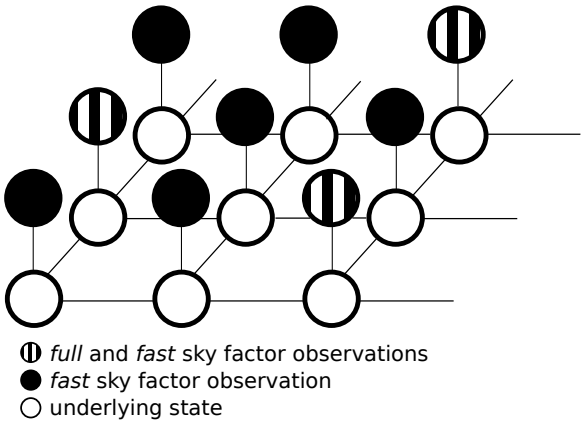


Figure 4.3 – Graph structure used for smoothing the sky factor approximation. Each node represents a pixel and all pixels contain a link to a *fast* estimate of the sky factor. A small subset of nodes also contain *full* estimates as given by the iterative sampling approach.

smoothing algorithms. However, the ability to define the graph structure and weights between each node in an MRF based on the data from different sensor modalities means that it offers greater control when it comes to smoothing near discontinuities.

Graph Structure

Smoothing the sky factor approximation is set up as an inference problem on an undirected graphical model as shown in Figure 4.3. Each pixel in the image is a node in the graph, with the underlying state $\mathbf{\Gamma}$ represented by the white circles. All nodes are connected to their four adjacent neighbours and have an associated *fast* estimate of the sky factor, shown by the black circles. At the training points selected in the iterative sampling procedure, *full* sky factor estimates are available and are represented by the striped circles. The state space of each node is between 0 and 1, and this is discretised into equal increments.

Each node consists of a unary potential ϕ_i that encodes the likelihood, and a pairwise potential $\psi_{i,j}$ that encodes a prior. The total energy in the model is calculated as:

$$E = \sum \log \phi_i + \sum \sum \log \psi_{i,j}, \tag{4.3}$$

and this must be minimised via an inference procedure.

Unary Potential

The *fast* estimation of the sky factor is available for all pixels in the dataset and since it is geometrically based, it provides an upper bound to the sky factor value. An example of this is seen when analysing the *fast* approximate for a horizontal and vertical plane. Ignoring occlusions, the horizontal plane is visible to the entire sky dome while the vertical plane can only see a maximum of half the sky dome. These are the approximations generated via the *fast* method and when factoring in occlusions, the underlying sky factor will always be below this value. Therefore, for all pixels that do not have an associated *full* sky factor estimate, the unary potential is described by:

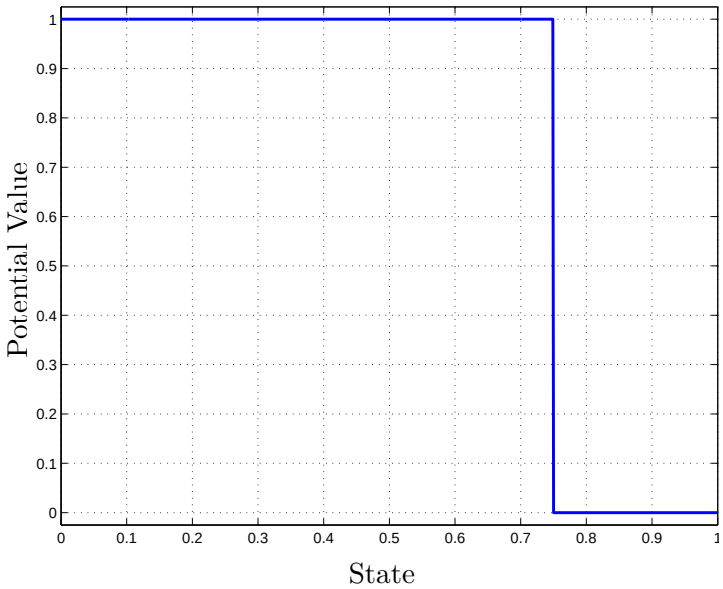
$$\phi_i(s) = \begin{cases} 1, & \text{if } s \leq \Gamma_{i,fast} \\ 0, & \text{otherwise,} \end{cases} \tag{4.4}$$

and an example is shown in Figure 4.4a for a point whose *fast* estimate is 0.75. The symbol s is used to indicate the state with the state space being discretised from 0 to 1 in increments of 0.025.

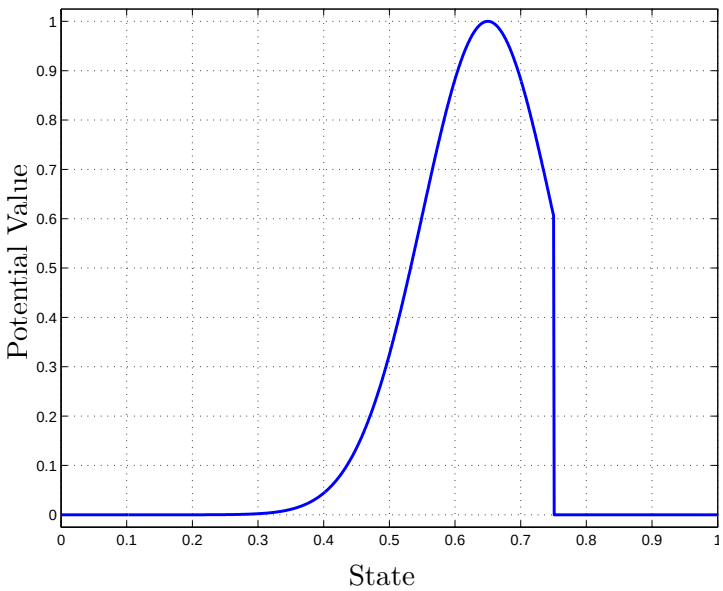
Performing the *full* computation of a sky factor calculation gives a confident estimate of the actual value. This certainty can be represented by a Gaussian distribution whose mean is given by the *full* estimate. Given that the upper bound is provided by the *fast* estimate, the Gaussian distributions are truncated at this value so that physically implausible estimates are not made. The unary potential can be described by:

$$\phi_i(s) = \begin{cases} \exp\left(\frac{-|s-\Gamma_{i,full}|^2}{0.02}\right), & \text{if } s \leq \Gamma_{i,fast} \\ 0, & \text{otherwise,} \end{cases} \tag{4.5}$$

and an example is shown in Figure 4.4b for a point whose *full* estimate is 0.65. For



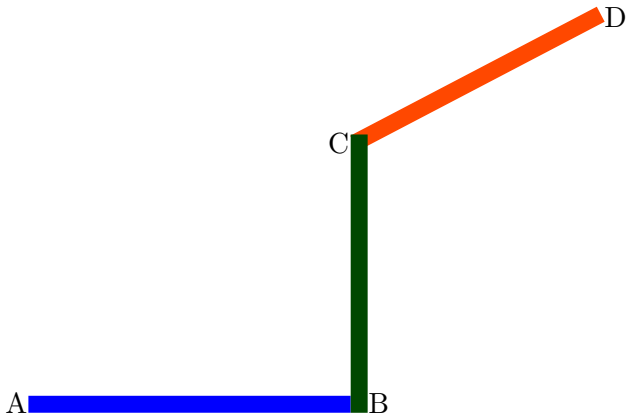
(a) The unary potential encodes the fact that the *fast* approximation provides an upper bound for the sky factor.



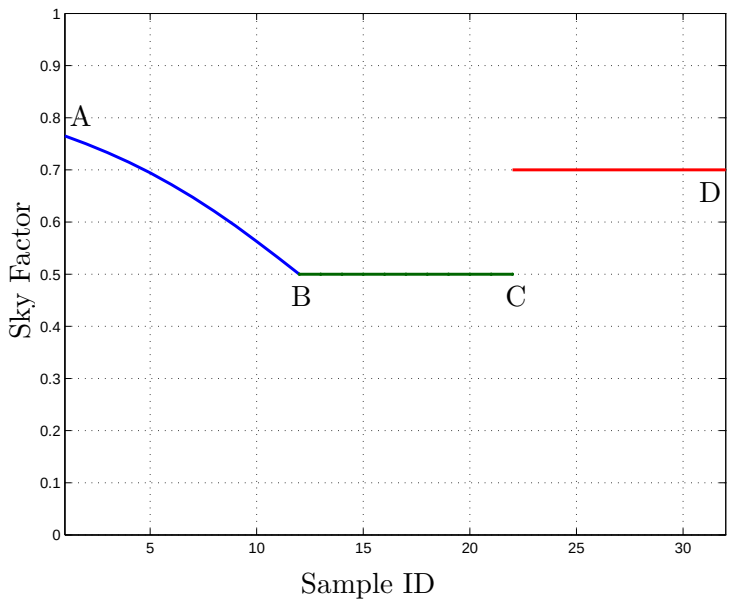
(b) The unary potential for the *full* calculation encodes the fact the estimate is close to the true value. The *fast* approximation is also used to enforce an upper bound.

Figure 4.4 – Unary potentials for a point with a *fast* estimate of 0.75, and *full* estimate of 0.65.

both cases, the unary potentials are normalised to have a maximum value of one.



(a) 1D scene structure.



(b) Sky factors for the 1D scene are seen to smoothly vary with the exception of a discontinuity.

Figure 4.5 – A 1D motivating example for spatial consistency in the sky factor image.

Pairwise Potential

The pairwise potential encodes the assumption that sky factors vary smoothly throughout the image. A 1D example is shown in Figure 4.5a, where the scene consists of a horizontal, vertical and slanted line. Taking samples of the sky factor as each component is traversed shows that as the sample position moves closer to the intersection of the horizontal and vertical components, the sky factor varies smoothly (Figure 4.5b). Along the vertical component (B-C), the sky factor is constant and there is a disconti-

nuity as the sample moves onto the slanted line (C-D). This motivates the smoothness prior between adjacent pixels while also incorporating discontinuities.

The first discontinuity observed from the motivating examples occurs due to a change in orientation between the different regions in the scene. The normal vector for each location in the point cloud can be used to identify changes in orientation. This is the advantage of using fused image and point cloud data, as they complement each other in the type of information available to the system. The cosine of the angle between a point and the sensor is calculated as the dot product between the normal and the vector towards the sensor. This value is projected into the image domain and a Sobel edge filter is used to extract sharp changes in orientation E_{ang}.

A second discontinuity exists due to objects appearing at different depths in the scene. For example, a vertical calibration board may appear in the image as being adjacent to the horizontal ground plane. If this were to be smoothed naïvely, then the sky factor for the area around the edge of the calibration board would gradually change from a high value of around 1 (obtained from the ground plane), to a value of 0.5 (for a vertical surface). However, this is incorrect as the sky factor should vary sharply due to the geometry. Once again, fused point cloud data helps in identifying these regions by analysing the range information.

Therefore, the pairwise potential encodes a smoothness prior in the image while also preserving discontinuities. It is expressed as:

$$\psi_{ij}(s_i, s_j) = \gamma \max(1 - \alpha|s_i - s_j|^2, 0), \tag{4.6}$$

where:

$$\gamma = \begin{cases} 0, & \text{if } |d_i - d_j| > 0.1 || E_{i,ang} = 1, \\ 1, & \text{otherwise,} \end{cases} \tag{4.7}$$

where $||$ is the logical OR operator. Here, d_i is the real world 3D distance between the sensor and point i, and α is a tuning parameter that dictates the strength of the pairwise potential compared to the unary potential.

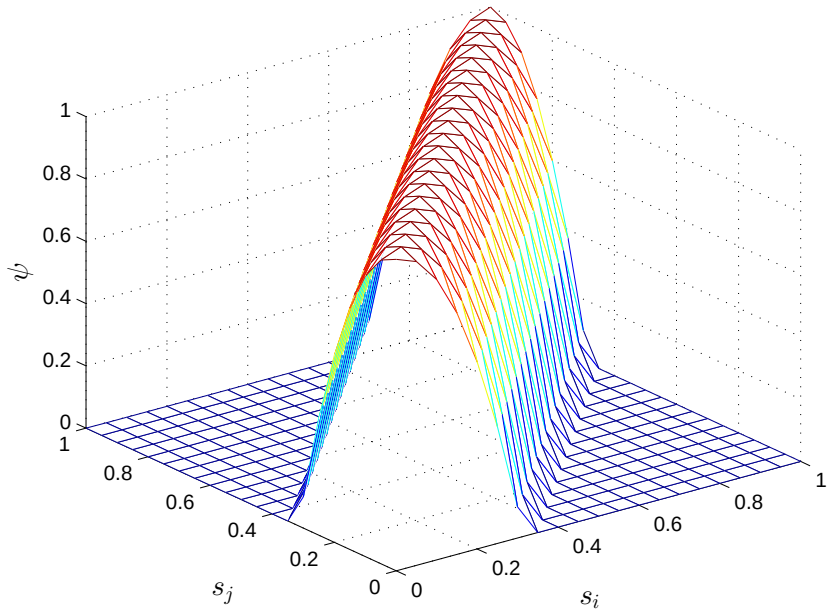


Figure 4.6 – Pairwise potential using $\alpha = 10$.

The advantage of using an MRF based approach to smoothing is that it allows the discontinuities to be preserved. This is compared to using techniques such as Gaussian or median filtering, which do not incorporate information from the point cloud data.

Energy Minimisation

Inference on a linear chain or tree structured graphical model can be performed exactly using belief propagation algorithms. However, the grid structure introduces loops which makes exact inference intractable. Fortunately, techniques such as Loopy Belief Propagation (LBP) [53] can be used to approximate the solution.

Through the use of an iterative sampling approach and performing inference on a grid structured MRF, high resolution estimates of the sky factors can be generated using a small amount of samples. This methodology is useful when encountering large scale scenes as is commonly the case in outdoor perception.

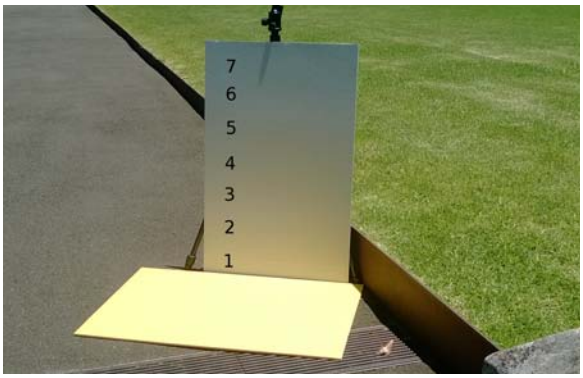


Figure 4.7 – An example of indirect illumination, where light striking the flat yellow board is reflected onto the upright grey board. The influence of indirect illumination depends on the geometry of the scene and in this example, the points lower down on the board (1 to 4), experience significantly more indirect illumination compared to higher points (5 to 7).

4.2 Indirect Illumination

When sensing in the outdoor environment, the most common illumination sources that are accounted for when performing radiometric calibration are terrestrial sunlight and diffuse skylight. However, indirect illumination is also present as an illumination source and can play a vital role in the appearance of the scene depending on the geometric structure.

In forward rendering, it is unknown which areas in the scene will receive a large amount of indirect illumination. However, since the final image is known and using the assumption of diffuse material reflection, inverse radiosity has the benefit of being able to identify sources of high indirect illumination as it is directly related to the pixel intensity. The radiometric normalisation method proposed in this thesis is focused on deriving material characteristics by utilising fused geometric and visual data. Therefore, this thesis proposes a sampling strategy that focuses calculations on areas where the change between material estimates with and without the inclusion of indirect illumination are greatest.

Indirect illumination occurs when the incident light upon a surface originates from light reflecting off another object. An example is shown in Figure 4.7, where light reflecting off the yellow board on the ground illuminates the upright gray board. Ar-

eas close to the base of the board are significantly affected by indirect illumination compared to the areas near the top of the board. In the computer graphics community, such a phenomenon is known as colour bleeding, while in the remote sensing community it can be considered as one source of non-linear mixing of substances in an image. In both cases, the colour captured by the camera does not accurately represent the underlying material and this can negatively impact high level algorithms such as classification and segmentation.

Radiometric normalisation can be reformulated as the inverse reflectometry problem found in computer graphics. Given the scene structure, illumination conditions and an image of the scene, the aim is to obtain the material characteristics of each object. In this scenario the scene structure is known from a LIDAR scan of the environment, the image is captured by a camera, and the terrestrial sunlight and skylight irradiance spectra is obtained using the methods presented in Section 4.3. In order to increase the accuracy of this approach, an inverse radiosity technique is used to incorporate indirect illumination into the model.

Inverse radiosity is chosen to calculate the indirect illumination for a number of reasons. It is relatively simple to implement and the assumption of diffusivity that has been used throughout this work fits within the framework of radiosity, which also has an underlying diffuse assumption of the world.

In order to calculate the indirect illumination using radiosity, form factors must first be calculated, which involves solving the double integral:

$$F_{ij} = \frac{1}{A_i} \int_{S_i} \int_{S_j} \frac{\cos(\Psi, N_x)\cos(-\Psi, N_y)}{\pi r_{xy}^2} V(x, y) dA_y dA_x. \tag{4.8}$$

With the exception of extremely simple scenes, this integral cannot be solved analytically and several numeric methods have been proposed [12]. The hemicube method [11] relies on a differential area to area approximation in order to generate the form factor. The method itself stems from the Nusselt analog which provides a geometrical perspective to understanding Equation (4.8). Initially, a hemisphere is placed above the surface i where the form factors are to be computed. An arbitrary surface j in

the scene is projected onto the hemisphere and this accounts for the solid angle calculation to the surface $\frac{\cos(-\Psi, N_y)}{r_{xy}^2}$ [12]. The area on the hemisphere is then projected onto the unit circle, accounting for the $\cos(\Psi, N_x)$ term. Finally, by dividing by the area of the circle π, the form factor is known assuming that the surface is completely visible from i.

Implementing a method using the Nusselt analog directly is computationally expensive and so an approximation was developed to allow faster rendering of scenes. The hemicube method [11] utilises a half cube placed above the surface i and discretises the surface of the cube into small pixel like regions. The form factors of each pixel is precomputed based on which side of the cube it is on. This takes advantage of a key outcome of the Nusselt analog, which is that two surfaces that project onto the same area on the hemisphere, will have equal form factors irrespective of their geometries. By ray tracing from a point on surface i through each discrete pixel on the hemicube, the form factors for surface i can be computed efficiently to all surfaces in the scene. This is because the number of rays required is completely dependent on the hemicube resolution, instead of the number of surfaces in the scene.

Although the hemicube method for estimating form factors is efficient, the large datasets involved in robotics and remote sensing means that it is still too computationally expensive to incorporate into the illumination model. Therefore, this thesis proposes an iterative method of sub-sampling and interpolation to estimate the indirect illumination in the scene using a small number of form factor calculations.

4.2.1 Iterative Sampling

When choosing the sampling strategy for calculating indirect illumination in the scene, a number of different factors must be taken into account. In a naïve implementation, the scene may be sampled using uniformly distributed points, but as shown previously in Section 4.1.1, this does not guarantee sufficient spatial sampling of the image. Therefore, grid based, stratified or best candidate sampling approaches may be used which provide greater spatial coverage of the image, thereby estimating the indirect

illumination with greater accuracy.

The issue with these approaches is that they do not consider the ultimate aim of calculating indirect illumination. In a forward rendering scenario, such an approach would be adequate, since the aim is to generate a realistic image. However, in the radiometric normalisation context where the aim is to estimate material reflectance properties, it is not actually necessary to calculate indirect illumination accurately throughout the entire image. Instead, indirect illumination calculation should be focused on areas where it plays a significant role in the estimation of reflection. This includes areas occluded from the sun, or those possessing low sky factors.

The change in reflectance estimate $\Delta\rho_i$ due to the inclusion of indirect sources in the illumination model, is described by:

$$\begin{aligned}\Delta\rho_i(\lambda) &= \frac{|\rho_{i,sun+sky}(\lambda) - \rho_{i,sun+sky+ind}(\lambda)|}{\rho_{i,sun+sky}(\lambda)} \\ &= \frac{\left|\frac{L_i(\lambda)\pi}{E_{sun}(\lambda)\tau(\lambda)\cos\alpha_i+\Gamma_i E_{sky}(\lambda)} - \frac{L_i(\lambda)\pi}{E_{sun}(\lambda)\tau(\lambda)\cos\alpha_i+\Gamma_i E_{sky}(\lambda)+E_{ind,i}(\lambda)}\right|}{\frac{L_i(\lambda)\pi}{E_{sun}(\lambda)\tau(\lambda)\cos\alpha_i+\Gamma_i E_{sky}(\lambda)}} \\ &= \frac{E_{ind,i}(\lambda)}{E_{sun}(\lambda)\tau(\lambda)\cos\alpha_i + \Gamma_i E_{sky}(\lambda) + E_{ind,i}(\lambda)},\end{aligned} \tag{4.9}$$

where the subscripts on the ρ_i term indicate the illumination sources that are being considered. The function to be sampled is therefore the proportion of the total illumination at each pixel whose source is indirect illumination. The dependence on wavelength λ is still present in $\Delta\rho_i$, so the final function to sample from is taken as the maximum over all wavelengths:

$$\Delta\rho_i = \max_{\lambda} \frac{E_{ind,i}(\lambda)}{E_{sun}(\lambda)\tau(\lambda)\cos\alpha_i + \Gamma_i E_{sky}(\lambda) + E_{ind,i}(\lambda)}. \tag{4.10}$$

This is necessary because indirect illumination influences each wavelength separately. An example of this is the influence of indirect illumination from grass is more pronounced in the near infra-red region of the electromagnetic spectrum, where it reflects more light compared to the visible regions.

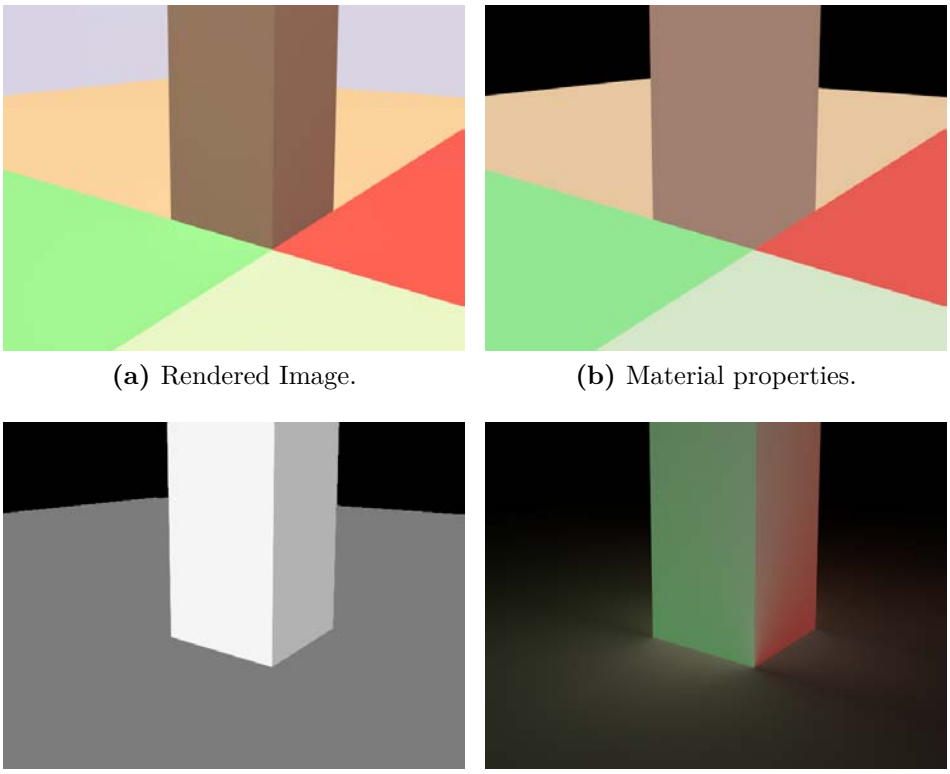
Similarly to the strategy developed in Section 4.1, two sampling strategies are used in

Algorithm 2 Approximating indirect illumination throughout the scene is performed via an iterative sampling process followed by guided filtering for all spectral channels.

```
1. Iterative Sampling
training set T ← ∅
validation set V ← ∅
error ← 1
indIll ← 0
for i:=1 to maxIterations do
    Calculate Δρ_{i−1} using Equation (4.10)
    T ← T ∪ importance sampling of max Δρ_{i−1}
    Calculate error_i
    T ← T ∪ V
    indIll_i ← linear interpolation of T for all channels
end for
2. Smoothing
Calculate camAng
for i:=1 to numSpectralChannels do
    indIll_i ← guided filtering( camAng, indIll_i )
end for
```

an iterative approach for estimating indirect illumination in the scene. The validation set is selected using best candidate sampling in order to provide sufficient spatial detail. Prior to the first iteration, N number of validation points are selected and split into several groups, each containing $N_{validation}$ number of points. Generating a large number of samples and then splitting them into $\frac{N}{N_{validation}}$ groups is chosen over selecting a smaller number of samples at each iteration. This is because best candidate sampling means the location of each sample is dependent on those already selected.

For each iteration, the $\Delta\rho_i$ function is calculated using the current estimate of indirect illumination in the scene, and $N_{training}$ points are selected by sampling from this function. These points are added to the training set and the indirect illumination estimate is generated by linear interpolation of all points within this set. The accuracy of this estimate is measured using the validation set by analysing the error between the estimate and the ground truth. Prior to the commencement of the next iteration, the validation set is moved into the training set in order to improve subsequent estimates.



(a) Rendered Image. **(b)** Material properties.

(c) Camera angle. **(d)** Indirect illumination.

Figure 4.8 – Simulated scene in Blender showing the different components of an image.

4.2.2 Smoothing

As shown in the indirect illumination component of the rendered example in Figure 4.8d, sharp changes occur at the intersection of objects or at changes in geometry. Therefore, using Gaussian or median filters to smooth the indirect illumination approximation is not appropriate as it does not incorporate the geometrical properties of the scene. Ideally, a similar method to the LBP smoothing approach used in Section 4.1.2 should be used, with the pairwise potential enforcing the discontinuities in the scene. However, LBP is computationally expensive and would need to be applied for each spectral channel since indirect illumination is wavelength dependent.

In this thesis, guided image filtering [36] is used in order to smooth the indirect illumination approximation. Guided filtering has greater edge preserving properties compared to standard image filtering techniques, while also being computationally

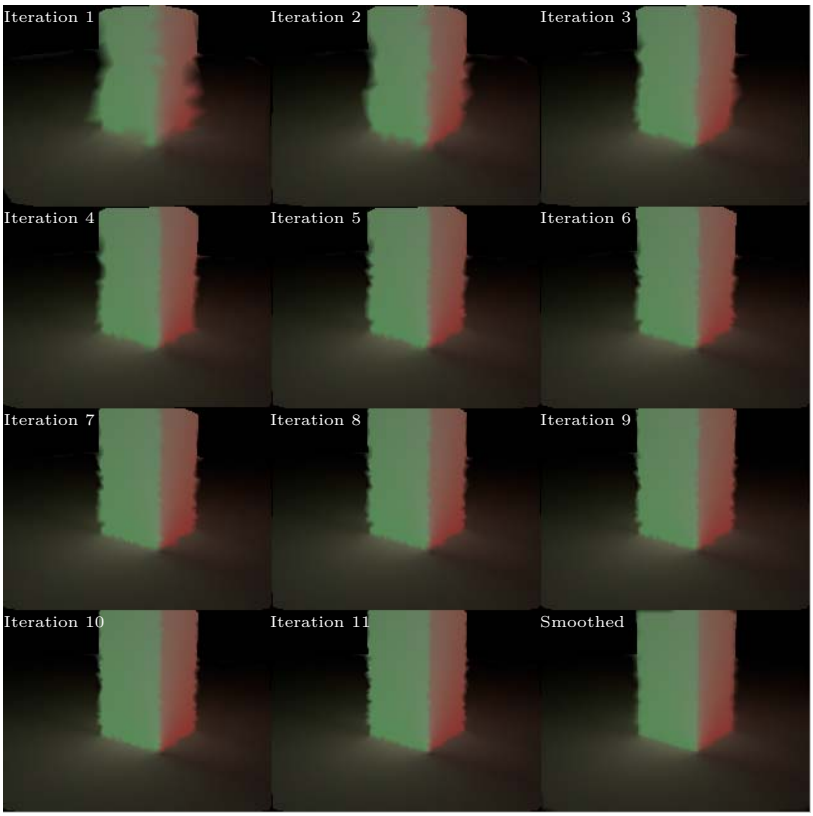


Figure 4.9 – At each iteration, new samples from the training set improve the indirect illumination approximation using linear interpolation. Smoothing using guided image filtering refines the approximation in the final step.

inexpensive. In this case, **camAng** - the angle between the normals at each point and the vector in the direction of the sensor, is used as the guidance image. The motivation for using the angle image stems from the observation that indirect illumination tends to change abruptly at the intersection of objects. This is seen in Figure 4.8d and is a result of adjacent objects being composed of different material colours and intensities.

An example of the sampling and smoothing process is shown in Figure 4.9 and demonstrates that at each successive iteration, the similarity between the estimate and the ground truth increases. The final smoothing stage is seen to sharpen the corners of the object to align it with the camera angle image.

In this section a sampling strategy for approximating the indirect illumination sources in large scale scenes was proposed. Combining this with the large scale sky factor approximation means that the parameters of the outdoor illumination model can be

known. This can potentially allow for increased accuracy during radiometric normalisation as it can account for the different illumination sources.

4.3 Radiometric Normalisation

The illumination invariant method presented in Chapter 3 compensates for geometrical variations in illumination due to occlusions and exposure to different amounts of the sky. However, the output image is still dependent on the skylight illuminant. This is because the scenario is ill posed and requires direct measurement of either the terrestrial sunlight or skylight spectra. Once these measurements are obtained, it is possible to perform radiometric normalisation that is invariant to illumination variation in the scene due to sunlight, skylight and indirect sources. In this section a new pipeline for illumination invariant radiometric normalisation is proposed that utilises multi-modal sensor data.

The proposed normalisation method is derived with the inclusion of indirect illumination in the outdoor illumination model (Equation 2.16). However, it is noted that this is an optional step and normalisation could also be approximated without the inclusion of the indirect illumination term. The computational requirements and the application which radiometric normalisation will be used dictate whether or not indirect illumination should be incorporated into the model.

With the inclusion of indirect illumination in the outdoor illumination model, the radiance of an arbitrary point i in the scene can be described as:

$$L_i(\lambda) = \frac{\rho_i(\lambda)}{\pi}[V_{i,sun}E_{sun}(\lambda)\tau(\lambda)\cos\theta_i + \Gamma_i E_{sky}(\lambda) + \sum_{j=1}^{N} F_{ij}L_j(\lambda)\pi], \tag{4.11}$$

where F_{ij} is the form factor representing the fraction of incident illumination reaching i that leaves j [17].

In order to perform normalisation, the wavelength dependent relationship between terrestrial sunlight and diffuse skylight is first obtained through the initialisation

points. This is followed by a measurement of the incident illumination via a hardware based method such as calibration boards or downwelling irradiance sensors [8]. This provides sufficient information for the illumination spectra to be resolved. Utilising the indirect illumination approximation for the entire scene developed in Section 4.2, radiometric normalisation can take place as all illumination sources in the outdoor illumination model have been approximated.

After determining the illumination source spectra, radiometric normalisation can take place by estimating the incident illumination at each pixel and solving Equation (4.11) for reflectance:

$$\rho_i(\lambda) = \frac{L_i(\lambda)\pi}{V_{i,sun}E_{sun}(\lambda)\tau(\lambda)\cos\theta_i + \Gamma_i E_{sky}(\lambda) + \underbrace{\sum_{j=1}^{N} F_{ij}L_j(\lambda)\pi}_{optional}}. \tag{4.12}$$

The proposed method of illumination invariant radiometric normalisation infers the illumination source and intensity from the scene and has several advantages over traditional and state of the art methods. Firstly, flat-field correction using the in situ hardware measurements leads to reflectance estimates that do not account for topographic variations in the scene. This is due to the fact that this method performs image wide normalisations. The proposed method on the other hand, estimates the incident illumination on a per-pixel basis. The most obvious advantage of this approach is that identical materials in varying illumination conditions will appear similar.

Secondly, there is no requirement for knowledge of the atmospheric conditions at the time of data acquisition. This is a very important advantage for practical applications since it allows operation in areas where these parameters are not available, compared to methods such as [10] and [40] that depend on models such as MODTRAN.

Finally, optimisation procedures such as those used in [29] require the selection of initialisation points from a known number of materials, which can be difficult and prone to error when performed by an operator. Through the use of an automatic

initialisation procedure, the proposed method is independent of the user and therefore does not suffer from incorrect selection of initialisation points.

4.3.1 Terrestrial Sunlight - Skylight Relationship

A terrestrial sunlight-skylight ratio was derived in Chapter 3 using two points from the same material; one exposed to sunlight and the other occluded. However, the inclusion of indirect illumination in the model causes this relationship to change and it must be re-derived. Given two points $\{A, A'\}$ from the same material with reflectance ρ, one exposed to sunlight and the other occluded, their radiance can be modelled as:

$$L_A(\lambda) = \frac{\rho(\lambda)}{\pi}[E_{sun}(\lambda)\tau(\lambda)\cos\alpha_A + \Gamma_A E_{sky}(\lambda) + E_{ind,A}(\lambda)], \tag{4.13}$$

$$L_{A'}(\lambda) = \frac{\rho(\lambda)}{\pi}[\Gamma_{A'} E_{sky}(\lambda) + E_{ind,A'}(\lambda)]. \tag{4.14}$$

Solving simultaneously gives the relationship between terrestrial sunlight and skylight:

$$\begin{aligned} E_{sun}(\lambda)\tau(\lambda) &= \frac{L_A(\lambda)E_{ind,A'}(\lambda) - L_{A'}(\lambda)E_{ind,A}(\lambda)}{L_{A'}(\lambda)\cos\alpha_A} \\ &\quad + \frac{L_A(\lambda)\Gamma_{A'} - L_{A'}(\lambda)\Gamma_A}{L_{A'}(\lambda)\cos\alpha_A} E_{sky}(\lambda) \\ &= \underbrace{T_1(\lambda)}_{optional} + T_2(\lambda)E_{sky}(\lambda). \end{aligned} \tag{4.15}$$

Therefore, with the inclusion of indirect illumination, a relationship involving an offset T_1 and gain factor T_2 between terrestrial sunlight and skylight can be found.

The automatic initialisation process developed in Chapter 3 generates multiple initialisation point pairs in the scene. Given that $T1$ is dependent on indirect illumination which varies throughout the scene, it is not possible to simply filter the estimates. Instead, each point pair used for initialisation produces an estimate of the terrestrial sunlight and diffuse skylight sources. The mean of the estimates is then used as the illumination source spectra in the model.

4.3.2 Illumination Source Measurement

To generate an estimate of the illumination sources, a direct measurement of the incident illumination in the scene is required. This is performed by measuring the incident illumination either directly through the use of a downwelling irradiance sensor, or indirectly through a calibration panel. The downwelling irradiance sensor is advantageous in that the sensor can be positioned on the sensor platform, so no hardware needs to be placed in the environment. However, the small field-of-view of the sensor means that there are often times of day where the terrestrial sunlight is not being measured due to a lack of line-of-sight visibility. Calibration boards can be easily orientated in a way that the incident illumination is a combination of terrestrial sunlight and skylight, however it requires placing the board in the scene. This may not be a viable option in hazardous situations. Once a measurement is obtained, an estimate of the illumination source can be obtained by combining the estimates at each initialisation point pair.

Calibration Board

When performing field based measurements, calibration boards such as those shown in Figure 4.10 are typically placed in the scene. These boards have a precisely known spectral reflectance curve and are therefore useful in providing a reference measurement in the scene. The radiance at a point p on the panel is described by:

$$L_p(\lambda) = \frac{\rho_p(\lambda)}{\pi}[V_{p,sun}E_{sun}(\lambda)\tau(\lambda)\cos\theta_p + \Gamma_p E_{sky}(\lambda) + \sum_{j=1}^{N} F_{pj}L_j(\lambda)\pi], \tag{4.16}$$

$$= \frac{\rho_p(\lambda)}{\pi}[V_{p,sun}E_{sun}(\lambda)\tau(\lambda)\cos\theta_p + \Gamma_p E_{sky}(\lambda) + E_{ind,p}(\lambda)], \tag{4.17}$$

where ρ_p is known from laboratory analysis of the panel.

Substituting Equation (4.15) into Equation (4.17) and solving for the skylight spectra

Figure 4.10 – Calibration boards of known reflectance are placed in the scene to provide reference measurements when calibrating the image.

gives:

$$E_{sky}(\lambda) = \frac{\frac{L_p(\lambda)\pi}{\rho_p(\lambda)}}{\Gamma_p + V_{p,sun}T_2(\lambda)\cos\alpha_p} - \underbrace{\frac{E_{ind,p}(\lambda) + V_{p,sun}T_1(\lambda)\cos\alpha_p}{\Gamma_p + V_{p,sun}T_2(\lambda)\cos\alpha_p}}_{optional}. \tag{4.18}$$

The terrestrial sunlight spectra can then be obtained by substituting Equation (4.18) into Equation (4.15).

Downwelling Irradiance Sensor

Another hardware based method of obtaining a measurement of the incident illumination in the scene are downwelling irradiance sensors. Typically, these sensors are situated at position s on the sensing platform and their radiometrically calibrated measurement can be described as:

$$E_s(\lambda) = [V_{s,sun}E_{sun}(\lambda)\tau(\lambda)\cos\alpha_s + \Gamma_s E_{sky}(\lambda)]. \tag{4.19}$$

In this formulation, indirect illumination is assumed to have a negligible influence on the measurement. This is a valid assumption as in most scenarios, the terrestrial sunlight and skylight will dominate the measurement, and indirect illumination from

the scene will enter the sensor at grazing angles leading to a low weighting.

Simultaneously solving Equation (4.15) and Equation (4.19) gives an expression for skylight spectra:

$$E_{sky}(\lambda) = \frac{E_s}{\Gamma_s + V_{s,sun}T_2(\lambda)\cos\alpha_s} - \underbrace{\frac{V_{s,sun}T_1(\lambda)\cos\alpha_s}{\Gamma_s + V_{s,sun}T_2(\lambda)\cos\alpha_s}}_{optional}, \tag{4.20}$$

and the terrestrial sunlight spectra can be obtained from Equation (4.15).

Therefore, given initialisation points and measurements from either a calibration board or downwelling irradiance sensor, the terrestrial sunlight and diffuse skylight illumination sources can be estimated in terms of colour and intensity. This allows illumination invariant radiometric normalisation to take place, which assists the performance of high level algorithms. Compared to performing flat-field correction using the hardware based measurements, the proposed method provides robustness against shadow artefacts that may appear in the image due to the surrounding geometry. The approach is flexible enough to allow the inclusion of indirect illumination as a light source, depending on whether the application requires it. It is important to note that the method is not limited to the use of calibration boards and downwelling irradiance sensors. As long as the measurement of the incident light can be modelled using the outdoor illumination model, then an explicit solution to the illumination sources can be obtained.

4.4 Summary

Radiometric normalisation is a key component of remote sensing systems as it allows measured data to be converted to a format where it can be compared against laboratory data. Accurate normalisation methods are required in for high level supervised algorithms such as clustering and classification to operate reliably.

Traditional normalisation methods include the use of calibration panels and downwelling irradiance sensors, which operate without any regard to the geometry of the

scene. This means that after normalisation, identical materials appear differently based on their incident illumination which is problematic for high level algorithms.

In this chapter, a large scale sky factor and indirect illumination approximation method was proposed that reduces the number of *full* calculations and form factors required. Finally, an illumination invariant radiometric normalisation procedure was proposed that can operate with or without estimates of indirect illumination.

The proposed novel pipeline utilises in situ hardware based approaches for direct measurement of the incident illumination and generates an illumination invariant normalised image through the use of fused image and point cloud data. The benefit of this is that identical materials appear similar, independent of their location and orientation in the scene. This is performed without the need for highly parameterised atmospheric radiative transfer models and user selection of a known number of materials by inferring the illumination sources from the image. The methods proposed in this thesis for attaining spatial and temporal illumination invariance in outdoor imagery, and illumination invariant radiometric normalisation are evaluated in Chapter 5 on a number of datasets.

Chapter 5

Experimental Results

The proposed system for attaining illumination invariance in outdoor imagery through accurate modelling of the physical processes involved is evaluated on a number of datasets. The experiments are designed to first assess the performance of the algorithm under ideal conditions such as simple geometry and narrow-band sensors, before testing on complex scenes and wideband sensors. The datasets are also gathered under different weather conditions in order to show the robustness of the proposed approach.

Section 5.1 describes the metrics used to evaluate the performance of the proposed approach, with a summary of the datasets being presented in Section 5.2. A brief overview of the implementation of the algorithm is given in Section 5.3. The experiments evaluate the spatial (Section 5.4) and temporal (Section 5.5) properties of the automatic relighting system presented in Chapter 3. Section 5.6 evaluates the large scale sky factor approximation approach. Finally, the illumination invariant radiometric normalisation approach proposed in Chapter 4 is evaluated in Section 5.7.

It should be noted that since relighting is performed on linear images with respect to diffuse skylight, visualisation of the images is hampered by the low intensity. Therefore, a gamma value of 1.5 is applied to the images to assist in visualisation, but the experimental validation of the proposed algorithms utilises the non-gamma corrected images.

5.1 Metrics

A variety of metrics are used in this thesis to evaluate the performance of the illumination invariance system. These metrics vary from those that are used to analyse the similarity between the colour and intensity of a pixel, to assessing the quality of high level algorithms.

5.1.1 Distance

The $\mathbf{l}_1$ and $\mathbf{l}_2$ distance metrics are used to determine the magnitude difference between two n dimensional vectors $\mathbf{a}$ and $\mathbf{b}$:

$\mathbf{l}_1$ distance

$$d(\mathbf{a}, \mathbf{b}) = |a(\lambda_1) - b(\lambda_1)| + |a(\lambda_2) - b(\lambda_2)| + ...|a(\lambda_N) - b(\lambda_N)|, \tag{5.1}$$

$\mathbf{l}_2$ distance

$$d(\mathbf{a}, \mathbf{b}) = \sqrt{[a(\lambda_1) - b(\lambda_1)]^2 + [a(\lambda_2) - b(\lambda_2)]^2 + ...[a(\lambda_N) - b(\lambda_N)]^2}. \tag{5.2}$$

In this thesis, the $\mathbf{l_2}$ metric (also referred to as the Euclidean distance) will typically be used to determine the difference between spectra, while the $\mathbf{l}_1$ metric will be used for single dimensional data such as sky factors.

5.1.2 Spectral Angle Mapper

While the Euclidean distance metric can discriminate based on colour intensities, it is not indicative of how similar their shape is. Spectral Angle Mapper (SAM) [85] is a measure of the angle between two vectors and is calculated as:

$$SAM(\mathbf{a}, \mathbf{b}) = \arccos \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}, \tag{5.3}$$

where $\cdot$ is the dot product operator. SAM can be seen to be partially illumination invariant, as the vectors **a** and **b** are normalised to have unit length. However, it is only invariant if each dimension is multiplied by the same constant value, which is not usually the case in the outdoor environment. The illumination source and strength varies as a function of wavelength on a per-pixel basis depending on the geometry. Therefore the appearance of identical materials changes in both shape and intensity depending on its location in the scene. For the SAM metric, a smaller angle between the two vectors indicates increasing spectral shape similarity.

5.1.3 Cluster Purity

Clustering is a type of unsupervised high level algorithm that combines pixels based on their similarity. As a means of assessing the quality of each cluster, the purity [61] can be evaluated:

$$purity(\Omega, \mathbb{C}) = \frac{1}{N} \sum_k \max_j |\omega_k \cap c_j|, \tag{5.4}$$

where Ω is the set of clusters $\{\omega_1, ..., \omega_k\}$, $\mathbb{C}$ is the set of classes $\{c_1, ..., c_J\}$, $\cap$ is the intersection between two sets, and N is the total number of data points. Essentially, the cluster purity is a measure of how homogeneous each cluster is in terms of its constituent classes. Higher purity values are desired as they indicate that the cluster is extracting a particular class from the data.

5.1.4 Earth Mover's Distance

When analysing the influence of variable illumination on the appearance of the scene, the intensity histograms of identical materials under the different conditions can be compared. An example of this would be that when a material is occluded from the sun, its histogram will typically be centred around a low intensity value. For the same material exposed to sunlight, the histogram will likely be centred around a higher value. Ideally, illumination invariance will allow these histograms to be similar and the Earth Mover's Distance (EMD) [66] provides a measure of this. It calculates

Figure 5.1 – The ColorChecker board contains coloured, uniform squares that represent colours typically found in nature.

how much work needs to be performed to transform one histogram to the other, with smaller values indicating that the histograms were close together.

5.1.5 ColorChecker Board

High spatial resolution ground truth data is difficult to obtain for large complex scenes, with material variability affecting the data. X-Rite ColorChecker calibration boards such as the one shown in Figure 5.1 offer a method for carrying out analysis on ground truth data on a small scale.

The board consists of 24 uniform patches of colours typically found in nature. These boards can be placed in the scene under different illumination conditions so the effect of illumination invariance can be evaluated by measuring the appearance of each colour patch.

5.1.6 Peak Signal to Noise Ratio

In order to measure image quality, the Peak Signal-to-Noise Ratio (PSNR) is calculated for an image with respect to a reference image and can take a value greater

than 0. The metric is calculated as:

$$PSNR(\mathbf{A}, \mathbf{B}) = 10 \log_{10} \frac{C}{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (A_{ij} - B_{ij})^2}, \tag{5.5}$$

where $\mathbf{A}$ and $\mathbf{B}$ are the two images of size $M \times N$, and C is the maximum value that the image could be (eg. 255 for an 8 bit image) [81]. Essentially, the metric divides the maximum capacity of the image by the mean squared error between the two images. Increasing similarity is obtained when the PSNR approaches infinity. The PSNR metric provides an indication as to how similar two images are. This is important when analysing the temporal illumination invariance aspects of the proposed system, with invariance being supported by a high PSNR for images captured from a static camera.

5.1.7 Structural Similarity

A second approach to measuring image similarity and quality is the Structural Similarity (SSIM) metric [81], which attempts to improve upon some of the weaknesses of the PSNR metric as it does not always match human perception :

$$SSIM(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_{\mathbf{x}}\mu_{\mathbf{y}} + C_1)(2\sigma_{\mathbf{xy}} + C_2)}{(\mu_{\mathbf{x}}^2 + \mu_{\mathbf{y}}^2 + C_1)(\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{x}}^2 + C_1)}. \tag{5.6}$$

The metric value ranges from 0 to 1 and is calculated using a sliding window approach, where $\mathbf{x}$ and $\mathbf{y}$ are patches of the two images of size 11×11. The global SSIM score for an image is the mean of all local values and as the score approaches 1, increasing similarity is achieved. μ_i and σ_i are the mean and standard deviation of patch i respectively.

5.2 Datasets

The illumination invariant system developed in this thesis was evaluated on a number of datasets that were gathered at the University of Sydney. These datasets were

Table 5.1 – Summary of the datasets used for experimental results. Datasets were captured under different weather conditions, ranging from overcast to sparse clouds, and with both hyperspectral and standard RGB cameras. The LIDAR resolution is equal in both the azimuth and elevation directions.

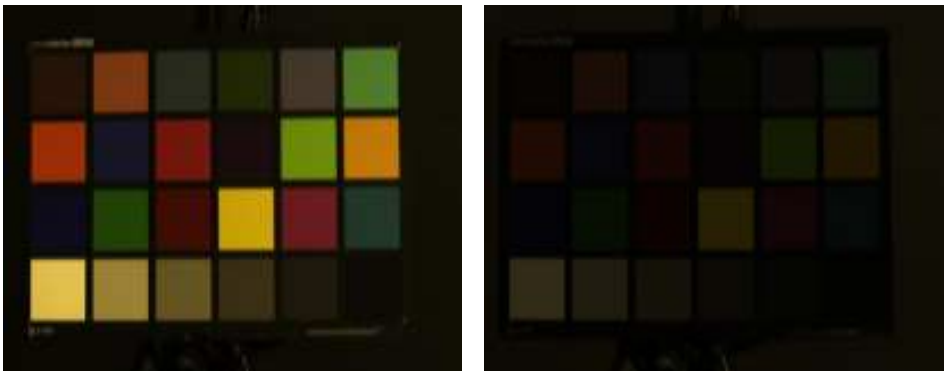
					Invariance	
ID	**Location**	**Camera**	**LIDAR Res. (deg)**	**Weather**	**Spatial**	**Temporal**
1	ColorChecker	VNIR	-	sunny	-	-
2	Shale	VNIR	0.040	partly cloudy	✓	×
3	USYD Hall	SWIR	0.050	partly cloudy	✓	×
4	Seymour Centre	RGB	0.041	sparse clouds	✓	×
5	Seymour Centre	RGB	0.041	sparse clouds	✓	✓
6	Seymour Centre	RGB	0.041	overcast	✓	✓
7	USYD Hall	RGB	0.030	sparse clouds	✓	×
8	ACFR Lawns	RGB	-	sunny/overcast	✓	✓

captured under a variety of weather conditions, with the scenes ranging in complexity. In this section a brief overview of each of the datasets is presented and the noteworthy aspects of each one are highlighted. A summary of the gathered dataset is provided in Table 5.1.

The cameras used in this thesis include SPECIM Visible Near Infra-Red (VNIR) $(400-970\text{nm})$ and Short Wave Infra-Red (SWIR)$(971-2516\text{nm})$ hyperspectral scanners, as well as a Canon PowerShot $A720$. Point clouds were obtained using a 3D RIEGL VZ1000 laser scanner and this was registered with the imagery using a markerless technique developed in [77]. The laser scanner has a maximum range of more than 1400m, accurate to within 8mm, and can rotate 100° and 360° in the vertical and horizontal directions respectively. Through the projection of the point cloud onto the image, the areas with no laser return (such as the sky) are removed and set to either a black or white colour.

5.2.1 Dataset 1 - VNIR ColorChecker

One of the initial requirements of the proposed illumination invariance system is that the terrestrial sunlight-skylight ratio be calculated from two points; one exposed



(a) Exposed to the sun. **(b)** Occluded from the sun.

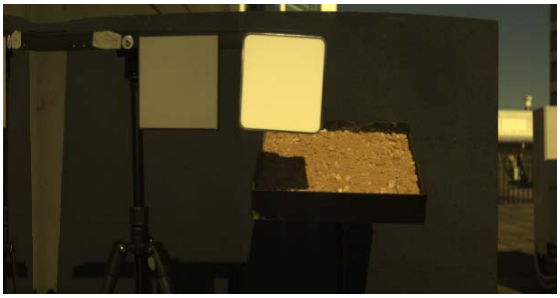
Figure 5.2 – The independence of the terrestrial sunlight-skylight ratio to the reflectance colour is shown by imaging a ColorChecker board under different illumination conditions.

to sunlight while the other is occluded. The derivation showed that this ratio was independent of the reflectance colour and to verify this, two VNIR hyperspectral scans of the ColorChecker board were captured as shown in Figure 5.2.

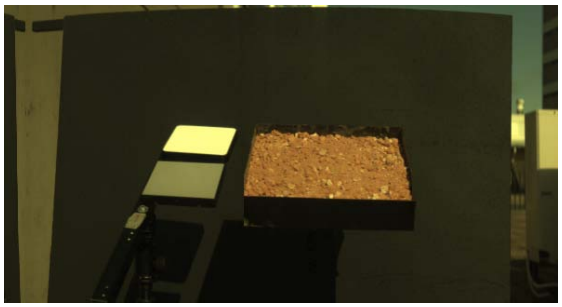
In the first image the board exposed to sunlight, while in the second image it was shadowed via the use of a small panel. This ensured that the sky factors in both images for the ColorChecker board remain approximately constant. Through this dataset, the sensitivity of the calculation of the terrestrial sunlight-skylight ratio with respect to the reflectance colour can be analysed.

5.2.2 Dataset 2 - VNIR & SWIR Shale

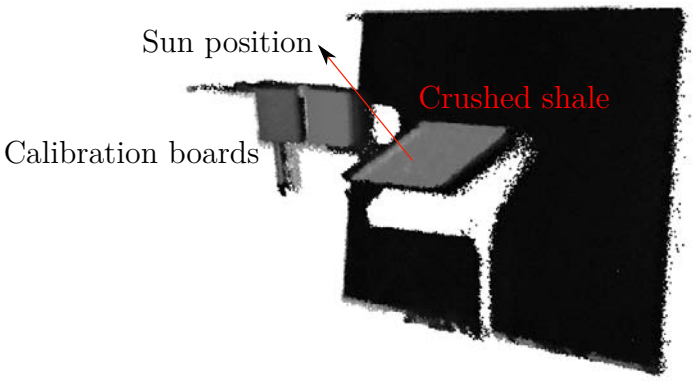
To validate illumination invariant relighting and radiometric normalisation, Dataset 2 provides a simple scene geometry with ground truth data available on a per-pixel scale. Within the scene lies a square container 70cm away from the camera containing powdered shale and a shadow is cast on it via a calibration board as shown in Figure 5.3a. The scene was imaged on a partially cloudy day, with the geometry captured by a LIDAR sensor. This dataset allows analysis of spatial illumination invariant relighting and radiometric normalisation processes to take place. A separate dataset was captured with no shadow being cast on the panel and the calibration



(a) Illumination variation is imposed on the scene via a calibration board casting a shadow onto the material.



(b) Ground truth reflectance data is obtained by positioning the calibration board at the same orientation as the panel and performing flat-field correction.



(c) Geometry of the scene.

Figure 5.3 – Dataset 2 consists of crushed shale and is captured using a VNIR and SWIR hyperspectral camera and fused with a LIDAR scan.

board situated adjacent to it in the same orientation. Ground truth reflectance data can then be obtained by performing flat-field correction with the calibration board measurement (Figure 5.3b).

5.2.3 Dataset 3 - SWIR USYD Great Hall

The first large scale complex scene dataset is of the Great Hall at the University of Sydney and consists of a sandstone building imaged using a SWIR hyperspectral scanner on a partly cloudy day. The building is approximately 100m wide, 17m tall and is 85m away from the sensor. The geometry of the scene is captured using a

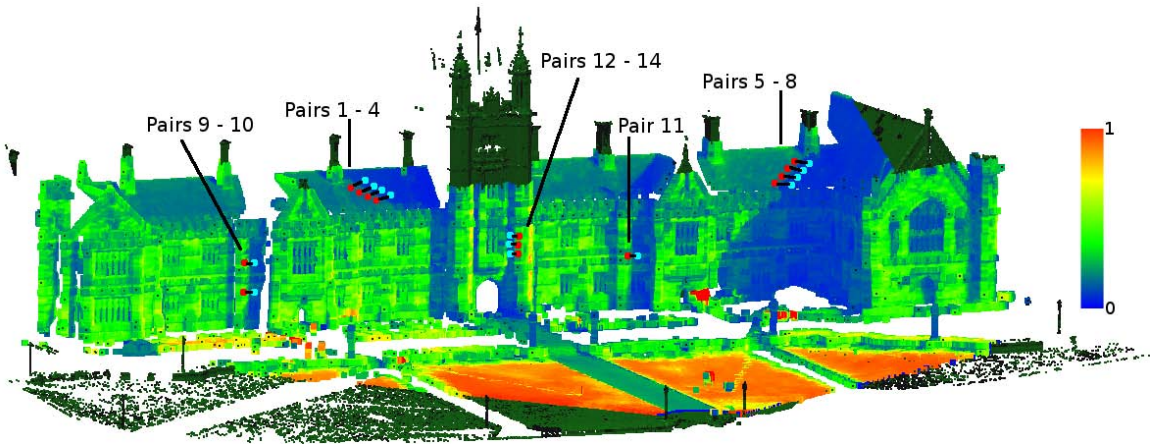


Figure 5.4 – Dataset 3 is captured using a SWIR hyperspectral scanner. Cast shadows in the scene affect appearance of the uniform front face of the building, as well as the roof. The displayed wavelength is 1067nm. Note that there is not 100% overlap between the point cloud (dark green) and image data. Pairs of points used for analysis from identical materials are shown in red (exposed to terrestrial sunlight) and cyan (shadowed).

LIDAR, yielding a point cloud of approximately $320,000$ points. The point cloud is fused with the hyperspectral data, with one channel being shown in Figure 5.4.

A number of interesting illumination phenomena are observed in the scene. Firstly, the uniform material that constitutes the front of the building is divided into two components; (i) high intensity when it is exposed to sunlight and skylight, and (ii) low intensity when it is occluded from the sun. A similar effect is seen for the roof. Other material classes within the scene include the road and grass in the foreground of the image.

A calibration panel of known reflectance is attached to a tripod and placed in the scene. A downwelling irradiance sensor is also connected to the hyperspectral scanner, providing an alternative method of measuring the incident illumination. This dataset allows both spatial illumination invariance and radiometric normalisation methods to be evaluated.

To assist in quantitative evaluation, 14 pairs of points along shadow boundaries were manually selected to generate a validation set. Each pair of points is specifically chosen to be obtained from a uniform material class.

Figure 5.5 – Dataset 4 was captured at the Seymour centre under sunny conditions. Cast shadows produce variable illumination on the uniform brick building.

5.2.4 Dataset 4 - RGB Seymour Centre

The second large scale dataset is taken using a consumer grade RGB camera at the Seymour Centre in the University of Sydney as shown in Figure 5.5. The dataset was captured under sunny conditions with sparse clouds present in the scene. The building is 17m tall and is 20m away from the sensor, The LIDAR scan of the scene yields a point cloud of approximately $500,000$ points and this is registered with the image.

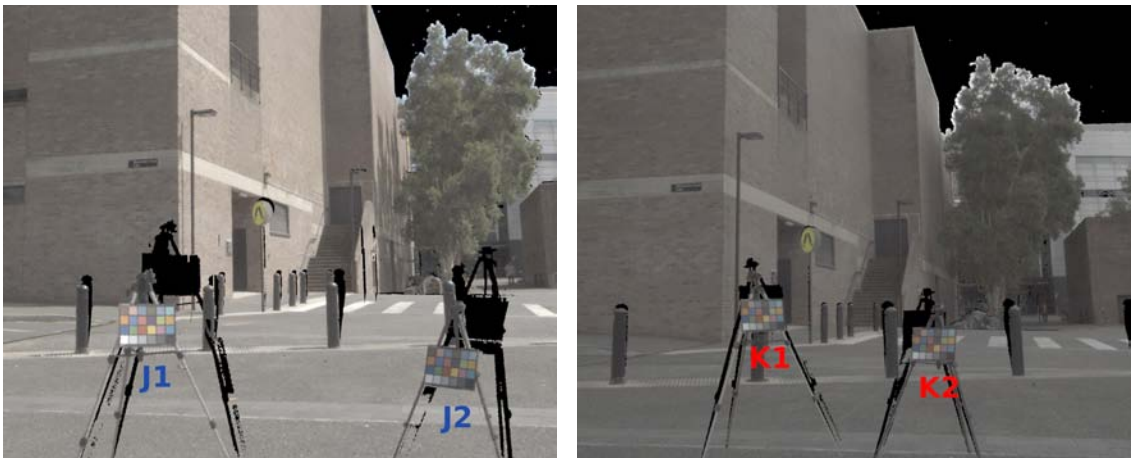
The building is predominantly made of a brown brick and there are three horizontal white concrete bands at various intervals. Due to the weather conditions, a strong shadow is cast on the building and this allows the performance of the proposed spatial illumination invariance algorithm to be analysed.

5.2.5 Dataset 5 and 6 - RGB Seymour Centre

The Seymour Centre was imaged under different illumination conditions as shown in Figure 5.6. For quantitative evaluation to take place, two ColorChecker boards were placed in the scene, with one exposed to sunlight ($J1$ and $K1$) and the other occluded ($J2$ and $K2$). This allows both spatial and temporal illumination invariance to be analysed using ground truth data, since there is a large amount of overlap between the scenes.

(a) Clear sky conditions. **(b)** Overcast conditions.

Figure 5.6 – Datasets 5 and 6 were captured at the same location under different illumination conditions.



(a) Dataset 5. **(b)** Dataset 6.

Figure 5.7 – Due to the varying illumination conditions, Dataset 5 appears brighter with high contrast between sunlit and shadowed regions due to the clear conditions it was imaged under. Dataset 6 is of low intensity due to the clouds present, which attenuate the terrestrial sunlight.

The two images are shown in Figure 5.7 and the appearance is significantly affected by the weather conditions present at time of capture.

5.2.6 Dataset 7 - RGB USYD Great Hall

This dataset consists of an image of the Great Hall at the University of Sydney under sunny conditions from a side on perspective using a consumer grade RGB camera as shown in Figure 5.8. Multiple laser scans were captured and combined in order to

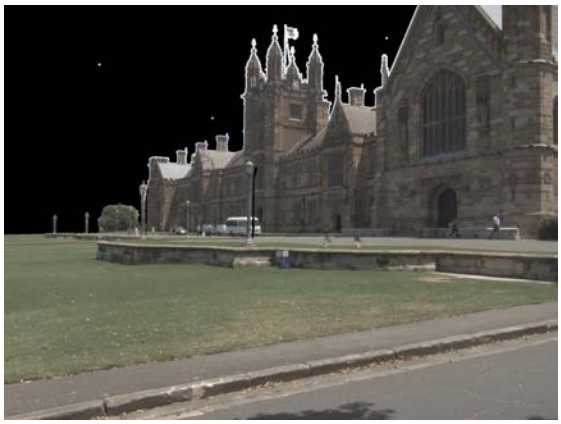


Figure 5.8 – Dataset 7 consists of an alternate view of USYD Great Hall captured using an RGB camera. The front face of the building is predominantly in shadow, while the side faces are exposed to sunlight.

(a) 9:30am. Sample patch regions are shown in red.

(b) 11:00am

(c) 2:00pm

Figure 5.9 – Dataset 8 is a time-lapse of a static scene taken at 30min. intervals. Ten sample patches of uniform material are used for evaluation.

gain a better geometrical representation of the scene. Through this dataset spatial illumination invariance can be evaluated.

5.2.7 Dataset 8 - RGB ACFR Lawns Timelapse

This dataset was obtained by positioning a camera on a tripod and capturing images at $30min$ intervals. Throughout the day, the sun positions varied, changing the position of shadows and incident illumination angles for each pixel as shown in Figure 5.9.

5.3 Implementation

Ray tracing is a computationally expensive task and is required for shadow detection, sky factor calculation and indirect illumination approximation. This is implemented using code written in C++, running on a computer with an Intel i7, quad core processor with multi-threading enabled. An octree [52] is used to increase the computational performance of the ray tracing algorithm.

The remaining components of the illumination invariance algorithm are implemented in MATLAB, with the exception of the Loopy Belief Propagation, which was also implemented in C++.

5.4 Spatial Illumination Invariance

A key contribution of this thesis is to develop a method to automatically generate illumination invariant images in the outdoor environment that compensates for spatial variations. The technique proposed in Section 3.2 is evaluated utilising a number of datasets in order to analyse its robustness to a number of scene properties.

The datasets are taken with hyperspectral and consumer grade RGB cameras under different weather conditions. They contain simple and complex geometries, and utilise the *full* and *fast* sky factor approximation methods.

5.4.1 Terrestrial Sunlight-Skylight Sensitivity

The terrestrial sunlight-skylight ratio (Equation 3.3) is seen to be independent of the material from which the pair of points was selected. To verify this, a VNIR hyperspectral image was captured of a ColorChecker board under two illumination conditions. Firstly, the board was exposed to terrestrial sunlight and then it was occluded. In both cases, the sky factor is approximately constant.

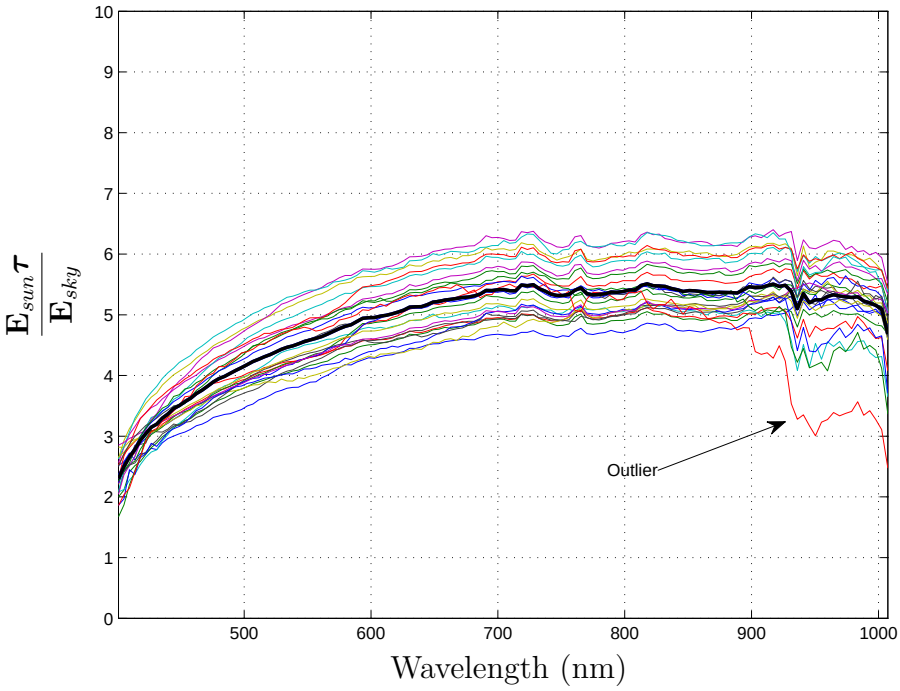


Figure 5.10 – Estimated terrestrial sunlight-skylight ratio using various reflectance colours.

Analysing Equation (3.3) shows that if the sky factors of the sunlit and shaded initialisation points are equal, the ratio is reduced to:

$$\begin{aligned}\frac{E_{sun}(\lambda)\tau(\lambda)}{E_{sky}(\lambda)} &= \frac{\Gamma_{A'}L_A(\lambda) - \Gamma_A L_{A'}(\lambda)}{L_{A'}(\lambda)\cos\alpha_A}, \\ &= \frac{\Gamma}{\cos\alpha_A}\left[\frac{L_A(\lambda)}{L_{A'}(\lambda)} - 1\right], \\ &\propto \frac{L_A(\lambda)}{L_{A'}(\lambda)} - 1. \end{aligned} \tag{5.7}$$

Therefore, to test the sensitivity of the ratio to the reflectance colour, the ratio between the pixel intensity values of the colour patches in the sunlit and shaded images is calculated. The colour patch values are found by manually segmenting each patch and calculating the average spectra. This is shown in Figure 5.10 and reveals that while there is approximately 20% variation in the magnitude, the spectral shape of the ratio is quite consistent. The magnitude variation exists due to factors such as indirect illumination and fact that the ColorChecker board itself is not actually

a diffuse reflector. Instead, each patch has some specular reflectance characteristics that are not catered for within the outdoor illumination model.

There are some outlier ratio estimates that occur from 800nm to the near infra-red region. These occur on patches 22 to 24 and are caused by the low reflectance in the infra-red region, which yield a low signal-to-noise ratio. Therefore, when using diffuse materials, the ratio between the dominant illumination sources is shown to be consistent in spectral shape, but with some magnitude variation depending on which material the ratio is derived from. This is important when employing the system on a complex outdoor scene, as obtaining the ratio estimate from a single material will only work well for the respective material. Taking the mean ratio value from a number of materials throughout the scene will lead to increased illumination invariance characteristics in the relit image, but it cannot be expected to be perfect due to several of the assumptions (diffuse material reflectance, no indirect illumination and no emission) being broken.

5.4.2 Simple Geometry

The proposed spatial illumination invariant approach is evaluated on Dataset 2, which provides a simple geometric scene observed using a VNIR hyperspectral camera. The original image (Figure 5.11a) consists of shale, with a region occluded from the sun. Visually, it is difficult to determine how many materials are present in the scene. After relighting the image with respect to diffuse skylight (Figure 5.11b), it is much clearer that the board consists of a single material. Along the borders of the shadow, there are some noticeable artefacts in the relit image due to the slight misalignment of the point cloud and the fact that shadows are not truly binary.

The colour difference between the two images is explained by analysing the mean spectra of the two regions (Figure 5.13). Following the application of relighting to generate a spatially illumination invariant image, the peak at approximately 550nm in the sunlit spectra becomes more pronounced. The cause of this is due to appearance being dependent on both reflectance and the illumination source. In this case the

(a) Original image with shadow being cast on the shale panel yielding an area of low intensity.

(b) Relit image with all pixels using skylight as their source of illumination. The relit image has a multiplicative gain factor applied to it for visualisation purposes.

Figure 5.11 – Relighting Dataset 2 with respect to a single common illuminant for all pixels allows the uniformity of the shale panel to be recovered.

illuminant is diffuse skylight, which also peaks around this region.

In order to quantitatively analyse the performance of the spatial illumination invariance algorithm, samples along a line transitioning from the shadowed to sunlit regions are obtained from the ground truth, original and relit images, as shown in Figure 5.14. Each sample spectra is compared against the mean spectra of the samples in terms of the spectral angle. The results demonstrate that for the ground truth image, the spectral angles are quite small, indicating that there is a large degree of similarity between the spectra along the line. This is as expected due to the relatively uniform nature of the material. The original image shows that there is distinct separation between samples obtained in the shadowed regions compared to the mean spectra as shown in Figure 5.13a. This is due to the illumination source, which is skewed towards shorter wavelengths, affecting the appearance of the material in accordance with the outdoor illumination model (Equation 2.16). After relighting, there is a noticeable improvement in terms of the spectral shape similarity in the shadowed region, with

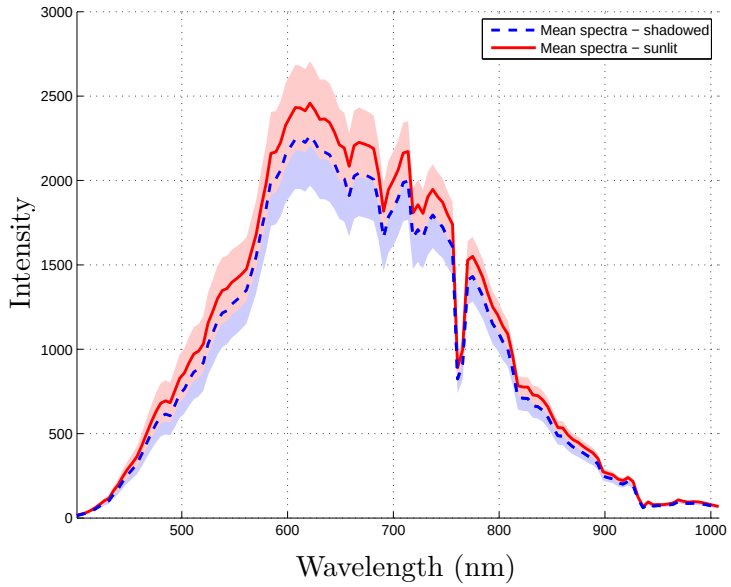


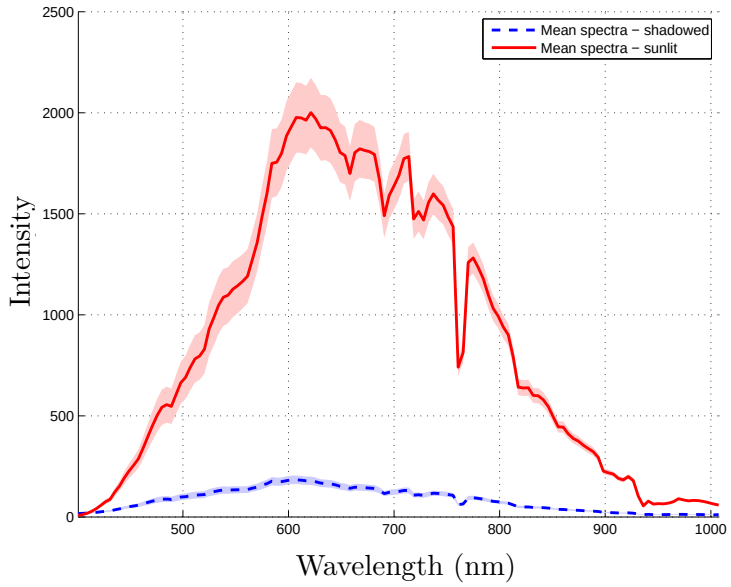
Figure 5.12 – Ground truth spectra of the shadowed (*blue*) and sunlit (*red*) areas are obtained by flat-field correction of the same scene (Dataset 2), with the calibration board at the same orientation as the panel. The shaded bands indicate one standard deviation from the mean.

the spectral angle closely matching those obtained for the ground truth image. At the transition point between the shadow and sunlit regions the spectral angle for the relit image increases dramatically due to the fact that shadows are not truly binary and the incorrect scaling factor has been applied.

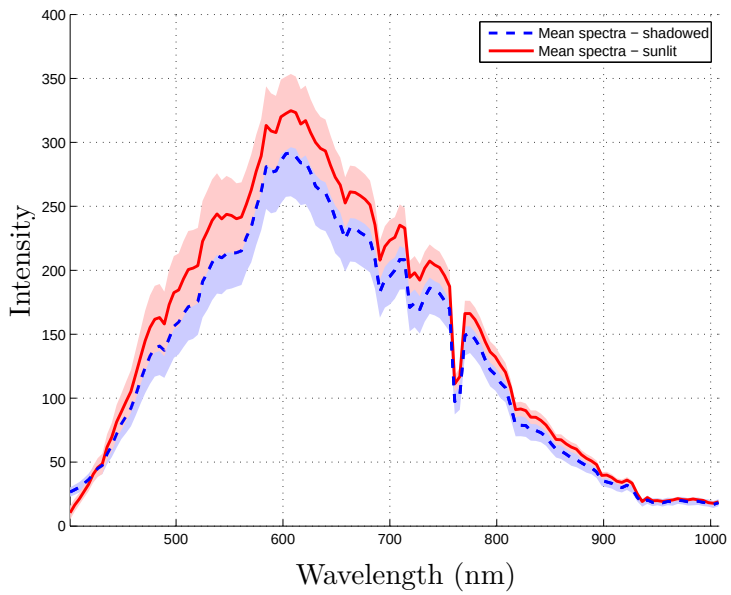
This experiment has demonstrated that when the scene closely matches the ideal conditions (simple geometry, one material and a strong shadow), spatial illumination invariance can be achieved using the approach proposed in this thesis. Relighting is seen to improve the spectral shape similarity between similar materials under different illumination conditions when measured using hyperspectral sensors.

5.4.3 Complex Geometry

Having shown that the proposed approach is able to achieve illumination invariance images for geometrically simple scenes, the algorithm is subsequently evaluated on a scene of high complexity. The spatial illumination invariance algorithm is run



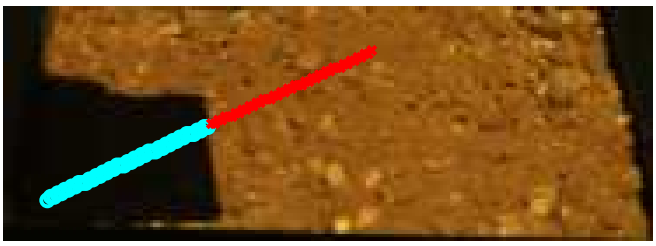
(a) Original image sample spectra.



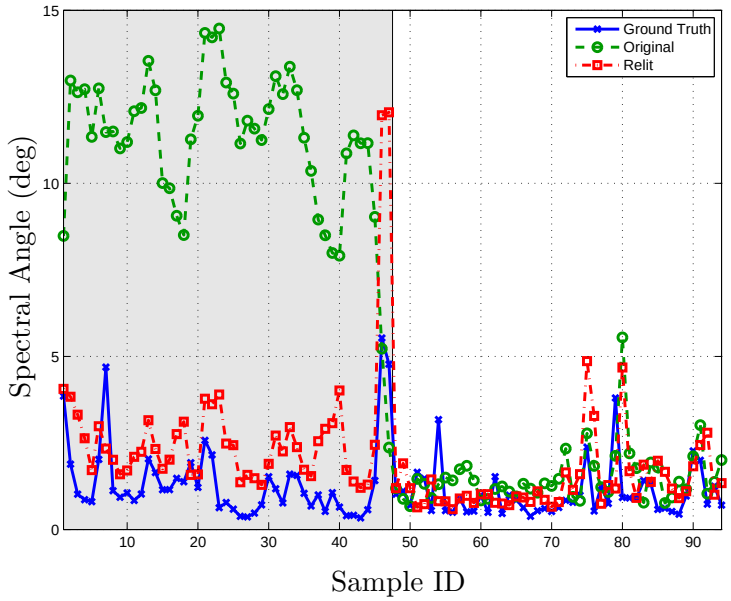
(b) Relit image sample spectra.

Figure 5.13 – Samples within the shadowed (*blue*) and sunlit (*red*) regions of Dataset 2 are shown in the original and relit images. While in both cases, there is still separability between the shadowed and sunlit regions, it is significantly less after relighting has been applied. Note the intensity scale for the relit image is low due to relighting with respect to skylight.

on Dataset 3 using the *full* sky factor calculation, with the results being shown in Figure 5.15b. The areas bounded by the red boxes highlight areas of high illumination



(a) Samples on a line that transitions from a shaded to sunlit region are taken on the ground truth, original and relit images.

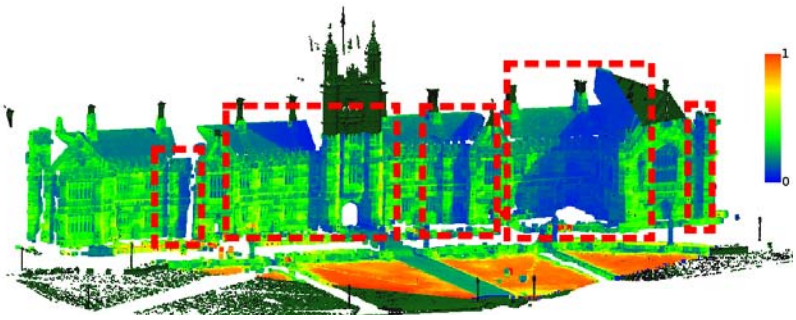


(b) Spectral angle between each sample and the mean spectra. The grey shading indicates that the sample is occluded from the sun.

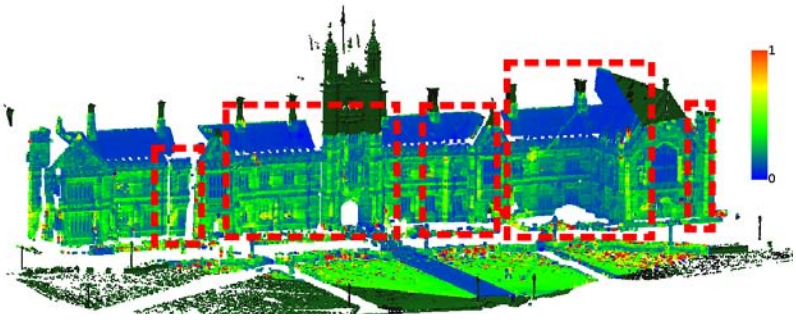
Figure 5.14 – Relighting Dataset 2 is seen to decrease the influence of illumination induced spectral variability.

variations due to occlusions. In the original image, the front face of the building is seen to have two distinctive intensities; (i) a bright intensity for areas exposed to sunlight and (ii) a low intensity for those occluded. In the relit image, the front face is of uniform colour which is more akin to the sandstone nature of the building. A similar improvement is seen for the roof, which in the relit image is also of uniform colour.

This is measured through the EMD metric where two areas in the original image are selected. These are obtained from the sandstone building under (i) terrestrial sunlight



(a) Original point cloud.



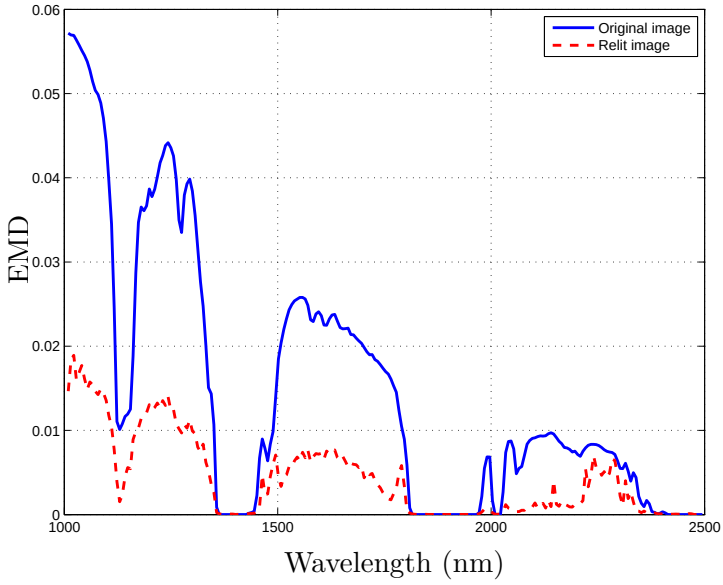
(b) Relit point cloud.

Figure 5.15 – Through relighting, illumination invariance for Dataset 3 is achieved as seen by the uniform nature of the building. Wavelength shown is 1067nm with the intensity return being normalised by the maximum value in the original image. The red rectangles indicate areas of interest which are characterised by strong shadows.

and (ii) shadowed conditions, as seen in Figure 5.16a. Histograms for the two areas are calculated on a per-wavelength basis and ideally, an invariant image would contain identical histograms for both areas. Figure 5.16b shows that for all wavelengths, the histograms are more closely aligned for the relit image as indicated by the significantly lower EMD metric value. This is a sign of the increased illumination invariance achieved by the proposed algorithm. Material variability, indirect illumination and the fact that the materials within the scene are not truly diffuse reflectors contribute to the fact that the algorithm is not able to achieve complete invariance.

A benefit of using a common illuminant for each pixel is that the resultant illumination invariant image recovers the relative colour ratios between different materials. This was verified by obtaining spectral samples of the building and grass using a spectrometer. A measurement of the incident illumination was taken using a calibra-

(a) Two regions of the same material under different illumination conditions are selected in order to evaluate the relighting system. The solid red box indicates sandstone being exposed to a combination of sunlight and skylight, while the dashed cyan box indicates that it is occluded from sunlight. The wavelengths $1010, 1250$ and 1300nm are used as the R, G and B channels respectively.



(b) EMD metric between building areas in sulight and in shadow shows that identical materials under different illumination conditions in the original image are much closer together for all dimensions.

Figure 5.16 – Comparing the similarity between identical materials under different illumination conditions in Dataset 3.

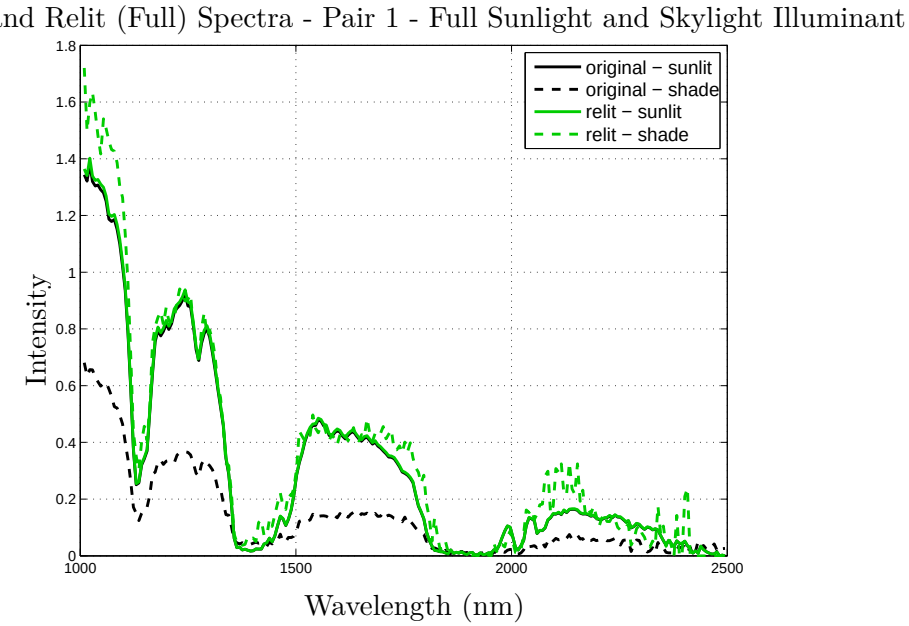
tion board and by performing flat-field correction, the reflectance spectra of the two material can be estimated. At the displayed wavelength of 1067nm, the reflectance of grass was approximately 0.6, compared to 0.5 for the sandstone building. This corresponds with the relit image where the grass in the foreground is seen to be of similar intensity as the building for the chosen wavelength. This is in contrast to the

original point cloud, where the grass was of high intensity in the image due to the small angle between the normal and the sun position at these points.

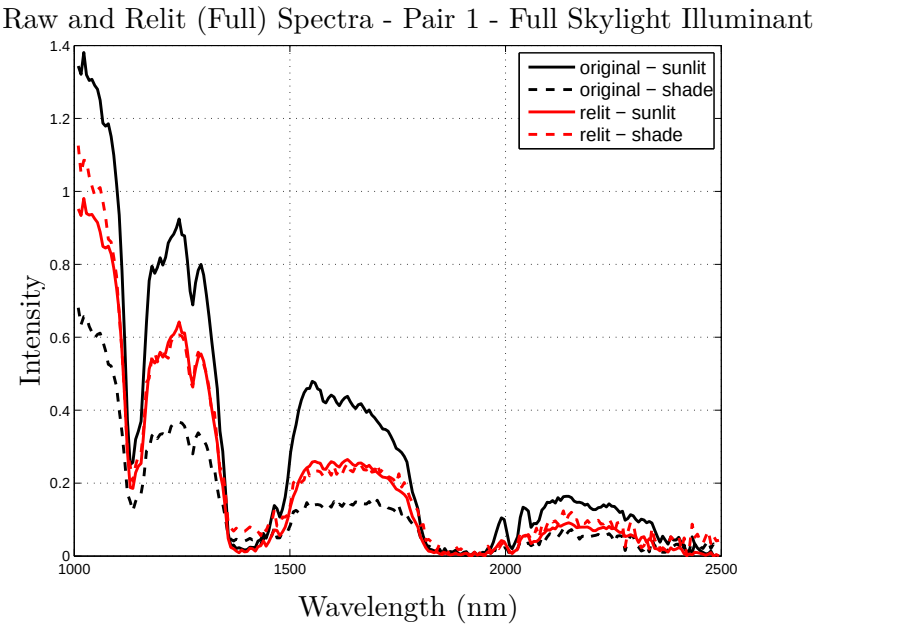
The decision to relight with respect to diffuse skylight, as opposed to a combination of terrestrial sunlight and diffuse skylight, is due to the fact that it will not amplify the already noisy shadow data (see Appendix B). This can be seen in Figures 5.17 and 5.18, where relighting with respect to full terrestrial sunlight and skylight introduces errors into the relit shadowed spectra.

Relighting with respect to skylight however, does not amplify the noise as the shadowed spectra is only multiplied by a constant. Only the points that are exposed to sunlight are multiplied by a wavelength dependent scaling factor. When using algorithms such as SAM to compare two spectra, additional noise negatively impacts the angle between them as it is a measurement of how similar the spectra are in term of shape. The spectral angle between the points in each pair of points in the validation set (Figure 5.19) is shown to decrease by more than 1° when relighting with respect to skylight. This is compared to the high noise relighting method when using terrestrial sunlight and skylight as the common illuminant, which actually decreases the similarity between the spectra.

The experiments upon Dataset 3 have shown that spatial illumination invariance can be achieved in complex scenes, where occlusions play a vital role in the appearance of the scene. It was demonstrated that the choice of diffuse skylight as the common illuminant did not significantly add noise to the measured spectra, which is beneficial over the methods developed in [29] and [35], especially at longer wavelengths. Further quantitative analysis of the validation set is provided in Section 5.4.6, where the relit spectra using the *full* and *fast* sky factor approximations are compared.

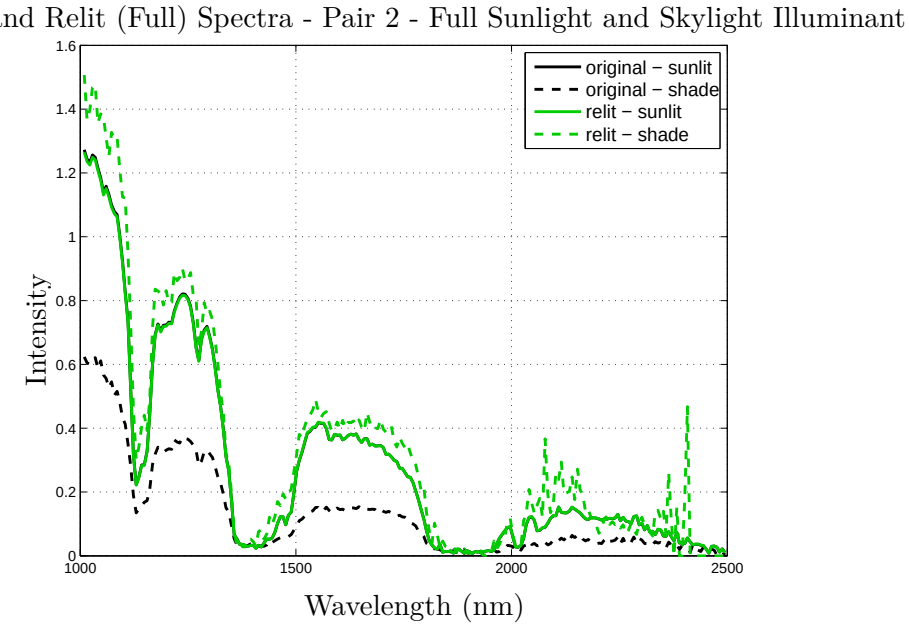


(a) The original observations are relit (*green*) using $\mathbf{E}_{sun}\boldsymbol{\tau} + \mathbf{E}_{sky}$ as the illuminant.

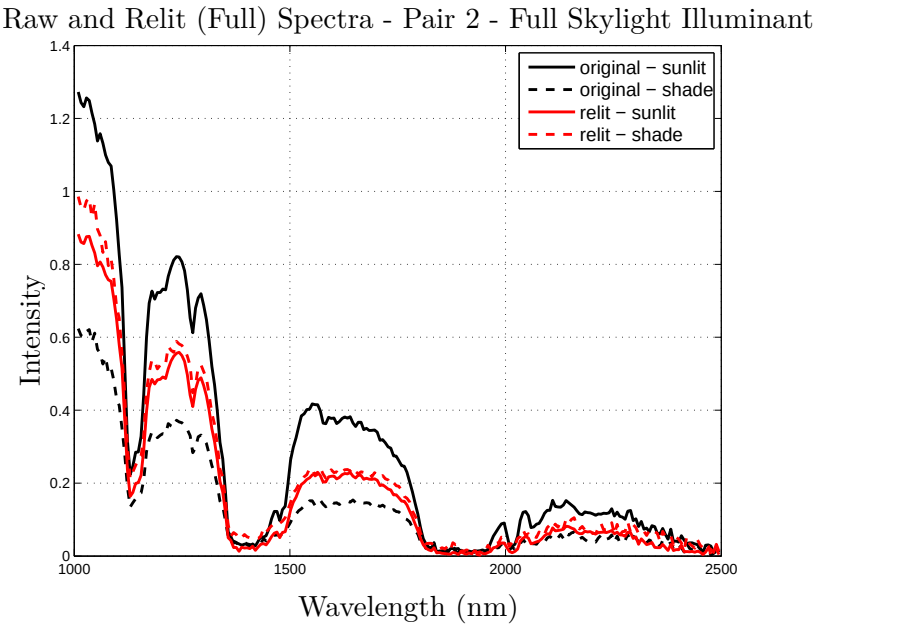


(b) The original spectra are relit (*red*) with $\mathbf{E}_{sky}$ as the common illuminant.

Figure 5.17 – Comparing the impact of relighting Dataset 3 with respect to $\mathbf{E}_{sun}\boldsymbol{\tau} + \mathbf{E}_{sky}$ vs. $\mathbf{E}_{sky}$. The graphs show a point pair, with one occluded from the sun (dashed) and the other exposed to sunlight (solid) obtained from a uniform material. In both cases, the original spectra are shown in *black* and the *full* sky factor estimate is used.



(a) The original observations are relit (*green*) using $\mathbf{E}_{sun}\boldsymbol{\tau} + \mathbf{E}_{sky}$ as the illuminant.



(b) The original spectra are relit (*red*) with $\mathbf{E}_{sky}$ as the common illuminant.

Figure 5.18 – Comparing the impact of relighting Dataset 3 with respect to $\mathbf{E}_{sun}\boldsymbol{\tau} + \mathbf{E}_{sky}$ vs. $\mathbf{E}_{sky}$. The graphs show a point pair, with one occluded from the sun (dashed) and the other exposed to sunlight (solid) obtained from a uniform material. In both cases, the original spectra are shown in *black* and the *full* sky factor estimate is used.

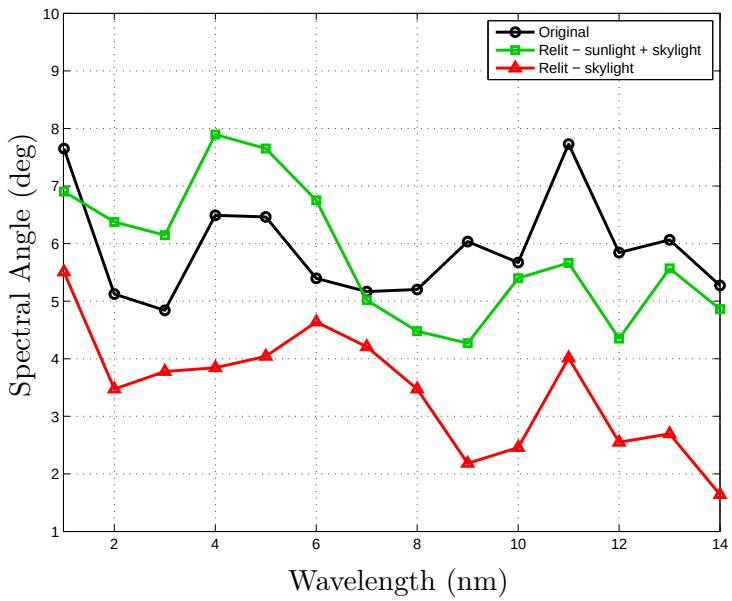


Figure 5.19 – The spectral angle between pairs of points in Dataset 3 is significantly reduced when relighting with respect to skylight, as opposed to a combination of sunlight and skylight.

5.4.4 Wideband Sensors

The original formulation of the invariant algorithm was derived with a dependence on wavelength, which implies that the sensor has a Dirac delta sensor response function. Hyperspectral cameras have narrow band responses and therefore closely fit this model, but consumer grade RGB cameras do not. The aim of this experiment is therefore to determine whether the proposed relighting algorithm can operate suitably upon cameras with wideband sensors. Linear images are obtained from the camera, with minimal post-processing and these are used to evaluate the invariant algorithm when using these sensors. To apply the proposed relighting approach, each channel is treated as having a Dirac delta sensor response function.

The original and relit image using the *full* sky factor calculation is shown in Figure 5.20 for Dataset 4. The uniform colour of the building is retained using relighting, as opposed to the bimodal nature of the original image. This is also seen for the road in the foreground of the image. Once again, the relative colours between the materials is recovered with the concrete road appearing with a low intensity in the relit image.

(a) Original image.

(b) Relit image.

Figure 5.20 – Relighting Dataset 4 using visual data from an RGB camera.

This is compared with the original image, where the high sun angle means that the sunlit road areas appear with a high intensity.

In this dataset, some artefacts arise due to the sky factor approximation. In the area of foliage, the relit image using the *full* calculation saturates for a significant percentage of the leaves in the tree. The saturation points arise due to the lack of inclusion of indirect illumination in the outdoor illumination model. The geometry of the tree means that indirect illumination plays a significant role in the appearance of the leaves that are shadowed. Therefore, due to the high amount of occlusions, the *full* calculation of the sky factor will be small, leading to the occluded scaling factor

(Equation 3.6) becoming large and saturating the relit image.

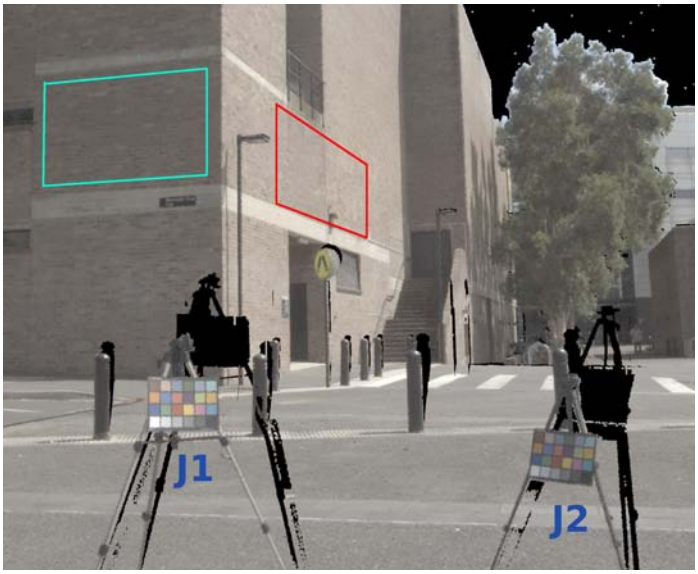
The proposed relighting algorithm is also applied to Dataset 7 (Figure 3.3), with similar qualitative results being attained. Through relighting Datasets 4 and 7, the illumination invariant system has been demonstrated to provide spatial invariance when the camera has a wideband sensor response. This is despite the fact that the scaling factors were derived with a dependence on wavelength. The caveat is that images should be used prior to the conversion to a colour space and gamma correction. This is required in order to minimise mixing of the colour channels and non-linear processing. These post-processing steps can be applied after the proposed illumination invariant algorithm in order to produce visually pleasing image. More extensive, quantitative results using wideband sensors under different weather conditions are presented in Section 5.4.5.

5.4.5 Variable Weather Conditions

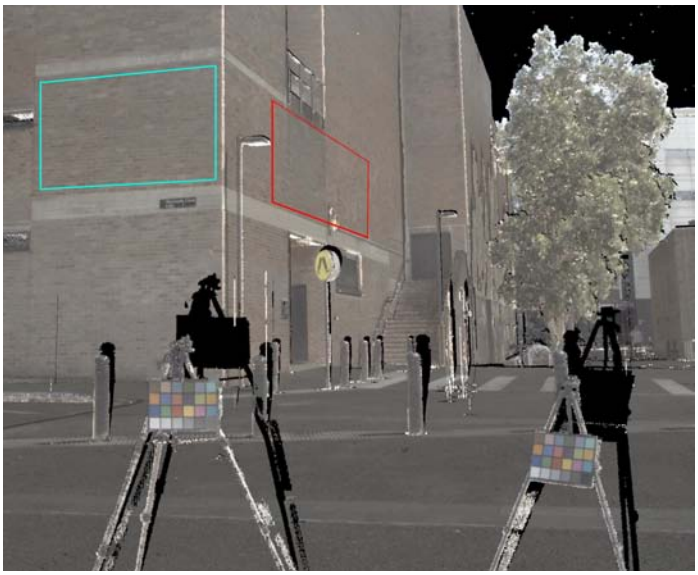
The efficacy of the proposed spatial illumination invariance system and its robustness to various weather conditions was evaluated by applying it to Datasets 5 and 6. Both scenes contain two ColorChecker boards, one exposed to sunlight, while the other is occluded. Therefore, spatial illumination invariance is quantitatively analysed by calculating the absolute intensity errors between the respective colour patches, before and after the application of relighting. The EMD metric is also calculated between a patch of the building exposed to the sun and another in shadow for all three camera channels.

The original and relit images for Dataset 5 is shown in Figure 5.21. Visually, the sunlit wall has been well compensated to match the colour of the adjacent shadowed walls. The relative colour between the various materials in the scene has also been recovered through relighting. The road section in the original image is of high intensity due to the high position of the sun. Through relighting, the low intensity nature of the material is revealed and it is shown to have similar intensity to the main building.

The EMD metric as shown in Table 5.2 reveals significant similarity between the



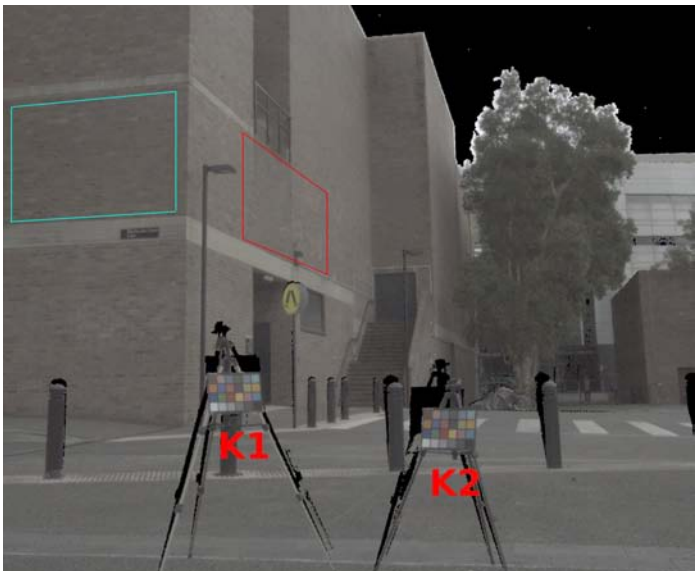
(a) Original image.



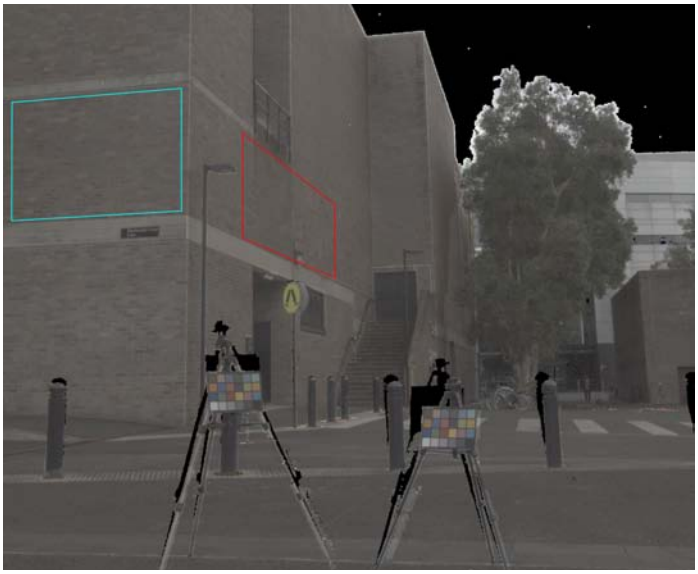
(b) Relit image.

Figure 5.21 – Relighting an RGB image from Dataset 5 under clear sky conditions. Areas exposed to terrestrial sunlight (red box) and areas in shadow (cyan box) are used for analysis.

histograms of the sunlit and shadowed regions of the building after the application of relighting. Evidence of spatial illumination invariance is also seen in the absolute intensity errors as shown in Figure 5.23 and Table 5.3. For all colour patches and across all channels, the error decreases by approximately a factor of 4.3.



(a) Original image.



(b) Relit image.

Figure 5.22 – Relighting an RGB image from Dataset 6 under overcast conditions. Areas exposed to terrestrial sunlight (red box) and areas in shadow (cyan box) are used for analysis.

Similar relighting results are seen for Dataset 6 as shown in Figure 5.22, which was captured under overcast conditions. It is worth noting that the intensity contrast observed in Dataset 5 between sunlit and shadowed region of the building is not as prevalent in this dataset. This is because terrestrial sunlight has been heavily

Table 5.2 – Earth Mover's Distance metric between points directly lit and in shadow. Lower values indicate greater similarity between the histograms

		Channel		
Dataset	**Image**	**Red**	**Green**	**Blue**
5	original	0.3289	0.2761	0.2276
	relit	**0.0274**	**0.0205**	**0.0182**
6	original	0.0569	0.0536	0.0485
	relit	**0.0185**	**0.0191**	**0.0213**

Table 5.3 – Absolute intensity error mean and standard deviation for all colour patches. The errors in the relit image are observed to be lower than that of the original image in both sunny and overcast conditions, indicating greater illumination invariance.

	Absolute Intensity Error					
	original image			relit image		
	R	G	B	R	G	B
Dataset 5 - J1 vs. J2						
μ	0.256	0.221	0.198	**0.068**	**0.050**	**0.042**
σ	0.178	0.154	0.152	**0.058**	**0.047**	**0.044**
Dataset 6 - K1 vs. K2						
μ	0.019	0.017	0.016	**0.005**	**0.007**	**0.008**
σ	0.012	0.011	0.011	**0.004**	**0.006**	**0.007**

attenuated by the clouds, thereby reducing the illumination variations in the scene.

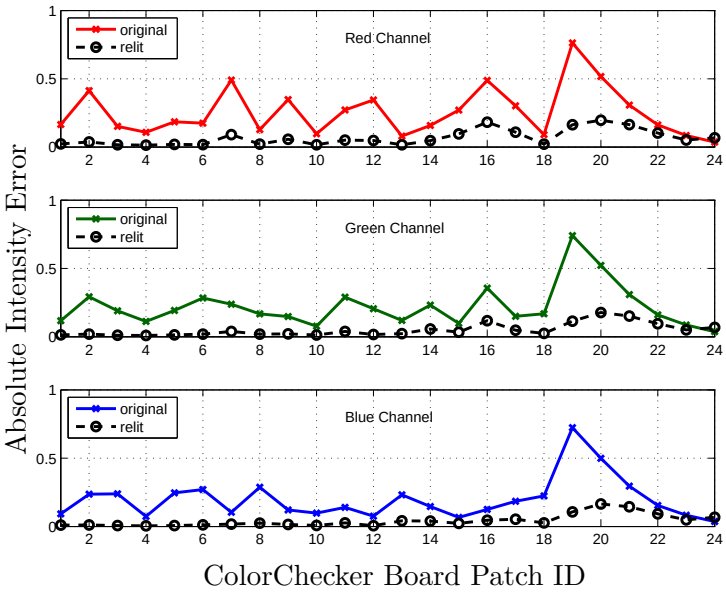
The building is coloured with different shades of brown between the relit datasets and this is due to the fact that when relighting, there is still a dependency on the skylight spectra present at the time of capture. On the overcast day, the skylight is of low intensity and this influences the appearance of the relit image.

The absolute intensity errors shown in Figure 5.23 reveal that while increased similarity is achieved through relighting, the magnitude difference is far less than that which was achieved in Dataset 5. The EMD metric shows an improvement by a mean factor of around 2.7, though the metric is already small for the original image. Once again, this is due to the small amount of illumination variation in the scene brought about by attenuation of terrestrial sunlight.

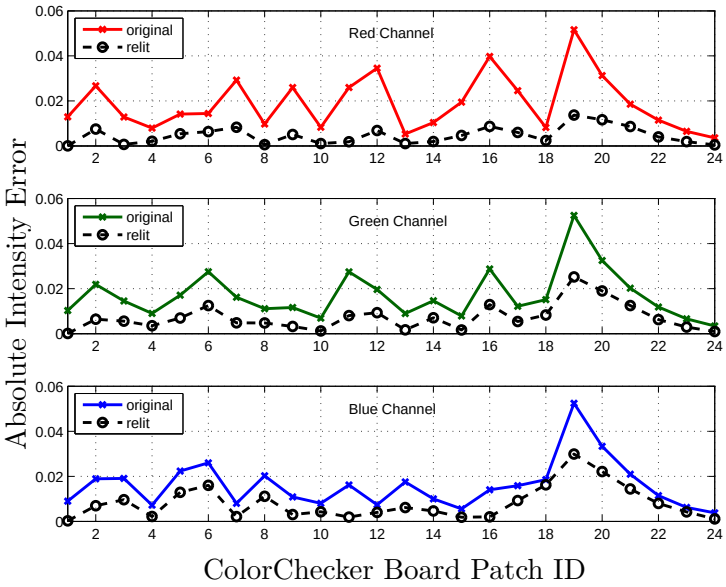
There are some noticeable artefacts in the relit images, especially on the corners of the edges of the building where it is relit incorrectly. This is because the calculation

to find the normals requires the surrounding points and is therefore smoothed over this region. This causes inconsistencies in the shadow detection algorithm which are minor and could be removed with some post-processing.

As shown with both the decrease in absolute intensity errors for ColorChecker boards under different illumination conditions, and smaller EMD metric values for regions of the building, the proposed physically based algorithm is able to achieve spatial illumination invariance within the scene. The method can operate in the presence of clouds and has been demonstrated to achieve invariance when using wideband sensors.



(a) After relighting has been performed, there is a significant decrease in the l_1 error between $J1$ and $J2$.



(b) Only a slight improvement is noticeable between $K1$ and $K2$, which is due to the fact that terrestrial sunlight is heavily attenuated by the clouds present in the scene. This makes the image already partially illumination invariant, with the remaining variations predominantly due to varying sky factors and indirect illumination.

Figure 5.23 – Absolute intensity error between ColorChecker boards under different illumination conditions in Datasets 5 and 6.

5.4.6 Fast Sky Factor Approximation

For applications such as robotics, computational speed is a high priority and as a result, the *fast* sky factor approximation was proposed in Section 3.6. In this section, the *full* calculation is used as the ground truth data. The method is assessed in terms of absolute error and computational time on a number of datasets. The spectral shape and intensity of identical materials under different illumination conditions is used to evaluate the proposed illumination invariant system when using the *full* and *fast* sky factor approximation methods.

In terms of accuracy, Figures 5.24 and 5.25 present the absolute error between the two methods. As expected, in areas where occlusions between a region and the sky dome are more prevalent, such as the intersection of the ground plane and the building wall, a larger amount of error is found. Other areas include the intersection of two orthogonal walls towards the right of Figure 5.24 and in the tree of Figure 5.25. This is because the LIDAR has sparsely sampled the tree, which is also dynamic in nature due to wind.

There is a significant reduction in the amount of computation required for the two methods. The proposed method can compute the sky factors for a point cloud of

Table 5.4 – Computation time for calculating sky factors using the *full* and *fast* methods. Using the *fast* approximation is seen to take a significantly shorter amount of time.

	Dataset 3	**Dataset 4**
Number of Points	321,880	574,980
Sky Factors - full	4.5h	9.6h
Sky Factors - fast	33s	58s

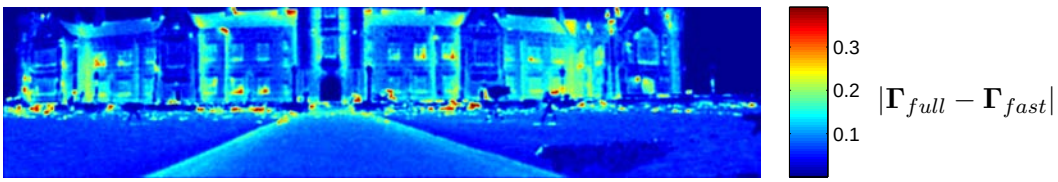


Figure 5.24 – Absolute error between *full* and *fast* approximation of sky factor for Dataset 3. Larger errors are noticeable at the intersections of walls.

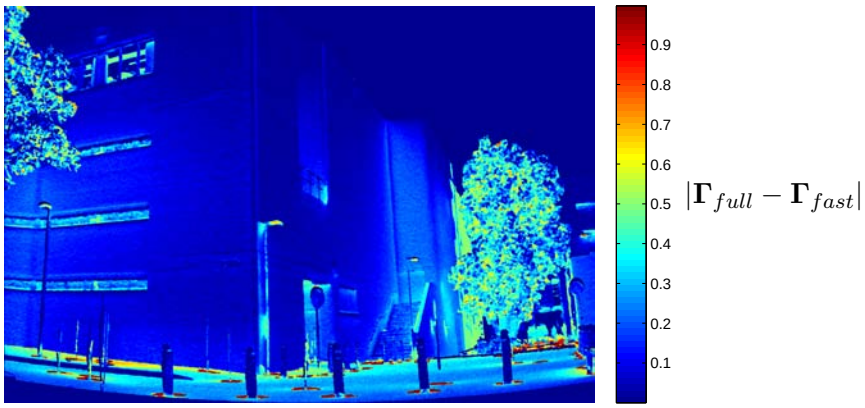


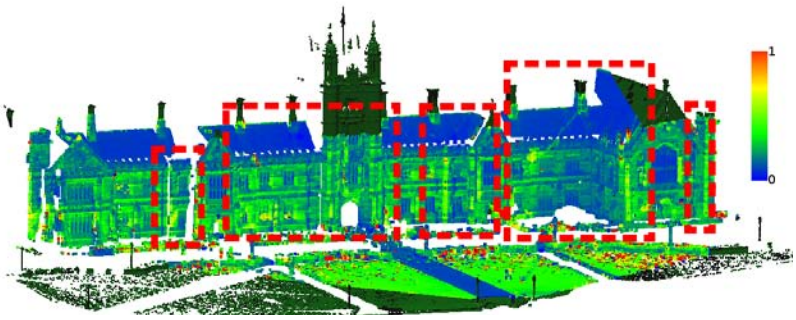
Figure 5.25 – Absolute error between *full* and *fast* approximation of sky factor for Dataset 4. Once again, larger errors are noticeable at the intersection of walls as well as in foliage.

more than 550,000 points in around 1 minute. This is compared to the 10 hours required for the *full* method and is due to the fact that ray tracing complex scenes, such as those captured from field based platforms, is an expensive task.

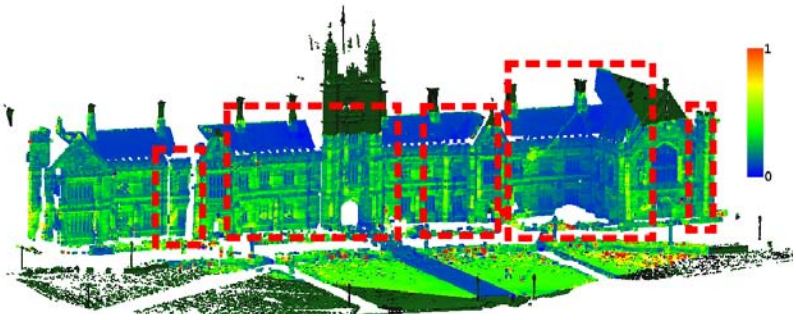
The evaluation of the *fast* sky factor approximation in terms of computation time has shown that significant improvements are achieved when ignoring occlusions in the scene. This makes it ideal for applications where real time performance is a requirement, such as robotics platforms which utilise the perception module for path planning. However, the assumption of no occlusions means that the accuracy of the outdoor illumination model decreases, since the diffuse skylight intensity will be modelled as higher than what it actually is. The effects of this trade off between computation time and accuracy are therefore evaluated on Dataset 3.

Visually, the relit results when using the *fast* approximation closely match those produced using the *full* calculation as shown in Figure 5.26. The difference between the two relit images is more pronounced on the right side of the building, where two orthogonal walls intersect. In this region, the *fast* approximation decreases in accuracy (see Figure 5.24) as occlusions play a greater role in determining how much of the sky is visible in this region.

Relighting using the *full* and *fast* sky factors is quantitatively compared using 14 manually selected pairs of points, with each pair obtained from the same material



(a) Relit point cloud - *full*



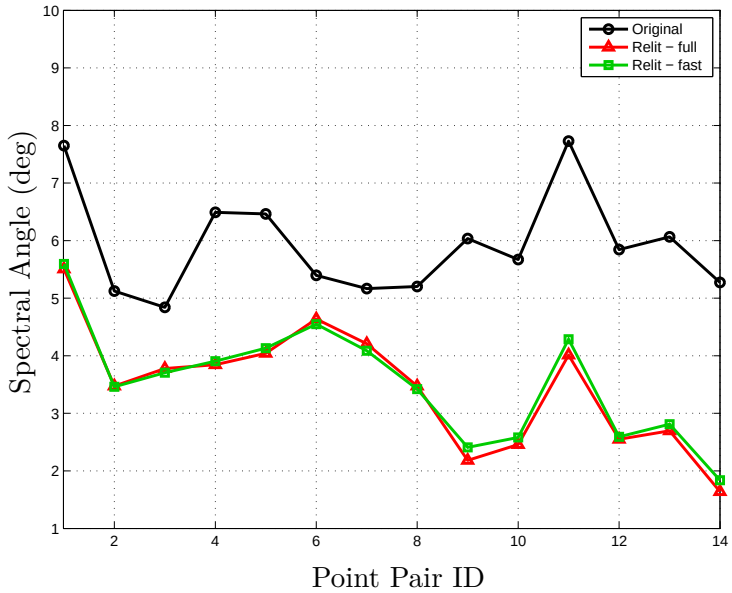
(b) Relit point cloud - *fast*

Figure 5.26 – Through relighting, illumination invariance for Dataset 3 is achieved as seen by the uniform nature of the building. Wavelength shown is 1067nm with the intensity return being normalised by the maximum value in the original image. Red boxes are used to outline areas of interest characterised by strong shadows in the original image.

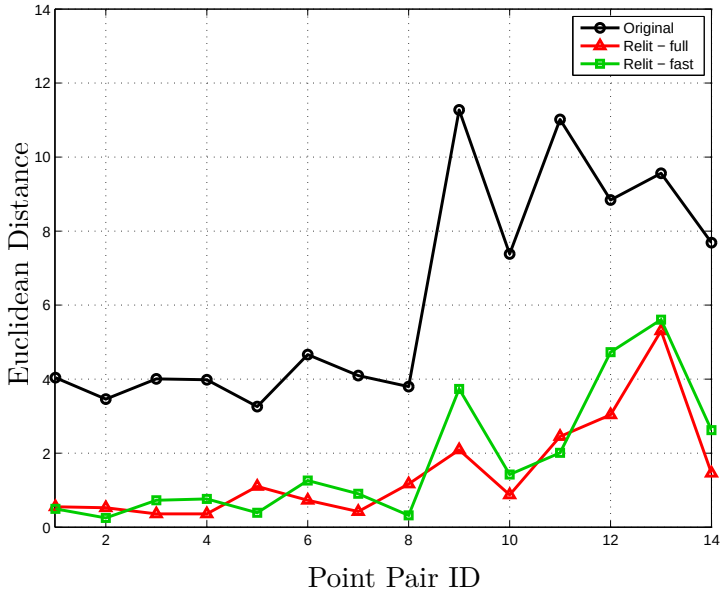
under sunlit and shadowed conditions (the full set of results is shown in Appendix C). The SAM and Euclidean distance metrics are used to evaluate their similarity in terms of spectral shape and intensity.

For all selected points, relighting increases the spectral shape similarity between the pairs as seen by the lower spectral angle in Figure 5.27. Using the *fast* approximate is seen to produce similar or identical results to relighting using the *full* calculation.

In terms of spectral intensity, large improvements are seen with the error decreasing significantly. When occlusions are present and the *fast* approximation decreases in accuracy, the error increases as would be expected. However, it is still much less than using the original data and is therefore beneficial for discriminating between material types in the scene. The larger magnitudes of error for points 9 to 14 are because the



(a) Spectral shape comparison using SAM metric.



(b) Spectral intensity comparison using Euclidean distance metric.

Figure 5.27 – Comparing the spectra of selected pairs of points, with each pair obtained from a uniform material under different illumination conditions. Relighting the scene using the *full* and *fast* methods is shown to decrease both SAM and Euclidean distance metrics.

original signal is more intense when compared to points 1 to 8.

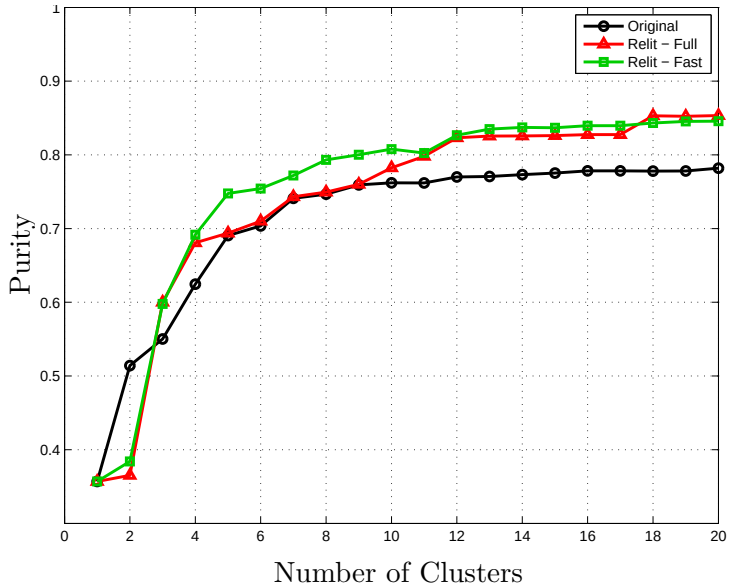
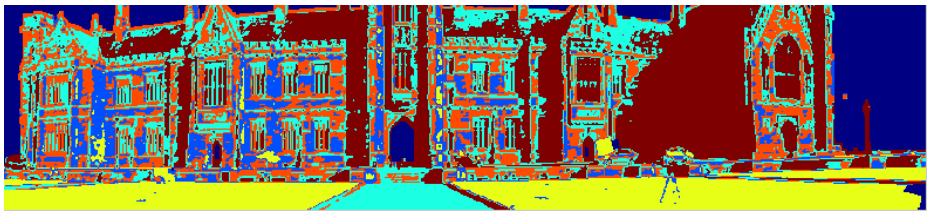
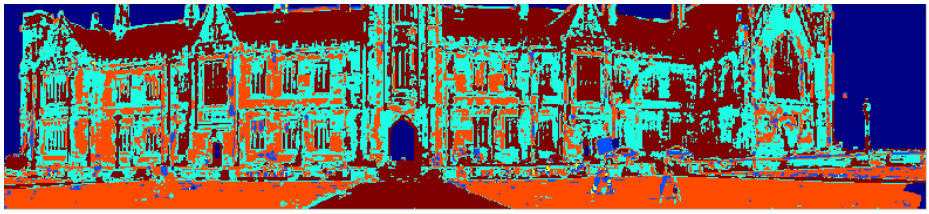


Figure 5.28 – Clustering purity is seen to improve following the relighting of Dataset 3, whether it be through the use of *full* or *fast* approximation of the sky factors.

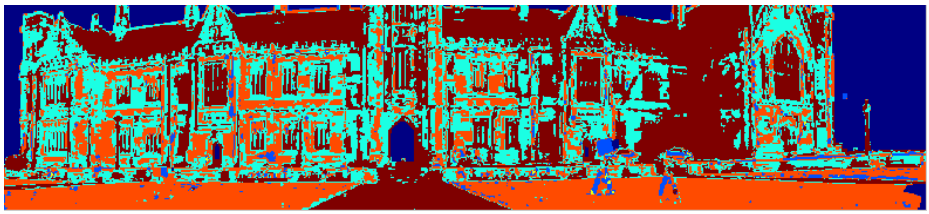
As a preliminary application for the proposed spatially illumination invariant algorithm, it is evaluated using k-means clustering. The data is hand labelled using the classes of building, roof, grass and road, with areas containing no laser return being excluded from analysis. The number of clusters is varied from 1 to 20 and the average purity of the clusters is calculated. A marked improvement is observed when using the relit images as shown by the higher cluster purity scores shown in Figure 5.28 for $k > 2$. This is mainly due to the shadowed regions of the building and roof being clustered together when using the original data as shown in Figure 5.29 which presents the clustering results for $k = 5$. For clusters 5 to 10, the *fast* method is seen to outperform relighting using the *full* method. This is due to the saturation of pixels when using the *full* method, leading to pixels of different classes being clustered together. After the number of clusters increases past 10, these saturated pixels are separated into individual clusters, meaning there is litle or no impact on the overall purity. Relighting using the *full* sky factor calculation is seen to combine the majority of the building into one cluster, indicated by the cyan colour. Similar attributes are observed when using the *fast* approximation to the sky factors, with the exception of areas near the right hand side of the building. In this area, the *fast* approximation de-



(a) Original image.



(b) Relit image - *full*



(c) Relit image - *fast*

Figure 5.29 – k-means clustering results using k=5 applied to Dataset 3. Clustering using the original data is seen to combine areas based on both material and illumination, such as the building and the roof. Using the relit data with *full* sky factor calculations allows the building face to be combined into predominantly one cluster. When using the *fast* approximation, the performance slightly decreases in areas where the sky factor approximation is less accurate.

creases in accuracy due to the occlusions present and the compensation performance decreases. Nevertheless, it still improves upon the performance of the original image.

A noteworthy feature of the clustering results of the original image is that it clusters the grass (yellow) and road (cyan) in the foreground particularly well. This is predominantly due to the fact that they are exposed to terrestrial sunlight, allowing these areas to be separated due to their high intensity when compared to the rest of the scene.

When utilising the *fast* approximation for sky factors on Dataset 4, similar results are achieved as seen in Figure 5.30. A noticeable difference is that there is a reduction in the saturation artefacts within the foliage region. This is because the *fast*

(a) Relit Image - *full*

(b) Relit Image - *fast*

Figure 5.30 – Relighting Dataset 4 from an RGB camera using the *full* and *fast* sky factor approximation.

approximation ignores the occlusions and utilises the upper bound for the sky factor. This is beneficial for this particular case as it prevents the occluded scaling factor (Equation 3.6) from becoming too large and causing the relit image to saturate.

The experiments on Datasets 3 and 4 have shown that relighting the scene using the proposed approach increases the spectral shape and intensity similarity between identical materials in the scene. The *fast* approximation to sky factors was demonstrated to provide reasonable invariance properties despite not taking account occlusions in the scene. It approximates the sky factors rapidly, making it suitable for online

applications where computation time is required to be kept to a minimum.

5.4.7 Summary

In this section, spatial illumination invariance was shown to be achieved through the use of the proposed automatic technique. The complementary nature of geometry and imagery was valuable in that it allowed each illumination source to be modelled and compensated for independently. The scene was relit with respect to diffuse skylight in the scene, meaning identical materials could appear similarly, regardless of their original incident illumination.

One of the key advantages of using a physically based approach is that the illumination invariant representation of the scene retains the full spectral dimensionality of the original data. This means there is no loss of valuable colour information that may be useful for high level algorithms to use when discriminating between materials.

The *fast* sky factor approximation was demonstrated to provide reasonable illumination invariance properties, despite the fact that it ignores occlusions in the scene. Not only is it able to compensate for illumination variations, but it does so rapidly when compared to the *full* calculation. Robotics and other real time applications would benefit from such an approximation as it allows the proposed illumination approach to be used online.

The experiments also demonstrated that spatial illumination invariance could be achieved in a variety of weather conditions. Heavy cloud attenuates sunlight, but due to the use of geometric data, shadow locations can still be identified and used to calculate the terrestrial sunlight-skylight ratio. This allows compensation to occur, with the main variability in the scene being due to varying exposure to diffuse skylight.

5.5 Temporal Illumination Invariance

While illumination invariance was shown to compensate for spatial variations in the scene, images taken under different lighting conditions will change in appearance. This is seen by analysing the outdoor illumination model (Equation 2.16), where variations in the transmittance and diffuse skylight spectra will cause the appearance of a material to change. This section aims to show the temporal illumination invariant properties achieved by the proposed system. Illumination variations due to changes in weather conditions and sun position are evaluated, and the approach is shown to provide significant improvement due to relighting the scene.

5.5.1 Variable Weather Conditions

To validate temporal illumination invariance Datasets 5 and 6 are used, which are images of the same area under sunlit and overcast conditions (Figure 5.31, left column). The temporal skylight scaling factor presented in Section 3.3.2 is found using the shadowed overlapping regions and the images are relit with respect to:

- $E_{sky,5}$ - the diffuse skylight present in Dataset 5. The relit results are shown by the middle column of Figure 5.31,
- $E_{sky,6}$ - the diffuse skylight present in Dataset 6. The relit results are shown by the right column of Figure 5.31.

Visually, the colour of the building is seen to be more consistent when using the temporal skylight scaling factor. The overcast conditions present in Dataset 6 mean that it is of lower intensity than Dataset 5.

Quantitative analysis is performed by using the ColorChecker boards present in each scene as they are identical and the incident illumination upon them can be controlled. Spatial illumination invariance was shown previously when comparing $J1$ against $J2$ and $K1$ against $K2$ in Section 5.4. For temporal invariance when relighting with

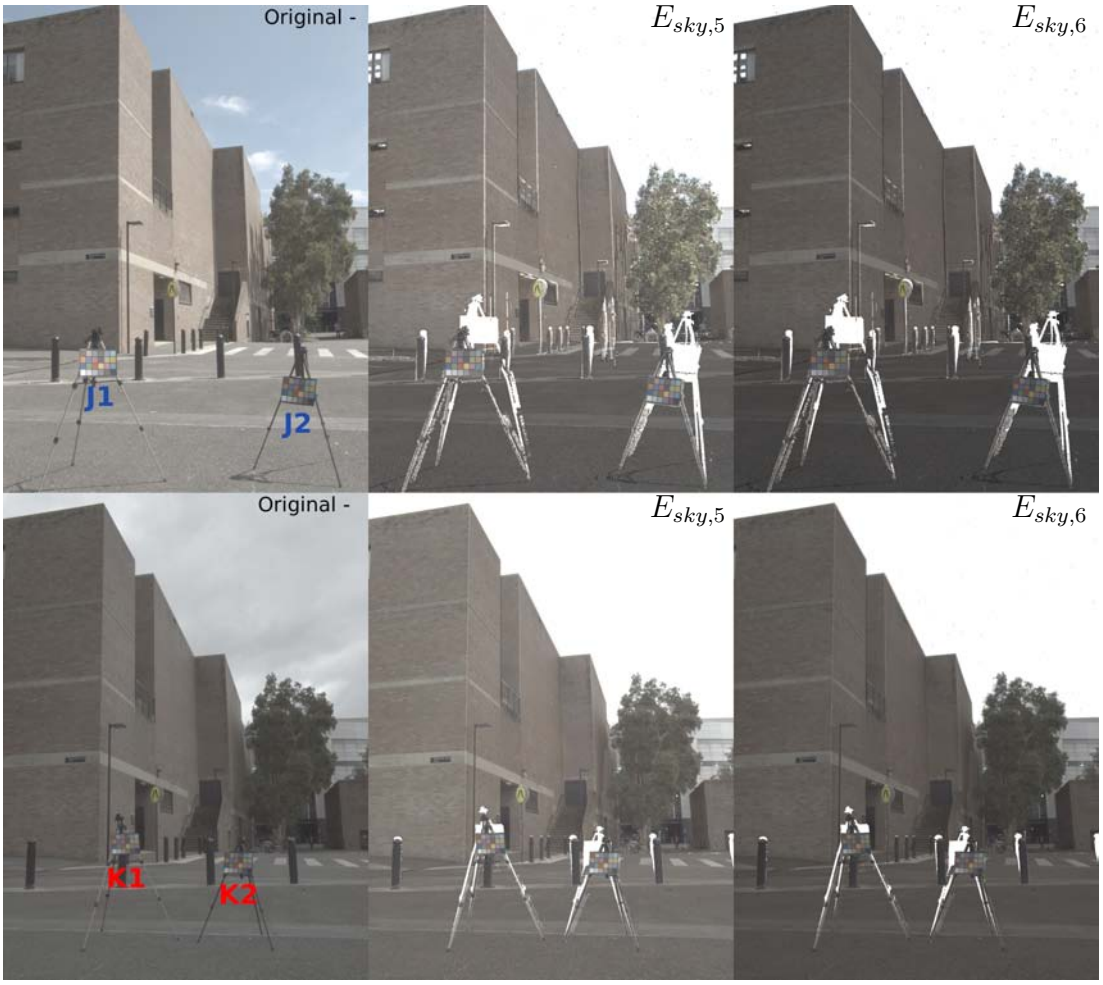


Figure 5.31 – Temporal illumination invariance is attained by relighting both images using the same illuminant. *From left column to right:* original images from Datasets 5 and 6 under sunlit and overcast conditions respectively; both images are relit with respect to $E_{sky,5}$; both images are relit with respect to $E_{sky,6}$.

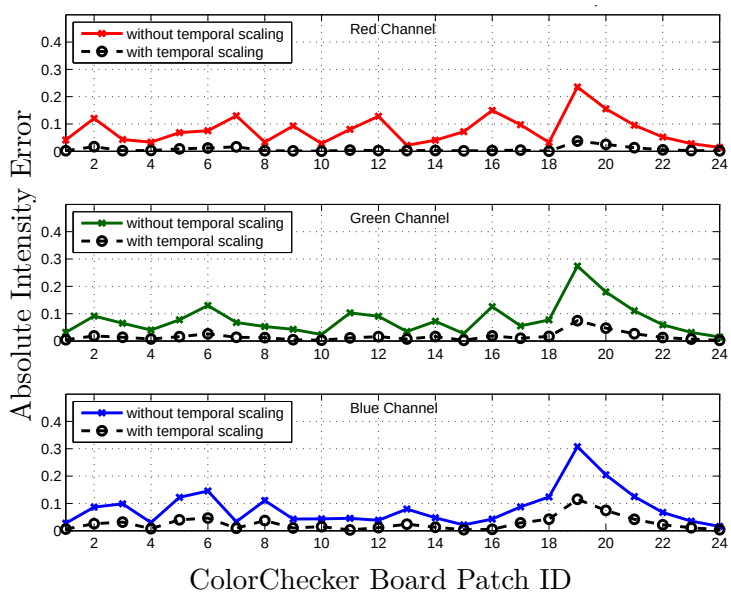
respect to $E_{sky,5}$, $K1$ and $K2$ are compared against $J2$. Similarly, when relighting with respect to $E_{sky,6}$, $J1$ and $J2$ are compared against $K2$. This comparison is performed using the relit images, before and after application of the temporal skylight scaling factor and is shown in Figures 5.32 and 5.33. The shadowed boards are considered as the reference measurements as the relighting scaling factor is a purely geometrical term that is constant across all channels and does not introduce noise into the image (see Equation 3.6).

Table 5.5 – Absolute intensity error mean and standard deviation for temporal illumination invariance. Application of the temporal skylight scaling factor is observed to lead to lower errors for all colour channels.

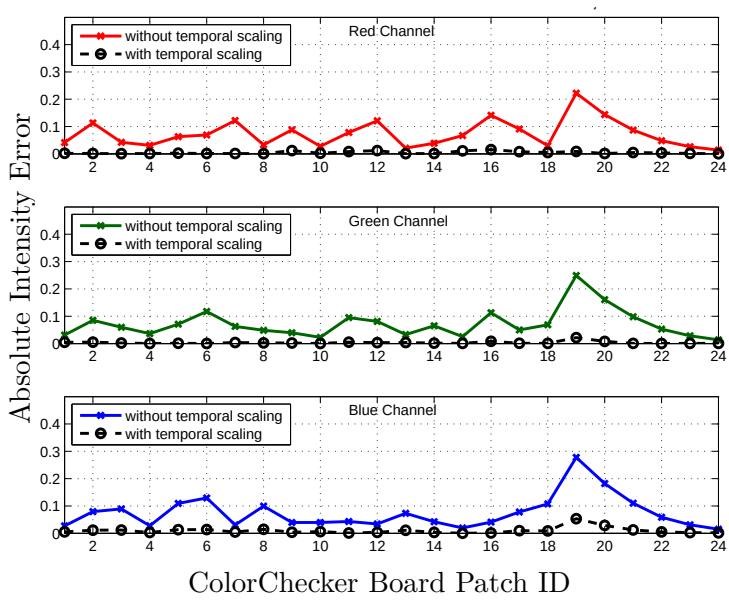
	Relighting with respect to $E_{sky,6}$					
	without temporal scaling			with temporal scaling		
	R	G	B	R	G	B
J1 vs. K2						
μ	0.141	0.121	0.116	**0.031**	**0.025**	**0.025**
σ	0.100	0.088	0.091	**0.028**	**0.024**	**0.025**
J2 vs. K2						
μ	0.073	0.071	0.074	**0.002**	**0.002**	**0.005**
σ	0.051	0.052	0.060	**0.002**	**0.002**	**0.005**
	Relighting with respect to $E_{sky,5}$					
	without temporal scaling			with temporal scaling		
	R	G	B	R	G	B
K1 vs. J2						
μ	0.078	0.078	0.082	**0.007**	**0.016**	**0.026**
σ	0.054	0.058	0.067	**0.009**	**0.016**	**0.026**
K2 vs. J2						
μ	0.073	0.071	0.074	**0.004**	**0.003**	**0.010**
σ	0.051	0.052	0.060	**0.005**	**0.005**	**0.011**

For all comparisons, the use of the temporal skylight scaling factor reduces the absolute intensity error over all colour patches. This is an indication of the large influence that illumination has on the appearance of materials in the scene. In this case, there is a significant change in both colour and intensity of the skylight illuminant used in both images due to the different weather conditions.

In both relighting cases, the error for the board exposed to sunlight is higher when compared to the board that is in shadow, reflecting the fact that the non-occluded scaling factor (Equation 3.8) is wavelength dependent (due to the terrestrial sunlight-skylight ratio term) and is more sensitive to small errors in the estimation of the terrestrial sunlight-skylight ratio and the angle towards the sun. Nevertheless, relighting the sunlit boards with the temporal skylight scaling factor provides improved similarity with the boards in the second image, compared to just using the relit image without temporal scaling.

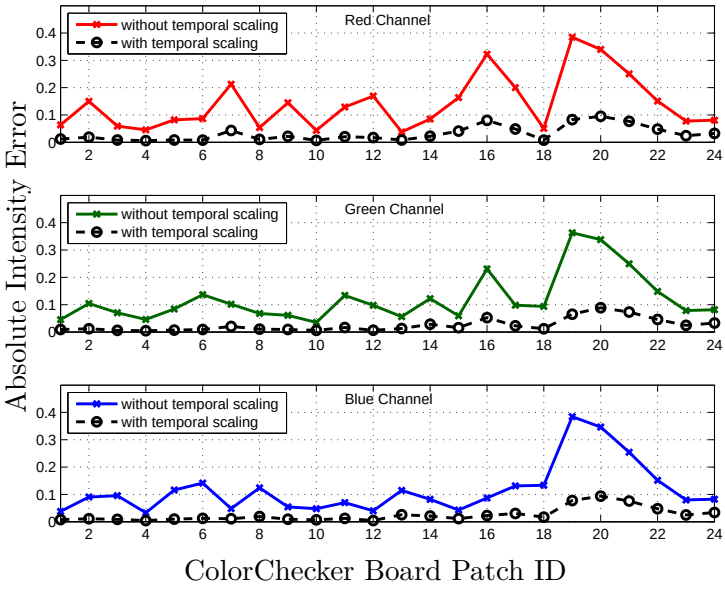


(a) The error is calculated between board $K1$ and $J2$ (images relit with $\mathbf{E}_{sky,5}$) and shows that for all three colour channels, the relit image experiences a decrease in error for all colour patches.

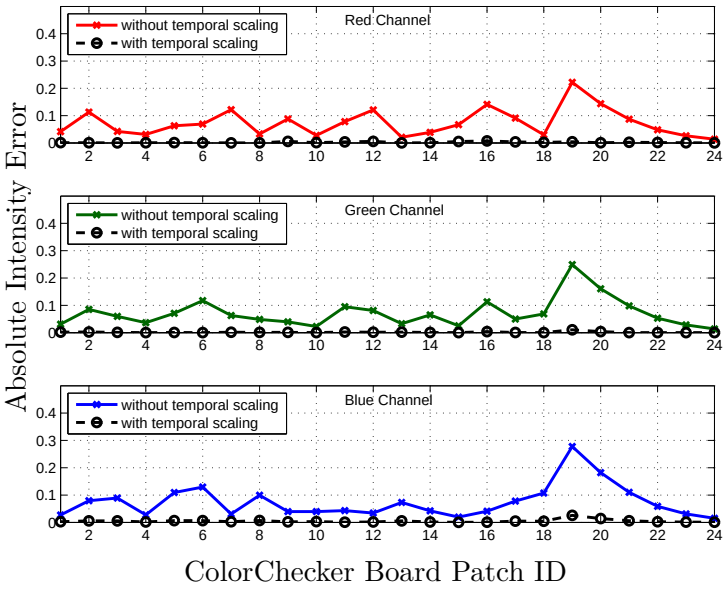


(b) The error between board $K2$ and $J2$ (images relit with $\mathbf{E}_{sky,5}$) is calculated and shows similar improvements.

Figure 5.32 – Datasets 5 and 6 are relit with respect to $E_{sky,5}$ and the absolute intensity error is calculated between the ColorChecker boards placed in each scene.



(a) The error between boards $J1$ and $K2$ (images relit with $\mathbf{E}_{sky,6}$) shows a decrease in error when using the relit image.



(b) Similar improvements are seen when comparing $J2$ and $K2$ (images relit with $\mathbf{E}_{sky,6}$).

Figure 5.33 – Datasets 5 and 6 are relit with respect to $E_{sky,6}$ and the absolute intensity error between the ColorChecker boards in each scene are calculated.

5.5.2 Variable Sun Position

In order to evaluate temporal illumination invariance throughout the day, Dataset 8 is used, which images a static scene as it varies in appearance due to changes in the sun position and overhead clouds.

Ten sample patches of uniform material (grass) were selected and used for validation. The appearance of each patch varies significantly throughout the day as shown in Figure 5.34a. For example, as sample patch 3 moves from being occluded to exposed to terrestrial sunlight ($t = 60$), its intensity increases dramatically. As the sun position changes and the patch remains exposed ($t = 60 \rightarrow 180$), its intensity decreases smoothly until it is once again shadowed ($t = 240$). There is a small, brief spike in intensity due to being exposed to sunlight at a large grazing angle ($t = 330 \rightarrow 360$).

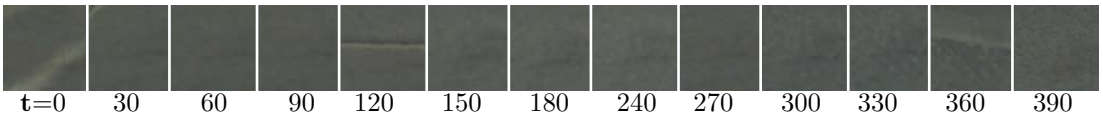
The absolute intensity error (l_1 metric) is obtained by calculated the mean intensity of a patch for each of the three channels at time k, and comparing it against the mean intensity of the respective patch at time 0. In the original image case, there is a large amount of variance in the error, as the appearance is influenced by the illumination sources. Patches 9 and 10 are observed to have less variability due to the fact that they remained in shadow for the majority of the time-lapse.

Relighting each image independently reduces the error significantly as shown in Figure 5.36a (note the change in scale for the y-axis compared to Figure 5.35). However, while spatial illumination invariance is achieved, temporal consistency is not enforced. This can be seen for example in patches 1 to 8, which contain outliers. This is due to the fact that the final image captured occurred when clouds appeared overhead, occluding the sun and changing the colour and intensity of the diffuse skylight illumination source.

Temporal illumination invariance is therefore required and a small overlapping region is selected in order to calculate the temporal skylight scaling factor. Each image is temporally relit utilising the proposed temporal illumination invariance procedure with respect to the diffuse skylight present at the beginning of the dataset. The absolute intensity error is shown in Figure 5.36b and shows that for almost all patches,

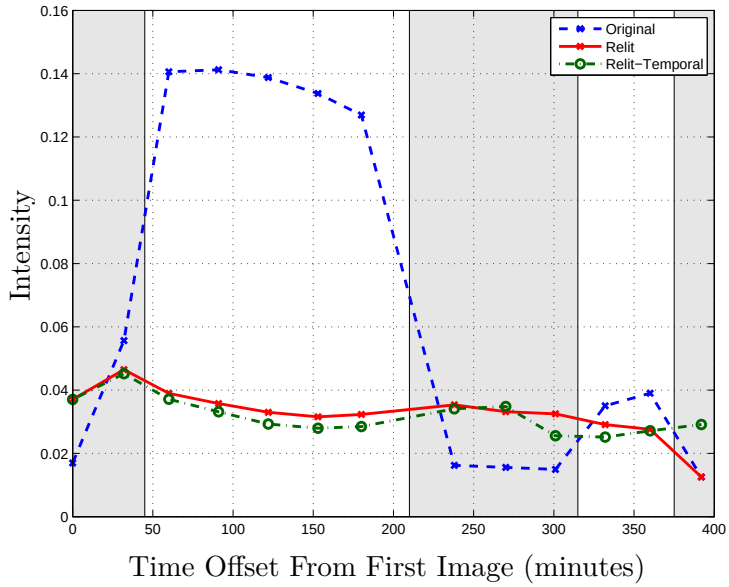
(a) Original image colour. The presence of shadows causes variability in the appearance of each individual patch and there is no colour consistency throughout the day.

(b) Relit image. Each patch shows an improvement in spatial illumination invariance, but there is a lack of consistency in the final two images as the sky colour dramatically changed.



t=0 30 60 90 120 150 180 240 270 300 330 360 390

(c) The use of the temporal skylight scaling provides consistency as at each time step, the image is relit with respect to the initial lighting conditions.



(d) The intensity of patch 3 as it varies over time is seen to be highly dependent on whether it is occluded from the sun (grey shading). Relighting the image reduces the variance throughout the day.

Figure 5.34 – Sample patch 3 in Dataset 8 as it varies throughout the day.

the outliers have been removed.

The quality of each image is measured with respect to the initial image through calculation of the PSNR and SSIM metrics, and is shown in Figure 5.37. The mean PSNR when using the original colour data is $25.9dB$, compared to $38.39dB$ and

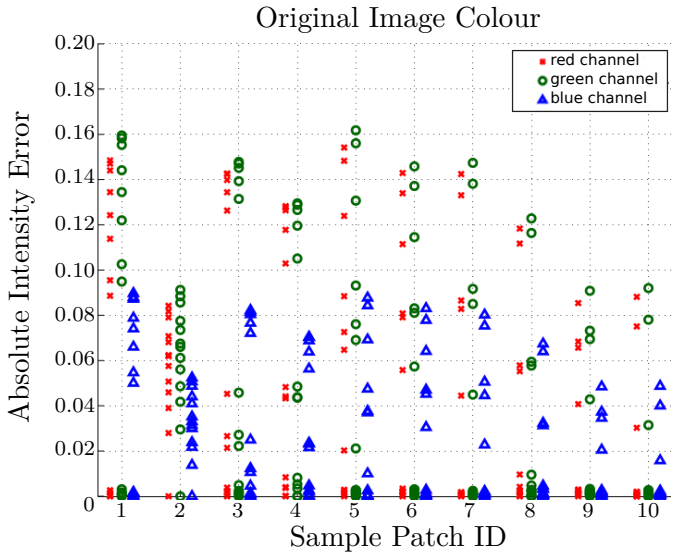
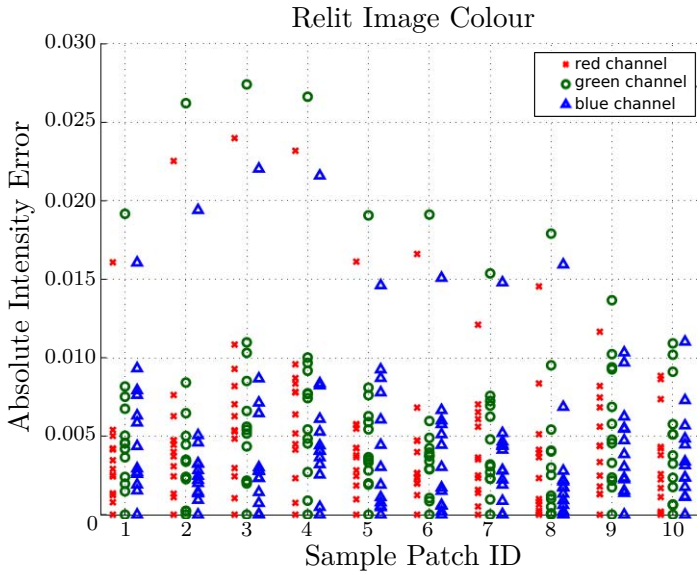


Figure 5.35 – Temporal illumination invariance in Dataset 8 over a period of $6.5hrs$. Absolute intensity errors are calculated by taking a sample area of uniform material and determining its mean colour for each image in the dataset. This is compared against the mean value in the first image. Large variations are observed due to illumination effects such as shadowing and varying sun angle.

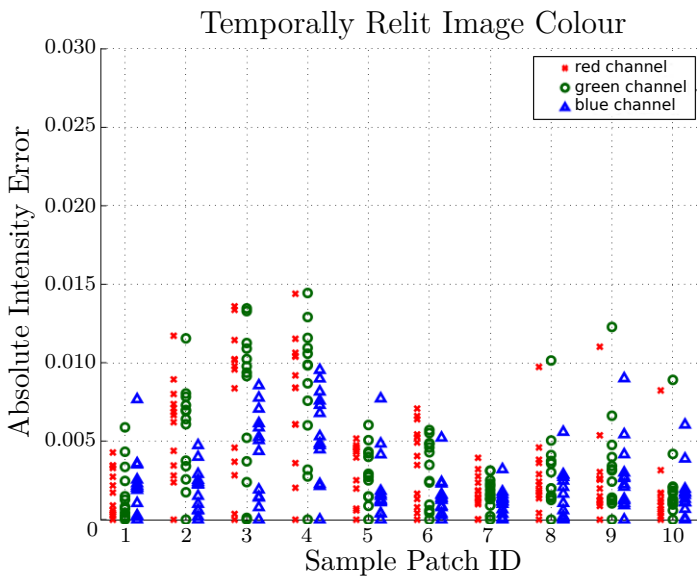
$37.83dB$ for the relit images with and without the temporal skylight scaling factor respectively. There is significantly more variation in the image quality when using the original data due to the illumination significantly affecting the appearance of the image. In the final image ($t = 375$), the relit image PSNR is seen to decrease due to a dramatic change in the skylight spectra. Through the use of the temporal skylight scaling factor, the PSNR is seen to remain stable as the image is relit with respect to the skylight illumination source present at the beginning of the dataset.

Similar stability in the image quality is seen for the relit images when using the structural similarity metric. For both metrics, using the original image colour data is seen to be correlated with the percentage of pixels occluded from the sun, highlighting the dependence that appearance has on the scene illumination conditions.

This experiment has demonstrated the temporal illumination invariance properties of the proposed system, when dealing with changes in illumination due to the movement of the sun. The relative position of the sun alters the incident illumination intensity of the terrestrial sunlight source, thereby affecting the appearance of a material.

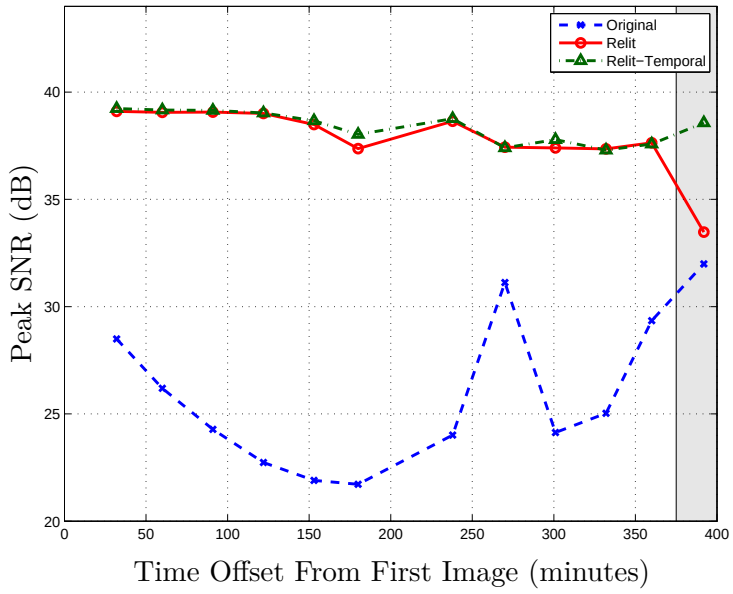


(a) Relighting each image individually reduces the error substantially. In the final image, the sun was occluded by clouds, leading to outliers in sample patches 1 to 8.

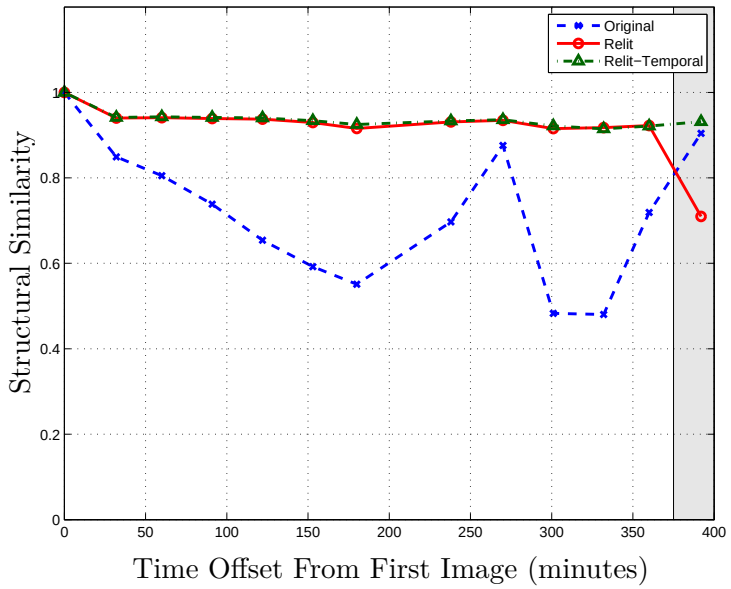


(b) Through temporal invariance, these outliers are reduced.

Figure 5.36 – Temporal illumination invariance in Dataset 8 over a period of $6.5hrs$. Absolute intensity errors are calculated by taking a sample area of uniform material and determining its mean colour for each image in the dataset. This is compared against the mean value in the first image. Note that the y-axis scale is significantly smaller than the one used in Figure 5.35.



(a) Peak SNR between each image and initial image.



(b) Structural similarity between each image and initial image.

Figure 5.37 – Image quality in Dataset 8 between each image and the initial image. Relighting is shown to improve the quality of the image when compared with the original image due to the reduction of illumination artefacts. The grey shading indicates that the image was captured under overcast conditions.

Throughout the day, regions change from being exposed to terrestrial sunlight to being in shadow and the proposed algorithm is able to compensate for its influence

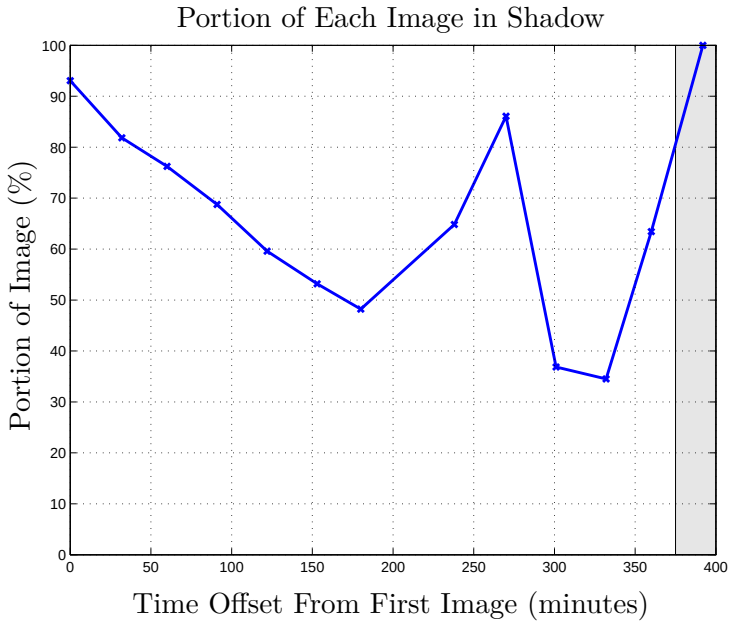


Figure 5.38 – The image quality metrics for the original images of Dataset 8 are correlated with the portion of each image that is shadowed. This is indicative of the dependence of appearance on the incident illumination.

on appearance.

5.5.3 Summary

As expected, relighting images independently of each other after there has been a change in illumination conditions was shown to produce spatially invariant, but inconsistently coloured results. Through the application of the temporal skylight scaling factor, the results have shown that for images taken at different times, whether it be over the period of a day or on separate days, improved relighting performance can be achieved by using the illuminant present in one of the images. This is vital for sensing over long periods of time, where it is desirable for objects in the scene to retain their appearance.

5.6 Large-Scale Sky Factor Estimation

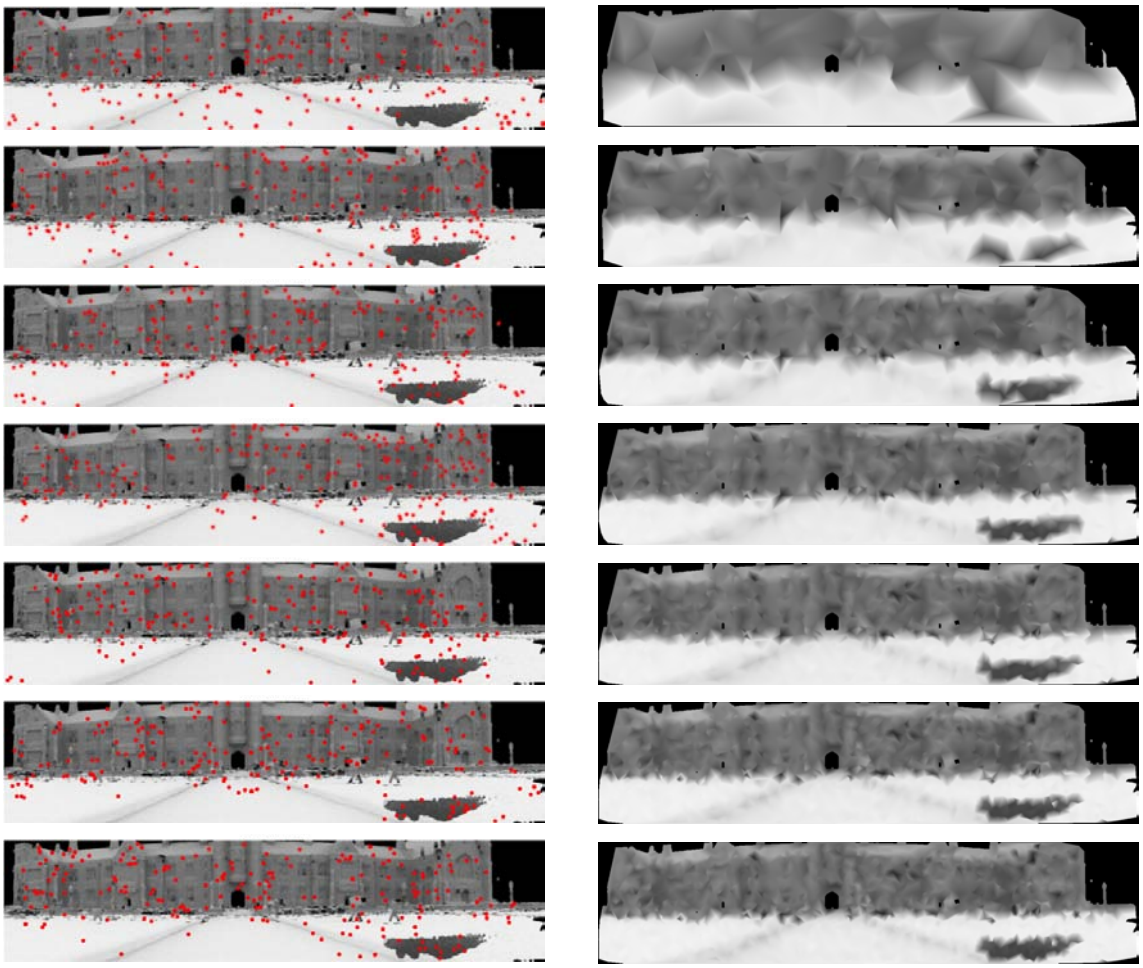
When operating in the outdoor environment over extended periods of time, the acquired point clouds gathered by a sensing platform increase in both size and complexity. Calculating sky factors using ray tracing can become computationally expensive, so this thesis proposed a Loopy Belief Propagation (Loopy Belief Propagation (LBP)) method in Section 4.1 which approximates the sky factors using a small number of samples. This method is compared against Linear Interpolation (Linear Interpolation (LI)) of the samples and the *fast* method presented in Section 3.6, using the *full* method as ground truth.

In this experiment, the threshold for the iterative sampling termination criteria κ is set at 0.001, and the number of validation points N_v and training points N_T used at each iteration is set to 200. The state space upon which inference is performed consists of increments of 0.025. The iterative sampling stage of the algorithm is run on Dataset 3 and the results are shown in Figure 5.39.

Of particular note here is that as the iteration number increases, the sampling procedure focuses onto the building. This is because the grass area in the foreground is geometrically simple and LI is able to estimate the sky factors accurately. Therefore, the error function is small in this area and the samples are directed towards more complex shapes such as the building, where the accuracy decreases. Visually, the LI sky factor approximation does not change significantly between iterations 5 and 7, and this causes the sampling procedure to terminate and move on to the smoothing stage.

The sky factor approximations using the *fast*, LI and proposed LBP methods are shown in Figure 5.40. To evaluate the three sky factor approximation methods, 5000 samples were selected using Mitchell's best candidate sampling. The absolute error between the approximation and the ground truth value (determined via ray tracing using the *full* method) was calculated and a histogram of the error is shown in Figure 5.41.

The *fast* approximation peaks at a higher error value of 0.065, compared to 0.005 for



(a) New samples are added at each iteration based on where the largest error in the approximation is.

(b) The estimate of the sky factors at each iteration is generated using linear interpolation of the samples.

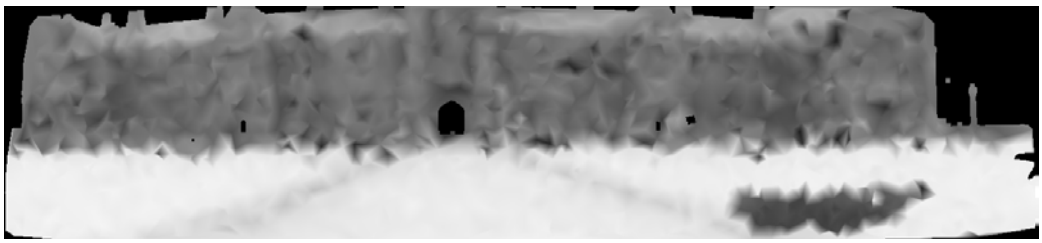
Figure 5.39 – Iterative sampling stage of the proposed large scale sky factor approximation method applied to Dataset 3.

the LI and LBP methods. This is expected since there are a number of occlusions in the scene due to the complex geometry of the building, which the *fast* method ignores. Both LI and LBP have similar error distributions, indicating that they both reconstruct the true, underlying values closely.

The real advantage of using the proposed LBP approach compared to LI is seen by visual inspection of the approximations. It is clear that using LBP generates an approximation that matches the geometry of the scene. LI on the other hand,

(a) *full* method.

(b) *fast* method.

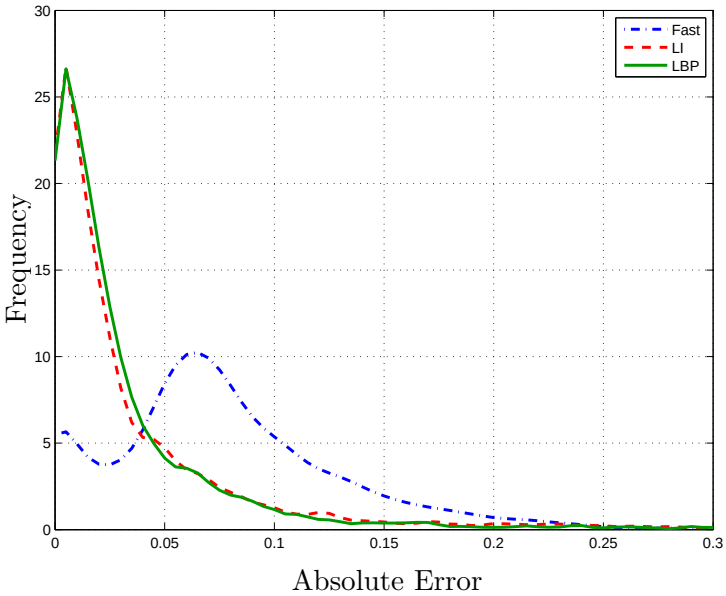


(c) *LI* method.

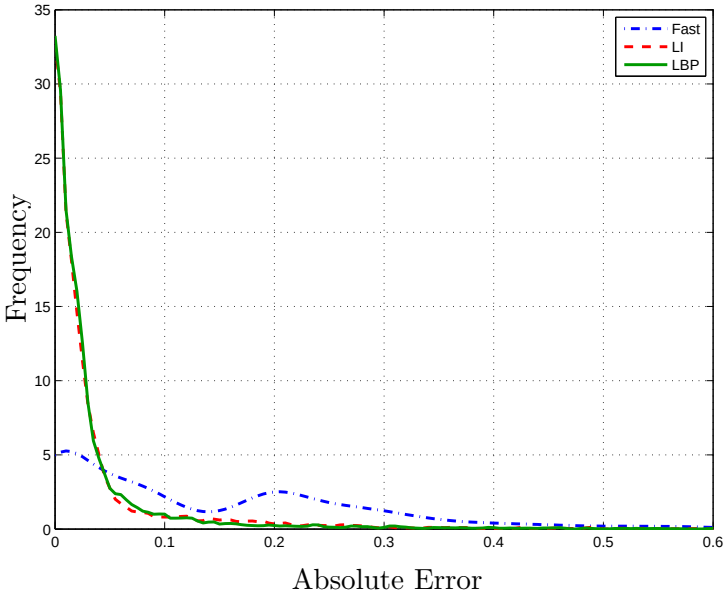
(d) *LBP* method.

Figure 5.40 – Large scale sky factor approximation of Dataset 3 using several different methods.

smooths over the geometric edges. This is clearly seen by focusing on the building and roof intersection, and yellow board regions in the scene as shown in Figure 5.42. In places where there is a depth discontinuity or a change in orientation between adjacent regions, the LBP approach produces superior approximations since it does not smooth across the boundaries as LI does. This is due to the fact that geometry was used to build a graph structure that prevents the propagation of information over



(a) Dataset 3 histogram for sky factor error.

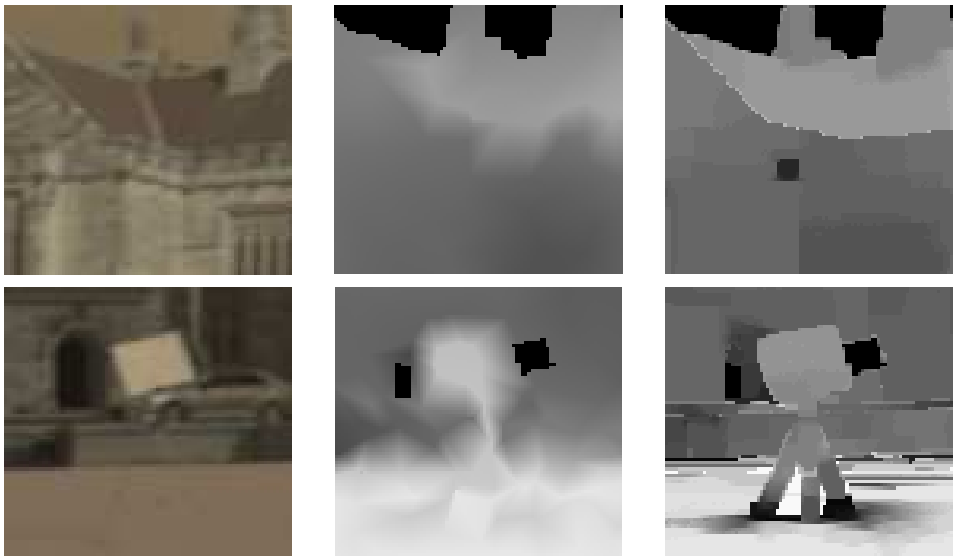


(b) Dataset 5 histogram for sky factor error.

Figure 5.41 – Histogram of absolute error for sky factor approximation using the *fast*, LI and LBP methods. Both LI and LBP have lower error peaks, indicating that they closely match the ground truth.

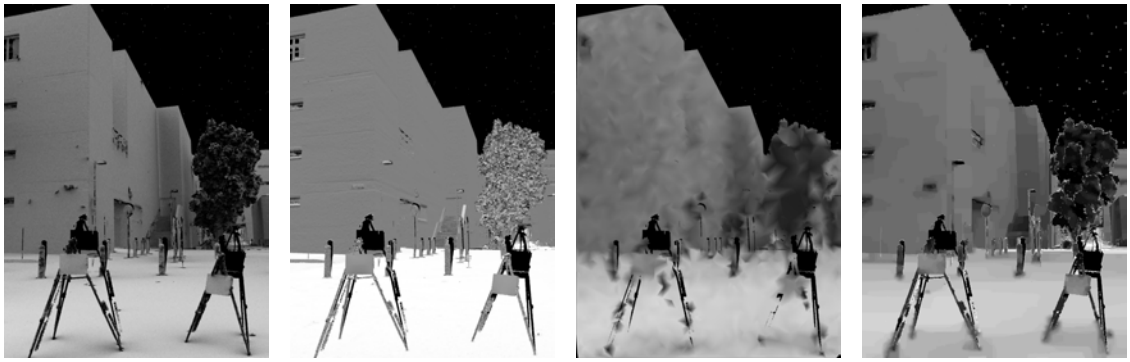
these boundaries.

Similar results are observed when running the algorithm on Dataset 5, the results of which are shown in Figure 5.43. While the *fast* approximation is seen to adhere to



(a) Original image. **(b)** *LI*. **(c)** *LBP*.

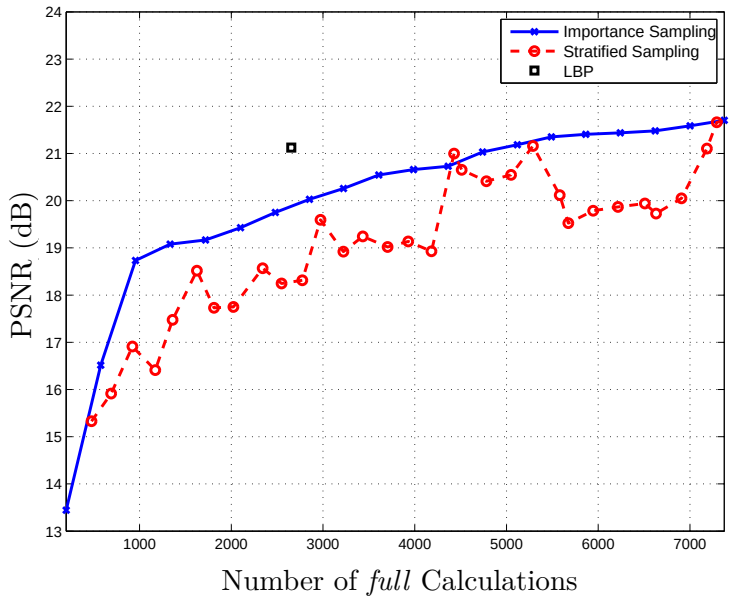
Figure 5.42 – Applying the *LBP* sky factor approximation to Dataset 3 is shown to produce estimates that are more aligned with the original scene structure compared to the *LI* method using the same samples.



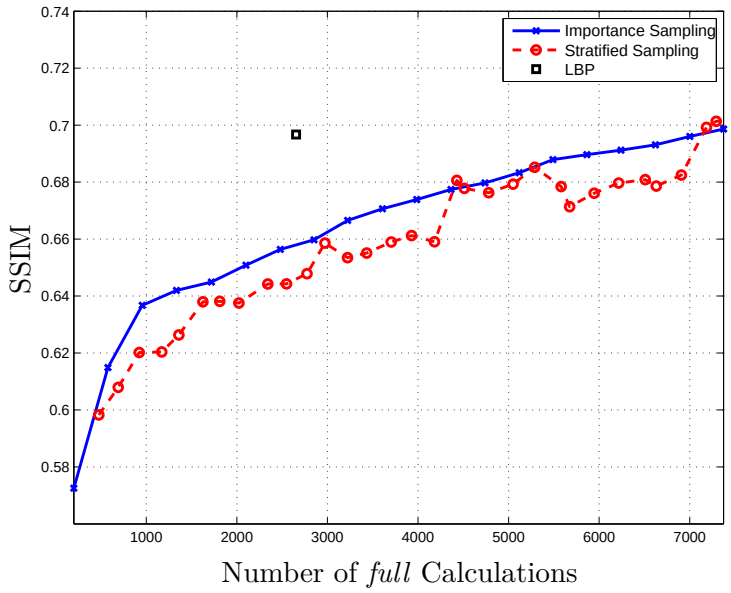
(a) *full*. **(b)** *fast*. **(c)** *LI*. **(d)** *LBP*.

Figure 5.43 – Large scale sky factor approximation of Dataset 5 using several different methods.

the geometry of the scene with the ColorChecker boards being clearly visible, it over estimates the sky factors in regions such as the intersection of the walls. Using the proposed iterative sampling method and LI shows that more accurate approximations are obtained in these areas, but smoothing over the geometric boundaries means it no longer resembles the scene structure. The proposed LBP method uses these samples



(a) PSNR for the different sampling methods.



(b) SSIM for the different sampling methods.

Figure 5.44 – The proposed LBP method outperforms both stratified and importance sampling method for approximating sky factors for Dataset 3. To attain the same PSNR and SSIM level as *LBP*, stratified and importance sampling require more than double the amount of *full* calculations.

in conjunction with the upper bounds generated by the *fast* method to reconstruct the sky factors accurately.

In terms of PSNR and SSIM, the use of importance sampling for the LI and LBP methods is seen to outperform stratified sampling as shown in Figure 5.44. The two metrics increase smoothly when using importance sampling as additional samples are added. This is not the case for stratified sampling and is explained by the fact that as additional samples are added, their position changes. The proposed LBP smoothing of the linearly interpolated function generated by importance sampling improves both the PSNR and SSIM metrics. In order for importance sampling to achieve the same PSNR and SSIM as the LBP method, an additional 2000 and 4000 samples would be required respectively. Compared to the LBP method requiring just 2600 samples, this represents a significant computational saving.

The limitations of the proposed sampling methods are the following. As the spatial sampling rate is below the spatial Nyquist frequency, situations can arise where the reconstruction of high frequency objects in the scene will not be accurate. If the termination criteria of the sampling stage is reached before a sample is acquired on a particular object, the LBP smoothing procedure will use the surrounding data to infer the sky factors, leading to an inaccurate approximation.

The primary aim of this thesis was to achieve illumination invariance in images taken in the outdoor environment. Using a model based approach means that the relative orientation between the scene and the illumination sources of terrestrial sunlight and diffuse skylight needs to be known. A critical parameter of this is the sky factor, which can be determined using a computationally expensive ray tracing approach.

In this section, an iterative sampling and smoothing approach was demonstrated to be able to reconstruct the sky factors using a small amount of samples. This means that in large complex environments, it is not necessary to determine sky factors for all points, thereby reducing the computation time. The proposed method outperforms the *fast* method due to the fact that it takes into account occlusions in the scene. It is also superior to the LI method as the graph structure is built in such a way that it restricts information flow across depth discontinuities and sharp orientation changes.

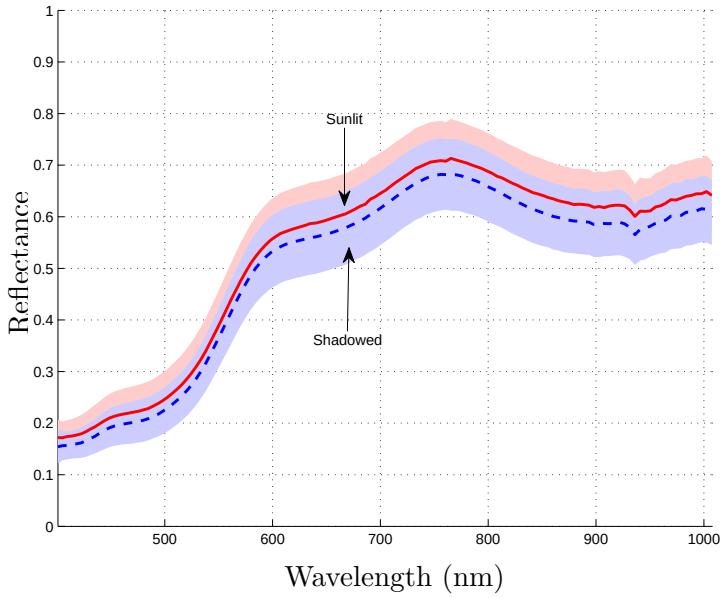
5.7 Radiometric Normalisation

The large scale sky approximation allows the scene radiance to be modelled accurately, which is beneficial for radiometric normalisation. Radiometric normalisation converts pixel intensity to reflectance, allowing the data to be compared against reference examples, such as spectral libraries obtained in laboratory environments. The proposed illumination invariant approach uses an in situ measurement of the incident illumination and a model of the physical processes involved in order to account for spatial variations. This method is evaluated on simple and complex geometry datasets, along with the addition of indirect illumination as a potential light source.

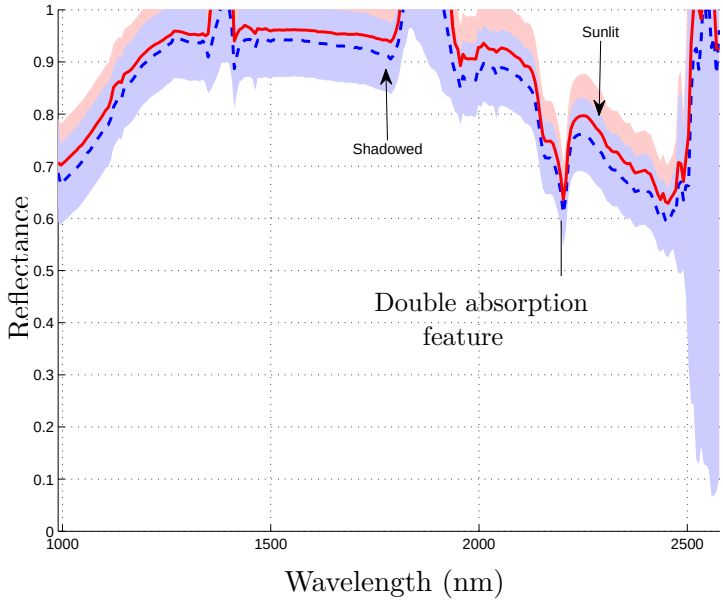
5.7.1 Simple Geometry

Illumination invariant radiometric normalisation using a calibration board is evaluated on Dataset 2, for which there is high spatial resolution ground truth data. In this experiment, the shadow cast onto the shale surface allows the algorithm to be initialised, and the terrestrial sunlight and diffuse skylight sources can be estimated from Equations (4.18) and (4.15). In this dataset indirect illumination is not considered as there are no major reflective surfaces illuminating the scene.

The standard method of obtaining reflectance estimates of the original image is to perform flat-field correction by normalising the spectra with the measurement off the calibration board. This is a common normalisation method used in remote sensing applications where quite often, the calibration board does not have the same orientation as the scene. For analysis, a region of shale exposed to sunlight and one in shade is selected, consisting of approximately 6000 and 1000 data points in each set respectively. The estimated reflectance using flat-field correction and the proposed method is compared on a per-pixel basis against ground truth reflectance data in terms of spectral shape. The ground truth reflectance is obtained by positioning a calibration board of known reflectance at the same orientation as the shale material and performing flat-field correction (Figure 5.45) [54].



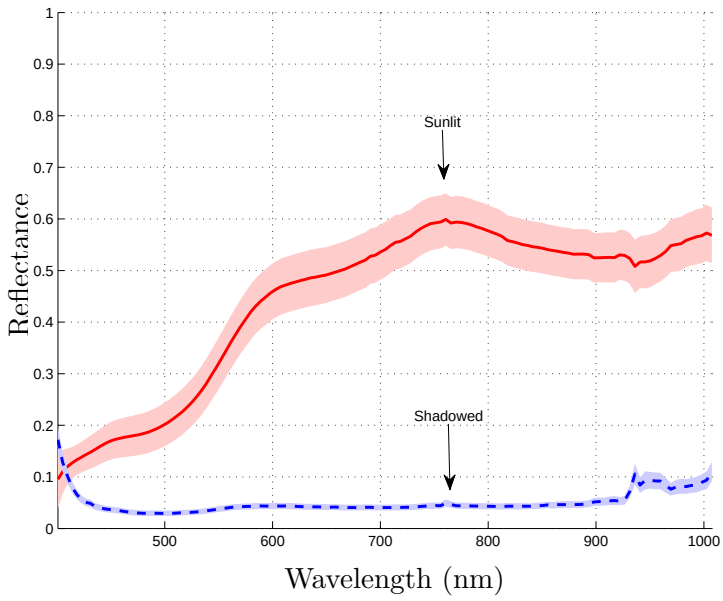
(a) VNIR.



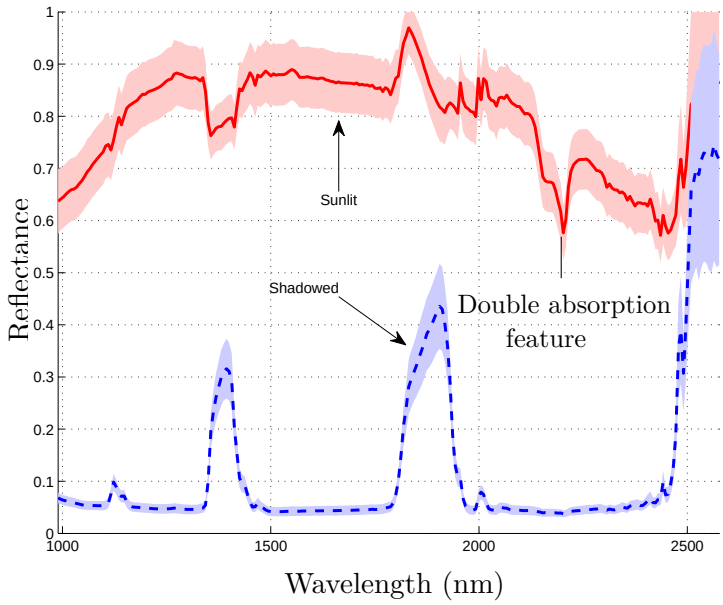
(b) SWIR.

Figure 5.45 – Ground truth reflectance of shale (Dataset 2) using flat-field correction by positioning a calibration board at the same orientation as the shale container. The mean spectra for sunlit regions is shown by the solid red line, while the dashed blue lines represent the shadowed areas.

As shown in Figure 5.46, the reflectance obtained using just the calibration board clearly separates the two classes. The spectral angle between the mean of the sunlit

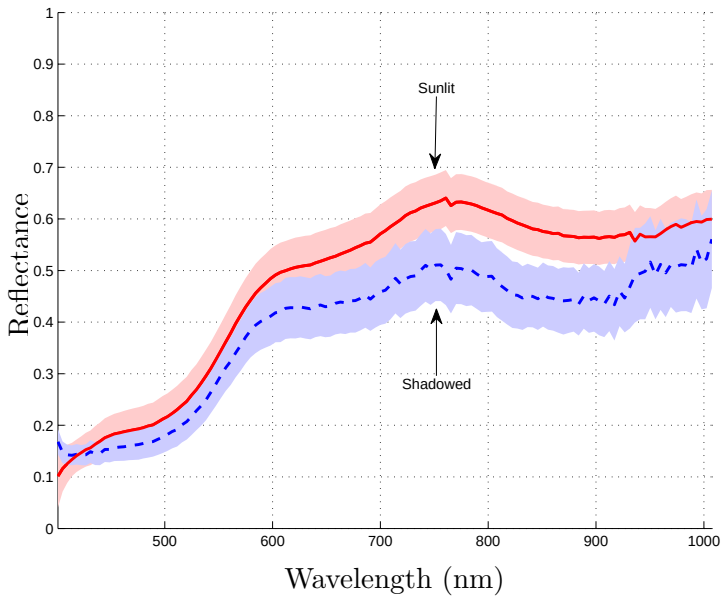


(a) VNIR.

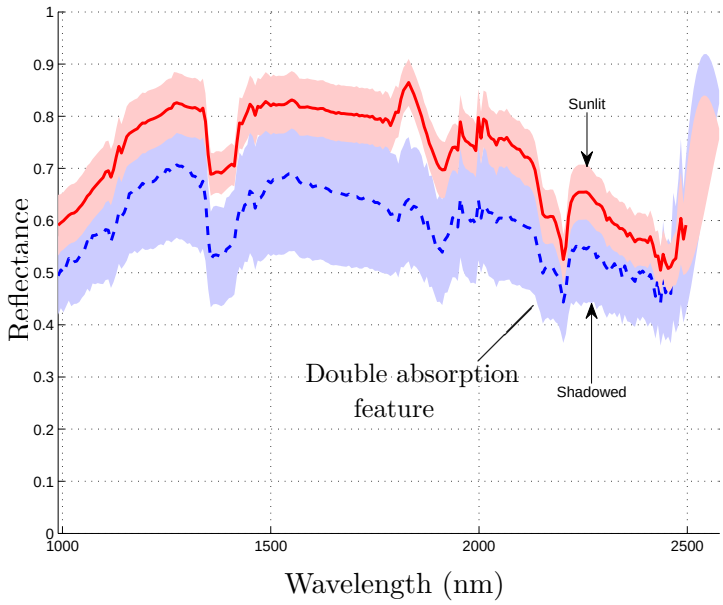


(b) SWIR.

Figure 5.46 – Reflectance estimates of shale (Dataset 2) using the calibration board for VNIR and SWIR hyperspectral data. The varying colour and intensity of the illuminant between the sunlit and shadowed regions means that the reflectance estimates are significantly different. The mean spectra for sunlit regions is shown by the solid red line, while the dashed blue lines represent the shadowed areas.



(a) VNIR.



(b) SWIR.

Figure 5.47 – Reflectance estimates of shale (Dataset 2) using the proposed illumination invariant method. Accounting for the geometry of the scene and normalising on a per-pixel basis means that the reflectance estimates are more closely aligned with the ground truth data. The mean spectra for sunlit regions is shown by the solid red line, while the dashed blue lines represent the shadowed areas.

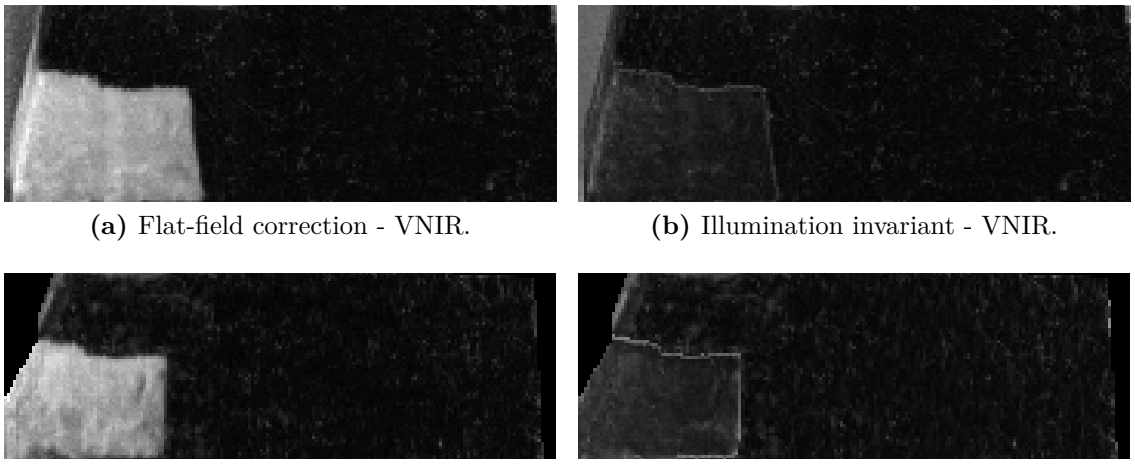
and shaded points is 29.82°, indicating a large spectral shape difference between the two. This is because in the VNIR region of the spectrum, diffuse skylight is skewed towards shorter wavelengths, thereby influencing the appearance of the scene dramatically.

Once relit, the shaded and sunlit reflectance estimates become closely aligned in terms of shape and magnitude as seen in Figure 5.47. Although there is still clear distinction between the two classes, the magnitude difference is significantly smaller compared to the original image. In terms of shape, the spectral angle between the mean of the sunlit and shaded points has decreased dramatically from 29.82° to 5.70°, indicating much higher similarity. This is as expected as the proposed method is calculating and compensating for the illumination on a per-pixel basis, as opposed to the image wide normalisation method used previously.

Two noteworthy features are observable in the proposed illumination invariant normalisation procedure. Firstly, at short wavelengths the reflectance using flat-field correction of the region in shadow is incorrectly estimated due to the diffuse skylight illuminant. By taking into account the variable illumination conditions, the method is able to closely match the underlying reflectance.

Another aspect to note is the presence of a double absorption feature at approximately 2159nm and 2202nm, which is critical for geological applications where the identification of kaolinite in the scene is required. This characteristic feature is only noticeable through the use of the proposed illumination invariant radiometric normalisation technique. However, it must be noted that this feature is only present in the mean reflectance spectra for the shadowed region and cannot be extracted on a per-pixel basis. This implies that the underlying absorption feature is still present in the data, but is masked by the measurement noise. Post-processing techniques such as super-pixels could be used as they allow the mean pixel value of a local neighbourhood to be used, thereby reducing the effects of noise.

For high level applications such as classification, it is important for the spectral shape of the estimated reflectance to be similar to that of the underlying material class. The SAM metric is used to compare the spectral shape on a per-pixel resolution for the



(a) Flat-field correction - VNIR. **(b)** Illumination invariant - VNIR.

(c) Flat-field correction - SWIR. **(d)** Illumination invariant - SWIR.

Figure 5.48 – SAM metric measuring spectral shape similarity between ground truth reflectance data and flat-field correction normalisation, and the proposed illumination invariant method for Dataset 2. The SAM metric for the VNIR portions of the spectrum has been normalised from 0° (black) to 46° (white), while the SWIR portion of the spectrum has been normalised to range from 0° (black) to 28° (white).

Table 5.6 – The mean SAM value for shadowed regions is dramatically smaller for VNIR and SWIR spectra when using the proposed illumination invariant radiometric normalisation method. There is a slight increase in SAM for sunlit areas but this is not of significant magnitude to cause concern.

	Shadowed		**Sunlit**	
Method	*VNIR*	*SWIR*	*VNIR*	*SWIR*
Calib. Board	29.82	18.02	2.48	1.59
Ill. Inv.	5.70	4.36	2.53	1.88

VNIR and SWIR data using the two normalisation procedures as shown in Figure 5.48.

Using the calibration board only for normalisation provides similar spectral shape similarity with the ground truth data for regions exposed to sunlight, but this performance severely degrades within the shadowed regions. For the SWIR spectra, the mean SAM value for areas in shadow is 18.0° when excluding the destructive water bands. This is compared to 4.36° for the illumination invariant method, indicating a high degree of similarity between the estimates and ground truth data. Within the sunlit regions, both methods perform similarly, with a SAM value of 1.59° and 1.88° respectively. SAM often plays a critical role in high level clustering and classification,

and by reducing the influence of illumination variation in the scene the proposed normalisation method can increase the performance of these algorithms by reducing the separation between identical materials in exposed to different illumination conditions.

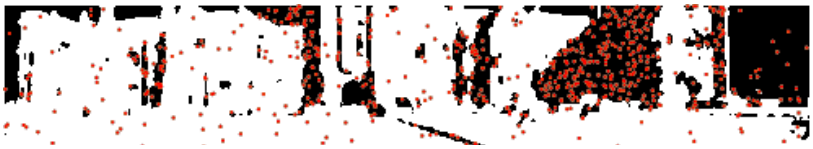
5.7.2 Complex Geometry

Having demonstrated the advantages of the proposed illumination invariant radiometric normalisation approach on a simple geometric scene, the algorithm is evaluated on a large scale, complex scene (Dataset 3) with the addition of indirect illumination into the model.

The iterative sampling and smoothing technique proposed in Section 4.2.1 is used to approximate the indirect illumination in the scene as shown in Figure 5.49. Bright regions in the approximation image are seen to be focused on the intersection of orthogonal walls and intuitively, this makes sense. This is because there will be a high amount of inter-reflections in these areas due to the geometry. The foreground of the scene where the grass is located is seen to not contain a substantial amount of indirect illumination. This is because the angle between the normal on the grass area, and the vector towards the building is large. This means that illumination along these vectors will only have a small impact due to their cosine weighting. Indirect illumination on the grass area is small as a fraction of the total illumination as the area is exposed to terrestrial sunlight and diffuse skylight. This means that the sampling function ($\max \Delta\rho_i$) will be small in these areas and this is reflected by the smaller number of samples used to estimate indirect illumination on the grass area.

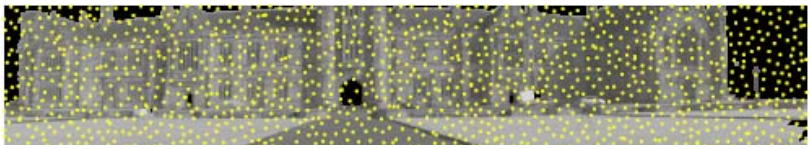
The first experiment investigates the performance of the proposed illumination invariant radiometric normalisation method by utilising the Downwelling Irradiance Sensor (DIS)[8]. This measurement of incident illumination is used in conjunction with Equations (4.15) and (4.20) to estimate the reflectance of the scene. The reflectance estimate obtained with and without the inclusion of indirect illumination sources is compared against the standard method of flat-field correction using the DIS measurement. This is evaluated by analysing the reflectance estimate of a calibration

(a) Original image.



(b) Initial samples.

(c) Training samples.



(d) Validation samples.

(e) Indirect illumination estimate shown using the wavelengths 1073, 1243 and 1319nm as the RGB channels respectively.

(f) Indirect illumination as a proportion of total illumination reveals that areas in shadow are influenced to a greater extent by indirect illumination.

Figure 5.49 – Estimated indirect illumination for Dataset 3 using the SWIR hyperspectral line scanner.

board of known reflectance in the scene.

The estimated reflectance of the calibration board in the scene using flat-field correction and the proposed illumination invariant radiometric normalisation method is

(a) Calibration board.

(b) Calibration board illumination invariant.

(c) Calibration board illumination invariant with indirect illumination.

Figure 5.50 – Radiometric normalisation of Dataset 3 using the calibration board.

Table 5.7 – Metrics for comparing the calibration board reflectance estimates using the DIS. Results are also shown in brackets when the water absorption and low signal to noise ratio bands are included in the analysis.

	Invariance				
Hardware	$\mathbf{E}_{sun}\boldsymbol{\tau}$	$\mathbf{E}_{sky}$	$\mathbf{E}_{ind}$	**SAM**	**Euc. Dist**
DIS	×	×	×	4.25 (16.11)	5.22 (6.50)
DIS	✓	✓	×	3.92 (13.62)	0.78 (3.81)
DIS	✓	✓	✓	2.72 (13.10)	1.43 (4.32)

shown in Figure 5.52. Through the utilisation of fused geometry and imagery, the proposed illumination invariant radiometric normalisation process is able to closely estimate the reflectance of the calibration board. Both the flat-field correction and proposed method have similar spectral angles when compared to the ground truth as shown in Table 5.7. However, the advantage of the proposed method is clear when analysing the spectral magnitude error through the Euclidean distance metric. This shows a significant drop in error due to the fact that it takes into account the incident illumination angle of terrestrial sunlight and the amount of skylight visible from the

(a) DIS.

(b) DIS illumination invariant.

(c) DIS illumination invariant with indirect illumination.

Figure 5.51 – Radiometric normalisation of Dataset 3 using the DIS.

board. Flat-field correction naively assumes that the calibration board has the same illumination source and strength as the DIS. As the sun position changes throughout the day, the reflectance estimate using the DIS measurement will fluctuate, while the proposed approach will be stable. The inclusion of indirect illumination into the outdoor illumination model has mixed results. While it reduces the spectral angle by more than 1°, the magnitude error actually increases.

For further evaluation of the proposed radiometric normalisation approach, four regions are selected; sunlit building, shadowed building, sunlit roof and shadowed roof. The reflectance estimates are shown in Figures 5.53 and 5.54 for the various normalisation methods. Note once again that the estimated reflectance within the destructive water band regions (shaded grey) cannot be used for material analysis. When performing flat-field correction using the calibration board and DIS methods, there is a clear separation in the reflectance estimate between identical materials under different illumination conditions. This is expected since the normalisation procedure does not take into account the illumination variation. The reflectance estimates using

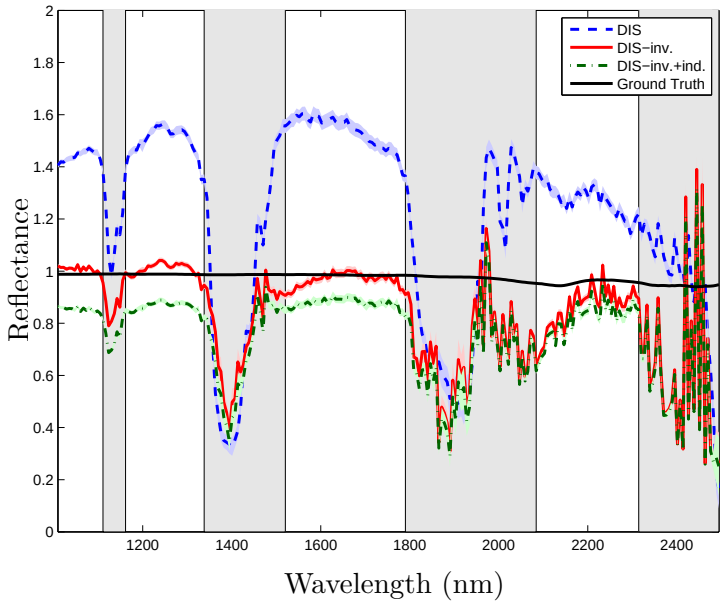
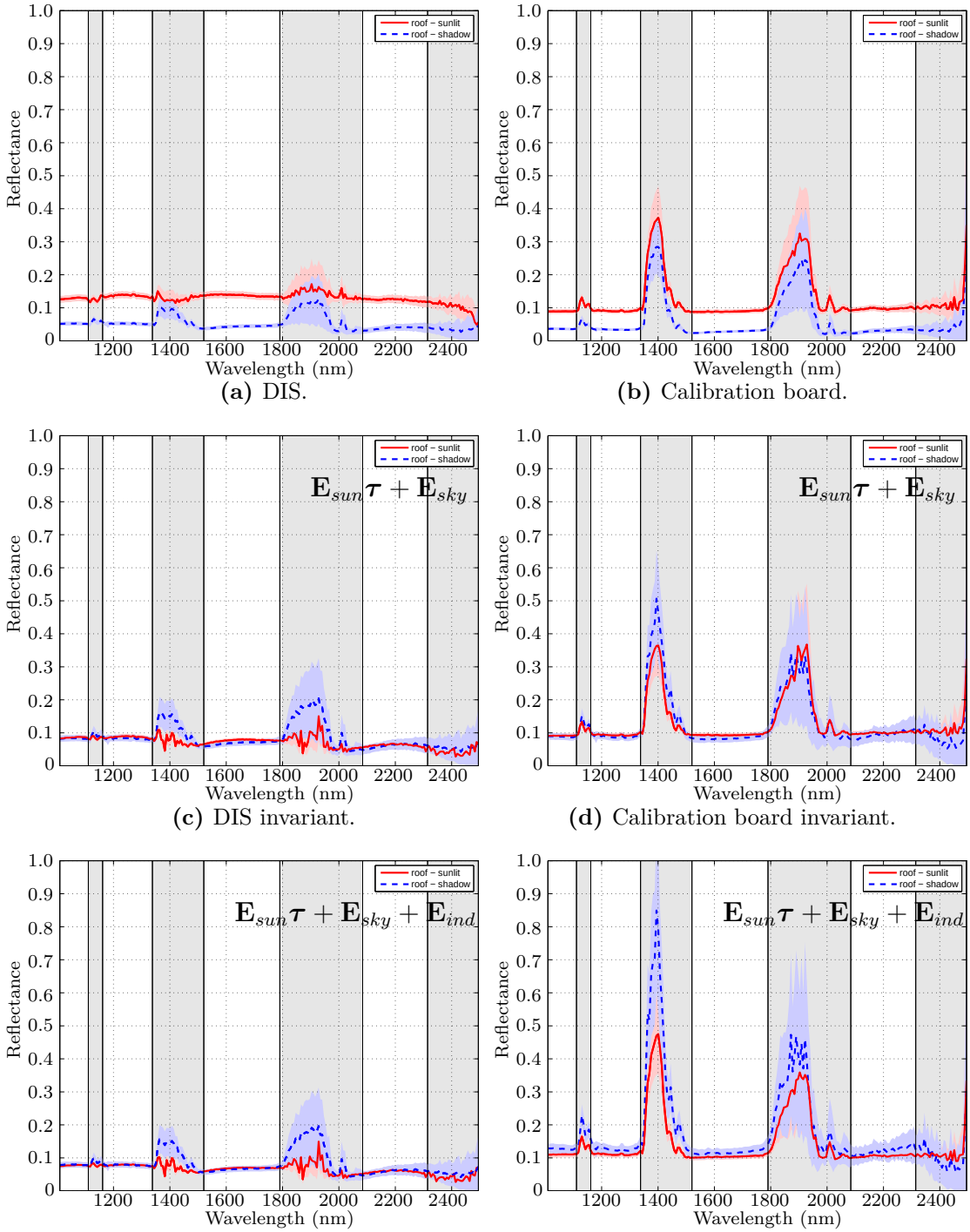


Figure 5.52 – Estimation of calibration board reflectance using the raw DIS measurement for Dataset 3, as well as the proposed illumination invariant radiometric normalisation method with and without the inclusion of indirect illumination sources. Grey areas indicate destructive water absorption bands with low SNR.

Table 5.8 – Metrics for comparing the reflectance estimates of the roof in the scene using different radiometric normalisation techniques. For the proposed methods, the inclusion of $\mathbf{E}_{sun}\boldsymbol{\tau}$, $\mathbf{E}_{sky}$ and $\mathbf{E}_{ind}$ in the initialisation of the algorithm and in the reflectance estimation are as indicated. Results are also shown in brackets when the water absorption and low signal to noise ratio bands are included in the analysis.

	Material: Roof							
	Initialisation			**Reflectance**				
Hardware	$\mathbf{E}_{sun}\boldsymbol{\tau}$	$\mathbf{E}_{sky}$	$\mathbf{E}_{ind}$	$\mathbf{E}_{sun}\boldsymbol{\tau}$	$\mathbf{E}_{sky}$	$\mathbf{E}_{ind}$	**SAM**	**EUC**
Panel	-	-	-	-	-	-	7.14 (19.04)	0.67 (1.01)
	✓	✓	×	✓	✓	×	3.19 (10.63)	0.09 (0.43)
	✓	✓	✓	✓	✓	✓	2.32 (10.71)	0.17 (0.99)
DIS	-	-	-	-	-	-	7.19 (20.67)	0.96 (1.23)
	✓	✓	×	✓	✓	×	3.45 (19.26)	0.07 (0.50)
	✓	✓	✓	✓	✓	✓	2.96 (17.99)	0.04 (0.48)

the proposed approach reduces this separation. As shown in Tables 5.8 and 5.9, the spectral shape and magnitude metrics decrease, indicating increased similarity due to per-pixel normalisation. This is evidence of reduced spectral variability which can be brought about due to multiple sources illuminating a scene.



(a) DIS.

(b) Calibration board.

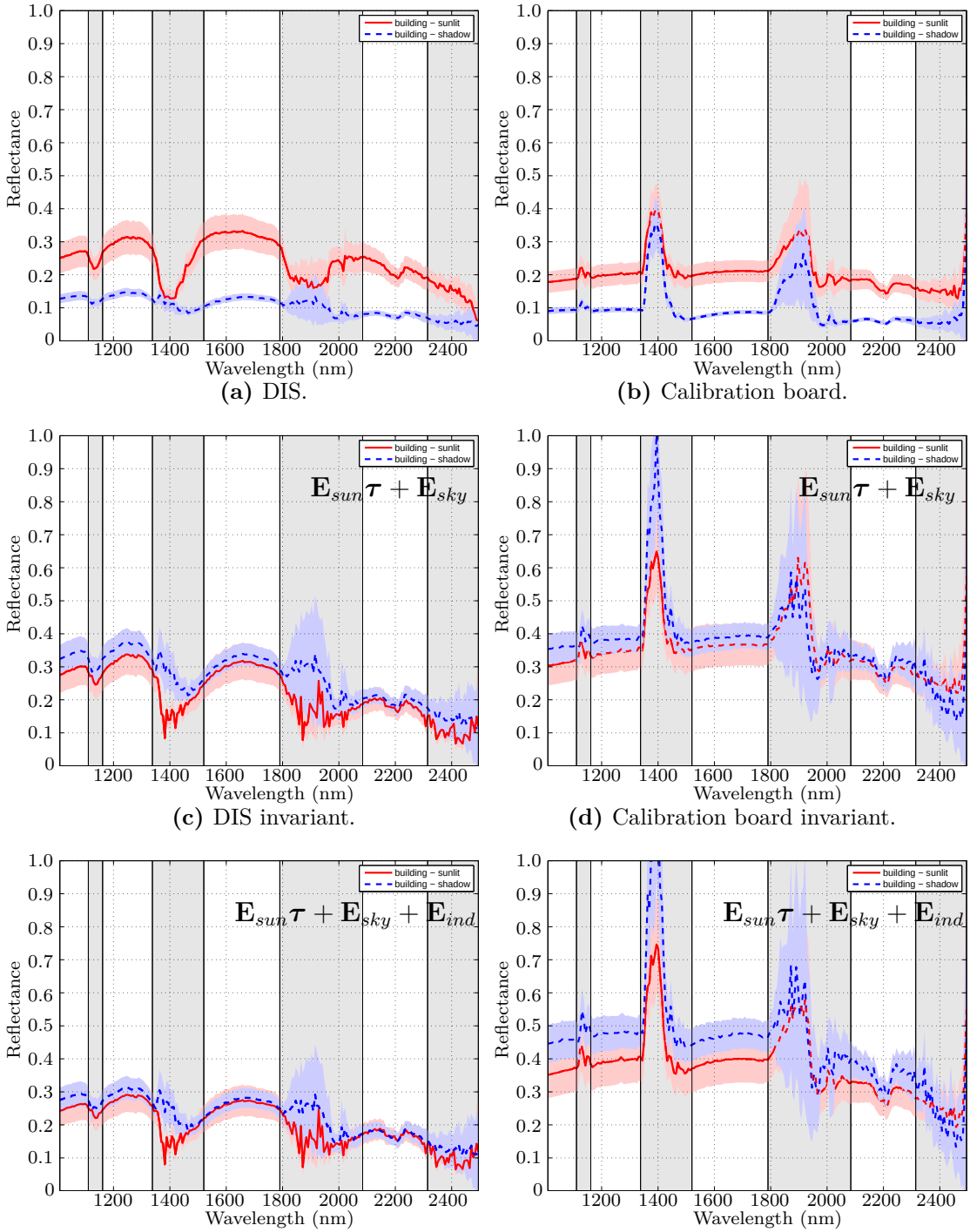
(c) DIS invariant.

(d) Calibration board invariant.

(e) DIS invariant + indirect illumination.

(f) Calibration board invariant + indirect illumination.

Figure 5.53 – Reflectance of the roof in sunlit and shaded regions in Dataset 3. Grey areas indicate destructive water absorption bands with low SNR.



(a) DIS.

(b) Calibration board.

(c) DIS invariant.

(d) Calibration board invariant.

(e) DIS invariant + indirect illumination.

(f) Calibration board invariant + indirect illumination.

Figure 5.54 – Reflectance of the building in sunlit and shaded regions in Dataset 3. Grey areas indicate destructive water absorption bands with low SNR.

Table 5.9 – Metrics for comparing the reflectance estimates of the building in the scene using different radiometric normalisation techniques. For the proposed methods, the inclusion of $\mathbf{E}_{sun}\boldsymbol{\tau}$, $\mathbf{E}_{sky}$ and $\mathbf{E}_{ind}$ in the initialisation of the algorithm and in the reflectance estimation are as indicated. Results are also shown in brackets when the water absorption and low signal to noise ratio bands are included in the analysis.

	Material: Building							
	Initialisation			**Reflectance**				
Hardware	$\mathbf{E}_{sun}\boldsymbol{\tau}$	$\mathbf{E}_{sky}$	$\mathbf{E}_{ind}$	$\mathbf{E}_{sun}\boldsymbol{\tau}$	$\mathbf{E}_{sky}$	$\mathbf{E}_{ind}$	**SAM**	**EUC**
Panel	-	-	-	-	-	-	7.19 (18.46)	1.24 (1.68)
	✓	✓	×	✓	✓	×	1.93 (8.49)	0.34 (0.96)
	✓	✓	✓	✓	✓	✓	1.62 (7.02)	0.81 (1.49)
DIS	-	-	-	-	-	-	7.05 (12.49)	1.81 (2.20)
	✓	✓	×	✓	✓	×	2.10 (10.24)	0.32 (0.94)
	✓	✓	✓	✓	✓	✓	2.29 (9.46)	0.18 (0.68)

A small region of ground truth is available in the form of a yellow board attached to a tripod in the scene. The reflectance spectra of the board is measured using a spectrometer so that it can be compared against the estimates calculated using various normalisation approaches. The yellow board is orientated in such a way that it is exposed to both terrestrial sunlight and skylight.

Flat-field correction using the calibration board and using the illumination invariant method without the inclusion of indirect illumination is seen to produce reflectance estimates with similar spectral shape and intensity error metrics (see Table 5.10). The performance of flat-field correction is explained by the fact that the calibration and yellow boards are similarly orientated. The inclusion of indirect illumination into the model has mixed results on the reflectance estimates. While there is a slight decrease in the spectral angle error, the magnitude error also increases.

When using the DIS measurements as the incident illumination measurement device, there is a more pronounced improvement in performance. The Euclidean distance metric significantly decreases when accounting for the geometry of the scene and this is due to the fact that the sensor is oriented differently to the yellow board. The spectral angle also decreases, with increased similarity being obtained through the inclusion of indirect illumination. The experiments show that the fusion of image and

Table 5.10 – Metrics for comparing the reflectance estimates of a yellow board in the scene using different radiometric normalisation techniques. For the proposed methods, the inclusion of $\mathbf{E}_{sun}\boldsymbol{\tau}$, $\mathbf{E}_{sky}$ and $\mathbf{E}_{ind}$ in the initialisation of the algorithm and in the reflectance estimation are as indicated. Results are also shown in brackets when the water absorption and low signal to noise ratio bands are included in the analysis.

	Material: Yellow Board							
	Initialisation			**Reflectance**				
Hardware	$\mathbf{E}_{sun}\boldsymbol{\tau}$	$\mathbf{E}_{sky}$	$\mathbf{E}_{ind}$	$\mathbf{E}_{sun}\boldsymbol{\tau}$	$\mathbf{E}_{sky}$	$\mathbf{E}_{ind}$	**SAM**	**EUC**
Panel	-	-	-	-	-	-	5.44 (8.44)	1.65 (3.20)
	✓	✓	×	✓	✓	×	5.09 (8.90)	1.86 (3.24)
	✓	✓	✓	✓	✓	×	2.54 (7.82)	3.29 (5.23)
	✓	✓	✓	✓	✓	✓	4.53 (8.42)	2.31 (3.91)
DIS	-	-	-	-	-	-	4.16 (14.14)	5.67 (6.74)
	✓	✓	×	✓	✓	×	4.06 (15.17)	0.78 (2.67)
	✓	✓	✓	✓	✓	×	3.51 (14.60)	0.74 (2.97)
	✓	✓	✓	✓	✓	✓	2.43 (13.91)	1.14 (3.20)

geometric data is vital in order to obtain illumination invariance in the radiometric normalisation process. This is in contrast to the standard flat-field correction methods which are unable to account for the change in illumination source within shadowed regions.

5.7.3 Summary

Radiometric normalisation of hyperspectral data using hardware based methods such as calibration boards or downwelling irradiance sensors experience a degradation in performance due to spatial illumination variations. The experiments showed that under different illumination conditions (e.g. when a material is exposed and occluded from terrestrial sunlight), the reflectance spectra obtained by performing flat-field correction is different in terms of spectral shape and intensity. This is due the fact this approach assumes constant illumination throughout the entire scene which is not the case when there is variable geometry and occlusions.

Through the utilisation of a per-pixel approach, which models the incident illumination at each pixel, the intra-class separability due to illumination variations is reduced,

and the estimates are closer to the ground truth. This has significant implications for high level algorithms such as classification and clustering, where the spectral signature can be used as the input feature.

When the parameters of the outdoor illumination model (Equation 2.16) are accurately known, as is the case for the simple geometry experiment, the spectral shape and intensity of the illumination invariant radiometric normalisation approach significantly outperformed the standard method of flat-field correction.

5.8 Limitations

In its current implementation, there are several limitations to the proposed illumination invariant technique. The first limitation is that shadows are not binary entities. Instead, there is a smooth transition across the shadow boundary and the visibility term in the illumination model could be replaced with a continuous parameter ranging from 0 to 1, indicating the degree of visibility. This will improve the invariance properties of the image along the shadow boundaries.

A second limitation is that the local neighbourhood of points is required in order to estimate the normals of the point cloud. This leads to the smoothing of sharp geometric changes in the scene, especially in urban scenarios. Purely image based approaches such as [26] do not have this problem since there is no need for registration between multiple sensor modalities. It is also possible to used the fused point cloud and imagery data in a Bayesian approach to smooth and correct the position of the normals. Edges in the image are representative of not only changes in illumination, material and texture, but also changes in geometry. An MRF can enforce consistency in the direction of the normals through an appropriate pairwise potential, with such an approach used to improve surface reconstruction in [16].

In applications where absolute atmospheric correction of spectral data is required, it may not be possible to use the proposed illumination invariant radiometric normalisation approach to identify weak absorption features. This is because the estimated

reflectance data is extremely noisy when the transmittance is small as seen in the destructive water band regions of the spectra. In these cases it is possible to used radiance transfer models to generate the terrestrial sunlight and diffuse skylight illumination spectra, and then use the indirect illumination estimate (Section 4.2) to compensate for inter-reflections. When performing this step, it is important to make sure that all illumination sources have the same units. Atmospheric radiative transfer models typically output the spectra in $W/m^2/nm$, while the image data is in digital values. It is therefore necessary to radiometrically calibrate the image prior to performing normalisation.

An advantage of using an image based approach is that LIDAR returns from structures such as windows and black coloured materials are noisy and inconsistent with the scene. Since cameras are a passive form of sensing, these artefacts are not induced in their measurements and this is an advantage in urban scenarios, where man-made objects such as buildings are prevalent.

5.9 Summary

The proposed illumination invariant approach for outdoor perception using multiple sensor modalities was extensively evaluated on a number of datasets. Simple small scale geometric scenes were used to analyse the performance when the outdoor illumination model parameters such as visibility, sky factors and the sun angle are accurately known. The results revealed that by accounting for the geometry of the scene, the intra-class variability due to illumination could be significantly reduced in terms of spectral shape and intensity. This was evidenced by reductions in the SAM and Euclidean distance metrics. Extending the evaluation to large scale, complex scenes showed similar benefits, with high level algorithms such as clustering benefiting from the invariant representation of the scene.

While the approach was derived using an assumption of narrowband sensor responses, experiments using consumer grade cameras showed that wideband imaging devices can also achieve illumination invariance using the proposed method. This invariance

can be used as a post-processing step, prior to colourspace conversion, white balancing and gamma corrections.

Using the *fast* sky factor approximation was also shown to increase the invariance properties of the image, with computation time being significantly reduced when compared to the *full* method.

The experimental results have demonstrated that the complementary nature of geometry and imagery is valuable for achieving illumination invariance in outdoor environments. Despite the assumption of diffuse material reflectance, the proposed approach was shown to produce high quality results in scenes where the materials have hybrid reflectance functions. The technique is able to recover invariant images, without reducing the dimensionality of the data, meaning both the colour and intensity of the material is retained. This is critical in situations where materials may have similarly shaped spectra or identical chromaticity, but differing intensities.

Chapter 6

Conclusion

The intent of this thesis was to investigate novel methods for generating illumination invariant representations of images taken in the outdoor environment. This was achieved through an automatic approach utilising the complementary nature of geometric and appearance data. The proposed system treats each illumination source independently, meaning spatial and temporal invariance can be achieved under a variety of weather conditions. This approach allows the full spectral dimensionality of the data to be retained, thereby providing high level algorithms with highly discriminative information upon which they can operate. The addition of a direct incident illumination measurement meant that illumination invariance could be extended to radiometric normalisation. It provides advantages over current methods as it does not require user input, multiple images or atmospheric radiative transfer models.

This chapter describes the thesis contributions in Section 6.1 and directions for future work in Section 6.2.

6.1 Summary of Contributions

For images taken in outdoor scenarios, the influence of the two main illumination sources (terrestrial sunlight and diffuse skylight) varies on a per-pixel basis, thereby

influencing the appearance of materials in the scene. The proposed method generates scaling factors for each pixel, thereby relighting the image using a common illuminant for all pixels. This method of illumination invariance allows all spectral channels of the data to be retained, which is important as colour information acts as a useful discriminative feature for high level algorithms. The technique is automatic and requires few tuning parameters, meaning that expert knowledge is not required for its operation. Through experimental validation, it was found that spatial and temporal illumination invariance was achieved, allowing identical materials to appear similarly, regardless of the original incident illumination. High level algorithms such as clustering and classification benefit from such a representation of the scene, as they can discriminate based on the underlying material properties.

Illumination invariant radiometric normalisation was also investigated in this thesis. Through the use of a model of the physical processes and the addition of an external hardware measurement of scene illumination, the incident irradiance at each pixel can be explicitly calculated. This allows reflectance estimation to occur on a per-pixel basis. The proposed method has the advantage over just using the hardware measurement, in that geometric variation are accounted for. The most obvious example of this is within shadowed regions of the scene and this was shown through the experimental results. Through the utilisation of the proposed method, there was a decrease in spectral angle and magnitude error when comparing estimated reflectance against ground truth data.

This thesis proposed a number of contributions, including

- An automatic approach to single and multiple image illumination invariance in the outdoor environment via the modelling of the physical processes involved. This method takes into account spatial and temporal illumination variations that influence the appearance of materials in the outdoor environment by relighting each pixel with respect to a common, but unknown illuminant. There is no reliance on external models of the atmosphere as it infers the illumination conditions from the measured data.

- An iterative sampling and smoothing approach to sky factor and indirect illumination approximation in large scale, complex scenes. The high degree of computational complexity required for the calculation of sky factors and indirect illumination in large scale scenes was reduced by introducing a strategy that allows these parameters to be approximated using a smaller number of explicit calculations.

- An illumination invariant radiometric normalisation approach. Standard methods for the radiometric normalisation of field based data involve using external hardware to measure the incident illumination. By utilising the geometry of the scene, the proposed method is able to estimate the illumination on a per-pixel basis, thereby reducing the influence of shadows.

- An approximation to sky factor calculation for applications requiring low computational time.

These methods exploit the geometry and imagery information provided by multimodal sensor suites commonly found on robotics and remote sensing platforms. The principled way in which the system was developed means that it is robust to changes in sun position and weather conditions as they are explicitly accounted for in the outdoor illumination model from which the method is derived.

6.2 Future Work

Some interesting extensions that could be undertaken include further work in automatic initialisation, spectral smoothing, an implementation for mobile platforms and investigation into the cloud cover shadowing.

Extension of Automatic Initialisation The LIDAR is an active sensing approach to perceiving the environment, with its measurements indicating the range to each point in the scene and the intensity of the return. In this thesis only range

data was utilised, however the intensity information is valuable as it is illumination invariant in nature and is indicative of the underlying material. The issue is that it only provides information at one particular wavelength, nevertheless, it can be used as an additional stage in automatic initialisation to remove outlier candidates from the selected pairs of points. This would allow greater accuracy to be achieved when calculating the terrestrial sunlight-skylight ratio, leading to improved relighting performance.

Spectral Smoothing The estimation of the terrestrial sunlight-skylight ratio, diffuse skylight and terrestrial sunlight spectra are calculated on a per-wavelength basis. However, the high spectral resolution of hyperspectral sensors means that there is a smooth variation between estimates at adjacent wavelengths. Therefore, a chain CRF could be used to reduce the noise in the estimates. This would be especially useful for long wavelengths in the SWIR region, where the SNR is extremely low due to the attenuation of light by the atmosphere. Several characteristic absorption features exist within this region and so a reduction in the amount of noise in the reflectance estimate would allow these features to be extracted.

Mobile Platform Implementation Robotic platforms typically move throughout the scene as they perceive the environment, often with lower resolution sensors when compared with 3D LIDARs. This thesis has dealt with rapid computation of sky factors for platforms, but has focused on scenarios where it moves to a position, stops and perceives the environment, and then continues. An interesting direction for future work would be to implement a real-time system for robotics that incorporates the continuous motion and low resolution observations.

Cloud Cover Shadowing A problem that is unable to be solved by the proposed illumination invariant system is the phenomena of clouds in the sky, which can limit the times at which sensing can take place. While the system is able to compensate imagery under different weather conditions, problems arise when clouds occlude direct

sunlight in parts of the scene. Ray tracing shadow detection methods only determine whether there is line of sight visibility in the direction of the sun and do not factor in cloud occlusions. In accordance with the belief that multiple sensor modalities is beneficial for generating a rich scene understanding, it may be worthwhile using a camera with a fish-eye lens pointed in the direction of the sky, in order to determine cloud cover and local weather conditions. Inclusion of cloud cover shadowing is beneficial as it makes the proposed illumination invariant system more robust to the variety of weather conditions encountered in the outdoor environment.

Bibliography

[1] Vivek Agarwal, Besma R Abidi, Andreas Koschan, and Mongi A Abidi. An Overview of Color Constancy Algorithms. In *Journal of Pattern Recognition Research*, volume 1, pages 42–54, 2006.

[2] Jonathan T Barron and Jitendra Malik. Shape , Albedo , and Illumination from a Single Image of an Unknown Object. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 334–341. IEEE, 2012.

[3] Jonathan T Barron and Jitendra Malik. Intrinsic Scene Properties from a Single RGB-D Image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 17–24. IEEE, 2013.

[4] Harry G Barrow and J Martin Tenenbaum. Recovering Intrinsic Scene Characteristics from Images. *Computer Vision Systems*, pages 2–25, 1978.

[5] Alexander Berk, Lawrence S Bernstein, and David C Robertson. MODTRAN: A moderate resolution model for LOWTRAN 7. Technical report, DTIC Document, 1987.

[6] Paul J Besl and Neil D McKay. Method for registration of 3-D shapes. In *Robotics-DL tentative*, pages 586–606. International Society for Optics and Photonics, 1992.

[7] Christoph C Borel, Siegfried A W Gerstl, and Bill J Powers. The Radiosity Method in Optical Remote Sensing of Structured 3-D Surfaces. *Remote Sensing of Environment*, pages 13–44, 1991.

[8] Marcus Borengasser, William S Hungate, and Russell Watkins. *Hyperspectral Remote Sensing: Principles and Applications*. CRC Press, 2007. ISBN 9781566706544.

[9] Yuri Y Boykov and Marie-Pierre Jolly. Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images. In *Proceedings. International Conference on Computer Vision*, number July, pages 105–112. IEEE, 2001.

[10] Joshua Broadwater and Amit Banerjee. Improved atmospheric compensation of hyperspectral imagery using LIDAR. In *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*, pages 2200–2203. IEEE, 2013.

[11] Michael F. Cohen and Donald P. Greenberg. The hemi-cube: a radiosity solution for complex environments. In *ACM SIGGRAPH Computer Graphics*, volume 19, pages 31–40. ACM, 1985.

[12] Michael F Cohen and John R Wallace. *Radiosity and realistic image synthesis*. Elsevier, 1993.

[13] Peter Corke, Rohan Paul, Winston Churchill, and Paul Newman. Dealing with shadows: Capturing intrinsic scene appearance for image-based outdoor localisation. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2085–2092. IEEE, 2013.

[14] Paul Debevec, Chris Tchou, Andrew Gardner, Tim Hawkins, Charis Poullis, Jessi Stumpfel, Andrew Jones, Nathaniel Yun, Per Einarsson, There Lundgren, Marcos Fajardo, and Philippe Martinez. Estimating surface reflectance properties of a complex scene under captured natural illumination. In *ACM Transactions on Graphics*, volume 19, pages 1–11, 2004.

[15] Mauricio Diaz and Peter Sturm. Estimating Photometric Properties from Image Collections. In *Journal of Mathematical Imaging and Vision*, volume 47, pages 93–107. Springer, 2013.

[16] James Diebel and Sebastian Thrun. An Application of Markov Random Fields to Range Sensing. In *NIPS*, volume 5, pages 291–298. MIT Press, 2005.

[17] Philip Dutre, Kavita Bala, Philippe Bekaert, and Peter Shirley. *Advanced Global Illumination*. AK Peters Ltd, 2006. ISBN 1568813074.

[18] Stephanie Fekete, Mark Diederichs, and Matthew Lato. Geotechnical and operational applications for 3-dimensional laser scanning in drill and blast tunnels. *Tunnelling and Underground Space Technology*, 25(5):614–628, 2010.

[19] Graham D Finlayson and Clement Fredembach. Fast Re-integration of Shadow Free Images. In *Color and Imaging Conference*, pages 117–122. Society for Imaging Science and Technology, 2004.

[20] Graham D Finlayson and Clement Fredembach. Hamiltionian path based shadow removal. In *Proceedings of the 16th British Machine Vision Conference*, pages 502–511, 2005.

[21] Graham D Finlayson and Steven Hordley. Improving Gamut Mapping Color Constancy. In *IEEE Transactions on Image Processing*, volume 9, pages 1774–1783, 2000.

[22] Graham D Finlayson and Steven D Hordley. Color constancy at a pixel. In *JOSA A*, volume 18, pages 253–264, 2001.

[23] Graham D Finlayson, Steven D Hordley, and Mark S Drew. Removing Shadows from Images. In *Computer Vision-ECCV 2002*, pages 823–836. Springer, 2002.

[24] Graham D Finlayson, Mark S Drew, and Cheng Lu. Intrinsic Images by Entropy Minimization. In *Computer Vision-ECCV 2004*, pages 582–595. Springer, 2004.

[25] Graham D Finlayson, Steven D Hordley, Cheng Lu, and Mark S Drew. On the removal of shadows from images. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, volume 28, pages 59–68. IEEE, 2006.

[26] Graham D Finlayson, Mark S Drew, and Cheng Lu. Entropy Minimization for Shadow Removal. *International Journal of Computer Vision*, pages 35–37, 2009.

[27] David A Forsyth. A Novel Algorithm for Color Constancy. In *International Journal of Computer Vision*, volume 5, pages 5–35. Springer, 1990.

[28] Clément Fredembach and Sabine Süsstrunk. Automatic and accurate shadow detection from (potentially) a single image using near-infrared information. Technical Report 165527, EPFL Tech Report 165527, 2010.

[29] Ola Friman, Gustav Tolt, and Jörgen Ahlberg. Illumination and shadow compensation of hyperspectral images using a digital surface model and non-linear least squares estimation. In *SPIE Remote Sensing*, pages 81800Q–81800Q. International Society for Optics and Photonics, 2011.

[30] Yongnian Gao and Wanchang Zhang. A simple empirical topographic correction method for ETM+ imagery. *International Journal of Remote Sensing*, 30:2259–2275, May 2002. ISSN 0143-1161.

[31] Peter V Gehler, Carsten Rother, Martin Kiefel, Lumin Zhang, and Bernhard Scholkopf. Recovering Intrinsic Images with a Global Sparsity Prior on Reflectance. In *Advances in neural information processing systems*, pages 765–773, 2011.

[32] Arjan Gijsenij, Rui Lu, and Theo Gevers. Color constancy for multiple light sources. In *Image Processing, IEEE Transactions on*, volume 21, pages 697–707. IEEE, 2012.

[33] Cindy M Goral, Kenneth E Torrance, Donald P Greenberg, and Bennett Battaile. Modeling the Interaction of Light Between Diffuse Surfaces. *Computer Graphics*, 18(3), July 1984.

[34] Christian Gueymard. SMARTS, A Simple Model of the Atmospheric Radiative Transfer of Sunshine: Algorithms and Performance Assessment. Florida Solar Energy Center, 1995.

[35] Ruiqi Guo, Qieyun Dai, and Derek Hoiem. Single-image shadow detection and removal using paired regions. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2033–2040. IEEE, 2011.

[36] Kaiming He, Jian Sun, and Xiaoou Tang. Guided Image Filtering. In *Computer Vision–ECCV 2010*, volume 35, pages 1–14. Springer, 2010.

[37] Lukas Hosek and Alexander Wilkie. Adding a Solar-Radiance Function to the Hošek-Wilkie Skylight Model. In *Computer Graphics and Applications, IEEE*, volume 33, pages 44–52. IEEE, 2013.

[38] Daniel F Huber and Nicolas Vandapel. Automatic 3D underground mine mapping. In *Field and Service Robotics*, pages 497–506. Springer, 2006.

[39] Christoph Hug. Combined use of laser scanner geometry and reflectance data to identify surface objects. In *Proceedings of the OEEPE Workshop*, pages 1–10, 1996.

[40] Emmett J. Ientilucci. Leveraging lidar data to aid in hyperspectral image target detection in the radiance domain. In *SPIE Defense, Security, and Sensing. International Society for Optics and Photonics*, pages 839007–839007. International Society for Optics and Photonics, 2012.

[41] T H Kurz, S J Buckley, and J A Howell. Close range hyperspectral imaging integrated with terrestrial lidar scanning applied to rock characterisation at centimetre scale. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 34, pages 417–422, 2012.

[42] Jean-François Lalonde, Alexei A Efros, and Srinivasa G Narasimhan. Detecting ground shadows in outdoor consumer photographs. In *Computer Vision–ECCV 2010*, pages 322–335. Springer, 2010.

[43] Jean-François Lalonde, Alexei A. Efros, and Srinivasa G. Narasimhan. Estimating the Natural Illumination Conditions from a Single Outdoor Image. In *International Journal of Computer Vision*, volume 98, pages 123–145, October 2011.

[44] Edwin H Land. *The Retinex Theory of Color Vision*, volume 237. Scientific America, 1977.

[45] Edwin H Land and John McCann. Lightness and Retinex Theory. In *JOSA*, volume 61, pages 1–11. Optical Society of America, 1971.

[46] Stuart E Lynch, Mark S Drew, and Graham D Finlayson. Colour Constancy from Both Sides of the Shadow Edge. In *Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on*, pages 899–906. IEEE, 2013.

[47] Takashi Machida, Naokazu Yokoya, and Haruo Takemura. Surface reflectance modeling of real objects with interreflections. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 170–177. IEEE, 2003.

[48] Takashi Machida, Naokazu Yokoya, and Haruo Takemura. Inverse global illumination rendering for dense estimation of surface reflectance properties. In *Proceedings of the 14th Scandinavian Conference on Image Analysis*, pages 790–801, 2005.

[49] Takashi Machida, Haruo Takemura, and Naokazu Yokoya. Inverse reflectometry for real objects with diffuse and specular interreflections. In *Electronics and Communications in Japan (Part II: Electronics)*, volume 90, pages 50–60, January 2007.

[50] Claus B Madsen and Brajesh B Lal. Estimating Outdoor Illumination Conditions Based on Detection of Dynamic Shadows. In *Computer Vision, Imaging and Computer Graphics. Theory and Applications*, pages 33–52. Springer, 2013.

[51] Colin Mcmanus, Winston Churchill, Will Maddern, Alexander D Stewart, and Paul Newman. Shady Dealings : Robust , Long-Term Visual Localisation using Illumination Invariance. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 901–906, Hong Kong, China, 2014. IEEE.

[52] Donald Meagher. Octree Encoding: a New Technique for the Representation, Manipulation and Display of Arbitrary 3-D Objects by Computer. Technical report, Rensselaer Polytechnic Institute, 1980.

[53] Kevin P Murphy, Yair Weiss, and Michael I Jordan. Loopy belief propagation for approximate inference: An empirical study. In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*, pages 467–475. Morgan Kaufmann Publishers Inc., 1999.

[54] R.J. Murphy, A.J. Underwood, T.J. Tolhurst, and M.G. Chapman. Field-based remote-sensing for experimental intertidal ecology: Case studies using hyperspatial and hyperspectral data for New South Wales (Australia). In *Remote Sensing of Environment*, volume 112, pages 3353–3365, August 2008.

[55] Juan I Nieto, Sildomar T Monteiro, and Diego Viejo. 3D geological modelling using laser and hyperspectral data. In *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, pages 4568–4571. IEEE, 2010.

[56] Melissa L Nischan, Rose M Joseph, Justin C Libby, and John P Kerekes. Active Spectral Imaging. In *Lincoln Laboratory Journal*, volume 14, pages 131–144, 2003.

[57] Alexandros Panagopoulos, Chaohui Wang, Dimitris Samaras, and Nikos Paragios. Simultaneous Cast Shadows , Illumination & Geometry Inference Using Hypergraphs. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, volume 35, pages 437–449. IEEE, 2013.

[58] Richard Perez, Robert Seals, and Joseph Michalsky. All-weather model for sky luminance distribution—preliminary configuration and validation. In *Solar energy*, volume 50, pages 235–245. Elsevier, 1993.

[59] Matt Pharr and Greg Humphreys. *Physically Based Rendering: From Theory to Implementation*. Morgan Kaufmann, 2nd edition, 2010.

[60] Arcot J. Preetham, Peter Shirley, and Brian Smits. A practical analytic model for daylight. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 91–100. ACM Press/Addison-Wesley Publishing Co., 1999.

[61] Prabhakar Raghavan, Christopher D Manning, and Hinrich Schütze. *Introduction to Information Retrieval*. Cambridge University Press, 2008.

[62] Rishi Ramakrishnan, Juan Nieto, and Steve Scheding. Shadow Compensation for Outdoor Perception. In *International Conference on Robotics and Automation*, pages 4835–4842, Seattle, 2015. IEEE.

[63] Ibrahim Reda and Afshin Andreas. Solar position algorithm for solar radiation applications. In *Solar energy*, volume 76, pages 577–589. Elsevier, 2004.

[64] R. Richter and D. Schläpfer. Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: Atmospheric/topographic correction. In *International Journal of Remote Sensing*, volume 23, pages 2631–2649, January 2002.

[65] R. Richter, D. Schläpfer, and A. Müller. An automatic atmospheric correction algorithm for visible/NIR imagery. In *International Journal of Remote Sensing*, volume 27, pages 2077–2085, May 2006.

[66] Yossi Rubner, Carlo Tomasi, and Leonidas J Guibas. The Earth Mover's Distance as a Metric for Image Retrieval. In *International Journal of Computer Vision*, volume 40, pages 99–121. Springer, 2000.

[67] Neda Salamati, Arthur Germain, and Sabine Susstrunk. Removing shadows from images using color and near-infrared. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 1713–1716. IEEE, 2011.

[68] Imari Sato, Yoichi Sato, and Katsushi Ikeuchi. Illumination Distribution from Shadows. In *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on.*, volume 1. IEEE, 1999.

[69] Imari Sato, Yoichi Sato, and Katsushi Ikeuchi. Illumination distribution from brightness in shadows: Adaptive estimation of illumination distribution with unknown reflectance properties in shadow regions. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, pages 875–882. IEEE, 1999.

[70] Yoichi Sato and Katsushi Ikeuchi. Reflectance analysis under solar illumination. In *Physics-Based Modeling in Computer Vision, 1995., Proceedings of the Workshop on*, pages 180–187. IEEE, 1995.

[71] Robert A. Schowengerdt. *Remote Sensing: Models and Methods for Image Processing*. Elsevier Science, 3rd edition, 2006. ISBN 9780080480589.

[72] Geoffrey M. Smith and Edward J. Milton. The use of the empirical line method to calibrate remotely sensed data to reflectance. In *International Journal of Remote Sensing*, volume 20, pages 2653–2662, January 1999.

[73] Randall Smith. Introduction to hyperspectral imaging. Technical report, MicroImages inc., 2012. URL `http://www.microimages.com/documentation/Tutorials/hyprspec.pdf`.

[74] Marshall F Tappen, William T Freeman, and Edward H Adelson. Recovering intrinsic images from a single image. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, volume 27, pages 1459–1472. IEEE, 2005.

[75] Zachary Taylor. Find 3D Normals and Curvature. Technical report, MATLAB Central File Exchange, 2014. URL `http://au.mathworks.com/matlabcentral/fileexchange/48111-find-3d-normals-and-curvature`.

[76] Zachary Taylor, Juan Nieto, and David Johnson. Automatic Calibration of Multi-Modal Sensor Systems using a Gradient Orientation Measure. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pages 1293–1300. IEEE, 2013.

[77] Zachary Taylor, Juan Nieto, and David Johnson. Multi-Modal Sensor Calibration Using a Gradient Orientation Measure. In *Journal of Field Robotics*, volume 32, pages 675–695. Wiley Online Library, June 2014.

[78] Gustav Tolt, M Shimoni, and Jorgen Ahlberg. A shadow detection method for remote sensing images using VHR hyperspectral and LIDAR data. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pages 4423–4426, 2011.

[79] Alejandro Troccoli and Peter Allen. Building Illumination Coherent 3D Models of Large-Scale Outdoor Scenes. In *International Journal of Computer Vision*, volume 78, pages 261–280. Springer, November 2008.

[80] Ben Upcroft, Colin Mcmanus, Winston Churchill, Will Maddern, and Paul Newman. Lighting Invariant Urban Street Classification. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 1712–1718, Hong Kong, China, 2014. IEEE.

[81] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P Simoncelli. Image Quality Assessment : From Error Visibility to Structural Similarity. In *Image Processing, IEEE Transactions on*, volume 13, pages 600–612, 2004.

[82] Yair Weiss. Deriving intrinsic images from image sequences. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, pages 68–75. IEEE, 2001.

[83] Weihua Xiong and Brian Funt. Stereo Retinex. In *Image and Vision Computing*, volume 27, pages 178–188. Elsevier, 2009. ISBN 0769525423.

[84] Yizhou Yu and Jitendra Malik. Recovering photometric properties of architectural scenes from photographs. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 207–217. ACM, 1998.

[85] R H Yuhas, A F H Goetz, and J W Boardman. Discriminisation among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In *Summaries of the third annual JPL airborne geoscience workshop*, pages 147–149. JPL Publication, 1992.

Appendix A

Disc Approximation for Surface Representation

Generating a well structured surface representation from a point cloud is often difficult and requires manual tweaking to form a model that is representative of the scene. This process is also computationally expensive and therefore a disc approximation is used in this thesis. Each return in the point cloud is represented by a disc whose radius is dependent on the resolution of the laser scan. This process is illustrated in Figure A.1.

An object is scanned with an angular resolution of ϕ in both the azimuth and elevation directions, leading to several points striking the surface. These points are approximated by discs encompassing the small surrounding region and we make some approximations in order to calculate the radius. The first approximation is that adjacent discs will have the same radius i.e. $r_i = r_j$. This approximation is only used for estimation purposes and is not enforced when assigning radius sizes. Application of the sine rule gives:

$$\frac{2r_i}{\sin\phi} = \frac{d_i}{\sin\gamma_i}, \tag{A.1}$$

where $\gamma = \frac{\pi}{2} + \alpha_i - \phi$ and r_i is the radius of disc i. Since the laser scans used typically

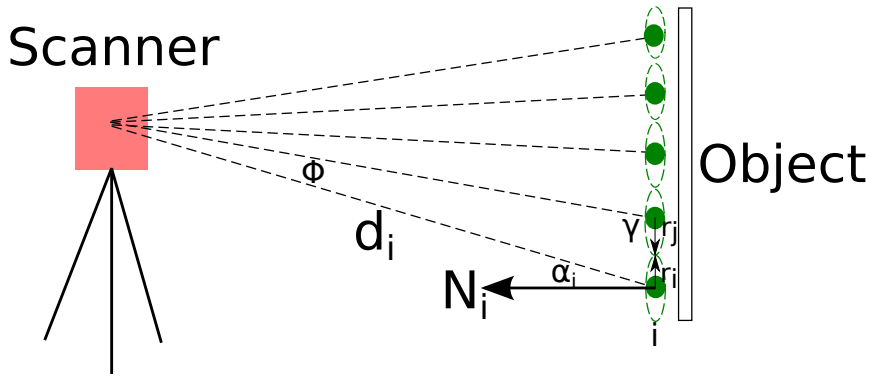


Figure A.1 – A disc approximation of the scene is used to represent the surrounding region around each element in the point cloud

have resolutions of less than 5°, ϕ can be approximated as a small angle. Following simplification, the radius can be calculated as:

$$r_i = \frac{d_i \phi}{2(\phi \sin \alpha_i + \cos \alpha_i)}. \tag{A.2}$$

In order to prevent the radius from becoming large as α approaches $\frac{\pi}{2}$, the denominator is clipped to 0.1. The result of this radius estimation is that discs that are further away from the scanner are larger as a consequence of the beam spreading. This approximation for the radius of each disc leads to both overlapping and the formation of gaps between adjacent discs. However, it was empirically found to produce a well structured representation of the scene from which the visibility can be computed. The advantage of the disc approximation is that ray-disc intersection algorithms are computationally efficient and no manual tweaking of the surface representation needs to be performed.

Appendix B

Relighting with Respect to Sunlight

The scaling factor for relighting the point B', which is occluded from the sun in the original image, with respect to full terrestrial sunlight and diffuse skylight exposure is calculated as:

$$\begin{aligned} L_{B'-relit}(\lambda) &= \frac{\rho_{B'}(\lambda)}{\pi}\left[E_{sun}(\lambda)\tau(\lambda) + E_{sky}(\lambda)\right], \\ &= \frac{\rho_{B'}(\lambda)}{\pi}\left[E_{sun}(\lambda)\tau(\lambda) + E_{sky}(\lambda)\right]\frac{\Gamma_{B'}}{\Gamma_{B'}}, \\ &= \frac{\rho_{B'}(\lambda)}{\pi}\left[\Gamma_{B'}E_{sky}(\lambda)\right]\left[\frac{\frac{E_{sun}(\lambda)\tau(\lambda)}{E_{sky}(\lambda)} + 1}{\Gamma_{B'}}\right], \\ &= L_{B'}(\lambda)\left[\frac{\frac{E_{sun}(\lambda)\tau(\lambda)}{E_{sky}(\lambda)} + 1}{\Gamma_{B'}}\right]. \end{aligned} \tag{B.1}$$

Therefore, the scaling factor is wavelength dependent due to the presence of the terrestrial sunlight-skylight ratio term.

Relighting a point C that is originally exposed to direct sunlight, mean the scaling factor is calculated as:

$$L_{C-relit}(\lambda) = \frac{\rho_C(\lambda)}{\pi}\left[E_{sun}(\lambda)\tau(\lambda) + E_{sky}(\lambda)\right],$$

$$
\begin{aligned}
&= \frac{\rho_C(\lambda)}{\pi}\left[E_{sun}(\lambda)\tau(\lambda) + E_{sky}(\lambda)\right]\left[\frac{E_{sun}(\lambda)\tau(\lambda)\cos\alpha_C + \Gamma_C E_{sky}(\lambda)}{E_{sun}(\lambda)\tau(\lambda)\cos\alpha_C + \Gamma_C E_{sky}(\lambda)}\right], \\
&= \frac{\rho_C(\lambda)}{\pi}\left[E_{sun}(\lambda)\tau(\lambda)\cos\alpha_C + \Gamma_C E_{sky}(\lambda)\right] \\
&\quad \left[\frac{E_{sun}(\lambda)\tau(\lambda) + E_{sky}(\lambda)}{E_{sun}(\lambda)\tau(\lambda)\cos\alpha_C + \Gamma_C E_{sky}(\lambda)}\right], \\
&= L_C(\lambda)\left[\frac{\frac{E_{sun}(\lambda)\tau(\lambda)}{E_{sky}(\lambda)} + 1}{\frac{E_{sun}(\lambda)\tau(\lambda)}{E_{sky}(\lambda)}\cos\alpha_C + \Gamma_C}\right]. \qquad \text{(B.2)}
\end{aligned}
$$

Therefore, the scaling factor for non-occluded point is wavelength dependent when relighting with respect to sunlight and skylight. The wavelength dependence for both the occluded and non-occluded cases is undesirable since the occluded areas typically have a low SNR and multiplication by these scaling factors will amplify the noise.

Appendix C

Spatial Illumination Invariance - Dataset 3

In this Appendix, the results for attaining spatial illumination invariance via relighting for Dataset 3 are presented. Relighting is performed using the *full* and *fast* methods for sky factor approximation in order to compare their efficacy.

In this dataset, 14 pairs of points were manually selected, with each pair obtained from a uniform material under different illumination conditions. This means the appearance of the points varies in terms of intensity and colour, and this compensated using the relighting algorithm proposed in Section 3.2. The approximation of the sky factors using the *full* and *fast* methods is shown in Figure C.1. The magnitude difference between the two approximations is indicative of the occlusions present in the scene. For example, in point pairs 1 to 8 there is an approximate difference of 0.15, due to the geometry reducing line of sight visibility to the sky dome.

Figures C.2, C.3, C.4, C.5 and C.6 present the relit spectra for each pair of points using the *full* and *fast* sky factor approximations. Greater spectral similarity is seen to be attained using both methods as indicated by the convergence of the spectra following relighting.

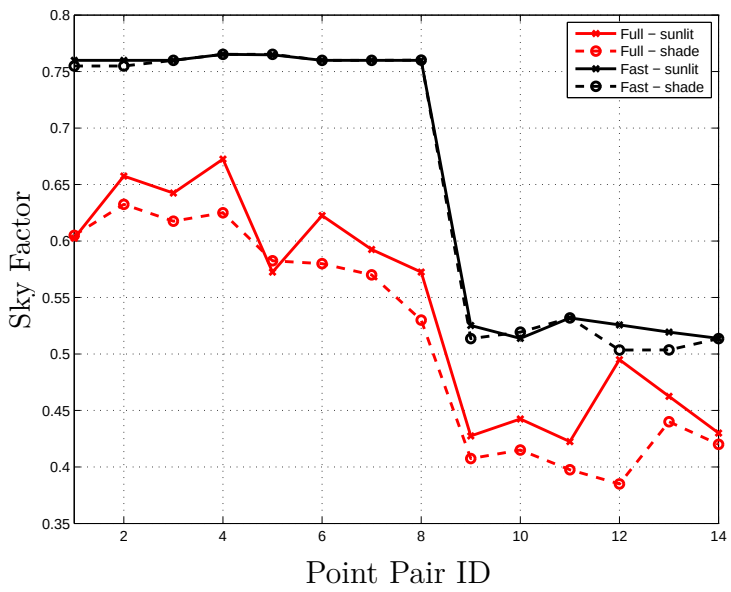
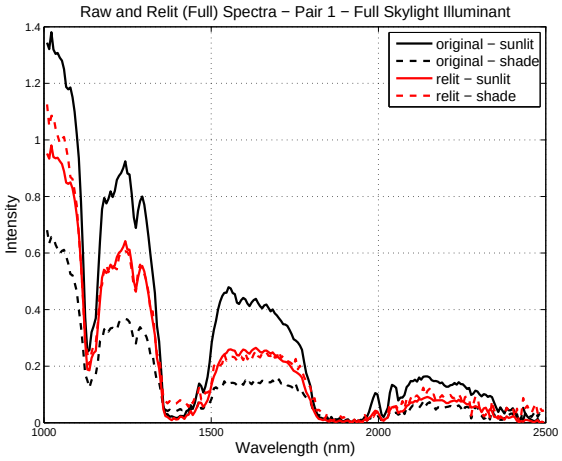
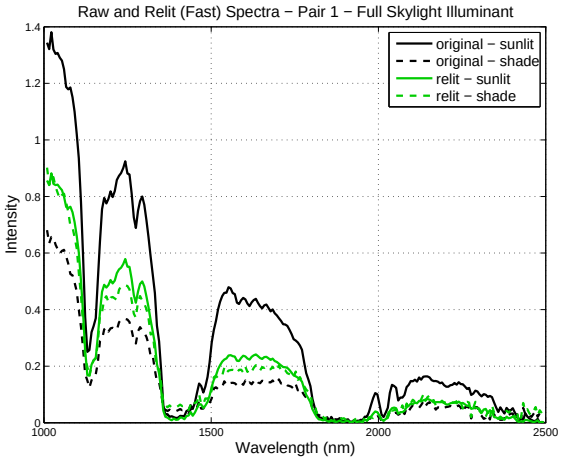


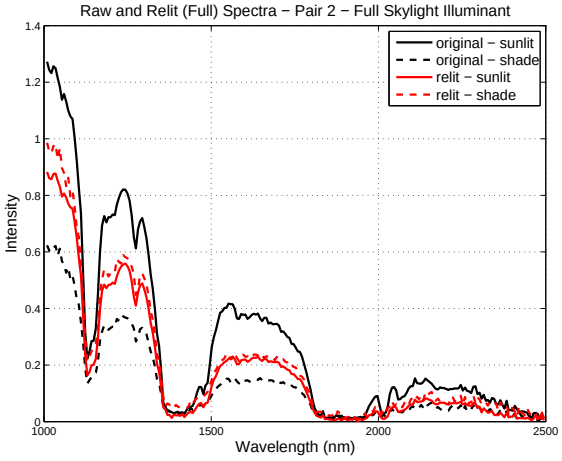
Figure C.1 – Sky factors calculated using the *full* and *fast* approximation.



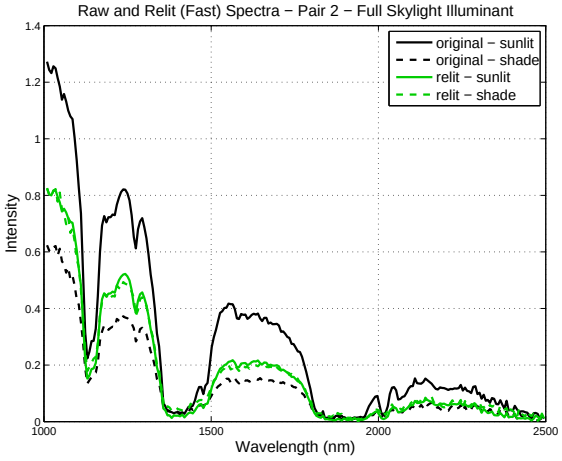
(a) Pair 1 - *full*



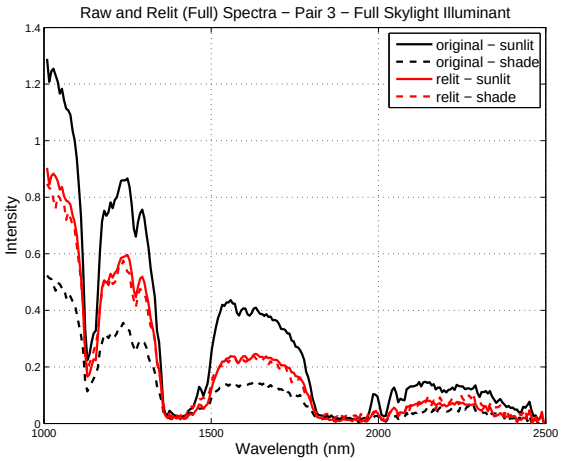
(b) Pair 1 - *fast*



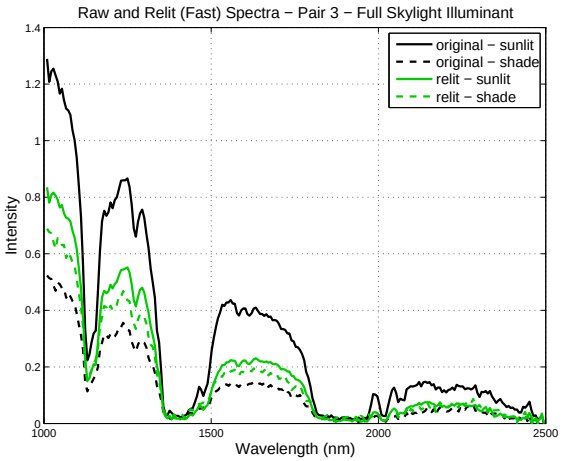
(c) Pair 2 - *full*



(d) Pair 2 - *fast*

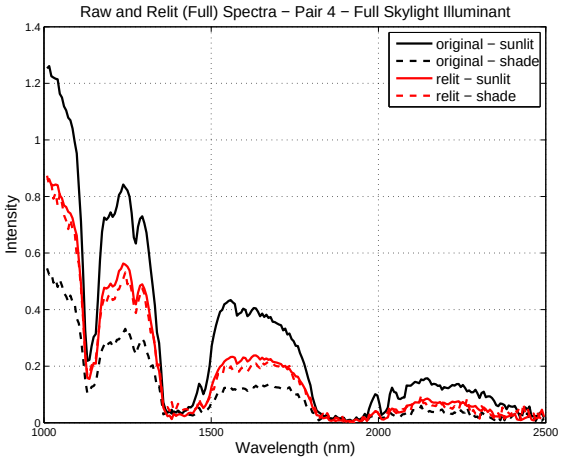


(e) Pair 3 - *full*

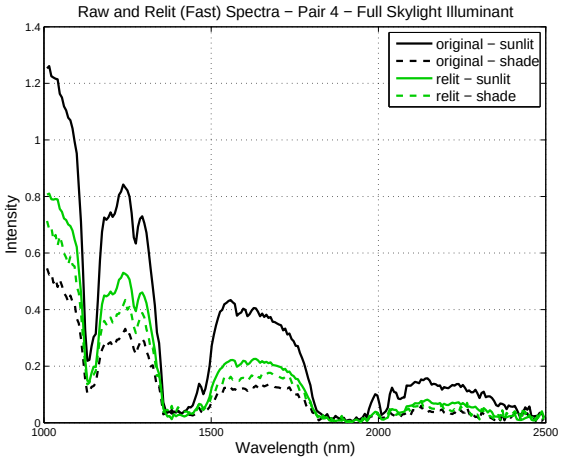


(f) Pair 3 - *fast*

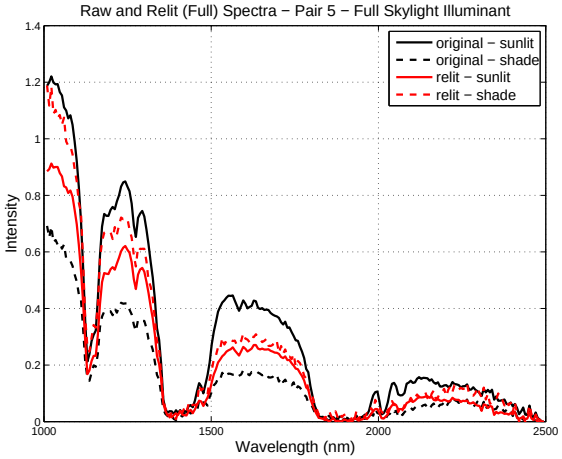
Figure C.2 – Original and relit pairs of points 1 to 3, using the *full* and *fast* sky factor approximation.



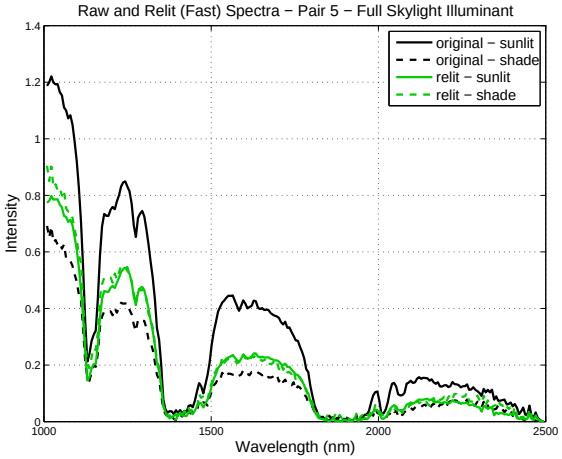
(a) Pair 4 - *full*



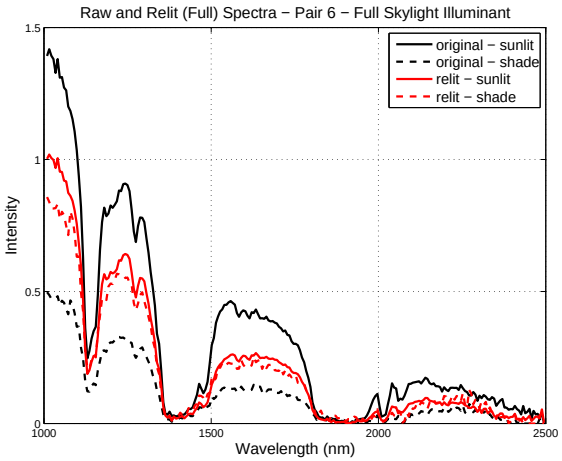
(b) Pair 4 - *fast*



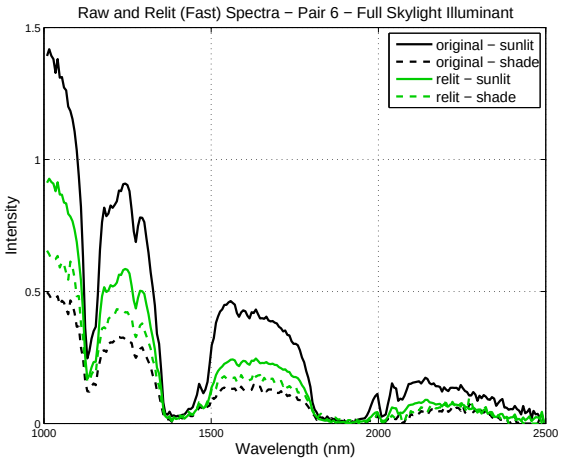
(c) Pair 5 - *full*



(d) Pair 5 - *fast*

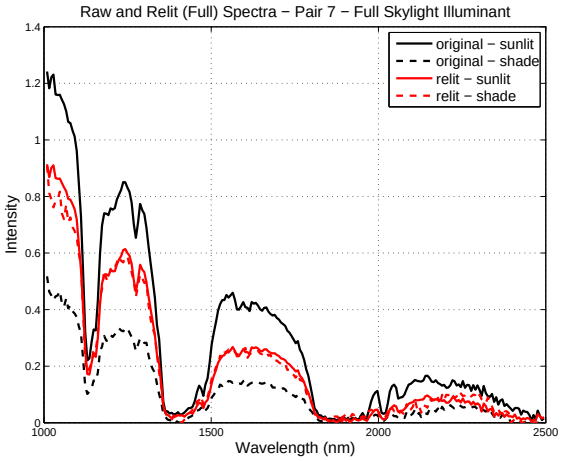


(e) Pair 6 - *full*

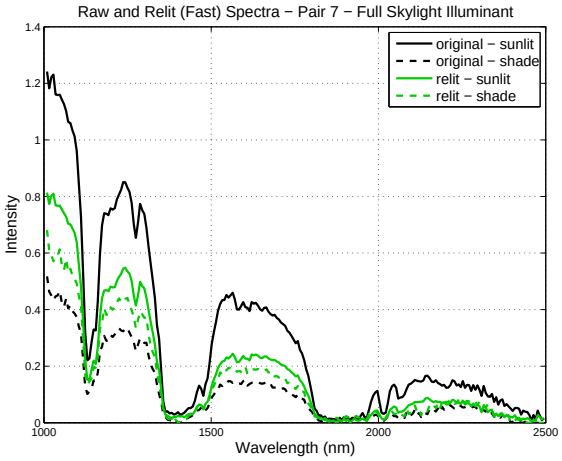


(f) Pair 6 - *fast*

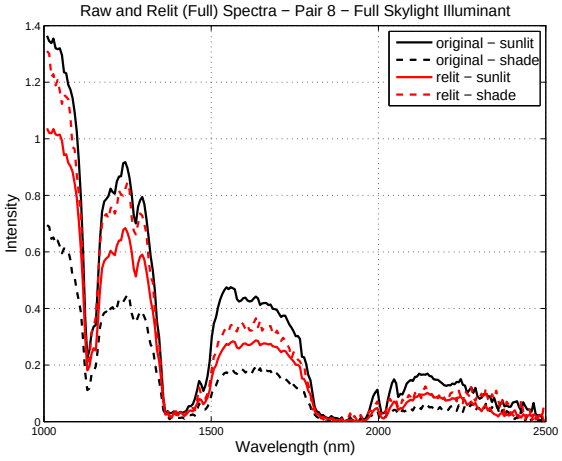
Figure C.3 – Original and relit pairs of points 4 to 6, using the *full* and *fast* sky factor approximation.



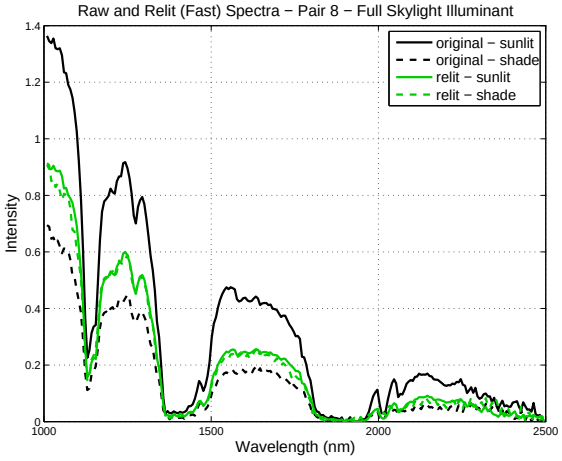
(a) Pair 7 - *full*



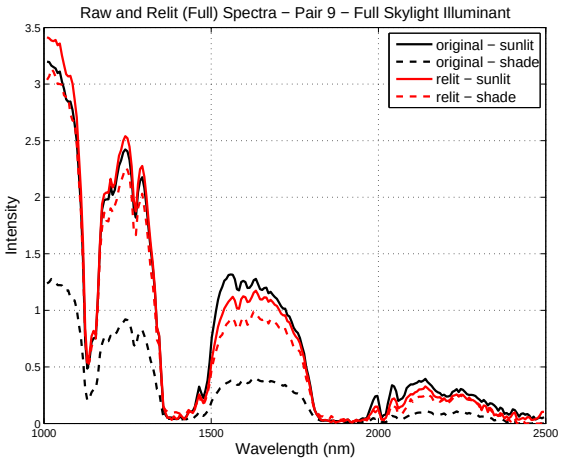
(b) Pair 7 - *fast*



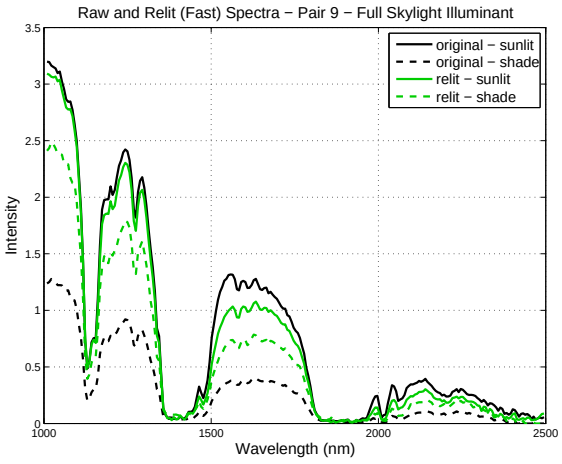
(c) Pair 8 - *full*



(d) Pair 8 - *fast*

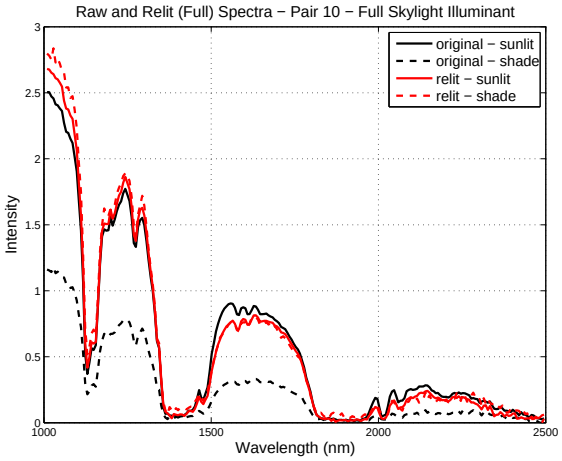


(e) Pair 9 - *full*

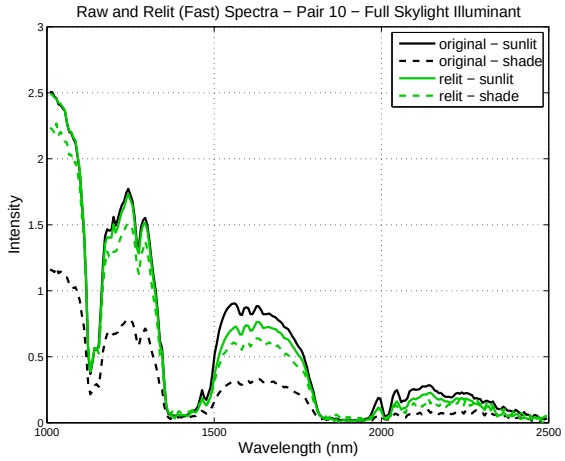


(f) Pair 9 - *fast*

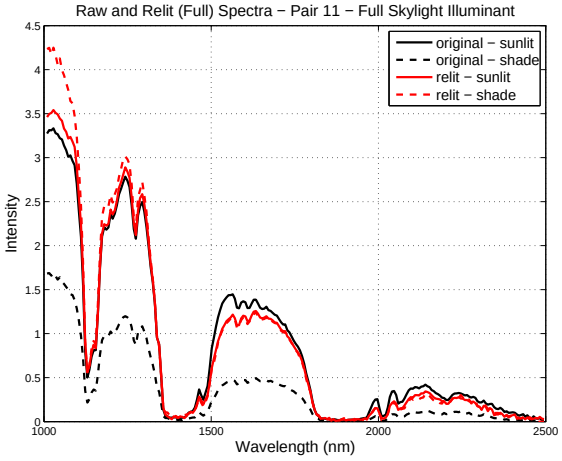
Figure C.4 – Original and relit pairs of points 7 to 9, using the *full* and *fast* sky factor approximation.



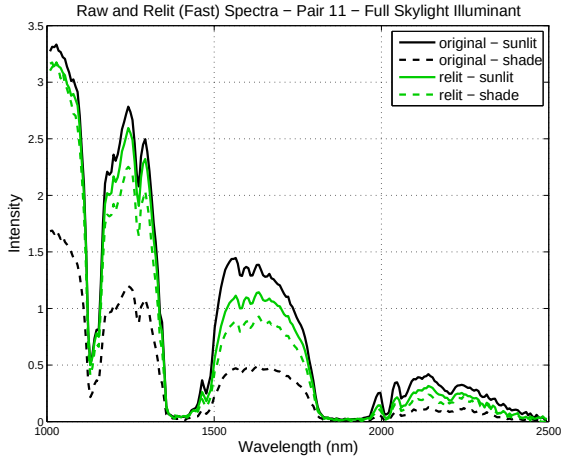
(a) Pair 10 - *full*



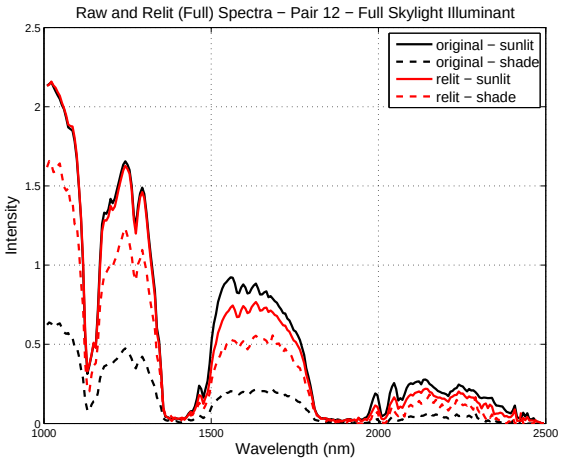
(b) Pair 10 - *fast*



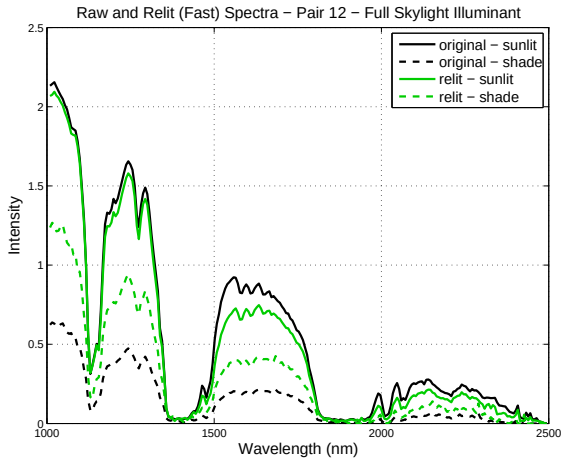
(c) Pair 11 - *full*



(d) Pair 11 - *fast*

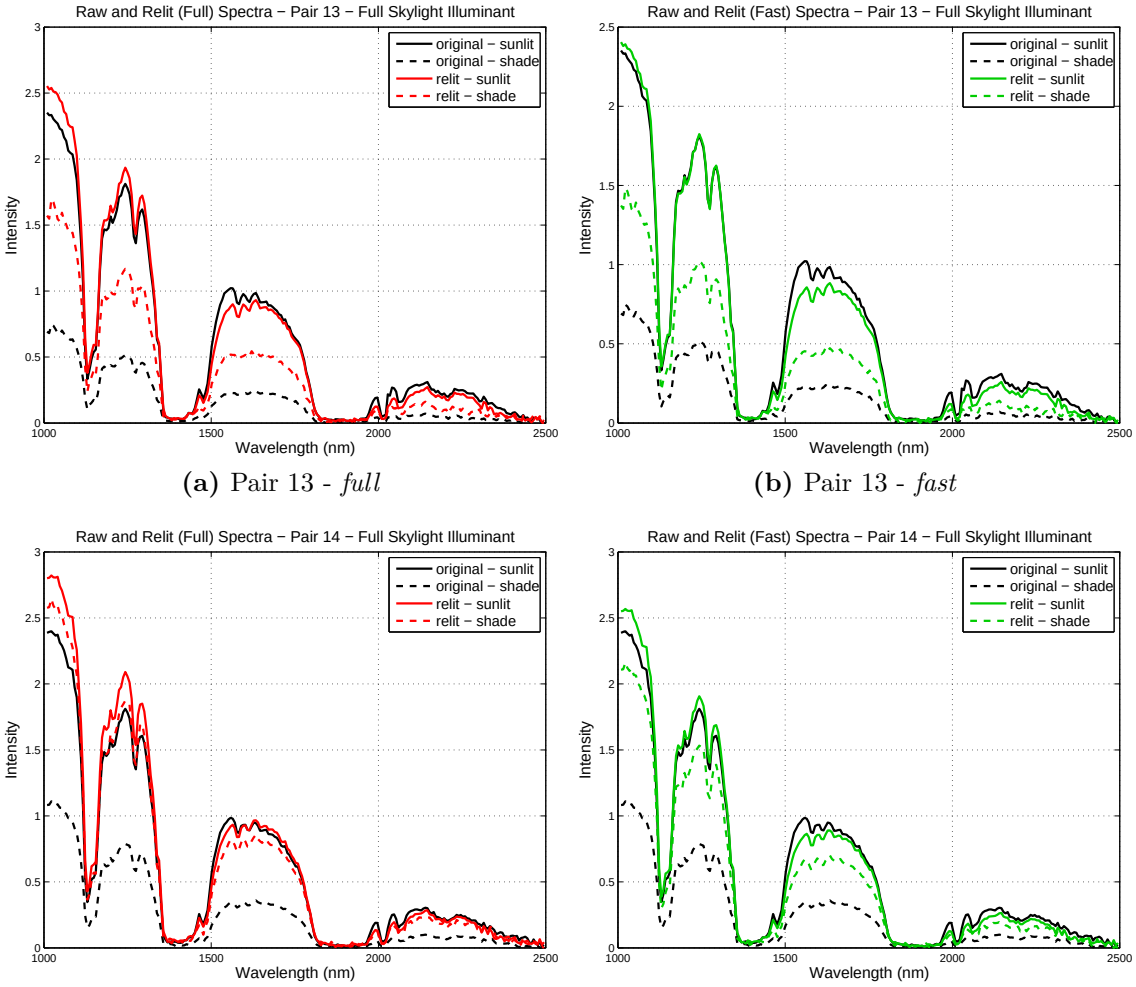


(e) Pair 12 - *full*



(f) Pair 12 - *fast*

Figure C.5 – Original and relit pairs of points 10 to 12, using the *full* and *fast* sky factor approximation.



(a) Pair 13 - *full*

(b) Pair 13 - *fast*

(c) Pair 14 - *full*

(d) Pair 14 - *fast*

Figure C.6 – Original and relit pairs of points 13 and 14, using the *full* and *fast* sky factor approximation.