Erasmus+

COSI Computational Colour and Spectral Imaging

Master in Computational Colour and Spectral Imaging (COSI)

UNIVERSITY OF EASTERN FINLAND · UNIVERSIDAD DE GRANADA · UNIVERSITÉ DE LYON · UNIVERSITÉ JEAN MONNET SAINT-ÉTIENNE · NTNU Norwegian University of Science and Technology

Estimation of Daylight Illumination Spectra and Relative Reflectance Recovery from Hyperspectral Radiance Cubes without Reference Targets

Master Thesis Report

Presented by

Maximilian Czech

and defended at the

Norwegian University of Science and Technology

September 2023

Academic Supervisors: Prof. Steven Yves Le Moan,
Norwegian University of Science and Technology, Norway

Prof. Javier Hernández-Andrés,
Universidad de Granada, Spain

Host Supervisor: M.Sc. Ben Müller,
Cubert GmbH, Germany

Jury Committee:

1. Sr. Researcher Pauli Fält,
   University of Eastern Finland, Finland
2. Prof. Pichayada Katemake,
   Chulalongkorn University, Thailand

Submission of the thesis: 10$^{\text{th}}$ August 2023
Day of the oral defense: 5$^{\text{th}}$ September 2023

Abstract

When working with hyperspectral images, a crucial step is to correctly identify the illumination conditions and remove their influence from the recorded scenes. While the most common approach is to capture a reference target of known reflectance as a ground truth measurement, this might not be sufficient for specific applications under certain conditions, such as unmanned aerial vehicle (UAV) based captures of forests and agricultural areas. Here, the proper placement of reference targets within the scene can be complicated or impossible, and time-separated captures for scene and calibration can be insufficient due to changes in cloud coverage or similar effects. A possible solution to this problem is the estimation of the illuminant from the scene capture directly without the use of a calibration target. Different approaches exist to tackle this underconstrained problem, mainly based on image statistics or by making use of the physical properties of the recorded objects. Nevertheless, each of those algorithms has its flaws, such as the dependency on objects with specific properties, for example convex surfaces or a large variety of different reflectance properties, while other approaches rely on the identification of specular highlights. These constraints might not be sufficient for applications involving vegetation such as agriculture or environmental monitoring, and therefore a need exists for the development of precise illumination estimation algorithms.

In this work, a novel image statistics-based approach is proposed to estimate the relative spectral power distribution (SPD) of daylight illuminants from a given hyperspectral radiance cube. The created method restricts the possible results in the spectral domain by applying an effective constraint to limit the search area by utilizing the information inherited in the image cube. Within the resulting search area, a set of random spectra are generated which represent the image-dependent plausible illuminants. These spectra are then transferred into a pre-trained, image-independent principal component space, where an intersection between these image-dependent plausible illuminants and a set of measured physically possible illuminants is calculated. The scene illuminant is then derived by transforming this intersection point back to the spectral domain, where the estimated illuminant can be used to recover the relative scene reflectance values. The framework created in this work is tested on a set of 150 hyperspectral radiance cubes simulated by using 25 spectral reflectance cubes and six representative daylight SPDs. The daylight illuminants are chosen based on their inverse correlated color temperature and cover a broad range from vertical daylight, and overcast sky to clear blue sky conditions. The estimated SPDs and the recovered reflectance spectra are evaluated using several well-adopted full reference metrics for spectral comparison. In addition, the performance of the proposed model is compared to several other image statistics-based illuminant estimation methods and it is demonstrated that the new model outperforms the competitors significantly. The results clearly show the ability of the proposed model to accurately estimate the SPD of a representative variety of common daylight illuminants in the VIS and NIR range.

Acknowledgment

I want to thank Cubert GmbH for giving me the opportunity to create this work by providing me with the latest technology, substantial data, and generous support.

Further, I express my gratitude to my supervisors for their valuable guidance and the constructive feedback they gave me along the way.

Contents

LIST OF ABBREVIATIONS

CCC computational color constancy. 19

CCT correlated color temperature. 45–47, 49, 90, 93

CCT^{-1} inverse correlated color temperature. 46, 48, 55, 58–62, 64, 65, 67, 90, 91, 93

CFA color filter array. 14

CGFC complemented GFC. 43, 52, 53, 55, 63–65

CNN convolutional neural network. 23

DN digital number. 15, 18

FOV field of view. 14, 19

GFC goodness of fit coefficient. 51, 52

HVS human visual system. 19, 20, 22

IRE integrated radiance error. 53, 55, 64, 65

LOWESS locally weighted scatter-plot smoothing. 30

NIR near infrared. 8

PC principal component. 34, 43

PCA principal component analysis. 31, 32, 34, 41, 42, 45, 67

PCS principal component space. 5, 25, 26, 32–39, 42–44, 48, 50, 55, 56, 64–66, 68, 90, 93

RANSAC random sample consensus. 36, 38

RMSE root mean square error. 51–53, 55, 63, 65

SAM spectral angle mapper. 52, 53, 55, 64, 65

SFA spectral filter array. 14

SNR signal-to-noise ratio. 13, 17, 18, 22

SPD spectral power distribution. 3, 4, 7, 12, 15, 16, 20–23, 25–27, 34, 35, 44–46, 48–51, 53, 55, 56, 58, 60–62, 64, 65, 75, 90, 91, 93, I

SVD singular value decomposition. 33, 34, 38

SWIR short wavelength infrared. 8, 66

UAV unmanned aerial vehicle. 4, 17–19, 26, 48, 90, I

UV ultraviolet. 8, 66

VIS visible light. 8, 22, 24, 46

1 INTRODUCTION

1.1 Contextualization of the Problem

In the last decades, multi- and hyperspectral imaging emerged to a promising technology beyond satellite-based earth observation. Especially with new developments in real-time spectral imaging and snapshot systems, a broad field of applications are becoming accessible, such as agricultural surveillance (Lu et al., 2020), food quality assessment (Ma et al., 2019) or cultural heritage (Jung, 2017). For most use cases, it is of particular interest to convert the captured scene from radiance data to reflectance values. To do so, the common practice is to capture a reference target within the same conditions, usually placed within the scene directly or right before or after the scene capture (Yao and Lewis, 2010). Since the reflective material properties of such a target are relatively uniform over a broad spectral range, the obtained data can be used as ground truth measurement for the SPD of the illuminant present in the scene. However, especially in uncontrolled conditions such as outdoor measurements, the use of a reference target can cause severe problems. Weather conditions and cloud coverage can change rapidly, causing a need for recapturing the ground truth data (Wendel and Underwood, 2017). In other cases, the correct positioning of the target might be complicated. One way to overcome these challenges is to eliminate the need for capturing a reference target with the data in the first place and estimate the illuminant directly from the captured data itself. Although several such estimation approaches were developed in the past, there is still room for significant improvement in terms of accuracy and robustness to a variety of input data. Since many existing algorithms require specific scene content such as specular reflections or objects with "flat" surface reflectances as discussed in Chapter 3, this topic is still considered ongoing research with room for finding new ways to estimate scene illuminant spectra.

1.2 Scope of this Work

The scope of this work is to develop a novel framework for the estimation of typical daylight illuminant spectra from normalized radiance hyperspectral cubes for *in situ* measurements. Thereby, the proposed model should not rely on any ground truth measurements such as the capture of a dedicated reference target of known reflectance. Such an algorithm is especially valuable for industrial applications where the proper capture of ground truth data is not possible. A relevant application in mind hereby are measurements for surveillance of forests and agricultural areas, for example by using a hyperspectral snapshot camera mounted on an UAV (Woo et al. (2021), Liu et al. (2020), Chen et al. (2019), Dash et al. (2017)). Of particular interest for vegetation analysis approaches spectral is thereby the VIS and NIR range (Marchi et al., 2022). Nevertheless, the proposed model should not be restricted to these applications. The focus of this work hereby lies on the development of a novel image statistics-based algorithm using general constraints and restrictions for the ill-posed problem of illuminant estimation. The possibility is explored to accurately estimate daylight spectra within an image-independent principal component space by finding the intersection of *image-dependent possible spectra* and a set of *physically plausible spectra*. It is evaluated by comparison of the estimated SPD with the corresponding ground truth illuminant, by comparing the illuminant estimation accuracy of the proposed model against competing image statistics-based algorithms, and by analyzing the reflectance images recovered by using the estimated illuminants against the ground truth values. To evaluate different properties of the estimated spectra, well-adopted metrics are used for comparison against the ground truth and each other.

1.3 Main Contributions

In this work, a novel approach is proposed to the unsolved problem of the estimation of the illuminant SPD from image data without ground truth measurements. The main contributions of the author are the development of the proposed framework both theoretically and in practice. For the practical implementation of the framework, the programming language Python version 3.9 was used.

The theoretical development of the algorithm included a broad-ranging review of existing literature in the field. Upon this foundation, a first approach for the estimation was created by using colorimetric information for the estimation of the illuminant as well as spectral correction factors as proposed by Derhak and Rosen (2006) to reconstruct the illuminant spectrum and to lessen the influence of metamerism on the results. Due to technical challenges with the reconstruction

of the illuminants, this approach was abandoned. Instead, the development of a method using effective but generic constraints as well as an image-independent principal component space were prioritized. The main benefit of this second approach is the utilization of the full potential of the available spectral data. The theoretical considerations went hand in hand with the practical implementation of the proposed method. Hereby, the practical implementation in Python does not only includes the main components of the algorithm itself but also the necessary data preprocessing and utility functions. Among others, functions were created for loading and saving hyperspectral data of different formats; the calculation of reflectance and radiance data; for cropping, interpolation, and normalization of the spectral data; for the calculation of colorimetric representations of the data cubes, the proper handling of image metadata; the implementation of existing algorithms and spectral comparison metrics for the evaluation and batch-processing as well as evaluation scripts for the creation and analysis of the results. A link to the relevant Python code that forms the practical part of this work can be found in Appendix A. In addition, Appendix B shows the structure of the code with all implemented classes, attributes, and functions in the form of UML diagrams. The designed method stands out from competing algorithms for two main reasons: The creation of an image-independent principal component space (PCS) in which the estimation is performed, as well as the estimation approach within this common space itself by combining information derived from image-dependent plausible and physically possible spectra.

1.4 Outline and Structure

This work is divided into five main parts. At first, Chapter 2 THEORY provides the necessary theoretical background on the topics of multi- and hyperspectral imaging, prevalent acquisition techniques, and the importance of illuminant invariant representation of captured imaging data. In addition, a comprehensive review is given of the related work done in the field of illuminant estimation from spectral data, categorized by the higher-level approaches established in the past. As a main part of this work, Chapter 3 METHODOLOGY then focuses on the proposed framework for illumination estimation. First, the assumptions and constraints are defined, followed by a detailed explanation of the novel framework itself. Afterwards, a concise summary is given of the main steps for the convenience of the reader, while the following sections focus on the experiments for evaluating the model. Here, the preparation of the training and test data is explained and the metrics used to evaluate the results are defined. Chapter 4 RESULTS then illustrates the obtained results before critically discussing the findings in Chapter 5 DISCUSSION. Last, Chapter 6 CONCLUSION summarizes the key findings and

gives an outlook on possible improvements that can be done to further develop the proposed illumination estimation algorithm in the future.

1.5 About the Use of Artificial Intelligence for the Creation of this Work

Within the process of creating this work in written form and practical implementation, Open AI's ChatGPT was used for minor tasks of debugging self-written Python code (such as syntax errors) to streamline the process of code creation. The author hereby assures that besides the above-mentioned tasks, neither ChatGPT nor other comparable tools were used for the creation of significant parts of the program code; the process of writing, editing, rephrasing, or creating this report; especially with respect to the development of the proposed ideas and techniques that are presented in this work.

2 THEORETICAL BACKGROUND AND RELATED WORK

This chapter aims to provide the relevant background on fundamental concepts related to this work. First, a brief introduction is given to basic terminology in the field of hyperspectral imaging. Then, an overview of different spectral data acquisition techniques is presented, followed by some details on how the captured raw data is processed after the acquisition process. Afterwards, existing work in the field of illumination and reflectance estimation is categorized and discussed.

At this point, some clarification must be given about the term illumination estimation, since its definition differs with the context. When talking about traditional color imaging and computational color constancy within the context of colorimetry, illumination estimation usually refers to the recovery of the chromaticity of a given light source present in the scene. However, in the context of spectral imaging, which applies to this work, illumination estimation describes the recovery of the light source's SPD, in other words, its power as a function of wavelength.

2.1 Hyper- and Multispectral Imaging

The following section focuses on fundamentals and basic principles within the field of spectral imaging, starting with some clarification about important terminology within the field. Then, a brief summary is given on what hyperspectral data is, how it is recorded using different acquisition techniques, and how it is post-processed to utilize the captured information.

2.1.1 Electromagnetic Spectrum and Spectral Range

The electromagnetic spectrum is caused by electromagnetic radiation of photons propagated as waves. Hereby, the distance between the peaks of such a wave is called the wavelength and describes the energy within the spectrum. It is possible to divide the electromagnetic spectrum into different wavelength-dependent regions, as

can be seen in 2.1. Hereby, the shorter the wavelength, the higher the frequency and energy of the corresponding electromagnetic radiation. The sensitivity of a spectral imaging system to a specific region of wavelengths within the electromagnetic spectrum is called its spectral range (Sun, 2010). The human ability to sense electromagnetic radiation through the visual system is thereby limited to the spectral range of around 380 - 780nm (Berns, 2019). This range is usually referred to as visible light (VIS). Other important spectral ranges in the field of spectral imaging are ultraviolet (UV) (around 100 - 400nm), near infrared (NIR) (around 780 - 1400 nm), and short wavelength infrared (SWIR) (around 1400 – 3000nm), whereas the boundaries of the respective ranges are only loosely defined.

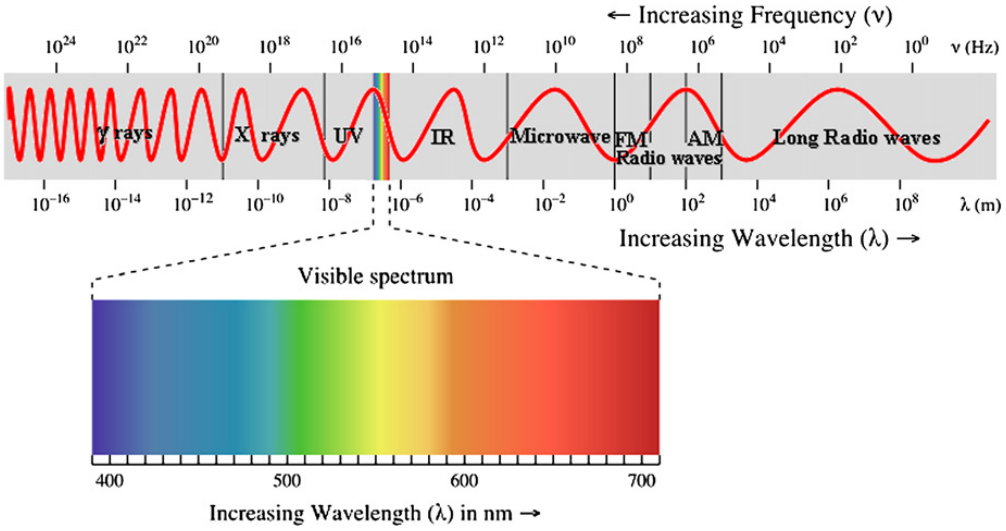


Figure 2.1: *The electromagnetic spectrum and its regions, from Sun (2010).*

2.1.2 Spatial and Spectral Resolution

Spatial resolution describes how detailed a scene can be captured geometrically and how fine the resulting capture is resolved by the image sensor (Liang and Wang, 2020). The main property which defines the spatial resolution is the pixel size and the number of pixels that built the sensor array. On the contrary, the spectral resolution is defined by the number of bands of a given imaging system as well as the bandwidth of each of the spectral bands (Liang and Wang, 2020). In general, the more spectral bands are used for a given wavelength range, the finer resolved are the measured spectra of a given device.

The number of bands can vary from three broadband channels when recording color images in the VIS range (low spectral resolution) up to hundreds of very narrow

bands of hyperspectral acquisition devices with a very high spectral resolution. Hereby, a trade-off exists between the spatial and spectral resolution, as discussed by Hagen and Kudenov (2013) and evaluated in depth by Jia et al. (2022). The reason for this trade-off is a decreased amount of effective energy collected by the system when reducing the pixel size, which is important to increase the spatial resolution of a sensor Jia et al. (2020). Less energy is then available for a precise detection in terms of spectral resolution. On the contrary, increasing the spectral resolution of a system leads to a decreased amount of energy per band, which causes a need for increased pixel size.

2.1.3 Multi- and Hyperspectral Image Data

In conventional digital color imaging, the data captured by a camera consists of three broadband channels red (R), green (G), and blue (B), which correspond to the respective long, medium, and short wavelength regions of the visible spectrum. In contrast, multispectral and hyperspectral data is recorded by using many more channels, each representing a narrower part of the electromagnetic spectrum. Although not strictly defined, the term multispectral image thereby usually refers to a capture which contains more than three and up to 10 bands (Lu and Fei, 2014). Hyperspectral images on the other hand can consist of up to several hundreds of narrow bands, where each channel refers to a very narrow region of the spectrum, resolving a scene with a much higher spectral resolution. Such a recorded image is usually referred to as a (hyper-)spectral image or cube and, dependent on the spatial and spectral resolution, can easily reach hundreds of megabytes up to several gigabytes of space needed for storage per captured scene. Whenever applicable and not explicitly stated otherwise, the term spectral imaging is used in this work to refer to hyper- and multispectral imaging likewise. With such imaging techniques, the goal is to preserve the spectral information of the captured scene. This emerging technology enables many useful applications in a broad range of industrial branches, such as earth observation and remote sensing (Bilal et al. (2019), Goetz (2009), Stuffler et al. (2007)), agriculture and food science (Abdelbaki et al. (2021), Lu et al. (2020), Rady et al. (2019), Adão et al. (2017), Aasen et al. (2015), Dale et al. (2013)), medical imaging (Cihan et al. (2022), More et al. (2019), Yang et al. (2018), Calin et al. (2013)), computer vision and object detection (Bajić and Bajić (2021), Serranti et al. (2011)) or cultural heritage (Cutajar et al. (2022), Grillini et al. (2021), Polak et al. (2017)).

2.1.4 Data Acquisition Techniques

Four main data acquisition techniques exist in the field of spectral imaging, namely point-scanning (whiskbroom), line-scanning (pushbroom), area-scanning (filter-wheel), and single shot (snapshot). Each acquisition technique has its advantages and disadvantages and different approaches when recording a hyperspectral cube. Figure 2.2 visualizes what data each method captures when triggering a single acquisition. The different techniques will be explained in more detail in the following sections.

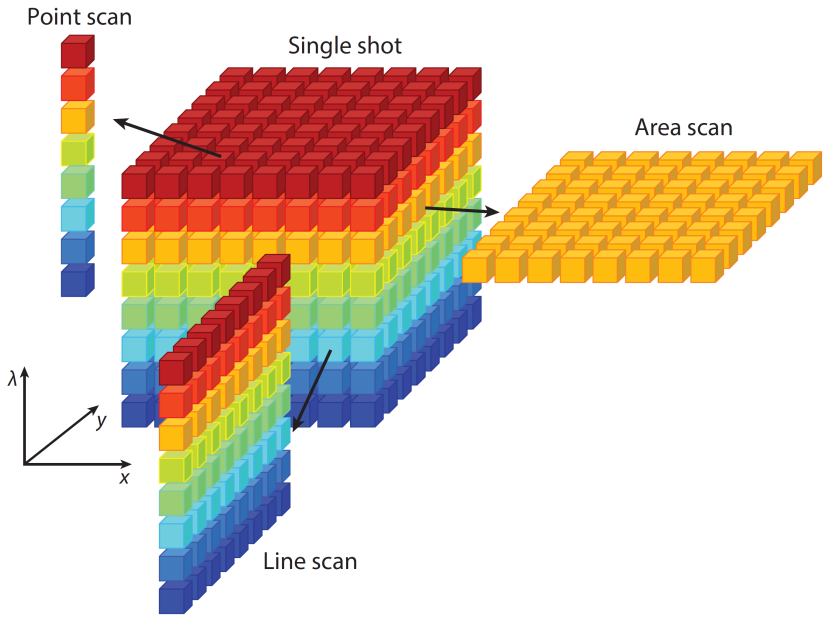


Figure 2.2: *Visualization of the data recorded per capture for the different data acquisition techniques in the field of spectral imaging. x and y represent the spatial dimensions of the datacube, while λ is the spectral dimension. Illustration from Ma et al. (2019).*

2.1.4.1 Whiskbroom / Point-Scanning

Whiskbroom acquisition systems, also called point-scanning devices or point-scanners record a single spectrum per acquisition. Point-scanning systems usually disperse the incoming radiance through a prism or grating (Li et al., 2013). The dispersed light then hits the linear image sensor array. Each pixel of the sensor records a different and usually narrow part of the spectrum. Such point-measuring devices can not record a scene as a spectral cube by default. To do so, either the device or the specimen is moved both in x and y direction over the area of interest to

take several measurements of each point of the scene (Ma et al., 2019). This leads to more complex hardware setups and a limitation to static samples. In addition, the acquisition process of capturing a spectral cube is very time-consuming compared to the other systems. Usually, one big advantage of whiskbroom systems is the large spectral resolution and the fast readout (Hagen and Kudenov, 2013). When there is no time constraint, the scene can be captured with a high spatial resolution. But since scanning large scenes with high spatial resolution takes a large amount of time when using the whiskbroom acquisition technique, there is a trade-off between acquisition time and spatial resolution.

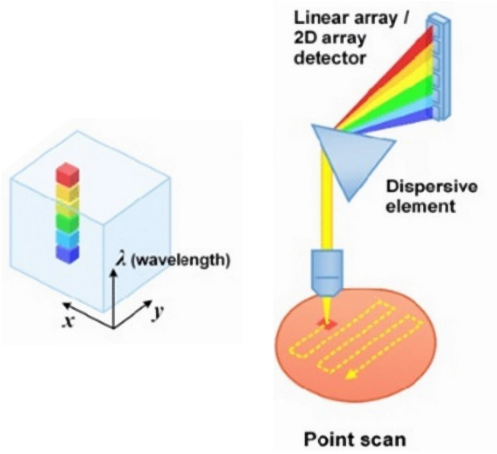


Figure 2.3: *Visualization of the working principle of a whiskbroom acquisition system, from simtrum.com, last accessed: 05.08.2023.*

2.1.4.2 Pushbroom / Line-Scanning

Pushbroom spectral acquisition systems, also referred to as line scanners, sample a scene in the spectral domain for one spatial dimension at the same time (Ma et al., 2019). The line is recorded by dispersing the slit image onto a 2D sensor, where one dimension of the sensor represents the spatial information and the second dimension inherits the spectral information. The sample or the detector is linearly moved in the perpendicular spatial dimension, and the acquisition trigger is synchronized with the moving speed and integration time. Several captures are then taken to cover the region of interest, creating the spectral cube with two spatial dimensions and a spectral dimension at each pixel. Thereby the spatial dimension is restricted in one axis but is theoretically infinite in the second axis defined by the movement of the sample or detector. The line-scanning method is not able to record samples in motion, apart from the motion introduced by the

scanning process itself. Line-scanning systems are especially suitable for airborne and spaceborne platforms and conveyor belt applications (Hagen and Kudenov, 2013). They are also excessively used in the field of cultural heritage, for example, to scan paintings for analysis and restoration as shown for example by Deborah et al. (2019).

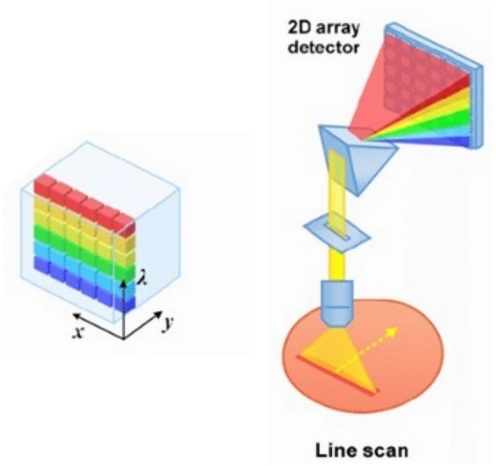


Figure 2.4: *Visualization of the working principle of a pushbroom acquisition system, from simtrum.com, last accessed: 05.08.2023.*

2.1.4.3 Filter Wheel / Area-Scanning

A filter wheel system typically consists of a camera, a light source with known SPD as well as a wheel with several fixed band-pass or linear variable filters (Li et al., 2013). The spectral bands of the system are determined by the transmittance of the filters and the quantum efficiency of the sensor. Two different kinds of filter wheel systems exist. The active type has a rotating wheel of different bandpass filters with known spectral transmittance in front of the camera's lens, directly in the optical path of the system. The passive type has no filter wheel in front of the imaging system, but in front of the light source instead. The multispectral cube is recorded by capturing a separate monochromatic 2D image for each of the filters, one spectral band at a time. Thereby, the number of spectral bands as well as the spectral sensitivity can be adjusted by using different filters and different filter combinations. Other approaches exist where multichannel sensors are used Shrestha and Hardeberg (2013) to reduce the amount of captures that have to be taken for the acquisition of the spectral cube. A variation of the passive approach is to use different light sources directly instead of filters, as shown for example by Trumpy et al. (2021) and Shrestha and Hardeberg (2013). With LED

technology becoming more advanced and cheaper in recent years, this method is now a promising alternative to the filter wheel in controlled conditions. Figure 2.5 shows an example of an active filter wheel system. In general, such area-scanning systems are easier and cheaper to build compared to line scanners or snapshot systems. In addition, the integration time can be varied for each band, to use more of the sensor's dynamic range and to increase the signal-to-noise ratio (SNR). Filter wheel systems are especially suitable for colorimetry applications, where multispectral approaches are sufficient and usually high spatial resolution is preferred. One of their biggest limitations is the need for capturing several images after another, trading a higher spectral resolution for longer acquisition times. Even when limiting the number of spectral bands and using an RGB instead of a monochromatic sensor, several captures have to be acquired. This makes area-scanning devices unsuitable for applications with moving samples, as well as for scenarios where a rapid or real-time capturing process is crucial. Instead of using a filter wheel, it is also possible to use mechanically tunable filters such as liquid-crystal tunable filters or acousto-optic tunable filters (Poger and Angelopoulou, 2001).

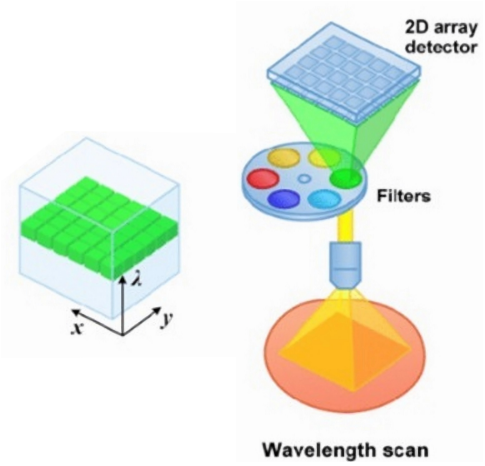


Figure 2.5: *Visualization of the working principle of an active filter wheel acquisition system, from simtrum.com, last accessed: 05.08.2023.*

2.1.4.4 Snapshot / Single Shot

Instead of capturing time-sequential 1D or 2D data as with the previously introduced techniques, snapshot or single shot acquisition systems are able to capture the whole datacube consisting of the two spatial dimensions as well as the spectral dimension with a single integration. This removes the temporal restriction of taking several exposures like when using whiskbroom, pushbroom, and area-scanning devices,

which makes this technology especially interesting for real-time applications with high spectro-temporal resolution Jung (2017). To be able to acquire a full capture in a single integration period, the sensor region is divided into multiple regions with each of the regions capturing 2D data, as described by Hagen and Kudenov (2013). The resulting data is then combined into a spectral datacube in postprocessing. Besides others, two approaches for creating a snapshot spectral imaging device are the use of a beamsplitter and the use of a detector array. Beamsplitter separate the incident light based on the wavelength. The different beams are then projected onto different sensors, recording the spectral bands. These systems are usually bulky and expensive as discussed by Lapray et al. (2014) and the number of spectral bands that can be recorded is limited in practice (Hagen and Kudenov, 2013). Designs based on detector arrays or spectral filter arrays (SFAs) on the other hand can overcome those limitations. SFAs are following the same principle as color filter arrays (CFAs) such as the Bayer pattern used in digital color cameras, where each of the different filters with unique transmission properties will correspond to one of the captured bands. The different filters are arranged in a predefined pattern, where the sensor region corresponding to each of the filters will usually record one part of the spectral range. Contrary to Bayer filter arrays, SFAs consist of a number of different filters > 3 that can be chosen to suit a specific use case, allowing them to sample spectral information rather than colorimetric information (Lapray et al., 2014). Usually, an additional array of lenslets is used in front of the filter array or between the filter array and the sensor to focus the incident radiance. The captured raw data must undergo a demosaicking process, where the sensor response is resolved to the multichannel datacube (Baone and Qi, 2006). Snapshot systems record the spectral image cubes faster than the other methods discussed. There also exist designs using linear varying bandpass filters instead of a filter array in combination with a microlens array as described by Hubold et al. (2018). The incoming radiance passes through different regions of the bandpass filter first. Afterwards, the filtered radiance passes the lens array, where each lens projects the entire field of view (FOV) at different filtered wavelengths onto different regions of the sensor. Their recording speed enables the real-time capture of spectral images and spectral videos, giving them a huge advantage over the other methods. In addition, snapshot systems are more robust due to the lack of moving parts (Descour and Dereniak, 1995) and have a much higher optical throughput (light collection) than pushbroom and whiskbroom systems, as discussed in detail by Hagen (2012). Caused by the trade-off between spatial and spectral resolution within the technology of today's snapshot systems, the quality (either spatial, spectral, or both) is usually lower than the ones of high-end systems of the other categories. Recent developments towards spatial (Hu et al., 2022) and spectral upsampling (Fotiadou et al., 2019) try to compensate for those limitations. Several

other technical approaches exist for building a snapshot hyperspectral imaging device, which are discussed in more detail in Hagen and Kudenov (2013).

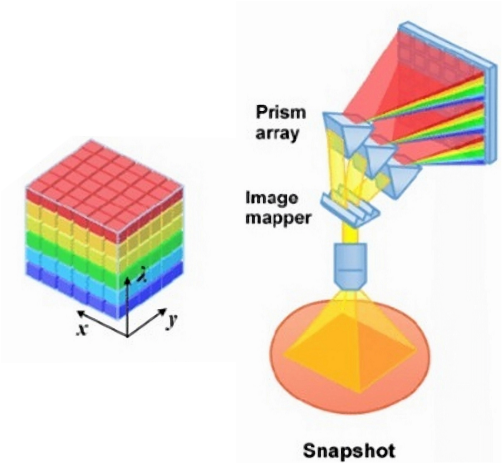


Figure 2.6: *Visualization of the working principle of a snapshot acquisition system, from simtrum.com, last accessed: 05.08.2023.*

2.1.5 Postprocessing of the Captured Spectral Data

The raw response ρ_i at channel i of a given imaging system in digital numbers (DNs) can be defined as seen in Equation 2.1, neglecting the spectral transmittance of the optical path. Hereby, $S_i(\lambda)$ denotes the spectral sensitivity at a given channel, $L(\lambda)$ represents the SPD of the illuminant and $R(\lambda)$ is the reflectance of the measured object. k is a normalizing constant, b_i represents the dark current noise and ϵ_i is the random noise. In practice, the equation is altered to work with discrete values rather than continuous functions, as shown in Equation 2.2

$$\rho_i = k \int_{\lambda_{min}}^{\lambda_{max}} S_i(\lambda) L(\lambda) R(\lambda) d\lambda + b_i + \epsilon_i \tag{2.1}$$

$$\rho_i = k \sum_{j=1}^{n} S_i(\lambda_j) L(\lambda_j) R(\lambda_j) \Delta\lambda_j + b_i + \epsilon_i \tag{2.2}$$

However, when working with multi- and hyperspectral data, the captured raw sensor response in DNs requires further processing. Although the raw response refers to the recorded incoming radiance at the system, effects such as the quantum efficiency of the sensor and other physical properties of the camera system influence

the resulting measurement (Yao and Lewis, 2010). Instead of this uncorrected radiance it is usually desired to retrieve physically meaningful information such as physical radiance or reflectance for further analysis as used for example by Liu et al. (2020) or classification applications discussed among others by Audebert et al. (2019) and Yan et al. (2019). To calculate physical radiance data ($mWcm^2sr^{-1}nm^{-1}$) including information about the scene illuminant, a radiometric calibration has to be carried out individually per camera. This is usually done by calculating a radiometric calibration coefficient for each spectral band and pixel of the camera using an integrated sphere and a standardized light source, as described in detail by Lucieer et al. (2014). After the calibration process, the errors introduced by the camera system are reduced and the captured spectral information is comparable to data measured from other calibrated spectral capturing devices.

In many cases, it is desired to convert the captured raw response to illuminant independent relative reflectance data. If so, the goal is to remove the impact of the illuminant spectrum on the obtained data completely, or at least to reduce its influence as much as possible. To obtain the illuminant invariant reflectance values, the common approach is to determine the SPD of the light source. This can be a single spectrum if it can be assumed that there is only a single light source which is spatially uniform over the whole scene captured by the sensor. In other cases, the light source might not be spatially uniform or even a mixture of different illuminants are present in the scene. If so, a single SPD might not be sufficient to represent the illumination conditions present in the scene, but rather a whole spectral cube is needed where an illumination spectrum for each pixel captured by the sensor is determined. Several main approaches exist to either measure or estimate the illuminant SPDs, which will be explained in Section 2.2. Once the illumination SPD is obtained, the raw sensor response can be corrected to reflectance values via flat-fielding. This process usually requires the measured or estimated illuminant and a dark current image and is described in detail in Section 2.2.1.

2.2 Illumination and Reflectance Estimation

In the field of illuminant estimation, much work focuses on the approximation of colorimetric information such as the illuminant chromaticity from color images Kaur and Sharma (2016). However, because the focus of this work is the estimation of illuminant spectra for reflectance recovery from radiance data, the term illumination estimation refers to the approximation of spectra rather than colorimetric information if not stated otherwise. Since this is an under-constrained problem, generally assumptions have to be made about the scene content or the acquisition environment to successfully narrow down the variety of possible solutions. A

common constraint used by many approaches is the assumption of a single spatially uniform source of illumination present in the recorded scene. While this assumption is valid for many situations where controlled conditions are possible, such single illuminant approaches might face issues when the scene is illuminated by multiple light sources or the illumination distribution is highly nonuniform.

The field of illumination estimation can be split into four main approaches. The most common is the target-based approach, which often is used as a ground truth measurement in practice. However, it is listed as an estimation approach here, since the measurement is either taken with a temporal difference in respect to the scene capture or without covering the full spatial nonuniformity of the light source. Image statistics-based approaches try to retrieve the illuminant by utilizing image-dependent information in combination with statistical assumptions. Physics-based approaches try to model the physical properties of a given scene for the illumination retrieval, whereas learning-based approaches use neural networks and large amounts of training data to fulfill their goal. In the following sections, each of the approaches is described in more detail, and existing work is discussed.

2.2.1 Target-Based (Ground Truth)

The target-based approach of retrieving the spectral reflectance values for a scene is the standard procedure to this day (Jablonski et al., 2016). Here, a standardized tile of known reflectance such as a Spectralon® is used, as described in detail by Yao and Lewis (2010). But also other, much cheaper materials such as Teflon™ can achieve good results (Koz, 2019). In any case, such a reference tile should inherit two important properties: First, its surface properties must be Lambertian. This means that the material is purely diffuse and therefore the distribution of the incident radiance which is reflected from its surface is distributed equally in all directions (Koppal, 2014). Hence, the reflected illumination is independent of the viewing angle. Second, it should be (nearly) 100% reflective and maintain a uniform reflectance over the whole spectrum of interest. To obtain a meaningful representation of the incident illumination of the whole scene that is to be captured, the reference tile should be placed approximately at the same distance as the object to be measured and at the same angle as the surface of the object to be measured. However, in many field applications where for example airborne or UAV-based measurements are taken, it is common practice to capture the reference white before and after the flight as described for example by Yang et al. (2017), Aasen et al. (2015) and Suomalainen et al. (2014). To keep track of nonuniform illumination, the reference should cover the whole spatial dimensions of the sensor. To increase the SNR, the exposure time should be set to obtain a sensor response that is as high as possible, while not introducing saturated pixels. But since the

behavior of imaging sensors becomes nonlinear when approaching the saturation point (Wang and Theuwissen, 2017) it is recommended to avoid these nonlinear regions (Aasen et al., 2015) or correct them as described for example in Czech et al. (2023). Equation 2.3 shows the flat-fielding process to obtain reflectance values R_λ from the raw camera response in DNs.

$$R_\lambda = \frac{S_\lambda - D_\lambda}{W_\lambda - D_\lambda} \times WC_\lambda \tag{2.3}$$

Where S_λ is the sample intensity of the captured scene at wavelength λ, W_λ is the measurement of the reference tile and D_λ is the dark current intensity. To measure D_λ, the sensor is covered light-tight and a capture is taken using the same integration time as for the scene recording. Yao and Lewis (2010) recommend to capture both calibration measurements W_λ and D_λ first and proceed with the scene capture afterwards. WC_λ refers to a correction factor to compensate for reference tiles with reflectance factors lower than 1.0. When using highly reflective reference targets such as a Spectralon® panel with a reflectance factor of 0.99 over the spectral range of 400 - 1500nm, WC_λ can be neglected (Yao and Lewis, 2010). Since the measured objects usually are much less reflective than the reference target, the objects of interest are often underexposed, leading to a smaller SNR. This problem can be avoided by using a gray target instead, as used for example by Khanna et al. (2017). A slightly different approach is to place the reference tile within the scene to be captured and acquire a single datacube with the reference tile close to and at the same angle as the surface of interest, as shown in Amziane et al. (2020). While no separate reference image has to be captured, the limitations are a decreased accuracy when calculating the reflectance if the assumption of a uniform illumination over the whole image does not hold, since the incident light is not captured on a per-pixel basis. In addition, the coverage of the scene by the tile reduces the spatial region available for capturing samples. As pointed out by Wendel and Underwood (2017), target-based approaches achieve accurate results in close-range settings under controlled lighting conditions but are also used in ground-based Kurz et al. (2012) and UAV-based (Abdelbaki et al., 2021) applications. Here, the disadvantages of the method become clearly visible: the data acquisition in uncontrolled environments is prone to rapid changes in illumination conditions, causing the need to recapture the reference target multiple times to maintain the desired accuracy of illumination and reflectance estimation. This process is time-consuming and usually involves manual labor, both being undesired. Although automation approaches exist as for example shown by Uto et al. (2013), they still require several reference measurements and several targets that must be manually placed beforehand and are not suited for a majority of scenarios. Other approaches such as an approach proposed by Amziane et al. (2020), where the authors permanently mount a reference target to cover a portion of the cameras

FOV can not represent the spatial nonuniformity of the light source at the sample and causes problems with vignetting. In addition, when using UAVs and airborne systems, it might not be at all possible to place a reference target properly as discussed by (Hakala et al., 2018) and also experienced by Arroyo-Mora et al. (2021), for example when the intention is to measure the tree canopy of a forest. Here, targets are often captured before the flight, which introduces inaccuracies if weather conditions change. A solution for these problems can be the estimation of illumination spectra without a reference target, using image statistics, physically- or learning-based methods, which will be described in the following sections.

2.2.2 Image Statistics-Based

Image statistics-based approaches try to make use of certain assumptions which can be made of a given scene to narrow down the number of possible estimation results. One way to estimate illuminant spectra is to adopt algorithms designed for the estimation of illuminant chromaticity of color images. They are based on the ability of the human visual system (HVS) to compensate for the influence of the surrounding illumination chromaticity when adapted to a scene. This allows humans to perceive the colors of an object in a consistent way, despite a change in the illumination conditions. This phenomenon is called color constancy and the adjustment to change in illumination itself is referred to as chromatic adaptation. It is an important part of our ability to identify objects (Smithson, 2005), since natural illumination conditions shift considerably throughout the day and with different weather conditions, as proven by Judd et al. (1964), Hernández-Andrés et al. (1999),Romero et al. (2002) and others. The mathematical modeling of this ability is called computational color constancy (CCC) and is of great interest in image processing and computer vision Krüger et al. (2013). The mechanisms behind this phenomenon and possible mathematical representations are described comprehensively by Ebner (2007), yet this topic is still an area of active research as shown for example by Seymour (2022) and discussed by Maule et al. (2023) and Foster (2011).

Widely used approaches to model CCC are the grayworld algorithm (Buchsbaum, 1980), the retinex theory or max-RGB algorithm (Land and McCann, 1971), the shades-of-gray algorithm (Finlayson and Trezzi, 2004) and the gray-edge algorithm (van de Weijer and Gevers, 2005). All of the aforementioned algorithms were expanded to work in the sensor domain for n-channels and therefore to estimate illuminant spectra instead of illuminant chromaticity by Khan et al. (2017c) with the goal to achieve an illuminant invariant representation of spectral data Khan et al. (2018). The gray-world algorithm assumes that the average response of a natural scene will result in a medium gray shade. Therefore the deviation of the mean

value per channel from that achromatic reference inherits information about the illuminant. This estimation method is computationally fast and simple to implement but also has major drawbacks, such as its high dependency on the image content. Scenes where the average reflectance is achromatic either require only objects with achromatic reflectance properties or a broad variety of different reflectances to be present. Both is not likely for many applications, such as agriculture or food quality assessment, and therefore the results are often inaccurate. Besides, several other algorithms for illuminant spectra estimation were adopted by Khan et al. (2017b) from chromaticity estimation algorithms used in the colorimetric domain, which will be presented in the following paragraph. The max-spectral algorithm for example is based on the max-RGB algorithm following the concept of Land and McCann (1971), where the authors state that the cone responses of the HVS built images for long, medium, and short wavelength regions each. The theory describes that these three channels are independent and are compared to each other by the HVS to generate color sensations. Now, by assuming that the reflectance properties of the brightest surface within a given scene usually correspond to a surface perceived as white, information about the light source can be extracted by determining the maximum value for each color channel. Transferring this concept to the spectral domain, the maximum response at each sampled wavelength will form the estimated SPD of the light source. Finlayson and Trezzi (2004) proposed an algorithm called shades-of-gray, which is shown in Equation 2.4, where F is the input data, p is the Minkowski norm, k is a constant and $\mathbf{e}$ is the estimated illuminant. The authors suggest a value of $p = 6$, as they found that for larger p-values, more weight is given to pixels with higher values. This equation can also represent the aforementioned grayworld for $p = 1$ and max-RGB algorithms for $p = \infty$ and their spectral counterparts by replacing the averaging operation with Equation 2.4.

$$\left(\frac{\int F^p dx}{\int dx} \right)^{1/p} = k\mathbf{e} \tag{2.4}$$

The gray-edge algorithm assumes that the average reflectance of the computed edges is achromatic for RGB cases as discussed by van de Weijer and Gevers (2005). First, each channel is smoothed using a Gaussian filter ($\sigma = 2$) first. Then, the edges are extracted computing the derivative in x and y direction. The average at each channel describes then the illuminant at each channel. As proposed by Finlayson and Trezzi (2004), the averaging operation can be replaced by the Minkowski norm. Here F_σ represents the smoothed input data, $p = 6$ is the order of the Minkowski norm, k is a constant between 0 (no reflectance) and 1 (total reflectance) and $\mathbf{e}$ is the resulting illuminant.

$$\left(\frac{\int \mid F_\sigma \mid^p dx}{\int dx} \right)^{1/p} = k\mathbf{e} \tag{2.5}$$

Khan et al. (2017c) found, that of the described algorithms max-spectra and spectral gray-edge gave the best results since both algorithms use bright pixels for the illuminant estimation.

Another approach was taken by Fredembach and Finlayson (2008). The authors used a six-channel system with known properties of the used filters. The transformation matrices for the sensor response to each SPD of an illuminant database were calculated and the illuminant SPD with the best fitting transformation was determined as the scene illuminant. The main drawback of this approach is the high dependency on the illuminant database used. In addition, for this approach to work the transmissive properties of the filters have to be known, which might not be feasible for some applications.

2.2.3 Physically-Based

Physically-based algorithms try to make use of the physical properties of a given scene to extract information about the present illumination conditions. A common approach is to estimate the emission spectrum of a given light source by examining the specular reflections it causes on the surfaces of objects in the scene. Thereby, it is assumed that specular highlights inherit information about the SPD of the source of illumination, while diffuse components mostly contain information about the spectral properties of the given surface material. In general, the radiance reflected by a dielectric material such as plastic can be expressed by using a basic dichromatic reflectance model as described by Shafer (1985) and Tominaga (1995). Here, the total reflected radiance $Y(\theta, \lambda)$ from a dielectric material illuminated by a light source equals a linear combination of its high-intensity interface (or specular) reflection $L_I(\lambda)$ and its lower intensity body (or diffuse) reflection $L_B(\lambda)$ components as determined in Equation 2.6. θ denotes the geometric parameters for the viewing angle, the angle of the incident illumination, and the phase angle. $L_I(\lambda)$ and $L_B(\lambda)$ are independent of the geometric angle. $c_I(\theta)$ and $c_B(\theta)$ are the respective geometric sale factors. An illustration is given in Figure 2.7.

$$Y(\theta, \lambda) = c_I(\theta)L_I(\lambda) + c_B(\theta)L_B(\lambda) \tag{2.6}$$

Illuminant estimation approaches based on dichromatic reflectance models usually divide the given hyperspectral image into specular and nonspecular regions first. To detect the highlight areas, different methods can be used, for example by assuming that the brightest pixels in channels of overall low intensity correspond

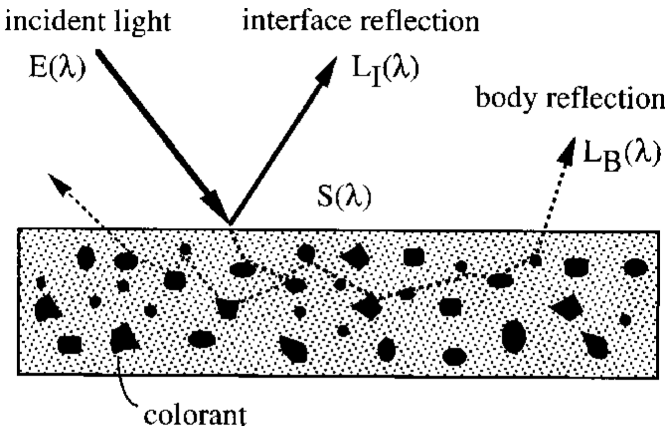


Figure 2.7: *Visualization of a dichromatic reflectance model, from Tominaga (1995).*

to specularity (An et al., 2015). A more advanced approach includes using HVS inspired methods such as the modeling of receptive fields (Goldstein, 2010) via convolution with two Gaussian distributions, as described by Tominaga et al. (2020). Once the highlight areas are detected, there are several methods to estimate the SPD of the light source. Tominaga et al. (2020) used a clustering approach of the spectral data after projecting it into a subspace of the first and second principal components and obtained the SPD by using an inverse. An et al. (2015) used an optimization framework based on several constraints as well as a dictionary of corresponding specular and diffuse surface areas of the materials in the scene to guide the estimation. Another approach is to identify objects with 'flat' spectral reflectance properties in a given scene and base the illumination estimation on the reflected lights coming from those, as proposed by Banerjee et al. (2009).

Dependent on the desired application, a major drawback of many physically-based models is their limitation to perform well only when used with surfaces that inherit dichromatic reflection properties, as in Tominaga et al. (2020) and An et al. (2015). Further, some models do not work without objects that inherit surface properties close to reference targets and are partially shaded and illuminated by sunlight at the same time (Banerjee et al., 2009), while others fail when used with monochromatic materials (An et al., 2015) or assume only convex surfaces (Tominaga et al., 2020). While most approaches only consider a single light source, algorithms also exist which try to compensate for multiple illuminants per scene such as Tominaga et al. (2020) and Tominaga et al. (2012), limited to the VIS-range. Estimation methods based on specular highlights, suffer from data captured by sensors with limited spectral range since setting the integration time for a high SNR might cause the specular highlights to be clipped, which means a significant

loss of information within the highlights in at least some of the channels. Another challenge is the proper detection of the specular highlights, which is highly based on constraints and assumptions and still an open topic of research Khan et al. (2017a). In addition, due to their dependency on highlights or flat surface reflectances, these algorithms struggle to estimate illuminants for scenes that predominantly consist of vegetation.

2.2.4 Learning-Based

The use of neural networks for the estimation of illuminants in the colorimetric domain was studied by many over the last two decades, as summarized by Sethu et al. (2023). In the field of hyperspectral imaging however, most of the machine learning-based applications are focused on classification tasks as discussed by Khan et al. (2022) and Ozdemir et al. (2020) or spectral unmixing, which describes the determination of spectral end-members within a scene and their proportional concentration (Grillini et al., 2021). Not much work exists on learning-based methods for the estimation of illuminant spectra from hyperspectral radiance data. Most likely, the reason for that is the massive amount of training data needed for the training of neural networks. In contrary to the amount of available datasets containing RGB-images that can be used for training and evaluating illuminant estimation networks (Gijsenij et al., 2011), there is no comparable amount of spectral data available including high variation and available ground truth measurements for the illuminant. Another reason might be the lack of standardization in the field of spectral imaging. Existing datasets show a high variation in acquisition techniques and calibration procedures, resulting in large differences in terms of the spectral range, as well as the spectral and spatial resolution. Robles-Kelly and Wei (2018) offering a combination of colorimetric and spectral estimation of the scene illuminant at a pixel-level basis. The main idea is to fine-tune a pre-trained convolutional neural network (CNN) similar in architecture to the model proposed by Snoek et al. (2012). The model consists of five convolution layers with pooling operations after the first three layers and rectifier linear units after each of the convolution layers. Randomly chosen tensors constructed from uniformly sampled local patches from the spectral cube at different scales computed via the use of a Gaussian kernel are used as input for the model. The size of each patch was determined with a spatial resolution of 10x10 pixels and the SPD at the center of each patch served as ground truth value. The model was then trained to predict the ground truth illuminant at the center of each patch. The authors state that the model is restricted to detecting "smooth spectra" (Robles-Kelly and Wei, 2018), irregular or "spiky" light sources like LEDs are not considered for the prediction. In addition, the spectral range of the estimated illuminants is restricted to the

VIS-range, in this case specifically to a range of 400 - 700nm with a spectral resolution of 10nm. It also has to be noted that the main focus of the authors lies on the colorimetric estimation of the illuminant and less on the spectral counterpart, which hinders an objective interpretation of the model quality in terms of the prediction of illuminant spectra.

3 METHODOLOGY

In this work, a novel framework is proposed for estimating daylight illumination spectra and calculating relative reflectance data directly from a hyperspectral radiance cube. Thereby no reference or ground truth target is needed for calibration as described previously in Section 2.2.1. Section 3.1 describes the proposed image statistics-based model itself as well as the general assumptions that were made to solve the under-constrained problem of illumination estimation in detail. Section 3.2 focuses on the evaluation of the experiments done.

3.1 The proposed Framework

The main idea developed in this work is to estimate the spectral power distribution of an unknown light source from a hyperspectral cube by using general assumptions about the input data to restrict the possible estimations from two perspectives: On one side, two image-dependent constraints define an area where the unknown illuminant spectrum is likely to be found in the spectral domain. Within this region, a set of random spectra are generated. They are referred to as *spectral candidates*. These spectral candidates define the variation of image-dependent plausible spectra which could represent the illuminant within the given hyperspectral cube. On the other side, the estimation result is restricted by a database of measured real-world illuminants. This data represents a set of physically possible illuminant spectra. Both restrictions are used to find a point in an image-independent PCS, where the image-dependent plausible candidates meet the physically possible measured illuminants. This intersection is then used to generate the estimated SPD by transforming back from the PCS to the spectral domain. The resulting estimated SPD can then be used to calculate reflectance values from the radiance input data. Although a physically possible set of illuminants is used to find the estimated illuminant spectra, the proposed algorithm for the illuminant estimation is classified as image statistics-based, since the illumination mostly depends on the generated spectral candidates.

In the following sections, the proposed model is explained in detail. At first, the

constraints for the model are defined and the generation of the image-dependent spectral candidates is described. Afterwards, the transformation to a common PCS to a broad range of training data is explained. Then, the illuminant estimation itself is explained in detail. Afterwards, the calculation of the resulting reflectance data is shown. At last, a summary of the steps is given for the convenience of the reader.

3.1.1 Assumptions and Constraints

Since the problem of illumination spectra estimation from radiance data without ground truth is an under-constrained problem, well-defined assumptions have to be made to obtain results that are accurate enough to justify their use over a ground truth-based approach. For the image statistics-based framework proposed in this work, uniform and diffuse illumination properties over the whole spatial domain of the captured scene is assumed. The focus lies on estimating daylight spectra for two main reasons: First, a special interest is taken in designing an illumination estimation approach for UAV-based applications. When working in outdoor environments within uncontrolled weather conditions, the capture of reference targets might introduce errors due to rapid changes in cloud cover and lighting conditions (Wendel and Underwood, 2017). In such time-critical scenarios, the use of estimation algorithms seems to be sufficient if the illuminant can be accurately recovered from within the captured scene data. In addition, the placement of delicate reference targets within the scene directly might not be sufficient as pointed out by Wendel and Underwood (2017), for example when monitoring forests or large agricultural areas. Laboratory-based measurements on the other hand do not have these restrictions, which makes the use of a target-based approach less challenging. Second, the availability of reference data of daylight illuminant spectra was several times larger in size and variation than other kinds, such as LEDs and incandescent light sources at the time this work is proposed. Nevertheless, the proposed framework is assumed to be adjustable in the future to estimate other light sources by using a different reference database for the estimation. Further, it is assumed that the captured scene does not contain objects which emit radiance themselves. Finally, the input data for the framework is assumed to be radiance data normalized to a range between 0 and 1. With the aforementioned assumptions in mind, several main constraints are defined for the framework.

1. Since the input data is normalized, the theoretically possible SPD with the highest maximum value at each wavelength is a spectrum of unity over the whole spectral range. This upper limit will be referred to as *maxConstraint*.

2. The lower limit of the theoretically possible SPD can be defined as the maximum value of the spectral cube at each wavelength of the input data. This constraint will from now on referred to as *minConstraint*.

3. The unknown SPD of the light source present in the radiance input data must therefore lie somewhere within an area defined by the *minConstraint* and *maxConstraint*.

3.1.2 Calculation of Constraints

To use the given information of the radiance input data to derive information about the illuminant, a simple but effective constraint was chosen. Given that the radiance input data is normalized to a range of 0 to 1, the theoretical upper limit for the illumination at any wavelength would be unity. Therefore, the maximum spectrum over the whole spectral range investigated can be described as a horizontal line with $y = 1.0$. This *maxConstraint* builds the upper limit of the search area for the unknown illuminant, as shown in Figure 3.1. Other maxConstraints were investigated, such as the maximum value at each wavelength of a given dataset of different illuminants. Although this would further restrict the search area for the illuminant, the idea was neglected to prevent any overfitting and instead use only the information available within the input data at this stage of the framework and not introduce external information from other datasets or measurements yet.

Next, to restrict the search area for the illuminant SPD further, a lower bound called the *minConstraint* was calculated by introducing a simple yet effective constraint: The incoming radiance at one point (x, y) at the sensor can be simplified as the product of the spectral power distribution of a given light source S_λ and the reflectance properties of the object present at the spatial coordinate R_λ as shown in Equation 3.1. Here, it is assumed that the radiance data is already corrected for dark current noise, vignetting, and the sensor sensitivities.

$$L_\lambda(x, y) = S_\lambda(x, y)R_\lambda(x, y) \tag{3.1}$$

Since the maximum response possible at each wavelength is 1.0 and the measured objects are assumed to be reflective and therefore only absorb incoming radiance but do not work as an emitter, the maximum value of the spectral cube at each wavelength band is the minimum possible value of the illuminant SPD. Therefore, the *minConstraint* C_{min} of the image-dependent area of plausible illuminants can be defined as an equivalent to the retinex theory proposed by Land and McCann (1971). Equation 3.2 shows the computation of the *minContraint* for a given channel.

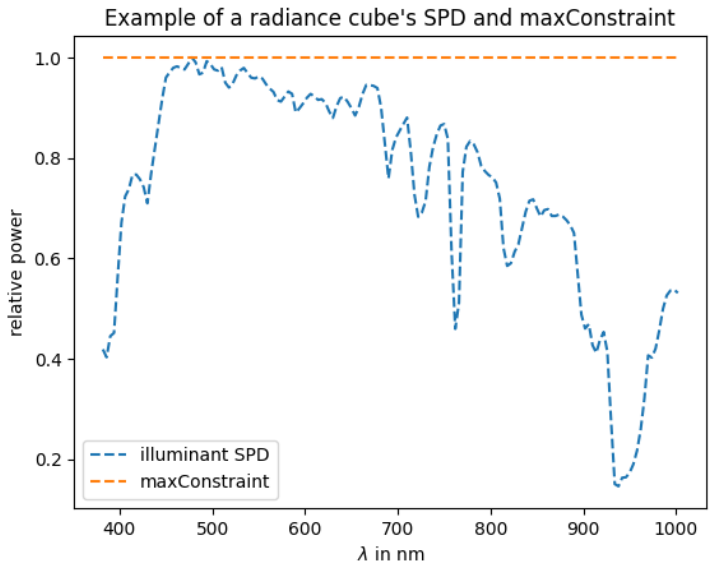


Figure 3.1: *Example of the maxConstraint as a spectrum of unity and an example SPD, normalized to a range between 0 and 1.*

$$C_{min}(\lambda) = \max\{L(\lambda)\} \tag{3.2}$$

It can be observed in Figure 3.2 that in this particular example, the *minConstraint* models the illuminant SPD fairly well by itself in the range of roughly 750 - 1000nm, but can not represent the shape for shorter wavelengths. The reason for that is the radiance cube which was used as input data for this visualization. Since the information about the constraint is derived from the image data directly, the quality of the *minConstraint* is highly dependent on a large variation of different reflectance spectra within the given scene.

Although a result as close as possible to the original illuminant would be preferable since this would restrict the search area as much as possible, a perfect representation would indicate the presence of a reference target or an object with similar properties within the scene, making the use of a complex framework for illuminant estimation redundant. Instead, the purpose of this constraint is not to work as an illuminant estimator itself but rather to be utilized as a restriction of the search area. This can be seen in Figure 3.3, where the *minConstraint* and *maxConstraint* define the search area over the spectral range of the input data. Per definition, the unknown illuminant SPD lies within this search area, which was the case for all tested scenarios. This area will be utilized in the next step.

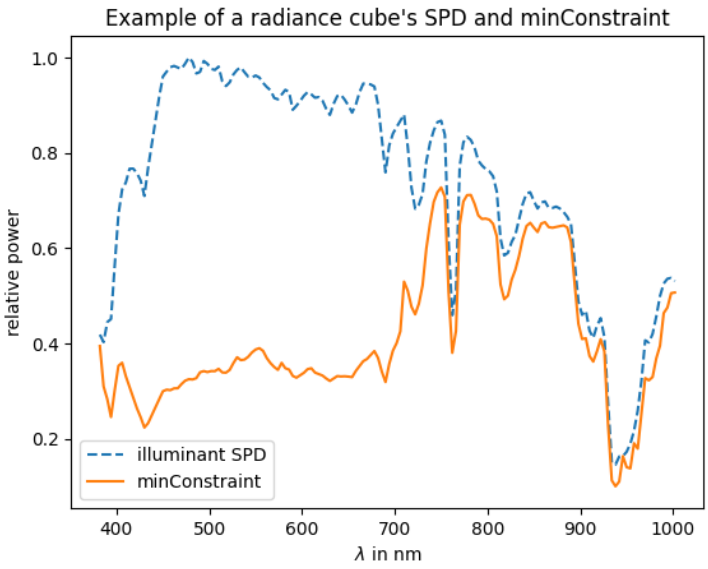


Figure 3.2: *Example of the minConstraint as maximum value of the cube at each wavelength and the original SPD, normalized to a range between 0 and 1.*

Figure 3.3: *minConstraint and maxConstraint define the search area for the unknown illuminant SPD.*

3.1.3 Calculation of the Spectral Candidates

With the upper and lower bound of the search area set, a series of random spectra in $\lambda_{min} \leq x \leq \lambda_{max}$ are generated, where the value y_i at each x_i lies withing the search

area defined by corresponding Y-values of the *minConstraint* and *maxConstraint*. To generate the spectral candidates, a modified version of Pearson's random walk algorithm (Pearson, 1905) with equal stepsize and variable amplitude is used. First, the Y-value of the starting point (x_1, y_1) is defined by a random number $0 \leq j \leq 1$. The values y_i of the remaining samples $x_i \leq x_n$ are then found by generating a different random number $0 \leq m_i \leq 1$ at each step, which determines the direction and amount of movement on the Y-axis compared to the previous sample. If $m \geq 0.5$, the resulting value y_i will be higher than its predecessor, resulting in a positive slope within the generated spectrum. Otherwise, the generated value will be lower, resulting in a negative slope of the generated spectrum. The result is a randomly generated "rough" spectrum in the desired spectral range with intensities between 0 and 1. Pseudocode for the generation of such a rough spectral candidate is given in Algorithm 1.

Algorithm 1 Pseudocode: Generation of spectral candidates

```
INPUT
minConstraint
spectralCandidate as empty list

PROCESS create spectralCandidate
value of first channel = random number between minConstraint_1 and 1
for all other channels of the image cube do
    create random number r_amount between 0 and 1        ▷ amount of movement
    create random number r_direction between 0 and 1     ▷ direction of movement
    if r_direction ≤ 0.5 then
        value at current channel = value at previous channel+r_amount
    else
        value at current channel = value at previous channel−r_amount
    end if
    append value of current channel to spectralCandidate
end for
return spectralCandidate
```

Once this initial spectrum is generated, a smoothing operation is carried out to ensure a more natural shape of the resulting spectral candidate. Based on the findings of Kosztyán and Schanda (2012), the locally weighted scatter-plot smoothing (LOWESS) first introduced by Cleveland (1979) is suitable for the smoothing of daylight spectra. Hereby, it offers locally adapting smoothing properties, robustness against outliers, and a good smoothing performance that can be adjusted using a single parameter. LOWESS is based on robust locally weighted regression. For the X-value x_i for each point (x_i, y_i) of the dataset, an initial weight w_k is assigned

to each of the remaining points x_k of the dataset with a parameter f determining the size of the neighborhood. The further away the abscissas x_k are from x_i, the smaller their weights. Now, a locally weighted regression is carried out, where weighted least squares are used to compute the initial smoothed values y'_i for each x_i using the previously calculated weights w_k. Afterwards, the residuals $y_i - y'_i$ are computed for each point. These residuals are now used as a second set of weights δ_i, where small residuals will cause larger δ_i than large residuals. The computation of all y'_i is performed repeatedly n times, using the new weights δ_i instead of w_k.

Since several hundreds of these spectral candidates are generated per input cube $n = 3$ was chosen to limit the computational cost for the smoothing operation while maintaining a good overall shape of the generated spectra. To assure that all generated values lie between the *minConstraint* and the *maxConstraint* but maintain their relative proportions, a normalization process is carried out at each x_i after the smoothing process using Equation 3.3. Some examples of randomly generated spectra can be seen in Figure 3.4. A whole set of hundreds of such generated spectra is generated for a given input cube.

$$y_i^{'} = \frac{y_i - y_{min}}{maxConstraint_i - minConstraint_i} + minConstraint_i \tag{3.3}$$

It has to be noted that the focus lies not on finding the actual illuminant by chance in one of the generated spectra. These random spectral candidates represent the image-dependent possible illuminant spectra, which are used further to estimate the unknown illuminant within an image-independent, multidimensional subspace. The creation of this subspace using principal component analysis (PCA) will be the main focus of the next section.

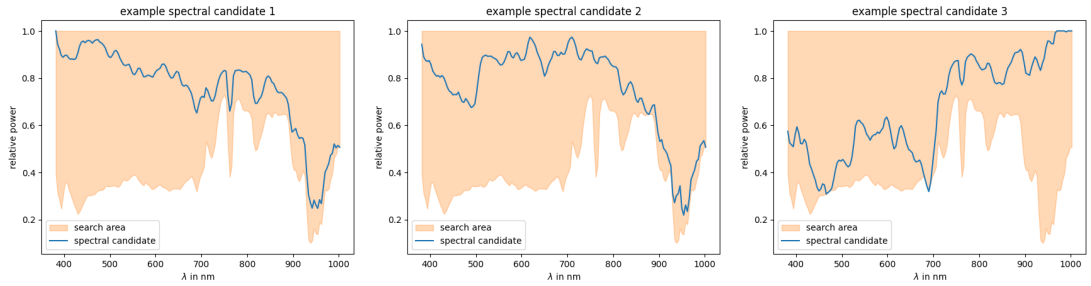


Figure 3.4: *Examples of randomly generated spectral candidates within the search area.*

3.1.4 Principal Component Analysis

The following sections focus on giving an overview of principal component analysis (PCA) with focus on the creation of an input-independent subspace in which the estimation of a scene illuminant from hyperspectral radiance data is performed. At first, Section 3.1.4.1 will summarize the general idea of what PCA is and how it is used to reduce the dimensionality of a given dataset while preserving most of the relevant data. Afterwards, Section 3.1.4.2 focuses on how an input-independent principal component space PCS can be created. This input-independent PCS will then serve as a multidimensional space for the illuminant estimation, which will be explained in detail in Section 3.1.5. A more in-depth explanation about PCA in general can be found in Principal Component Analysis, Second Edition by Jolliffe (2002).

3.1.4.1 Dimensionality Reduction using Principal Component Analysis

Principal component analysis is a widely used technique to reduce the dimensionality of a given dataset while preserving most of its meaningful information in a least-square sense. It is used in many fields and different applications and is of particular interest in the spectral imaging domain, since it can be used for compressing the large amount of data that often has to be saved and processed when working with hyperspectral images (Du and Fowler, 2007), or for analysis (Yang et al., 2018). This section focuses on summarizing the fundamentals of PCA while the following section will explain how it was used to create an input-independent multidimensional space for the illuminant estimation.

The general idea of PCA is to project the given data onto a new multidimensional orthogonal coordinate system. The axes of this new coordinate system, the principal component space, are sorted by decreasing variance of the data represented at these axes. It can therefore be assumed that the data along the first axis represents more of the important information about the input data as the second one and so on (Richards, 2022). A visual example is given in Figure 3.5.

In reverse, this means that most of the data variation can be preserved when performing an inverse transformation (or reconstruction) only using the first few principal components, with 6 components required for accurate spectral reproductions as investigated by Hernández-Andrés et al. (2001a).

Before performing the PCA itself, the first step is to standardize the input data. In particular, each input sample is centered to a mean value of 0 and scaled with respect to unit variance. This process is called z-transform and ensures that all features contribute equally to the PCA. The formula to perform the standardization is given by Equation 3.4, where x is an input sample, u denotes the mean of the whole

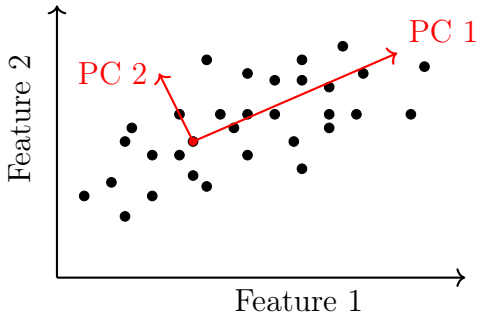


Figure 3.5: *Visualization of the first two principal components as new orthogonal axes for a given set of input data. Most of the variance of the data is captured by PC 1. The variance accounted for decreases with each additional PC.*

dataset, s represents the standard deviation and z is the resulting standardized sample. The centering ensures that the data will be distributed around the origin, and the offset of the data from the center is removed.

$$z = \frac{(x - u)}{s} \tag{3.4}$$

Next, a singular value decomposition (SVD) is performed on the covariance matrix of the input data. Let $\mathbf{M}$ be a $m \times n$ matrix, where m represents the number of spectral samples and n represents the number of features, with the features being the measured points of the spectral range between λ_{min} and λ_{max} at an interval of λ_{Δ}. SVD decomposes $\mathbf{M}$ into three matrices as shown in Equation 3.5, where $\mathbf{S}$ contains the explained variance, or singular values, within its diagonal in decreasing order. $\mathbf{U}$ and $\mathbf{V}^T$ are orthonormal matrices, meaning the result of $\mathbf{U}\mathbf{U}^T = \mathbf{I}$ and $\mathbf{V}\mathbf{V}^T = \mathbf{I}$, where $\mathbf{I}$ denotes an identity matrix. They contain the left singular vectors and right singular vectors respectively, with the left singular vectors $\mathbf{U}$ representing the principal components of the data.

$$\mathbf{M}_{m \times n} = \mathbf{U}_{m \times m}\mathbf{S}_{m \times n}\mathbf{V^T}_{n \times n} \tag{3.5}$$

The explained variance ratio σ_i for a given principal component $\mathbf{U}_i$ can be calculated by using Equation 3.6. σ inherits information about how much variance of the data is explained by each principal component in particular. By sorting $\mathbf{S}$ and $\mathbf{U}$ by the explained variance ratio σ in descending order, a ranking from the most important to the least important principal component is performed.

$$\sigma_i = \frac{\mathbf{S}_i^2}{\sum \mathbf{S}^2} \tag{3.6}$$

To transform the input data $\mathbf{M}$ into the PCS, Equation 3.7 is used, where $\cdot$ represents the dot product. The inverse transform for the scaled reconstruction

from PCS back to the scaled data is given by Equation 3.8. After performing an inverse scaling on $\mathbf{M_{rec}}$, the final reconstructed values of the input data are obtained.

$$\mathbf{M}^{'} = \mathbf{M} \cdot \mathbf{U} \tag{3.7}$$

$$\mathbf{M_{rec}} = \mathbf{M}^{'} \cdot \mathbf{U}^{\mathbf{T}} \tag{3.8}$$

3.1.4.2 Fitting an Input-independent Principal Component Space

For the illuminant estimation approach in this work, PCA was used to generate an input-independent subspace in which the illuminant estimation is performed. This has two advantages: First, when transforming an illuminant SPD into the previously fitted PCS, it is geometrically represented by a n-dimensional point, where n denotes the number of discrete wavelengths that define the spectrum. This is a key concept for the estimation approach explained in Section 3.1.5. Second, the use of a PCS allows to perform a dimensionality reduction of the estimated illuminant at the same step when converting back from the principal component (PC) domain to the spectral domain. This leads to more robustness against noise within the data, which is assumed to be of more influence in the lower-ranked components.

To create the input independent PCS, a representative set of radiance training data was simulated using a set of varying reflectance spectra in combination with a set of measured illuminant SPDs as described in Section 3.2.1. Both datasets were first cropped and linearly interpolated to a common spectral range of the input cubes from which the illuminant information is later estimated. Here, it is preferred to preserve the range and interval of the radiance input cubes, if possible, and rather adjust the ranges and interval of the reflectance and illuminant dataset. It is important that all ranges and sampling intervals match exactly, otherwise the transformation to a common PCS will not give the desired results. Once a common spectral range for all components is found, Equation 3.1 is used to simulate radiance data out of the reflectance and illuminant sets. Then, a normalization is performed to obtain relative spectral radiance in the range of $[0-1]$. The resulting data is from now on referred to as the training dataset, which represents a variety of different Lambertian surfaces under a broad amount of different daylight illuminants. The training dataset is centered and scaled, using Equation 3.4 and a PCS is fitted for the scaled data training, using SVD as previously described in Section 3.1.4.1.

3.1.5 Illuminant Estimation in the Principal Component Space

Now, the relevant data for the illuminant estimation is transformed into the common subspace by applying the same centering and scaling as well as the same transformation using 3.7, where U are the principal components already calculated for the input independent space. The relevant data consists of two parts: A dataset of physically possible illuminant spectra such as the preprocessed Granada daylight spectral database (Section 3.2.1.1), as well as the input-dependent plausible spectral candidates calculated as described in Section 3.1.3. Once converted into principal component space, each spectrum is represented as a multidimensional point. The goal is now to estimate the illuminant in the PCS by calculating the intersection of the physically possible illuminants and the image-dependent plausible region defined by the spectral candidates, as described in the following sections.

3.1.5.1 Fitting a Line to the Illuminant Dataset using RANSAC

After visually examining as in Figure 3.6, it can be assumed that the distribution of the points within the illuminant dataset can be approximated by a line.

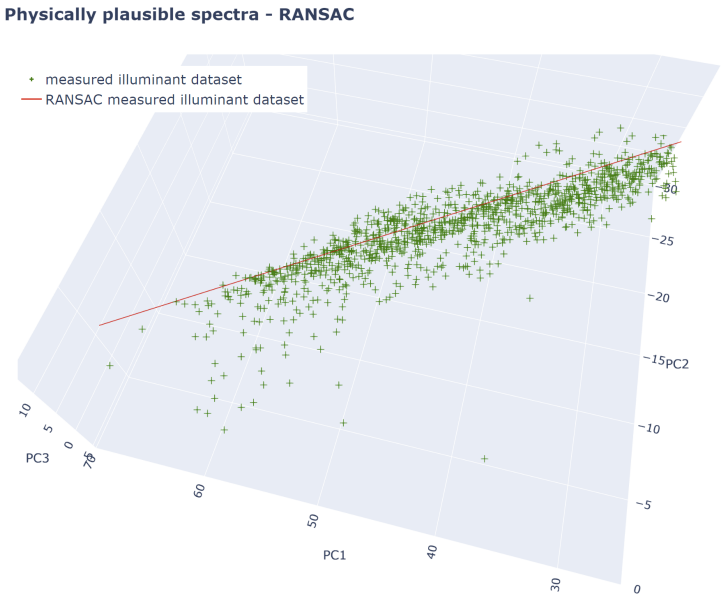


Figure 3.6: *The general trend of the illuminant dataset for the first principal components can be represented by fitting a line. Each green dot represents a SPD of the Granada daylight spectral database. The red line was fitted using RANSAC.*

This property can be utilized by fitting a line that represents a majority of the physically possible daylight illuminants. To do so, the well-established and widely used random sample consensus (RANSAC) algorithm developed by Fischler and Bolles (1981) is used, since it is robust against outliers and does not add significant complexity. Two random points of the dataset are picked and a linear model is fitted to them. Then, the distance for each point of the dataset to the current model is calculated. All points of the dataset below the residual threshold, which is set to 0.5 for the calculations of this work, were chosen as inliers. This process is now repeated for a set number of iterations, 15.000 in this case, and the model which produced the most inliers is chosen as the final one. This fitted linear model now represents the illuminant database in the image-independent PCS. The general idea of RANSAC that was used in this work is further expressed in the form of pseudocode in Algorithm 2.

Algorithm 2 Pseudocode: RANSAC

```
INPUT
data as a set of points
maxIterations > 0
maximumInliersRequired > 0

PROCESS RANSAC previousModel = empty
while iterations < maxIterations do
    currentModel (Line) = two random points from data
    for all other points in data do
        calculate distanceToLine
        if distanceToLine < threshold then
            increase numberInliers by 1
        end if
    end for
    if numberInliers >= maximumInliersRequired then
        bestModel = currentModel
    end if
end while
if numberInliers > previousnumberInliers then
    finalModel = currentModel
else
    finalModel = previousModel
end if
return finalModel
```

3.1.5.2 Fitting a Hyperplane to represent the Spectral Candidates

Once the n-dimensional line is fitted to the illuminant dataset, the spectral candidates are calculated as described in Section 3.1.3 and converted into the common PCS. Figure 3.7 shows the ground truth illuminant within the region where the illuminant data and the spectral candidates intersect.

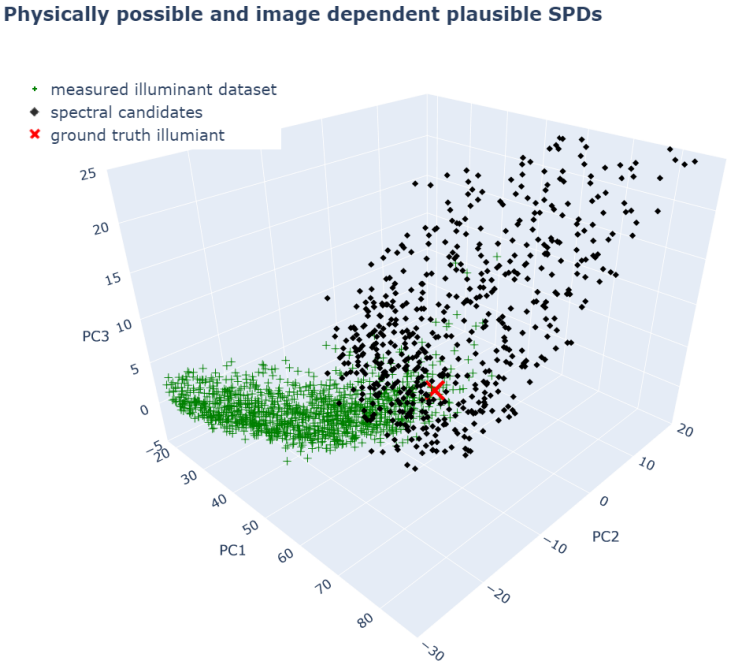


Figure 3.7: *The physically possible illuminants (green dots) meet the image-dependent possible candidates (black squares) within the input independent PCS. The red 'X' lies in the region where both sets intersect and marks the ground truth illuminant.*

The spectral candidates are now represented by a hyperplane (Wang, 2012). The hyperplane hereby serves as a simplified representation of the image-dependent 'plausible' spectral candidates. This step is necessary to be able to estimate the illuminant as a single, n-dimensional point of intersection between the illuminant dataset represented by the n-dimensional line and the spectral candidates represented by the hyperplane as described below.

The hyperplane of the given set of points can be expressed by its centroid C and its normal vector N. Let matrix M of shape $m \times n$ be defined by a set of m input points in n dimensions, in this case, the representation of the spectral candidates as points in the PCS. The centroid C is defined as a n-dimensional

point, which is calculated by the mean value of all m points of the dataset in each dimension (Equation 3.9).

$$C = \frac{1}{N} \sum_{i=1}^{N} v_i \tag{3.9}$$

To calculate the normal vector N, SVD is used once again (see Section 3.1.4.1). First, M is decomposed into the three matrices U, S, and V^T. The normal N of the hyperplane can now be defined as the last unitary vector of V^T. Hereby, the last unitary vector is determined by finding the vector of V^T which corresponds to the smallest singular value of the diagonal matrix S. It represents - in a least squares sense - the direction with the least variation for the set of input points in M.

3.1.5.3 Calculating the Estimated Illuminant as Intersection Point in PCS

A rough illuminant estimation in the fitted PCS can now be achieved by calculating the intersection between the line representing the physically possible data and the hyperplane representing the image-dependent plausible candidates.

To calculate the intersection point, the hyperplane is represented by its centroid C and its normal N as discussed previously. The line, defined in parameterized form by two points P_1, P_2 in n-dimensions found via RANSAC, and parameter t, is shown in Equation 3.10, where $P(t)$ represents a point on the line. A given point $P(t)$ of the line lies on the plane if Equation 3.11 is satisfied, with $\cdot$ representing the dot product. By substituting the line Equation 3.10 into the equation of the hyperplane (Equation 3.11), Equation 3.12 is obtained. It is possible to solve for t as shown in Equation 3.13. Substituting t back into Equation 3.10 and solving for $P(t)$ now gives intersection point P_{est}, which represents the estimated illuminant.

$$P(t) = t * P_1 + (1 - t) * P_2 \tag{3.10}$$

$$N \cdot (P(t) - C) = 0 \tag{3.11}$$

$$N \cdot ((t * P_1 + (1 - t) * P_2) - C) = 0 \tag{3.12}$$

$$t = (N \cdot C - N \cdot P_2)/(N \cdot P_1 - N \cdot P_2) \tag{3.13}$$

The process is visualized in Figure 3.8, where the red line represents the illuminant dataset, the plane represents the spectral candidates, the red cross

marks the intersection point between both, as calculated using Equations 3.10 - 3.13, while the blue dot represents the ground truth illuminant. This solution works well, particularly in the first three dimensions.

The obtained rough estimation of the input cubes illuminant is now validated, refined and transformed back into the spectral domain as described below.

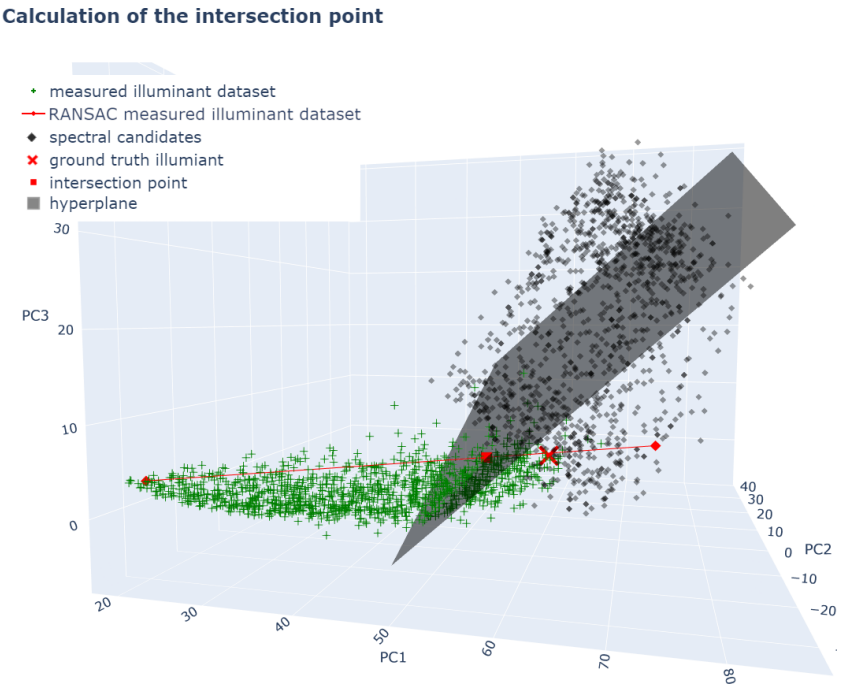
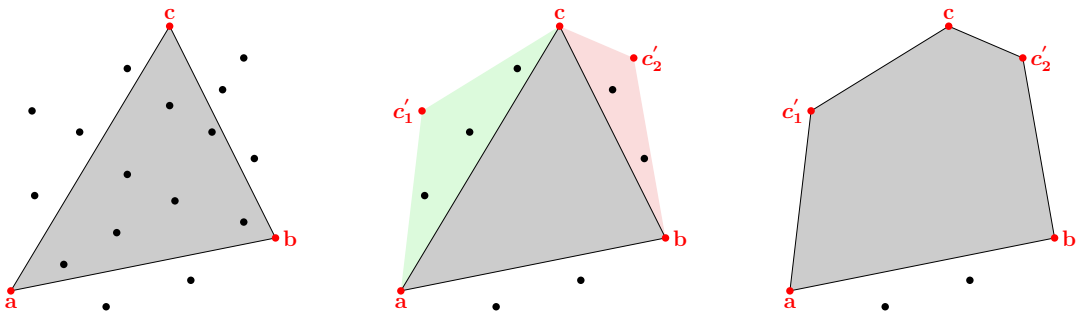


Figure 3.8: *The red line is fitted using RANSAC on the illuminant dataset (green dots), the plane representing and the the spectral candidates are shown in gray. The intersection point is marked as red square, whereas the ground truth is shown as red 'X'.*

3.1.5.4 Validation and Refinement of the Estimated Illuminant

The next step is to ensure that the intersection point which represents the estimated illuminant lies within a valid region of the PCS. To do so, a n-dimensional convex hull is created, enclosing a volume that represents the intersection $A \cap B$ for the illuminant training set and the set of spectral candidates. It is then ensured that the estimated illuminant lies inside that convex hull.

A convex hull of a set of points is defined as the smallest convex shape which encloses all points of the set (Chadnov and Skvortsov, 2004). Different algorithms



(a) *Initial hull as triangle defined by line ab and the farthest point c. All points inside abc are discarded.*

(b) *Repeating the process for subset S_1' (points to the left side of ac) and S_2' (points to the right side of bc).*

(c) *Hull after the first iteration. This process is repeated iteratively until all points of the set are addressed.*

Figure 3.9: *Determination of the convex hull H of a set of points S using the Quickhull algorithm in two dimensions, where the hull is represented by the black lines, the points that represent the hull are marked as red dots and the area inside the hull is shaded in gray.*

exist to compute a convex hull, as summarized for example by Jayaram and Fleyeh (2016) and Chadnov and Skvortsov (2004). In this work, the *Quickhull* algorithm as proposed by Barber et al. (1996) was used to calculate the convex hulls. To demonstrate the general idea of the Quickhull algorithm and give an impression of its functionality, a two-dimensional example is given as follows. To find the convex hull of a set of points S, first, the leftmost point a (smallest x-value) and the rightmost point b (largest x-value) of the set are determined. Next, point c is defined which is the farthest point from line ab within S. All points of S that lie within the triangle abc are now omitted since they lie inside the hull. Then, the residual points are divided into two subsets based on their positions. The first subset S_1' contains all remaining points to the left of the line ac, and the second subset S_2' contains all remaining points to the right in respect to the line bc. Now, the point c_1' with the largest distance to line ac is found and the points inside the triangle acc_1' are omitted. the same procedure is executed for S_2' in respect to line bc, by finding point c_2', discarding the points inside bcc_2'. This process is repeated iteratively with the remaining points. The same procedure is then used to calculate the boundary points below the line ab. graphical visualization is given in Figure 3.9.

The valid region for the illuminant estimation is defined by a convex hull $H_{A\cap B}$, which represents the intersection between the set of spectral candidates and the set of illuminants. To calculate $H_{A\cap B}$, first a convex hull H_{sc} representing the spectral candidates is defined. All points of the illuminant set within H_{sc} are added to the set $S_{A\cap B}$ that defines the valid region. This process is repeated in the opposite

configuration, determining all spectral candidates inside the convex hull of the illuminant dataset H_{ill} and adding them to $S_{A\cap B}$. The convex hull $H_{A\cap B}$ can now be derived from the combined set of inliers.

Next, it is checked if the intersection point P_{est} calculated previously lies within $H_{A\cap B}$. This can be done by adding P_{est} to the set of vertices that define $H_{A\cap B}$, and then computing a new convex hull $H'_{A\cap B}$. If P_{est} is part of the vertices that form $H'_{A\cap B}$, it lies outside of the convex hull, otherwise it lies inside. A sketch of this process is given in Algorithm 3 in the form of pseudocode.

Algorithm 3 Pseudocode: Point inside ConvexHull

INPUT
P_{est}, the previously calculated intersection point
$H_{A\cap B}$, the Convexhull

PROCESS point_inside_ConvexHull
add P_{est} to the vertices of $H_{A\cap B}$ to create a new set of points
create $H'_{A\cap B}$ from the new set
if vertices of $H_{A\cap B}$ are equal to vertices of $H'_{A\cap B}$ **then**
    return point_inside $= True$
else
    return point_inside $= False$
end if

If the estimated illuminant lies within $H_{A\cap B}$, it is proceeded with the inverse transformation of the PCA to transfer the estimated illuminant back to the spectral domain. But if the estimated point lies outside of $H_{A\cap B}$, the Euclidean distance of P_{est} to all points of the convex hull $H_{A\cap B}$ are calculated as shown in Equation 3.14, where d is the euclidean distance, and p, n are the points in n dimensions. The new estimated illuminant P'_{est} is defined as the point of $H_{A\cap B}$ with the smallest d to P_{est}.

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_1 - q_1)^2 + ... + (p_n - q_n)^2} \tag{3.14}$$

Not all components are used for transforming the estimated point back into the spectral domain. To save computation time, these calculations are done only for the dimensions that are used for the conversion, usually the first n-components, where n is dependent on the configuration of the reconstruction as described in the next section.

3.1.6 Obtaining and further Refining the Estimated Illuminant in the Spectral Domain

Once the intersection point is determined in the PCS, the estimated point is converted back into the spectral domain by using the inverse transform of the PCA first and then the inverse scaling and centering described in Section 3.1.4.1. The number of components for the reconstruction is determined beforehand. Hernández-Andrés et al. (2001b) suggest the use of a minimum of six components for a spectrally accurate reconstruction. After the estimated spectrum is obtained, it is ensured that the estimated spectral power distribution still meets the initial considerations as described in Section 3.1.2. For each spectral band, the estimated curve is checked to meet the criteria of being equal to or above the curve defined by the minConstraint. If the estimated illuminant lies below the minConstraint, it is modified to satisfy the constraint. Else, the initial estimated values are preserved. Algorithm 4 illustrates the procedure in the form of pseudocode.

Algorithm 4 Pseudocode: minConstraint adjustment

INPUT
estimated illuminant
minConstraint

PROCESS minConstraint adjustment
for λ **in** spectral range **do**
  if estimated illuminant$(\lambda) \geq$ minConstraint(λ) **then**
    estimated illuminant$'(\lambda)$ = estimated illuminant(λ)
  else
    estimated illuminant$'(\lambda)$ = minConstraint(λ)
  end if
end for
return estimated illuminant$'$

3.1.7 Optimization of the Input Parameters

To find the model parameters which give the most accurate results for the illuminant estimation, an optimization process was carried out. complemented GFC (CGFC) as explained in Section 3.2.2 was chosen as a metric for the optimization because of its ability to compare the overall shape of two spectra, which was chosen as the most important property of the reconstruction. A set of radiance cubes were calculated using Equation 3.1 on three test images excluded from the final test dataset as well as 20 randomly chosen daylight illuminants. Then different configurations were tested, varying several input parameters of the model. The altered input parameters and their values are listed in Table 3.1.

Table 3.1: *Parameters chosen for optimization of the proposed method*

Input Parameter	**Values**
number of spectral candidates	500, 1000, 1500, 2000
smoothing method	none, Median, LOWESS
Median kernel size smoothing LOWESS (f)	3, 5, 7 0.03, 0.08, 0.15
number of PC for reconstruction	3, 6

The results of the optimization process show that in some cases 500 spectral candidates are not enough to define a meaningful volume when computing the convex hull for the validation process in six dimensions as described in Section 3.1.5.4. This can be caused by input images where the minConstraint does not significantly narrow down the search area in the spectral domain, as shown in Figure 3.10. In the worst case, none of the spectral candidates intersects with the set of physically possible illuminants within the PCS. A number of 1500 for the creation of spectral candidates was chosen in the end to provide a balance between creating a detailed volume for the estimation and saving computational costs.

In order to investigate the performance of LOWESS smoothing as described in Section 3.1.3, it was tested against a Median filter at different configurations. Thereby, both algorithms were used with three different amounts of smoothing. As expected, the LOWESS algorithm provides a more natural shape of the generated curve while preserving local variation. In the end, a smoothing factor of $f = 0.03$ was determined to give the best results overall.

Figure 3.10: *In this example, the minConstraint does not restrict the search area significantly.*

3.1.8 Step by Step Framework Summary

In this section, a concise overview is given of the different steps that built the proposed model as described comprehensively in the previous sections.

1. Define the image-dependent constraints to restrict the possible illumination estimation results to a plausible region.
2. Randomly generate spectral candidates that lie within the plausible region for a given radiance cube.
3. Fit a common PCS using a set of reflectances and a representative database of daylight illuminants SPDs, such as the Granada daylight spectral database created by Hernández-Andrés et al. (2001b).
4. Transform the illuminant database as well as the spectral candidates into the PCS.
5. In the PCS, fit an n-dimensional line that represents the illuminant database and a hyperplane that represents the spectral candidates.
7. Estimate the illuminant within the PCS by calculating the intersection of line and hyperplane.
8. Ensure a valid estimation by comparing it against the volume of possible points.

9. Convert the estimated illuminant back to the spectral domain using the PCAs inverse transform.

10. Ensure the estimated illuminant lies within the initial constraints.

11. Use the estimated illuminant spectrum to calculate reflectance values for the input radiance cube.

3.2 Experiments

To test the performance of the proposed image statistics-based framework, the model was evaluated regarding its ability to estimate illuminant spectra and to reconstruct relative reflectance from the input radiance cubes. First, a set of normalized radiance cubes was simulated from ground truth reflectance cubes and representative ground truth daylight illuminants. These radiance cubes served as input for the model. The framework as proposed in the previous sections then predicted the illuminant spectrum for each cube from the input data without the use of a reference target. Afterwards, the reconstructed illuminants were evaluated using different full reference metrics explained in detail in Section 3.2.2. In addition, the proposed illuminant estimation was compared to other statistics-based illuminant estimation algorithms, namely spectral grayworld, spectral gray-edge, and max-spectral, referred to as spectral constancy algorithms.

In addition, the estimated SPD was used to recover the relative reflectance values for each of the input cubes, which were then compared to the ground truth reflectance. The same reference metrics as for the illuminant estimation were then applied to evaluate the ability of the proposed framework in terms of its reflectance recovery capabilities.

The following sections focus in detail on how the training and test datasets were created and which full reference metrics were chosen to evaluate the reconstructed data. Additionally, an overview is given of the competing estimation algorithms used for comparing the proposed method against other work that was done in the field as described in Section 2.2.2.

3.2.1 Data Preparation

3.2.1.1 Reducing the Bias within the Illuminant Database

A simple and well-adopted way of specifying a given source of daylight by using a single parameter is to express the spectrum in terms of its correlated color temperature (CCT) as discussed by McCamy (1992). The CCT in Kelvin indicates

the perceived color of a given illumination spectrum by referring to the temperature of a Planckian radiator with approximately the same color under specified viewing conditions Hunt (1978). Due to metamerism, caused by the trichromatic properties of the human visual system this concept is based on, infinite spectra exist that can result in the same CCT. To determine the CCT of a SPD, the radiation of a Planckian body at different temperatures is expressed as unique CIE 1931 x,y chromaticity coordinates. These chromaticity coordinates, whose computation is described in detail by Wyszecki and Stiles (2000) and others, are a way to describe the quality of a given color in terms of hue and saturation. A locus can be built within CIE 1931 chromaticity space using these chromaticity coordinates, called the Planckian or blackbody locus. The computed chromaticity coordinates of many natural and artificial light sources lie on or near by this Planckian locus in CIE 1931 x,y chromaticity space. The point on the Planckian locus which is closest to the chromaticity coordinates of a given illuminant hereby defines its CCT. The CCT of a given daylight spectrum can be obtained by using Equation 3.15 and Equation 3.16 alongside the corresponding values from Table 3.2 from Hernández-Andrés et al. (1999). For other light spectra, it is recommended to use the method proposed by Robertson (1968). Hereby, x and y denote the CIE 1931 chromaticity coordinates. A perceptually more uniform scale to describe the color of a given light source is to express the illuminant spectrum by the inverse correlated color temperature (CCT^{-1}) instead. The CCT^{-1} can describe the perceived color temperature of a given spectral power distribution in reciprocal mega-Kelvin (MK^{-1}) and can be a good estimate for the classification of irradiance spectra in the VIS range. It can be calculated by computing the CCT as mentioned before and then using Equation 3.17.

As stated by the authors (Hernández-Andrés et al., 2001b), the Granada daylight spectral database[1] is biased towards illuminants of an CCT^{-1} of around 175 to 180 MK^{-1}. To reduce the bias on the computation of the principal components significantly, the dataset was divided into different regions with a range of $5\mathrm{MK}^{-1}$ each, spanning from 5 to $270\mathrm{MK}^{-1}$. Up to 40 spectra were then randomly selected from the dataset for each region to simulate a more uniform distribution. At the extremes of the range, only a few spectra were available for selection. The distribution of illuminant spectra with respect to their CCT^{-1} before and after bias reduction are shown in Figure 3.11. The result was a dataset of 1326 daylight spectra, which replaces the initial dataset in all calculations that are described from now on.

$$n = (x - x_e)/(y - y_e) \tag{3.15}$$

[1] available at `https://colorimaginglab.ugr.es/pages/Data`, last accessed 09.08.2023

$$CCT = A_0 + A_1 e^{(-n/t_1)} + A_2 e^{(-n/t_2)} + A_3 e^{(-n/t_3)} \tag{3.16}$$

$$CCT^{-1} = 10^6 / CCT \tag{3.17}$$

Table 3.2: *Constants and colorimetric epicenters (x_e, y_e) for the calculation of CCT values within the valid range of $3000 - 8 \times 10^5$ Kelvin, as defined by Hernández-Andrés et al. (2001b)*

Constants	$3000K - 50.000K$	$50.000K - 8 \times 10^5 K$
x_e	0.3366	0.3356
y_e	0.1735	0.1691
A_0	−949.86315	36284.48953
A_1	6253.80338	0.00228
t_1	0.92159	0.07861
A_2	28.70599	5.4535×10^{-36}
t_2	0.20039	0.01543
A_3	0.00004	
t_3	0.07125	

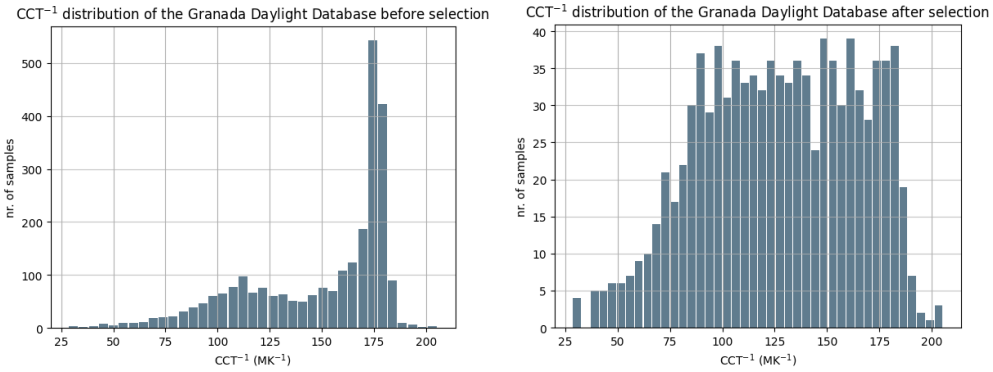


Figure 3.11: *Left: initial distribution of all 2600 illuminant spectra of the Granada daylight spectral database in respect to their inverse CCT.*
Right: Dataset after bias reduction, consisting of a total of 1326 spectra.

3.2.1.2 Simulating Radiance Cubes for Testing

A test dataset of relative radiance consisting of a total of 150 radiance cubes in the range from 0 to 1 was calculated by the combinations of 25 reflectance cubes as well as six representative illuminant spectra. Since homogeneous illumination conditions are assumed in this work, Equation 3.1 was used for the computation. The reflectance cubes were recorded using a Cubert Ultris X50 as well as a Cubert Ultris X20 hyperspectral snapshot camera mounted on an UAV, as seen in Figure 3.12. Both cameras capture a spectral range of 350 - 1002nm with a spectral sampling of 4nm, resulting in 164 channels with a bit depth of 12 bit. While the Ultris X20 provides a native spatial resolution of 410x410 pixels, the spatial resolution of the Ultris X50 is 570x570 pixels. The image cubes of the Ultris X50 were cropped to 550x550 pixels to remove artifacts that were present in some of the images.

Figure 3.12: *Cubert Ultris X20P, including an additional panchromatic sensor, mounted on a UAV mid-flight.*

From the altered illuminant dataset described in Section 3.2.1.1, six representative spectra were chosen based on their CCT^{-1}. The chosen spectra range from 75 to 200 MK^{-1} in steps of 25 MK^{-1} and plots of their relative SPDs are shown in Figure 3.13. Their variety with respect to common daylight scenarios is visualized in Figure 3.14. The spectra used for the calculation or radiance test images were removed from the dataset since they were used to calculate the training data for the PCS, as described in Section 3.2.1.3. The illuminant data was cropped and interpolated to match the spectral bands recorded by the Cubert Ultris cameras.

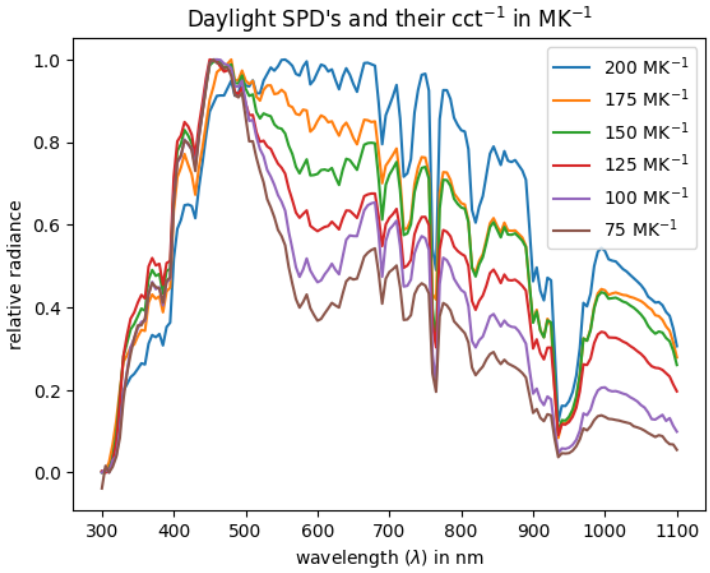


Figure 3.13: *Chosen SPDs of the Granada daylight spectral dataset based on their inverse CCT.*

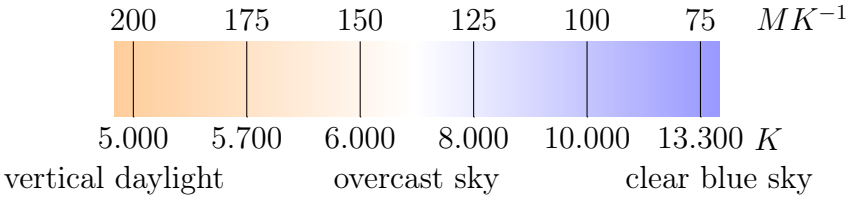


Figure 3.14: *Inverse correlated color temperature of the chosen SPDs shown above the bar and proximate CCT values shown below. The color of the bar indicates the perceived shift in chromaticity from warmer to cooler colors for the respective illuminants.*

Colorimetric thumbnail example images rendered in sRGB using the CIE D65 standard illuminant at a spectral resolution of 5nm and the CIE 1964 color matching functions at 5nm resolution are shown in Figure 3.15 below. With a focus on drone-based applications, the images show a variation of vegetation and crops as well as soil, dirt, and tarmac roads as well as other man-made objects like cars and houses. But also several cubes captured under controlled conditions in the lab were included.

Figure 3.15: *Example images, rendered as sRGB images using CIE standard illuminant D65 and CIE 1964 color matching functions at 5nm spectral resolution.*

3.2.1.3 Simulating Radiance Data for Training of the PCS

Besides preparing radiance cubes for testing the framework, radiance spectra were simulated to serve as a training dataset to fit the image-independent PCS. The training dataset contains a total of 316.800 radiance spectra. It was created by combining 240 reflectance spectra and a total of 1320 illuminant spectra. The reflectance spectra, kindly provided by the University of Granada, were obtained by measuring a Greta McBeth ColorChecker DC using a Photoresearch PR745 spectroradiometer within a spectral range of 380 - 1080nm. The spectra of the Granada daylight spectral database that were not used for creating the radiance test cubes served as SPDs for the calculation, which was carried out using Equation 3.1 for all combinations of reflectance and radiance spectra, resulting in 316.800 training spectra in a range of 380 - 1080nm.

Afterwards, it was assured that all data involved in the illuminant estimation was sharing a common spectral range and spectral resolution. This was necessary to make it possible to fit the common PCS, transfer the input data into it, calculate the intersection and convert the data back to the spectral domain. Therefore, the

training and the test dataset as well as the illuminant database were all cropped to a common spectral range of 382 - 1002 nm. Then, the training and illuminant datasets were linearly interpolated to match the spectral bands of the radiance data.

3.2.2 Evaluation Metrics

The following paragraphs explain the full-reference metrics used to evaluate the results of the estimated SPD and reflectance values, by comparing the estimated spectra against the respective ground truth data. As stated by Imai et al. (2002), there is no single metric recommended for the evaluation of a spectral match. Therefore, a variety of metrics are chosen to reflect the model's performance in different aspects.

3.2.2.1 Root Mean Square Error (RMSE)

The root mean square error (RMSE) is a well-established metric that can be used to evaluate the average magnitude of the differences between an estimated spectrum and its original counterpart. Equation 3.18 shows how to calculate the RMSE of two given spectra, where S is the original spectrum, S_R is the estimated spectrum and n is the number of bands.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S(\lambda_i) - S_R(\lambda_i))} \tag{3.18}$$

with $0 \geq \text{RMSE} \geq 1$ for input data in a range between 0 and 1. A RMSE value of 0 indicates a perfect match, whereas 1 indicates the worst match (or ∞, if the input data is not normalized). Shifts in amplitude and scale affect the score of the metric, but offsets by a constant do not.

3.2.2.2 Goodness of Fit Coefficient (GFC)

The goodness of fit coefficient (GFC), based on Schwartz's inequality and proposed by Romero et al. (1997) can be used to compare two spectra, usually a reference spectrum and its reproduction. The GFC can be calculated as shown in Equation 3.19, where S is the original, measured spectrum, S_R is the estimated spectrum and n indicates the number of bands:

$$\mathrm{GFC} = \frac{\left|\sum_{i=1}^{n} S(\lambda_i) S_R(\lambda_i)\right|}{\left|\sum_{i=1}^{n} [S(\lambda_i)]^2\right|^{1\backslash 2} \left|\sum_{i=1}^{n} [S_R(\lambda_i)]^2\right|^{1\backslash 2}} \tag{3.19}$$

The resulting value lies between 0 and 1, where 1 indicates a perfect match of the two spectra. As described by Romero et al., a GFC ≥ 0.99 indicates acceptable results, a GFC ≥ 0.999 is considered of very good reproduction quality and a GFC ≥ 0.9999 can be considered an almost exact mathematical match. These ratings were proposed in the context of dimensionality reduction of given spectra, not in terms of illuminant estimation. Nevertheless, these numbers can also give a good indication of the quality of an estimated spectrum. It has to be noted that GFC is insensitive to shifts in scale. An adjustment of this method was proposed by Viggiano (2004). For the convenience of matching other metrics like RMSE, GFC can be subtracted from unity, resulting in 0 being the best fit and 1 being the worst possible spectral match. Viggiano (2004) refer to this method as CGFC. GFC and CGFC are good metrics to compare the overall match in shape of a sample spectrum against a reference. For the convenience of the reader, CGFC is chosen over GFC for presenting and discussing the results in this work. By doing so, a perfect match between two spectra corresponds to a value of 0 for all metrics presented in this section.

3.2.2.3 Spectral Angle Mapper (SAM)

The spectral angle mapper (SAM) as described by Kruse et al. (1993) is another widely used metric to compare the spectral similarity of a reference spectrum and a sample spectrum. SAM treats the two input spectra as n-dimensional vectors and calculates the angle between them. The mathematical expression is given in Equation 3.20, where S is the reference spectrum and S_R is the spectral sample to compare. SAM is the inverse cosine of the GFC. The smaller the resulting value, the more similar the compared spectra. However, SAM does not consider the overall shape of the spectra, but rather their similarity in a higher dimensional space. Since the lengths of the vectors are not taken into account, SAM is insensitive to differences in intensity.

$$\mathrm{SAM} = \cos^{-1}\left(\frac{\left|\sum_{i=1}^{n} S(\lambda_i) S_R(\lambda_i)\right|}{\left|\sum_{i=1}^{n} [S(\lambda_i)]^2\right|^{1\backslash 2} \left|\sum_{i=1}^{n} [S_R(\lambda_i)]^2\right|^{1\backslash 2}}\right) \tag{3.20}$$

3.2.2.4 Integrated Radiance Error (IRE)

The integrated radiance error (IRE) metric as described by Michalsky (1985) calculates the sum of the absolute values of difference at each wavelength and normalizes the result by dividing by the integrated reference spectrum. The mathematical notation is shown in Equation 3.21, where S is the reference spectrum, S_R is the reproduction, and n is the number of spectral bands.

$$\text{IRE} = \frac{\left|\sum_{i=1}^{n} S(\lambda_i) - S_R(\lambda_i)\right|}{\sum_{i=1}^{n} S(\lambda_i)} \tag{3.21}$$

The best possible match is 0, whereas the worst match is ∞. The metric can give a good indication of how well the two compared spectra match overall, but spectra that produce an identical area will result in the same IRE. The metric is highly sensitive to a change of scale.

3.2.3 Comparison to other Statistics-based Illumination Estimation Algorithms

In addition to comparing the estimated illuminant and the recovered reflectance cubes against their respective ground truth data, the estimated SPDs is also compared against three other image statistics-based illumination estimation algorithms, namely spectral grayworld, spectral gray-edge and max-spectral as proposed by Khan et al. (2017c) and already described in Section 2.2.2. Since these methods are adopted from their color constancy algorithm counterparts, they will be referred to as spectral constancy algorithms when talking about them as a set of algorithms. For each of the radiance cubes used as test data for the proposed algorithm, the illumination is estimated using each of the spectral constancy algorithms. Their performance is evaluated by comparison against the ground truth illuminant by using the spectral full reference metrics CGFC, RMSE, IRE and SAM described in Section 3.2.2.

4 RESULTS

In this chapter, the results are represented. Hereby, the chapter is split into two main parts. The first one focuses on the results obtained from the SPD estimation. Tables are presented that show the performance of the presented model when estimating the SPD present in the given input data by comparing it to the ground truth SPD via the previously described metrics. Hereby, a perfect match between a sample and the ground truth with respect to a given metric will always result in a value of 0. Results are given for the performance of the proposed model on the whole dataset, which are then compared to competing statistics-based illuminant estimation algorithms. Then, results are displayed that provide information about the performance of the proposed algorithm to predict illuminants with a certain CCT^{-1}. Alongside, plots are given as an example for the estimation results for a given CCT^{-1}. The second part focuses on the performance of relative reflectance recovery for the whole dataset.

4.1 Estimated Illuminants

After estimating the SPD for each of the 150 radiance input cubes in the common PCS and reconstructing the illuminant spectra using three principal components, the error metrics explained in the previous section were computed. The results in terms of the mean CGFC, RMSE, IRE and SAM over all input cubes are shown in Table 4.1. In addition, the table shows the performance compared to the spectral constancy algorithms described in Chapter 2.

After presenting the results for the general performance over a broad range of different daylight illuminants, the capabilities of the proposed method to estimate daylight illuminants of a certain CCT^{-1} are shown below. Figure 4.1 consists of six plots showing an estimated relative SPD against the corresponding ground truth for each CCT^{-1} used for the experiments. Below, Table 4.2 shows the ability of the proposed model to predict daylight illuminants with a certain CCT^{-1}. In addition, a plot of the trendlines for estimating different CCT^{-1} is shown in Figure 4.2.

Table 4.1: *Results of the estimation of the illuminant SPD using three components for reconstruction from PCS, mean values for all 150 estimated illuminants of the test dataset.*

	proposed	**grayworld**	**max-spectral**	**gray-edge**
CGFC				
min	0.0026	0.0074	**0.0011**	0.0049
mean	**0.0219**	0.2537	0.1279	0.1539
max	**0.0832**	0.5297	0.3259	0.3527
90^{th}pctl	**0.0508**	0.4141	0.2675	0.2910
RMSE				
min	0.0568	0.0829	**0.0368**	0.0735
mean	**0.1593**	0.4156	0.3183	0.3465
max	**0.3685**	0.6117	0.6533	0.6166
90^{th}pctl	**0.2795**	0.5430	0.4629	0.4995
IRE				
min	0.0022	0.0038	0.0019	**0.0003**
mean	**0.2043**	0.2961	0.2815	0.3154
max	0.6775	**0.5782**	0.7870	0.7262
90^{th}pctl	**0.4464**	0.4859	0.5824	0.5456
SAM				
min	0.0726	0.1213	**0.0474**	0.0993
mean	**0.1889**	0.6988	0.4689	0.5207
max	**0.4109**	1.0812	0.8310	0.8668
90^{th}pctl	**0.3200**	0.9448	0.7488	0.7827

Surprisingly, a reconstruction from the PCS back to the spectral domain using three components yielded better results than using at least six components, as recommended by Hernández-Andrés et al. (2001b). A table showing a comparison of illuminant reconstruction results using three and six principal components is shown in Appendix C, Table C.1. A possible reason for lower estimation accuracy using 6 components will be addressed within Chapter 5, DISCUSSION.

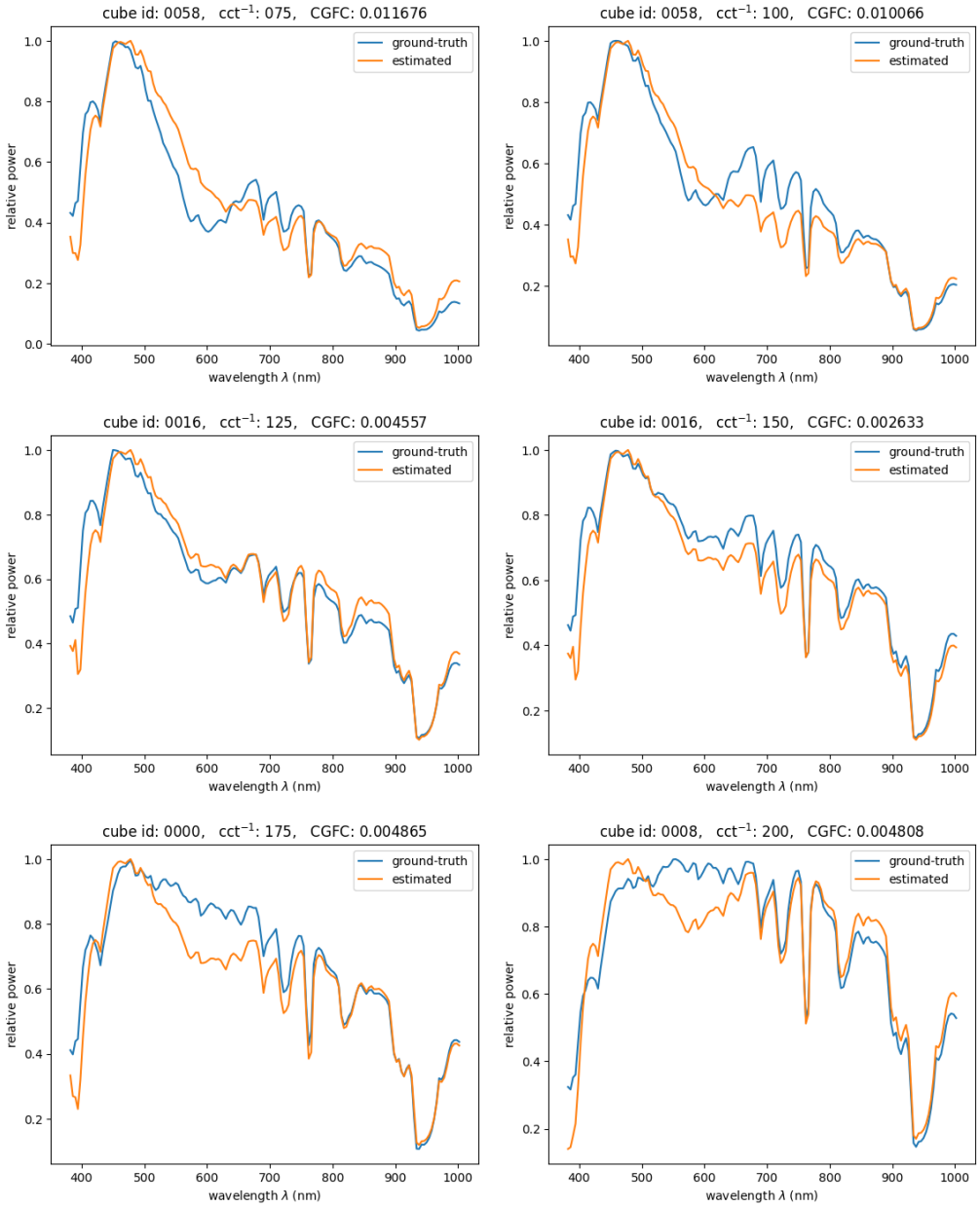


Figure 4.1: *Individual reconstruction results, one for each chosen CCT^{-1}. The blue curve represents the ground truth illuminant used for creating the radiance image, the orange line is the reconstructed illuminant from three principal components using the proposed estimation model.*

Table 4.2: *Results of the illuminant estimation, subdivided by the CCT^{-1} of the ground truth SPD.*

	75MK^{-1}	**100MK^{-1}**	**125MK^{-1}**	**150MK^{-1}**	**175MK^{-1}**	**200MK^{-1}**
CGFC						
min	0.0117	0.0101	0.0046	**0.0026**	0.0048	0.0046
mean	0.0566	0.0339	0.0175	0.0081	0.0079	**0.0072**
max	0.0832	0.0503	0.0277	0.0209	**0.0118**	0.0144
90^{th}pctl	0.0745	0.0454	0.0260	0.0113	**0.0104**	0.0118
RMSE						
min	0.0818	0.0821	0.0597	**0.0568**	0.0740	0.0776
mean	0.2698	0.2168	0.1525	0.1020	**0.1006**	0.1141
max	0.3685	0.2963	0.2237	0.1845	**0.1486**	0.1881
90^{th}pctl	0.3340	0.2720	0.2125	0.1314	**0.1175**	0.1695
IRE						
min	0.0443	0.0611	0.0123	0.0094	**0.0022**	0.0112
mean	0.4848	0.3276	0.1719	0.0878	**0.0674**	0.0862
max	0.6775	0.4638	0.2803	0.2466	**0.1799**	0.1927
90^{th}pctl	0.6099	0.4238	0.2631	0.1177	**0.1051**	0.1702
SAM						
min	0.1530	0.1420	0.0955	**0.0726**	0.0984	0.0964
mean	0.3329	0.2559	0.1807	0.1225	0.1242	**0.1176**
max	0.4109	0.3184	0.2359	0.2047	**0.1540**	0.1698
90^{th}pctl	0.3884	0.3024	0.2283	0.1506	**0.1442**	0.1533

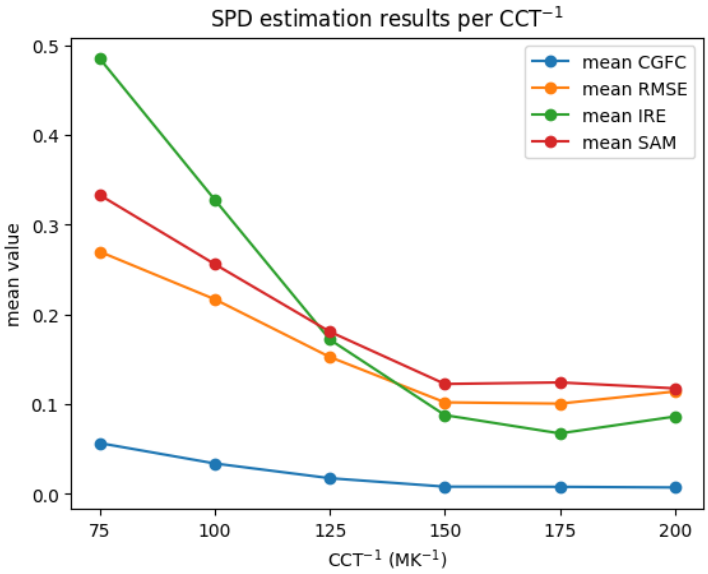


Figure 4.2: *Trendlines for the mean estimation accuracy of the proposed model per CCT^{-1} of all evaluation metrics.*

4.2 Reconstructed Reflectance Cubes

Comparing ground truth reflectance against reconstructed reflectance using the illuminant SPD estimated by the proposed model. First, the reflectance cubes are reconstructed by dividing each radiance cube by its corresponding estimated illuminant. Then, for each point spectrum of all 150 reflectance cubes (more than 45.3 Million point spectra in total), the reconstructed reflectance is compared against the ground truth using the metrics discussed in Section 3.2.2. From these sets of values, the minimum, mean, maximum, and 90^{th} percentile are calculated for each metric to evaluate the overall performance of the reflectance reconstruction. The results are shown in Table 4.3. Then, detailed results of the performance per tested CCT^{-1} are shown in Table 4.4. Figure 4.3 then shows a comparison between the errors of the SPD estimation and the reflectance recovery, summarizing the general findings of the comparison of Table 4.2 and Table 4.4.

Table 4.3: *Results of the relative reflectance reconstruction. The results are calculated using all 45.3 million spectra of the 150 test cubes.*

CGFC	
min	0.0002
mean	0.0172
max	0.3364
90^{th}pctl	0.0461

RMSE	
min	0.0003
mean	0.0470
max	0.3783
90^{th}pctl	0.1178

IRE	
min	0.0000
mean	0.2312
max	0.7199
90^{th}pctl	0.5110

SAM	
min	0.0200
mean	0.1520
max	0.8451
90^{th}pctl	0.3048

Table 4.4: *Results of the reflectance recovery, subdivided by the CCT^{-1} of the ground truth SPD.*

	75MK^{-1}	**100MK^{-1}**	**125MK^{-1}**	**150MK^{-1}**	**175MK^{-1}**	**200MK^{-1}**
CGFC						
min	0.0072	0.0043	0.0004	**0.0002**	0.0005	**0.0002**
mean	0.0470	0.0305	0.0114	0.0067	0.0069	**0.0033**
max	0.3364	0.2619	0.2155	0.1764	0.1299	**0.0537**
90^{th}pctl	0.0915	0.0554	0.0272	0.0171	0.0147	**0.0094**
RMSE						
min	0.0005	0.0004	0.0004	**0.0003**	**0.0003**	**0.0003**
mean	0.0867	0.0709	0.0497	0.0338	0.0355	**0.0180**
max	0.3783	0.3327	0.2580	0.2006	0.1775	**0.1010**
90^{th}pctl	0.1857	0.1607	0.1171	0.0876	0.0884	**0.0418**
IRE						
min	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
mean	0.4457	0.3425	0.2367	0.1586	0.1679	**0.0747**
max	0.7199	0.5923	0.4456	0.4576	0.5749	**0.2321**
90^{th}pctl	0.6162	0.5378	0.4150	0.2932	0.2849	**0.1400**
SAM						
min	0.1201	0.0929	0.0282	**0.0200**	0.0326	0.0216
mean	0.2941	0.2368	0.1242	0.0933	0.1047	**0.0720**
max	0.8451	0.7405	0.6689	0.6031	0.5155	**0.3291**
90^{th}pctl	0.4312	0.3344	0.2339	0.1851	0.1718	**0.1369**

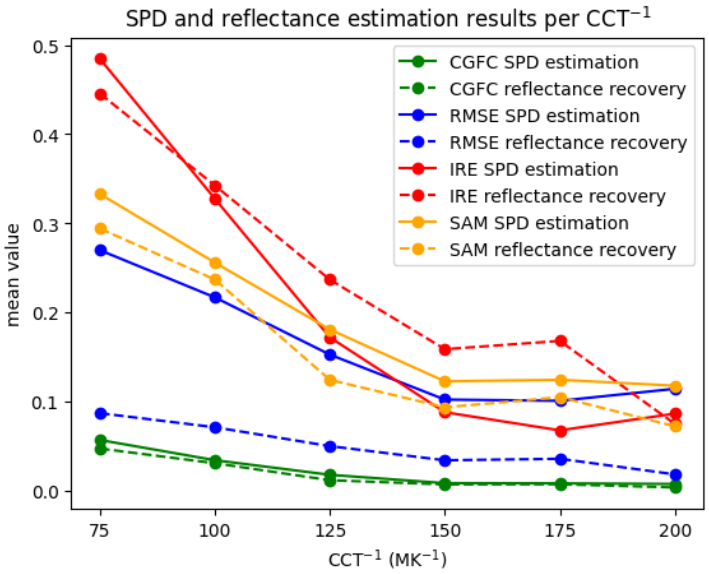


Figure 4.3: *Comparison of trendlines for the mean results in terms of illuminant SPD estimation and reflectance recovery of the proposed model per CCT^{-1} of all evaluation metrics. First three principal components used.*

5 DISCUSSION

This chapter discusses the results presented in the previous section as well as the general abilities and restrictions of the proposed framework. It is thereby divided into three main parts. The first section focuses on the capabilities of the designed algorithm in terms of its illumination estimation capabilities, whereas the second part then elaborates on the reflectance recovery results. The third part considers general remarks about the model, its limitations, and possible improvements.

5.1 Evaluation of Illuminant Estimation Capabilities and Comparison to Competing Algorithms

The numerical results of the illuminant estimation presented in Table 4.1 clearly indicate the ability of the model to successfully estimate a variety of representative daylight illuminant spectra to a reasonable precision. A CGFC mean value of 0.02 and a value of 0.05 for measuring the 90^{th} percentile over all images indicates a strong correlation of the estimated spectrum with the ground truth. Following the lead of Romero et al. (1997), the threshold for a 'good' spectral reproduction of CGFC-value ≤ 0.01 is not achieved over the whole dataset, indicating room for improvement. On the other hand, the best illuminant reconstruction achieves an CGFC value of 0.0026, which is very promising. In addition, one has to keep in mind that these threshold values were defined for dimensionality reduction. Considering the ill-posed problem of no-reference illumination estimation as this work does, the perfect estimation of a broad variety of ground truth spectra over a large variety of different scenes is unlikely when using a single algorithm. Therefore, the evaluation against the aforementioned threshold values can be thought of as a guideline rather than a definitive categorization.

Examining the results in terms of RMSE, the proposed algorithm shows decent estimation qualities. A mean RMSE of 0.1593 and 0.2795 when considering the 90^{th} percentile, over all estimated illuminants is moderately accurate. The ability

of the model to accurately estimate SPDs is visually underlined by Figure 4.1, showing the selected ground truth SPDs and a corresponding estimation result for different CCT^{-1}.

When comparing the results against the other tested image statistics-based algorithms, it can be seen that the proposed model outperforms its competitors with respect to most metrics used. Although max-spectral and spectral gray-edge achieve better peak performance in single cases, the proposed algorithm is able to achieve significantly better estimation precision over a larger variety of different input data, which can be seen when comparing the mean values and 90^{th} percentiles of the given metrics in Table 4.1. The maximum error over all metrics besides IRE is lowest when using the proposed algorithm, strongly indicating that fewer outliers with very large reproduction errors are produced. In addition, the created algorithm still shows the best results in terms of IRE mean and 90^{th} percentile, indicating that it is able to estimate the overall normalized power better than the competitors over a variety of input data. Hereby, IRE compares the integrals under the sample and ground truth curves and does not take into account the general shape of the spectra. Comparing the different algorithms in terms of the resulting CGFC values, it can be seen that the proposed model significantly outperforms the other algorithms. The same can be observed when evaluating SAM results, where the error between estimation and ground truth SPDs can be substantially reduced by using the proposed algorithm.

A reason for the good overall performance of the newly created model is the restriction of the estimated illuminant in PCS to a volume of intersection between the image-dependent plausible and physically possible measured spectra of the dataset. This prevents the estimation to be far off of any plausible daylight illumination spectrum within the broad variety represented by the Granada daylight illuminant database. The competing algorithms lack such restrictions in terms of the plausibility of their estimation results. The good peak performances of the max-spectral algorithm are most likely achieved for images where at least one pixel captured a specular reflection or an object with high reflective and relatively uniform spectral properties and therefore is able to predict the SPD accurately. In such cases, the proposed model restricts the estimation in a way that it can not lie below its maxConstraint, which is similar to estimation using the max-spectral algorithm. Hereby, it is possible that the created model overestimates the relative radiance of the SPD, which then results in slightly higher errors. Nevertheless, such cases are not very common for the intended application in forest areas and agricultural environments.

It is noted, that among the three spectral-constancy algorithms tested, the max-spectral and spectral gray-edge algorithms showed better results than spectral grayworld. This correlates with the findings of the authors who originally proposed

these estimation methods.

Comparing the SPD estimation results subdivided by their CCT^{-1}) as presented in Table 4.2 shows that on average, the proposed model predicts illuminants with an CCT^{-1} of 150 MK^{-1} and above with higher precision. This behavior is clearly visible when evaluating Figure 4.2, where a trend over all metrics can be observed. Since the bias of the illuminant dataset was reduced before simulating the training data by a large amount, this is not expected to be the reason.

When comparing the illuminant estimation results using the first three principal components against the first six for the reconstruction of the estimated illuminant spectrum from the PCS, the results in Table C.1 indicate a difference in performance. It is noticeable that the performance gets worse when using six principal components. Since the physically possible daylight illuminants are not represented precisely enough by a line in the input independent PCS for dimensions higher than three, the calculated intersection point is not accurate enough to precisely estimate the illuminant. Hernández-Andrés et al. (2001a), however, concluded that at least six components should be used for accurate spectral reconstruction. This indicates that refining the assumption of a linear approximation within the PCS might increase the estimation accuracy of the framework significantly in the future. Possible improvements could be a different technique when calculating the intersection or different training data to fit the common PCS. Nevertheless, this work shows that the presented method can achieve good results in illuminant estimation accuracy using only the first three principal components for reconstructing the SPD from the principal component space.

5.2 Evaluation of Reflectance Recovery Results

The evaluation of the reflectance recovery in Section 4.2 shows excellent reconstruction qualities in the best cases with a minimum CGFC value of 0.0002, minimum RMSE of 0.0003 and an almost perfect match in terms of IRE. Also, as expected, the overall reflectance estimation results increase slightly in comparison to the illumination estimation. This improvement is visualized in Figure 4.3, where the mean results of all metrics are plotted for the mean SPD estimation and the reflectance recovery for each CCT^{-1}. It can be seen that the overall trend of the curves is the same, but the error is slightly smaller in terms of SAM and significantly smaller in terms of RMSE over all CCT^{-1}. Since the illuminant estimation from image data without ground truth measurements is always dependent on the information present in the image, some regions of the SPDs spectral range might not be properly represented by the input data. There always exists the possibility that the objects

within the scene barely reflect any radiance in certain ranges of the measured spectrum. Since the illuminant information is recovered from a particular scene measurement directly, the uncertainties of the estimation results will most likely be higher for the same regions as well. However, this means in reverse that even if the illuminant can not be predicted with high accuracy over the whole spectral range, using that estimation might not introduce large errors when predicting reflectance values of the given radiance cube.

5.3 General Considerations

After showing that the created method is clearly capable of accurately estimating daylight illuminants, some general considerations will be discussed in this section. Although the spectral range of 382 - 1002nm for the illuminant estimation is suitable for many applications, especially when examining biochemical and biophysical plant parameters, it is not limited to a specific spectral range, as long as sufficient training data is available for creating the independent PCS. This could for example enable an expansion of the spectral range for estimation to the UV and SWIR range as well. Hereby, it is expected that the general structure of the model must not undergo any changes and can simply be extended by using different training data. Although the Granada daylight spectral database can be considered a representative collection for daylight in (southern) Europe, the estimation results might benefit from an extended database consisting of illuminant measurements from different sites all over the world.

The major limitation of the proposed method is, like with most other methods in this field, the assumption of a single, uniformly distributed light source. When considering cloudless illumination conditions at the time of the scene capture, the estimation of a single light source might not be sufficient enough, considering the effect of direct and indirect sunlight as well as shadow areas on the captured data. A possible solution to this problem will be given in Section 6.2, Possible Improvements and Future Work. Nevertheless, this work not only proposes a completely new approach for the estimation of illumination spectra, but it also lays the foundation for further development of illumination estimation algorithms.

6 CONCLUSION

6.1 Key Findings

This work aimed to develop a framework capable of precisely estimating daylight illuminant spectra from hyperspectral radiance data without the use of simultaneous ground truth measurements such as reflectance targets. A special interest was taken in the application of drone-based surveillance of agricultural and forestry applications. Unlike other work, the proposed method restricts the possible illuminant spectra by defining image-dependent plausible spectra and physically possible spectra. The image-dependent plausible spectra are generated by restricting the search space within the spectral domain by using generic assumptions and calculating spectral candidates using a random walk approach. The physically possible spectra are represented by a set of measured daylight illuminants. Within a pre-trained, input-independent principal component space, the intersection point between those two sets of data is calculated. By converting the intersection point back to the spectral domain, the estimated illuminant is derived.

The evaluation of the obtained results demonstrates that the proposed model is capable of accurately estimating a representative set of illuminant spectra over a large set of 150 input images. It thereby significantly outperforms competing statistics-based approaches over the tested images.

It is evident that the key assumption, the possibility to estimate a given daylight illuminant as an intersection between these restrictions within an image-independent PCA, can yield promising results. Especially illuminants corresponding to an CCT^{-1} of $150 - 200\mathrm{MK}^{-1}$ are predicted accurately. In addition, the proposed framework is able to accurately recover reflectance values from the radiance input data.

At its current state, the model has been validated to estimate illuminants within a spectral range of $382 - 1002\mathrm{nm}$, making it suitable for many vegetation-based analysis approaches. Nevertheless, the proposed model is not restricted to these applications.

6.2 Possible Improvements and Future Work

Although the framework proposed in this work can estimate a broad range of daylight illuminants with acceptable to very good performance, some key points can be identified to further improve the proposed method in the future. First, the illuminant estimation within the PCS can be improved, especially for dimensions higher than three. This can be done by using another way to calculate the estimation result, for example by using other, gamut-based methods instead of the intersection between line and plane. Another possibility would be an extension and optimization of the training data to create the common PCS. This might lead to a different representation of the physically possible daylight spectra within the dimensions greater than three, which can be utilized for the estimation in a better way.

Since daylight spectra can be highly dependent on the geographical location, the use of a worldwide database of measured illuminants could lead to accurate estimation results for an even broader range of daylight scenarios. In addition, possibilities should be explored to extend and test the proposed framework for the estimation of illumination spectra other than those of typical daylight conditions, such as incandescent light or LED. A possible approach would be the use of different reference databases for different illuminant categories, which can be interchanged based on the desired application.

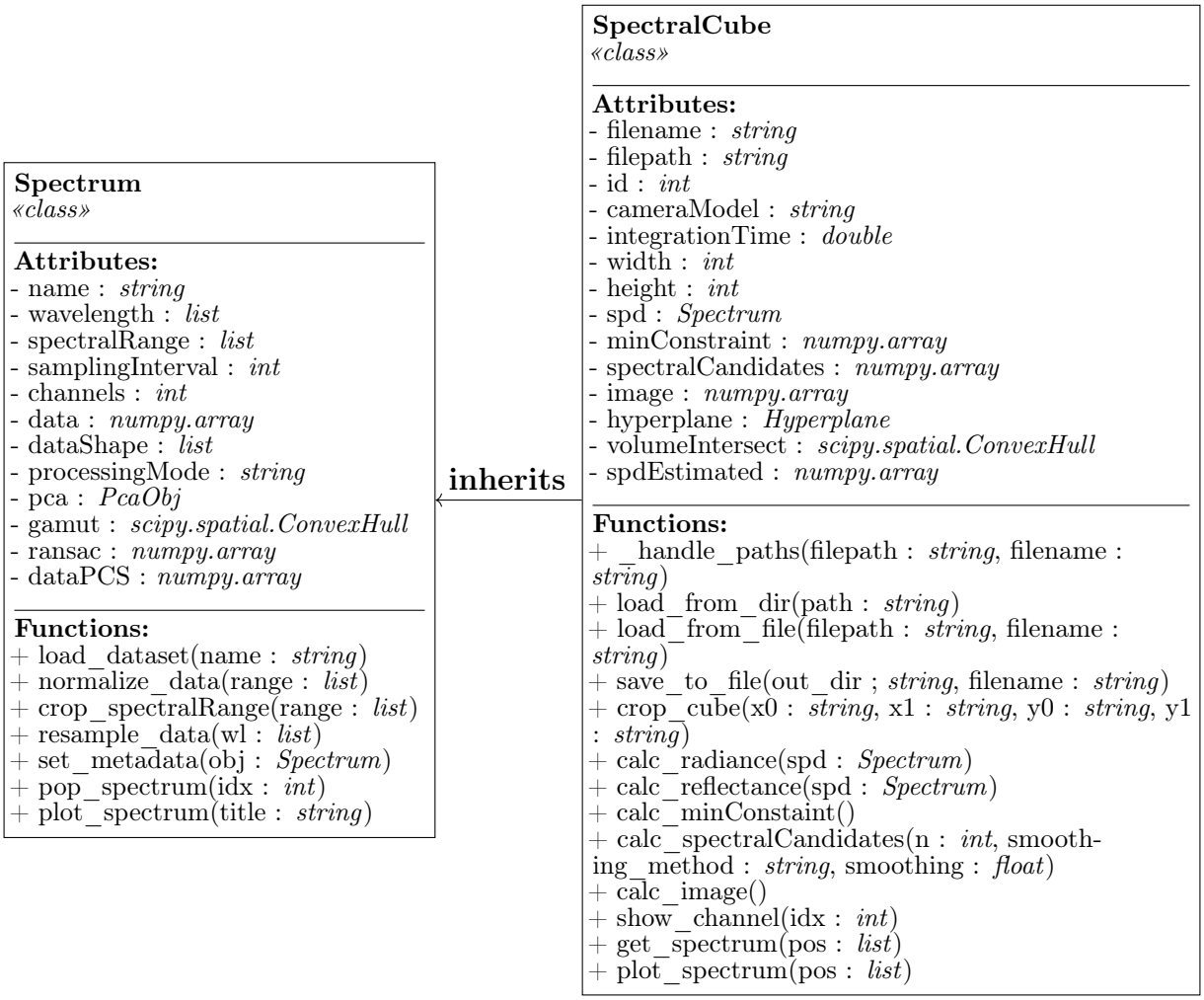
Last, in this work it is considered that the scenes are illuminated by a single, uniformly distributed light source. While this is a common restriction for most illumination estimation algorithms, this assumption does not hold for many practical applications. The estimation of spatially nonuniform illumination conditions can be achieved by dividing the scene into smaller portions in a grid-wise manner and deriving local estimates, for example, as proposed for three-channel color images by (Gijsenij et al., 2012). It would then be possible to interpolate these local results to create a map that considers the spatial distribution of several illuminants. A more advanced approach could include the generation of 3D hyperspectral information to guide the local illuminant estimation process.

A Appendix

The corresponding Python code of the proposed framework is available on request at:

`https://github.com/CaptainCornflakes/PCA_based_illuminant_estimation_framework`

B | Appendix


Spectrum
«class»
Attributes:
- name : string
- wavelength : list
- spectralRange : list
- samplingInterval : int
- channels : int
- data : numpy.array
- dataShape : list
- processingMode : string
- pca : PcaObj
- gamut : scipy.spatial.ConvexHull
- ransac : numpy.array
- dataPCS : numpy.array
Functions:
+ load_dataset(name : string)
+ normalize_data(range : list)
+ crop_spectralRange(range : list)
+ resample_data(wl : list)
+ set_metadata(obj : Spectrum)
+ pop_spectrum(idx : int)
+ plot_spectrum(title : string)
inherits
SpectralCube
«class»
Attributes:
- filename : string
- filepath : string
- id : int
- cameraModel : string
- integrationTime : double
- width : int
- height : int
- spd : Spectrum
- minConstraint : numpy.array
- spectralCandidates : numpy.array
- image : numpy.array
- hyperplane : Hyperplane
- volumeIntersect : scipy.spatial.ConvexHull
- spdEstimated : numpy.array
Functions:
+ _handle_paths(filepath : string, filename : string)
+ load_from_dir(path : string)
+ load_from_file(filepath : string, filename : string)
+ save_to_file(out_dir ; string, filename : string)
+ crop_cube(x0 : string, x1 : string, y0 : string, y1 : string)
+ calc_radiance(spd : Spectrum)
+ calc_reflectance(spd : Spectrum)
+ calc_minConstaint()
+ calc_spectralCandidates(n : int, smoothing_method : string, smoothing : float)
+ calc_image()
+ show_channel(idx : int)
+ get_spectrum(pos : list)
+ plot_spectrum(pos : list)

PcaObj
«class»

Attributes:
- name : *string*
- processingMode : *string*
- data_train : *numpy.array* - pca : *sklearn.decomposition.PCA*
- pcaSscaler : *sklearn.preprocessing.StandardScaler*

Functions:
+ pcaScaler_fit(data : *numpy.array*)
+ pcaScaler_transform_data(data : *numpy.array*) : *numpy.array*
+ pcaScaler_inverse_transform(data : *numpy.array*) : *numpy.array*
+ pca_fit(data : *numpy.array*)
+ pca_transform_data(data : *numpy.array*) : *numpy.array*
+ pca_inverse_transform(data : *numpy.array*) : *numpy.array*

Hyperplane
«class»

Attributes:
- centroid : *numpy.array*
- normal : *numpy.array*
- points : *numpy.array*
- dims : *int*

Functions:
+ calc_from_points(points : *numpy.array*)
+ calc_intersection(p1 : *numpy.array*, p2 : *numpy.array*) : *numpy.array*
+ point_inside(point : *numpy.array*, hull : scipy.spatial.ConvexHull)

utils_COLOR
«package»

Functions:
- calc_XYZ(obj : *Spectrum*, cmfs : *numpy.array*, spd : *Spectrum*) : numpy.array
- calc_cct_from_spectra(obj : *Spectrum*) : float
- calc_inverse_cct_from_spectra(obj : *Spectrum*) : *double*
- cct_to_inverse_cct(cct : *double*) : *double*
- inverse_cct_to_cct(inv_cct : *double*) : *double*

utils_GEO
«package»

Functions:
- fit_line_ransac(X : *numpy.array*, y : *numpy.array*, max_trials : *int*) : *numpy.array*
- line_hyperplane_intersection(data : *numpy.array*) : *numpy.array*
- find_nearest_point(p : *numpy.array*, set_p : *numpy.array*) : *numpy.array*

utils_EVAL
«package»

Functions:
- GFC(sample : *numpy.array*, reference : *numpy.array*) : *double*
- CGFC(sample : *numpy.array*, reference : *numpy.array*) : *double*
- RMSE(sample : *numpy.array*, reference : *numpy.array*) : *double*
- IRE(sample : *numpy.array*, reference : *numpy.array*) : *double*
- SAM(sample : *numpy.array*, reference : *numpy.array*) : *double*

colorSpace_transforms
«package»

Functions:
- sRGB2sRGBlin(data : *numpy.array*) : *numpy.array*
- sRGBlin2sRGB(data : *numpy.array*) : *numpy.array*
- sRGB2XYZ(data : *numpy.array*) : *numpy.array*
- XYZ2sRGB(data : *numpy.array*) : *numpy.array*
- XYZ2xyY(data : *numpy.array*) : *numpy.array*
- xyY2XYZ(data : *numpy.array*) : *numpy.array*

spectral_constancy_algorithms
«package»

Functions:
- spectral_grayworld(data : *numpy.array*) : *numpy.array*
- max_spectral(data : *numpy.array*) : *numpy.array*
- spectral_grayedge(data : *numpy.array*) : *numpy.array*

C | Appendix

Table C.1: *Results of the estimation of the illuminant SPD using three components for reconstruction from PC-space, mean values for all 150 estimated illuminants of the test dataset.*

	3 principal components	6 principal components
CGFC		
min	**0.0026**	0.0038
mean	**0.0219**	0.0616
max	**0.0832**	0.1271
90^{th}pctl	**0.0508**	0.1051
RMSE		
min	**0.0568**	0.0918
mean	**0.1593**	0.2325
max	**0.3685**	0.4149
90^{th}pctl	**0.2795**	0.3234
IRE		
min	0.0022	**0.0017**
mean	0.2043	**0.1923**
max	0.6775	**0.6125**
90^{th}pctl	0.4464	**0.3599**
SAM		
min	**0.0726**	0.0874
mean	**0.1889**	0.3341
max	**0.4109**	0.5097
90^{th}pctl	**0.3200**	0.4625

Bibliography

Aasen, H., Burkart, A., Bolten, A., and Bareth, G. (2015). Generating 3d hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance. *ISPRS Journal of Photogrammetry and Remote Sensing*, 108:245–259. (cited on pages 9, 17, and 18)

Abdelbaki, A., Schlerf, M., Retzlaff, R., Machwitz, M., Verrelst, J., and Udelhoven, T. (2021). Comparison of Crop Trait Retrieval Strategies Using UAV-Based VNIR Hyperspectral Imaging. *Remote Sensing*, 13(9):1748. (cited on pages 9 and 18)

Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., and Sousa, J. (2017). Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. *Remote Sensing*, 9(11):1110. (cited on page 9)

Amziane, A., Losson, O., Mathon, B., Dumenil, A., and Macaire, L. (2020). Frame-based reflectance estimation from multispectral images for weed identification in varying illumination conditions. In *2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA)*. IEEE. (cited on page 18)

An, D., Suo, J., Wang, H., and Dai, Q. (2015). Illumination estimation from specular highlight in a multi-spectral image. *Optics Express*, 23(13):17008. (cited on page 22)

Arroyo-Mora, J. P., Kalacska, M., Løke, T., Schläpfer, D., Coops, N. C., Lucanus, O., and Leblanc, G. (2021). Assessing the impact of illumination on UAV pushbroom hyperspectral imagery collected under various cloud cover conditions. *Remote Sensing of Environment*, 258:112396. (cited on page 19)

Audebert, N., Saux, B. L., and Lefevre, S. (2019). Deep Learning for Classification of Hyperspectral Data: A Comparative Review. *IEEE Geoscience and Remote Sensing Magazine*, 7(2):159–173. (cited on page 16)

Bajić, M. and Bajić, M. (2021). Modeling and Simulation of Very High Spatial Resolution UXOs and Landmines in a Hyperspectral Scene for UAV Survey. *Remote Sensing*, 13(5):837. (cited on page 9)

Banerjee, A., Burlina, P., and Broadwater, J. (2009). Hyperspectral video for illumination-invariant tracking. In *2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*. IEEE. (cited on page 22)

Baone, G. A. and Qi, H. (2006). Demosaicking methods for multispectral cameras using mosaic focal plane array technology. In Rosen, M. R., Imai, F. H., and Tominaga, S., editors, *SPIE Proceedings*. SPIE. (cited on page 14)

Barber, C. B., Dobkin, D. P., and Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software*, 22(4):469–483. (cited on page 40)

Berns, R. S. (2019). *Billmeyer and Saltzman's Principles of Color Technology, 4th Edition*. Wiley & Sons, Incorporated, John. (cited on page 8)

Bilal, M., Nazeer, M., Nichol, J. E., Bleiweiss, M. P., Qiu, Z., Jäkel, E., Campbell, J. R., Atique, L., Huang, X., and Lolli, S. (2019). A Simplified and Robust Surface Reflectance Estimation Method (SREM) for Use over Diverse Land Surfaces Using Multi-Sensor Data. *Remote Sensing*, 11(11):1344. (cited on page 9)

Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1):1–26. (cited on page 19)

Calin, M. A., Parasca, S. V., Savastru, D., and Manea, D. (2013). Hyperspectral Imaging in the Medical Field: Present and Future. *Applied Spectroscopy Reviews*, 49(6):435–447. (cited on page 9)

Chadnov, R. and Skvortsov, A. (2004). Convex hull algorithms review. In *Proceedings. The 8th Russian-Korean International Symposium on Science and Technology, 2004. KORUS 2004.* IEEE. (cited on pages 39 and 40)

Chen, Y., Hou, C., Tang, Y., Zhuang, J., Lin, J., He, Y., Guo, Q., Zhong, Z., Lei, H., and Luo, S. (2019). Citrus Tree Segmentation from UAV Images Based on Monocular Machine Vision in a Natural Orchard Environment. *Sensors*, 19(24):5558. (cited on page 4)

Cihan, M., Ceylan, M., and Ornek, A. H. (2022). Spectral-spatial classification for non-invasive health status detection of neonates using hyperspectral imaging and deep convolutional neural networks. *Spectroscopy Letters*, 55(5):336–349. (cited on page 9)

Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, 74(368):829–836. (cited on page 30)

Cutajar, J. D., Babini, A., Deborah, H., Hardeberg, J. Y., Joseph, E., and Frøysaker, T. (2022). Hyperspectral Imaging Analyses of Cleaning Tests on Edvard Munch's Monumental Aula Paintings. *Studies in Conservation*, 67(sup1):59–68. (cited on page 9)

Czech, M., Trumpy, G., and Syed, A. R. (2023). Do-It-Yourself LUT-based Linearization of Image Sensors. In *Archiving Conference (publication pending)*. (cited on page 18)

Dale, L. M., Thewis, A., Boudry, C., Rotar, I., Dardenne, P., Baeten, V., and Pierna, J. A. F. (2013). Hyperspectral Imaging Applications in Agriculture and Agro-Food Product Quality and Safety Control: A Review. *Applied Spectroscopy Reviews*, 48(2):142–159. (cited on page 9)

Dash, J. P., Watt, M. S., Pearse, G. D., Heaphy, M., and Dungey, H. S. (2017). Assessing very high resolution UAV imagery for monitoring forest health during a simulated disease outbreak. *ISPRS Journal of Photogrammetry and Remote Sensing*, 131:1–14. (cited on page 4)

Deborah, H., George, S., and Hardeberg, J. Y. (2019). Spectral-divergence based pigment discrimination and mapping: A case study on *The Scream* (1893) by Edvard Munch. *Journal of the American Institute for Conservation*, 58(1-2):90–107. (cited on page 12)

Derhak, M. and Rosen, M. (2006). Spectral Colorimetry using LabPQR: An Interim Connection Space. *Journal of Imaging Science and Technology*, 50(1):53–63. (cited on page 4)

Descour, M. and Dereniak, E. (1995). Computed-tomography imaging spectrometer: experimental calibration and reconstruction results. *Applied Optics*, 34(22):4817. (cited on page 14)

Du, Q. and Fowler, J. E. (2007). Hyperspectral Image Compression Using JPEG2000 and Principal Component Analysis. *IEEE Geoscience and Remote Sensing Letters*, 4(2):201–205. (cited on page 32)

Ebner, M. (2007). *Color Constancy*. Wiley. (cited on page 19)

Finlayson, G. D. and Trezzi, E. (2004). Shades of gray and colour constancy. In *Color and Imaging Conference*, volume 2004, pages 37–41. Society for Imaging Science and Technology. (cited on pages 19 and 20)

Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus. *Communications of the ACM*, 24(6):381–395. (cited on page 36)

Foster, D. H. (2011). Color constancy. *Vision Research*, 51(7):674–700. (cited on page 19)

Fotiadou, K., Tsagkatakis, G., and Tsakalides, P. (2019). Spectral Super Resolution of Hyperspectral Images via Coupled Dictionary Learning. *IEEE Transactions on Geoscience and Remote Sensing*, 57(5):2777–2797. (cited on page 14)

Fredembach, C. and Finlayson, G. (2008). Bright Chromagenic Algorithm for Illuminant Estimation. *Journal of Imaging Science and Technology*, 52(4):40906–1–40906–11. (cited on page 21)

Gijsenij, A., Gevers, T., and van de Weijer, J. (2011). Computational Color Constancy: Survey and Experiments. *IEEE Transactions on Image Processing*, 20(9):2475–2489. (cited on page 23)

Gijsenij, A., Lu, R., and Gevers, T. (2012). Color Constancy for Multiple Light Sources. *IEEE Transactions on Image Processing*, 21(2):697–707. (cited on page 68)

Goetz, A. F. (2009). Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sensing of Environment*, 113:S5–S16. (cited on page 9)

Goldstein, E. B. (2010). *Sensation and perception*. Wadsworth, Cengage Learning. (cited on page 22)

Grillini, F., Thomas, J.-B., and George, S. (2021). Comparison of Imaging Models for Spectral Unmixing in Oil Painting. *Sensors*, 21(7):2471. (cited on pages 9 and 23)

Hagen, N. (2012). Snapshot advantage: a review of the light collection improvement for parallel high-dimensional measurement systems. *Optical Engineering*, 51(11):111702. (cited on page 14)

Hagen, N. and Kudenov, M. W. (2013). Review of snapshot spectral imaging technologies. *Optical Engineering*, 52(9):090901. (cited on pages 9, 11, 12, 14, and 15)

Hakala, T., Markelin, L., Honkavaara, E., Scott, B., Theocharous, T., Nevalainen, O., Näsi, R., Suomalainen, J., Viljanen, N., Greenwell, C., and Fox, N. (2018). Direct Reflectance Measurements from Drones: Sensor Absolute Radiometric Calibration and System Tests for Forest Reflectance Characterization. *Sensors*, 18(5):1417. (cited on page 19)

Hernández-Andrés, J., Lee, R. L., and Romero, J. (1999). Calculating correlated color temperatures across the entire gamut of daylight and skylight chromaticities. *Appl. Opt.*, 38(27):5703–5709. (cited on pages 19 and 46)

Hernández-Andrés, J., Romero, J., and Lee, R. L. (2001a). Colorimetric and spectroradiometric characteristics of narrow-field-of-view clear skylight in Granada, Spain. *Journal of the Optical Society of America A*, 18(2):412. (cited on pages 32 and 65)

Hernández-Andrés, J., Romero, J., Nieves, J. L., and Lee, R. L. (2001b). Color and spectral analysis of daylight in southern europe. *Journal of the Optical Society of America A*, 18(6):1325. (cited on pages 42, 44, 46, 47, 56, and 93)

Hu, J.-F., Huang, T.-Z., Deng, L.-J., Dou, H.-X., Hong, D., and Vivone, G. (2022). Fusformer: A Transformer-Based Fusion Network for Hyperspectral Image Super-Resolution. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5. (cited on page 14)

Hubold, M., Berlich, R., Gassner, C., Brüning, R., and Brunner, R. (2018). Ultra-compact micro-optical system for multispectral imaging. In *MOEMS and Miniaturized Systems XVII*. SPIE. (cited on page 14)

Hunt, R. W. G. (1978). Colour terminology. *Color Research & Application*, 3(2):79–87. (cited on page 46)

Imai, F. H., Rosen, M. R., and Berns, R. S. (2002). Comparative study of metrics for spectral match quality. In *Conference on colour in graphics, imaging, and vision*, volume 2002, pages 492–496. Society for Imaging Science and Technology. (cited on page 51)

Jablonski, J., Durell, C., Slonecker, T., Wong, K., Simon, B., Eichelberger, A., and Osterberg, J. (2016). Best practices in passive remote sensing VNIR hyperspectral system hardware calibrations. In Bannon, D. P., editor, *Hyperspectral Imaging Sensors: Innovative Applications and Sensor Standards 2016*. SPIE. (cited on page 17)

Jayaram, M. and Fleyeh, H. (2016). Convex hulls in image processing: a scoping review. *American Journal of Intelligent Systems*, 6(2):48–58. (cited on page 40)

Jia, J., Chen, J., Zheng, X., Wang, Y., Guo, S., Sun, H., Jiang, C., Karjalainen, M., Karila, K., Duan, Z., Wang, T., Xu, C., Hyyppa, J., and Chen, Y. (2022). Tradeoffs in the Spatial and Spectral Resolution of Airborne Hyperspectral Imaging Systems: A Crop Identification Case Study. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–18. (cited on page 9)

Jia, J., Wang, Y., Chen, J., Guo, R., Shu, R., and Wang, J. (2020). Status and application of advanced airborne hyperspectral imaging technology: A review. *Infrared Physics & Technology*, 104:103115. (cited on page 9)

Jolliffe, I. T. (2002). *Principal component analysis for special types of data*. Springer. (cited on page 32)

Judd, D. B., MacAdam, D. L., Wyszecki, G., Budde, H. W., Condit, H. R., Henderson, S. T., and Simonds, J. L. (1964). Spectral Distribution of Typical Daylight as a Function of Correlated Color Temperature. *J. Opt. Soc. Am.*, 54(8):1031–1040. (cited on page 19)

Jung, A. (2017). HYPERSPECTRAL IMAGING. In *Digital Techniques for Documenting and Preserving Cultural Heritage*, pages 217–220. Arc Humanities Press. (cited on pages 3 and 14)

Kaur, H. and Sharma, S. (2016). A comparative review of various illumination estimation based color constancy techniques. In *2016 International Conference on Communication and Signal Processing (ICCSP)*. IEEE. (cited on page 16)

Khan, A., Vibhute, A. D., Mali, S., and Patil, C. (2022). A systematic review on hyperspectral imaging technology with a machine and deep learning methodology for agricultural applications. *Ecological Informatics*, 69:101678. (cited on page 23)

Khan, H. A., Thomas, J.-B., and Hardeberg, J. Y. (2017a). Analytical Survey of Highlight Detection in Color and Spectral Images. In *Lecture Notes in Computer Science*, pages 197–208. Springer International Publishing. (cited on page 23)

Khan, H. A., Thomas, J. B., and Hardeberg, J. Y. (2017b). Multispectral constancy based on spectral adaptation transform. In *Image Analysis*, pages 459–470. Springer International Publishing. (cited on page 20)

Khan, H. A., Thomas, J.-B., Hardeberg, J. Y., and Laligant, O. (2017c). Illuminant estimation in multispectral imaging. *Journal of the Optical Society of America A*, 34(7):1085. (cited on pages 19, 21, and 53)

Khan, H. A., Thomas, J.-B., Hardeberg, J. Y., and Laligant, O. (2018). Spectral Adaptation Transform for Multispectral Constancy. *Journal of Imaging Science and Technology*, 62(2):20504–1–20504–12. (cited on page 19)

Khanna, R., Sa, I., Nieto, J., and Siegwart, R. (2017). On field radiometric calibration for multispectral cameras. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. (cited on page 18)

Koppal, S. J. (2014). *Lambertian Reflectance*, pages 441–443. Springer US, Boston, MA. (cited on page 17)

Kosztyán, Z. and Schanda, J. (2012). Smoothing spectral power distribution of daylights. *Color Research & Application*, 38(5):316–321. (cited on page 30)

Koz, A. (2019). Ground-Based Hyperspectral Image Surveillance Systems for Explosive Detection: Part II—Radiance to Reflectance Conversions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(12):4754–4765. (cited on page 17)

Kruse, F., Lefkoff, A., Boardman, J., Heidebrecht, K., Shapiro, A., Barloon, P., and Goetz, A. (1993). The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44(2-3):145–163. (cited on page 52)

Krüger, N., Janssen, P., Kalkan, S., Lappe, M., Leonardis, A., Piater, J., Rodriguez-Sanchez, A. J., and Wiskott, L. (2013). Deep Hierarchies in the Primate Visual Cortex: What Can We Learn for Computer Vision? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1847–1871. (cited on page 19)

Kurz, T. H., Buckley, S. J., and Howell, J. A. (2012). Close-range hyperspectral imaging for geological field studies: workflow and methods. *International Journal of Remote Sensing*, 34(5):1798–1822. (cited on page 18)

Land, E. H. and McCann, J. J. (1971). Lightness and Retinex Theory. *Journal of the Optical Society of America*, 61(1):1. (cited on pages 19, 20, and 27)

Lapray, P.-J., Wang, X., Thomas, J.-B., and Gouton, P. (2014). Multispectral Filter Arrays: Recent Advances and Practical Implementation. *Sensors*, 14(11):21626–21659. (cited on page 14)

Li, Q., He, X., Wang, Y., Liu, H., Xu, D., and Guo, F. (2013). Review of spectral imaging technology in biomedical engineering: achievements and challenges. *Journal of Biomedical Optics*, 18(10):100901. (cited on pages 10 and 12)

Liang, S. and Wang, J. (2020). *Advanced Remote Sensing Terrestrial Information Extraction and Applications*. Elsevier Science & Technology, 2 edition. (cited on page 8)

Liu, M., Zhang, Z., Liu, X., Yao, J., Du, T., Ma, Y., and Shi, L. (2020). Discriminant Analysis of the Damage Degree Caused by Pine Shoot Beetle to Yunnan Pine Using UAV-Based Hyperspectral Images. *Forests*, 11(12):1258. (cited on pages 4 and 16)

Lu, B., Dao, P., Liu, J., He, Y., and Shang, J. (2020). Recent Advances of Hyperspectral Imaging Technology and Applications in Agriculture. *Remote Sensing*, 12(16):2659. (cited on pages 3 and 9)

Lu, G. and Fei, B. (2014). Medical hyperspectral imaging: a review. *Journal of Biomedical Optics*, 19(1):010901. (cited on page 9)

Lucieer, A., Malenovský, Z., Veness, T., and Wallace, L. (2014). HyperUAS-Imaging Spectroscopy from a Multirotor Unmanned Aircraft System. *Journal of Field Robotics*, 31(4):571–590. (cited on page 16)

Ma, J., Sun, D.-W., Pu, H., Cheng, J.-H., and Wei, Q. (2019). Advanced Techniques for Hyperspectral Imaging in the Food Industry: Principles and Recent Applications. *Annual Review of Food Science and Technology*, 10(1):197–220. (cited on pages 3, 10, 11, and 89)

Marchi, M. D., Diantini, A., and Pappalardo, S. E. (2022). *Drones and Geographical Information Technologies in Agroecology and Organic Farming Contributions to Technological Sovereignty*. Taylor & Francis Group. (cited on page 4)

Maule, J., Skelton, A. E., and Franklin, A. (2023). The Development of Color Perception and Cognition. *Annual Review of Psychology*, 74(1):87–111. (cited on page 19)

McCamy, C. S. (1992). Correlated color temperature as an explicit function of chromaticity coordinates. *Color Research & Application*, 17(2):142–144. (cited on page 45)

Michalsky, J. (1985). Estimation of continuous solar spectral distributions from discrete filter measurements: II. a demonstration of practicability. *Solar Energy*, 34(6):439–445. (cited on page 53)

More, S. S., Beach, J. M., McClelland, C., Mokhtarzadeh, A., and Vince, R. (2019). In Vivo Assessment of Retinal Biomarkers by Hyperspectral Imaging: Early Detection of Alzheimer's Disease. *ACS Chemical Neuroscience*, 10(11):4492–4501. (cited on page 9)

Ozdemir, A., , and and, K. P. (2020). Deep Learning Applications for Hyperspectral Imaging: A Systematic Review. *Journal of the Institute of Electronics and Computer*, 2(1):39–56. (cited on page 23)

Pearson, K. (1905). The Problem of the Random Walk. *Nature*, 72(1865):294–294. (cited on page 30)

Poger, S. and Angelopoulou, E. (2001). Multispectral sensors in computer vision. Technical report, Stevens Institute of Technology. (cited on page 13)

Polak, A., Kelman, T., Murray, P., Marshall, S., Stothard, D. J., Eastaugh, N., and Eastaugh, F. (2017). Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication. *Journal of Cultural Heritage*, 26:1–11. (cited on page 9)

Rady, A., Guyer, D., Kirk, W., and Donis-González, I. R. (2019). Prediction of the Leaf Primordia of Potato Tubers Using Sensor Fusion and Wavelength Selection. *Journal of Imaging*, 5(1):10. (cited on page 9)

Richards, J. A. (2022). *Remote Sensing Digital Image Analysis*. Springer International Publishing. (cited on page 32)

Robertson, A. R. (1968). Computation of Correlated Color Temperature and Distribution Temperature. *Journal of the Optical Society of America*, 58(11):1528. (cited on page 46)

Robles-Kelly, A. and Wei, R. (2018). A Convolutional Neural Network for Pixelwise Illuminant Recovery in Colour and Spectral Images. In *2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE. (cited on page 23)

Romero, J., García-Beltrán, A., and Hernández-Andrés, J. (1997). Linear bases for representation of natural and artificial illuminants. *Journal of the Optical Society of America A*, 14(5):1007. (cited on pages 51, 52, and 63)

Romero, J., Hernández-Andrés, J., Nieves, J. L., and García, J. A. (2002). Color coordinates of objects with daylight changes. *Color Research & Application*, 28(1):25–35. (cited on page 19)

Serranti, S., Gargiulo, A., and Bonifazi, G. (2011). Characterization of post-consumer polyolefin wastes by hyperspectral imaging for quality control in recycling processes. *Waste Management*, 31(11):2217–2227. (cited on page 9)

Sethu, S., Devaraj, J., and Wang, D. (2023). A Comprehensive Review of Deep Learning based Illumination Estimation. *Preprints.org (publication pending)*. (cited on page 23)

Seymour, J. (2022). Color inconstancy in CIELAB: A red herring? *Color Research & Application*, 47(4):900–919. (cited on page 19)

Shafer, S. A. (1985). Using color to separate reflection components. *Color Research & Application*, 10(4):210–218. (cited on page 21)

Shrestha, R. and Hardeberg, J. Y. (2013). Multispectral imaging using LED illumination and an RGB camera. In *Color and Imaging Conference*, volume 2013, pages 8–13. Society for Imaging Science and Technology. (cited on page 12)

Smithson, H. (2005). Sensory, computational and cognitive components of human colour constancy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1458):1329–1346. (cited on page 19)

Snoek, J., Larochelle, H., and Adams, R. P. (2012). Practical Bayesian Optimization of Machine Learning Algorithms. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'12, page 2951–2959, Red Hook, NY, USA. Curran Associates Inc. (cited on page 23)

Stuffler, T., Kaufmann, C., Hofer, S., Förster, K., Schreier, G., Mueller, A., Eckardt, A., Bach, H., Penné, B., Benz, U., and Haydn, R. (2007). The EnMAP hyperspectral imager—an advanced optical payload for future applications in earth observation programmes. *Acta Astronautica*, 61(1-6):115–120. (cited on page 9)

Sun, D.-W. (2010). *Hyperspectral Imaging for Food Quality Analysis and Control*. Elsevier. (cited on pages 8 and 89)

Suomalainen, J., Anders, N., Iqbal, S., Roerink, G., Franke, J., Wenting, P., Hünniger, D., Bartholomeus, H., Becker, R., and Kooistra, L. (2014). A Lightweight Hyperspectral Mapping System and Photogrammetric Processing Chain for Unmanned Aerial Vehicles. *Remote Sensing*, 6(11):11013–11030. (cited on page 17)

Tominaga, S. (1995). Using reflectance models for surface estimation. In *Color and Imaging Conference*, volume 1995, pages 29–33. Society for Imaging Science and Technology. (cited on pages 21, 22, and 89)

Tominaga, S., Hirai, K., and Horiuchi, T. (2020). Spectral Estimation of Multiple Light Sources based on Highlight Detection. *Journal of Imaging Science and Technology*, 64(5):50408–1–50408–9. (cited on page 22)

Tominaga, S., Horiuchi, T., and Kato, Y. (2012). Scene illuminant estimation of multiple light sources. In *Color and Imaging Conference*, volume 2012, pages 47–51. Society for Imaging Science and Technology. (cited on page 22)

Trumpy, G., Hardeberg, J. Y., George, S., and Flueckiger, B. (2021). A multispectral design for a new generation of film scanners. In Groves, R. and Liang, H., editors, *Optics for Arts, Architecture, and Archaeology VIII*. SPIE. (cited on page 12)

Uto, K., Seki, H., Saito, G., and Kosugi, Y. (2013). Characterization of Rice Paddies by a UAV-Mounted Miniature Hyperspectral Sensor System. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2):851–860. (cited on page 18)

van de Weijer, J. and Gevers, T. (2005). Color constancy based on the Grey-edge hypothesis. In *IEEE International Conference on Image Processing 2005*. IEEE. (cited on pages 19 and 20)

Viggiano, J. (2004). Metrics for evaluating spectral matches: a quantitative comparison. In *Conference on Colour in Graphics, Imaging, and Vision*, volume 2004, pages 286–291. Society for Imaging Science and Technology. (cited on page 52)

Wang, F. and Theuwissen, A. (2017). Linearity analysis of a CMOS image sensor. *Electronic Imaging*, 29(11):84–90. (cited on page 18)

Wang, J. (2012). *Geometric Structure of High-Dimensional Data and Dimensionality Reduction*. Springer. (cited on page 37)

Wendel, A. and Underwood, J. (2017). Illumination compensation in ground based hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, 129:162–178. (cited on pages 3, 18, and 26)

Woo, H., Acuna, M., Madurapperuma, B., Jung, G., Woo, C., and Park, J. (2021). Application of Maximum Likelihood and Spectral Angle Mapping Classification Techniques to Evaluate Forest Fire Severity from UAV Multi-spectral Images in South Korea. *Sensors and Materials*, 33(11):3745. (cited on page 4)

Wyszecki, G. and Stiles, W. S. (2000). *Color Science: Concepts and Methods, Quantitative Data and Formulae*, volume 40. John Wiley & Sons. (cited on page 46)

Yan, Deng, Liu, and Zhu (2019). Application of UAV-Based Multi-Angle Hyperspectral Remote Sensing in Fine Vegetation Classification. *Remote Sensing*, 11(23):2753. (cited on page 16)

Yang, G., Li, C., Wang, Y., Yuan, H., Feng, H., Xu, B., and Yang, X. (2017). The DOM Generation and Precise Radiometric Calibration of a UAV-Mounted Miniature Snapshot Hyperspectral Imager. *Remote Sensing*, 9(7):642. (cited on page 17)

Yang, Q., Sun, S., Jeffcoate, W., Clark, D., Musgove, A., Game, F., and Morgan, S. (2018). Investigation of the Performance of Hyperspectral Imaging by Principal

Component Analysis in the Prediction of Healing of Diabetic Foot Ulcers. *Journal of Imaging*, 4(12):144. (cited on pages 9 and 32)

Yao, H. and Lewis, D. (2010). Spectral Preprocessing and Calibration Techniques. In *Hyperspectral Imaging for Food Quality Analysis and Control*, pages 45–78. Elsevier. (cited on pages 3, 16, 17, and 18)

List of Figures

LIST OF FIGURES

List of Tables