Multispectral constancy for illuminant invariant representation of multispectral images

Haris Ahmad Khan

THÈSE

Pour l'obtention du grade de
Docteur de l' Université Bourgogne Franche-Comté
Cotutelle: Norwegian University of Science and Technology (Gjøvik, Norvège)

Présentée par:
Haris Ahmad Khan

Multispectral constancy for illuminant invariant representation of multispectral images

Soutenue le 9 Octobre 2018

Jury:

Président	Nathalie Destouches	Professor, Université Jean Monnet
Rapporteur	Theo Gevers	Professor, University of Amsterdam
Rapporteur	Markku Hauta-Kasari	Professor, University of University of Eastern Finland
Examinateur	Ali Alsam	Associate Professor, Norwegian University of Science and Technology
Co-directeur	Olivier Laligant	Professor, Université Bourgogne Franche-Comté
Co-directeur	Jon Yngve Hardeberg	Professor, Norwegian University of Science and Technology
Co-encadrant	Jean-Baptiste Thomas	Associate Professor, Université Bourgogne Franche-Comté

Cofinancé par l'Union Européenne

Haris Ahmad Khan

Multispectral constancy for illuminant invariant representation of multispectral images

Thesis for the degree of Philisophiae Doctor

Gjøvik, October 2018

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

Université Bourgogne Franche-Comté
Laboratoire Electronique, Informatique et Image

NTNU
Norwegian University of Science and Technology

UBFC
UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ

Thesis for the degree of Philisophiae Doctor

Joint degree between

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

Université Bourgogne Franche-Comté
Laboratoire Electronique, Informatique et Image

ISBN 978-82-326-3371-5 (printed version)
ISBN 978-82-326-3370-8 (electronic version)
ISSN 1503-8181

Doctoral theses at NTNU, 2018:290

Printed by NTNU-Skipnes

Abstract

A conventional color imaging system provides high resolution spatial information and low resolution spectral data. In contrast, a multispectral imaging system is able to provide both the spectral and spatial information of a scene in high resolution. A multispectral imaging system is complex and it is not easy to use it as a hand held device for acquisition of data in uncontrolled conditions. The use of multispectral imaging for computer vision applications has started recently but is not very efficient due to these limitations. Therefore, most of the computer vision systems still rely on traditional color imaging and the potential of multispectral imaging for these applications has yet to be explored.

With the advancement in sensor technology, hand held multispectral imaging systems are coming in market. One such example is the snapshot multispectral filter array camera. So far, data acquisition from multispectral imaging systems require specific imaging conditions and their use is limited to a few applications including remote sensing and indoor systems. Knowledge of scene illumination during multispectral image acquisition is one of the important conditions. In color imaging, computational color constancy deals with this condition while the lack of such a framework for multispectral imaging is one of the major limitation in enabling the use of multispectral cameras in uncontrolled imaging environments.

In this work, we extend some methods of computational color imaging and apply them to the multispectral imaging systems. A major advantage of color imaging is the ability of providing consistent color of objects and surfaces across varying imaging conditions. In this work, we extend the concept of color constancy and white balancing from color to multispectral images, and introduce the term *multispectral constancy*.

The validity of proposed framework for consistent representation of multispectral images is demonstrated through spectral reconstruction of material surfaces from the acquired images. We have also presented a new hyperspectral reflectance images dataset in this work. The framework of multispectral constancy will make it one step closer for the use of multispectral imaging in computer vision applications, where the spectral information, as well as the spatial information of a surface will be able to provide distinctive useful features for material identification and classification tasks.

Translations of the Abstract

The following translations of the abstract have been kindly provided by the supervisors, Jon Yngve Hardeberg (Norwegian translation) and Jean-Baptiste Thomas (French translation).

Sammendrag (Norwegian translation)

Et konvensjonelt fargebasert avbildningssystem kan gi høy romlig oppløsning men lav spektral oppløsning . Et multispektralt avbildningssystem, derimot, kan gi både spektral og romlig informasjon om en scene i høy oppløsning. Et multispektralt avbildningssystem er komplekst, og det er ikke lett å bruke det som en håndholdt enhet for opptak av data under ukontrollerte forhold. Bruken av multispektral avbildning for maskinsynanvendelser har startet nylig, men er ikke veldig effektivt på grunn av disse begrensningene. Derfor er de fleste maskinsynssystemer fortsatt basert på tradisjonell fargebildeteknologi og potensialet for multispektral avbildning for disse anvendelsene har ennå ikke blitt utnyttet.

Gjennom teknologiske fremskritt innen sensorteknologi er håndholdte multispektrale avbildningssystemer i ferd med å komme på markedet. Et eksempel er kameraer basert på multispektrale filtermatriser. Så langt har datainnsamling fra multispektrale avbildningsssystemer blitt gjort under kontrollerte avbildningsforhold, og bruken av dem er derfor begrenset til noen få applikasjoner, inkludert fjernnmåling og laboratoriebaserte systemer. Informasjon om belysningen ved multispektral avbildning er et av de viktige forholdene. Innen fargeavbildning håndterer fargekonstansalgoritmer dette problemet, men mangelen på et slikt rammeverk for multispektral avbildning er en av de største begrensningene for å muliggjøre bruk av multispektrale kameraer under ukontrollerte avbildningsforhold.

I dette arbeidet foreslår vi å utvide eksisterende metoder for fargekonstansberegning og anvende dem for multispektral avbildning. En stor fordel ved fargeavbildning er evnen til å gi konsistent farge på gjenstander og overflater på tvers av varierende bildeforhold. I dette arbeidet utvider vi konseptet fargekonstans og hvitbalanse fra farge til multispektrale bilder, og introduserer begrepet multispektral konstans. Gyldigheten av det foreslåtte rammeverket for konsistent representasjon av multispektrale bilder er demonstrert gjennom spektral rekonstruksjon av materialoverflater fra de multispektrale bildene. Vi har også presentert et nytt hyperspektralt bildedatasett i dette arbeidet. Rammeverket for multispektral konstans vil muliggjøre bruk av multispektral bildebehandling i maskinsynapplikasjoner, hvor spektral informasjon, sammen med romlig informasjon av en overflate, vil kunne gi spesifikke nyttige egenskaper for materialidentifikasjon og klassifiseringssoppgaver.

Résumé (French translation)

En imagerie couleur, un système d'acquisition capture une scène avec une haute résolution spatiale mais une résolution spectrale limitée. L'imagerie hyperspectrale permet d'acquérir la scène avec une grande résolution spectrale. Un système d'acquisition hyperspectrale est un ensemble complexe et il est difficile de l'utiliser pour acquérir des données dans une situation où les conditions d'imageries ne sont pas contrôlées. De plus, ces systèmes sont chers et souvent encombrants ou difficiles à manipuler. À cause de ces problèmes, l'utilisation de l'imagerie hyperspectrale n'a pas encore été beaucoup utilisée en vision assistée par ordinateur, et la plupart des systèmes de vision utilise l'imagerie couleur.

L'imagerie multispectrale propose une solution intermédiaire, elle permet de capturer une information moins résolue selon la dimension spectrale, comparée à l'hyperspectrale, tout en préservant la résolution spatiale. Ces systèmes sont moins encombrants et moins difficiles à maitriser grâce aux récentes avancées technologiques, et arrivent sur le marché en tant que produits commerciaux. On peut citer les matrices de filtres spectraux (spectral filter arrays) qui permettent l'acquisition en temps réel d'images multispectrales grâce à l'utilisation d'une caméra de complexité similaire à une caméra couleur. Jusqu'ici, les informations capturées par ces systèmes étaient considérées de la même manière que les imageurs hyperspectraux en champ proche, c'est à dire que pour utiliser l'information au mieux, les conditions d'acquisitions devaient être connues et le système calibré, en particulier pour l'éclairage de la scène et la dynamique de la scène.

Afin d'élargir l'utilisation de l'imagerie multispectrale pour la vision par ordinateur dans des conditions générales, je propose dans cette thèse de développer les méthodes calculatoires en imagerie couleur (computational color imaging) et de les adapter aux systèmes d'imagerie multispectraux. Une caractéristique très puissante de l'imagerie couleur est de proposer un rendu constant des couleurs de la surface d'un objet à travers différentes conditions d'acquisition via l'utilisation d'algorithmes et divers traitements de l'information.

Dans cette thèse, j'étends la notion de constance des couleurs et de balance des blancs de l'imagerie couleur à l'imagerie multispectrale. J'introduis le terme de constance de l'information spectrale (multispectral constancy).

Je propose la construction d'un ensemble d'outils permettant la représentation constante de l'information spectrale à travers le changement d'éclairage. La validité de ces outils est évaluée à travers la reconstruction de la réflectance spectrale des objets lorsque l'éclairage change. Nous avons également acquis de nouvelles images hyperspectrales et multispectrales mises à disposition de la communauté.

Ces outils et données permettront de favoriser la généralisation de l'utilisation de l'imagerie multispectrale en champ proche dans les applications classiques utilisant traditionnellement l'imagerie couleur et de sortir ce mode d'imagerie des laboratoires. L'avantage en vision par ordinateur est une meilleure analyse de la réflectance de la surface des objets et donc un avantage certain dans les tâches de classification et d'identification de matériaux.

Acknowledgments

First of all, I want to express my gratitude to my supervisors from both NTNU Gjøvik and Université Bourgogne Franche-Comté: Professor Jon Y. Hardeberg, Associate Professor Jean-Baptiste Thomas, and Professor Olivier Laligant. Thank you for the invaluable feedback, support, and opportunities given to me throughout my study. I have learnt a lot from all of you, not only academically but also as a person.

Being in a co-tutelle program has exposed me to two different cultural norms at The Norwegian Colour and Visual Computing Laboratory (Colorlab) and Laboratoire Electronique, Informatique et Image (Le2i). It has certainly been a great experience to have in such a short time. I am thankful to the management of both institutes for helping me in the complicated matters related to dual degree program. Thank you, because without your help, I would not have been able to work on my research in peace. I am thankful for the lecturers whose courses I had attended: Professor Marius Pedersen, Professor Phil Green, Professor Stephen Wolthusen and Professor Einar Snekkenes. Thank you for the guidance and invaluable discussions. The challenging and highly demanding courses are worth the knowledge I gained from them.

I met and made great friends in both places. Thanks Jessica and Margarita for helping me get settled in the initial days in France and your help in dealing with French administration. Thanks to Danish and Babar for welcoming me in Dijon, and Ahmed, Mohib and Zeeshan for being my travel buddies in Norway. I would like to mention and thank Malik Anas Ahmad specially, since he is the one who helped me during my Bachelor's final year project and guided me in the selection of track for Masters. I am thankful to Nabila, who motivated me to work in computer vision applications at the same time. I am thankful to all my friends who kept in contact and did not let me feel alone.

I am extremely blessed to have these people as my family: My father, Mama, Hafsah, Hassan and Maryam. My parents have been constantly supporting me throughout my entire life and I am in the current position because of them. I am thankful to Maryam for understanding me, being helpful in the time when I was stressed, and for her sacrifices.

Last but not least, this PhD research work has been made possible by the funding from Conseil Régional de Bourgogne, France and the Research Council of Norway. Thanks to Jon, Jean-Baptiste, Olivier and Marius for taking care of the financial aspects during my research.

Contents

Chapter 1

Introduction

1.1 Background

Digital imaging has evolved tremendously in the last few decades. The amount of visual information acquired in the form of digital images is huge and imaging devices are easily available due to the advancement in computational power, optics, sensor technology, memory management and several other factors. The importance of imaging has been recognized throughout the history of mankind, as the visual sense is regarded as one of the dominant senses. That's why there is the famous saying that *a picture is worth thousand words*.

Most digital imaging systems are based on a small number of channels that are sensitive to relatively wide bands along the visible wavelength spectrum. A system with only one channel can be used to capture grayscale image of a scene with high spatial resolution but no spectral information. By using three channels in an imaging system, color images are be acquired, providing high spatial resolution information of the scene.

Although color images are able to provide details of the objects and surfaces in a scene and enable humans to detect and recognize objects, the spectral information content is limited. Spectral reflectance is the intrinsic property of materials and surfaces. It depends on the physical and chemical properties of a material, color, roughness and the geometric properties of a surface. The reflectance from a surface is unique for each material and is called the "spectral signature". This property makes it possible to identify the material surfaces. The electromagnetic energy reflected back from a surface is called radiance. This radiance is captured by imaging devices and is composed of a mixture of the spectral reflectance of

material and the incoming energy. For computation of spectral reflectance of a surface, the imaging system is generally calibrated for the incoming radiation and the system required to be re-calibrated when the imaging conditions are changed.

The human visual system is able to interpret grayscale and color images, therefore majority of imaging systems focus on these two types. Both of these imaging techniques tend to work well in applications where the spectral information, color accuracy and color control are not important.

Multispectral imaging provides more spectral information as compared to color images. With the advancement in sensor technology and imaging techniques, devices are developed that are able to capture more information as compared to color images, while still providing high resolution spatial information of the scene. A multispectral imaging system can acquire spectral data in the visible, ultraviolet and infrared regions. These systems use a set of filters over the imaging sensor to capture scene information in a particular spectral region. Due to the higher number of channels, output of multispectral images is often interpreted and stored as a *datacube*. The output of each channel is stacked over each other and the information for one pixel provide the spectral characteristics of captured radiance from the imaged surface.

1.2 Motivation

Our research focuses on multispectral imaging and the aim is to enable the use of multispectral cameras in uncontrolled imaging environments. Most of the use for multispectral cameras is in remote sensing and indoor imaging in controlled imaging conditions. The reason for such limitation is the requirement of system's calibration for specific imaging environment. One of the factors that cause change in imaging conditions is the illumination in a scene, and therefore an illuminant invariant imaging system is desired when the reflectance information from a scene is required.

Color imaging systems are able to operate in uncontrolled imaging conditions because of the processing techniques of the captured data. One of the key elements among them is white balancing, which deals with the correction of color casts that occur due to the influence of illumination color. In this way, the objects with white color are rendered white in the final image and the colors are corrected in a way that the human visual system expects to see. This correction of colors is called *color constancy* and is widely used in digital color images. The ability of providing illuminant invariant imaging data is one of the key factors that has made color imaging widely popular for photography and for computer vision applications, where the data is used.

Our motivation for this research is to enable the use of multispectral imaging in the same way as color imaging for uncontrolled imaging environments. We want to develop a framework for exploiting the rich spectral information obtained from multispectral imaging. To do so, the techniques of computational color imaging are extended for multispectral imaging. So far, multispectral imaging systems are calibrated for specific imaging conditions and our aim is to introduce a dynamic calibration system in the imaging pipeline. As described above, one of the factors contributing to the wide acceptability of color imaging is the ability of representation of consistent colors in uncontrolled imaging environments. The aim of this thesis is to provide the framework for enabling the use of multispectral imaging outside the laboratory conditions and to obtain consistent representation of multispectral data across different imaging environments.

1.3 Research Objectives

The main objective of this study is to enable the use of multispectral imaging for uncontrolled imaging conditions. A change in the illumination of a scene causes change in the output of camera data for the same imaging system and surfaces being captured. Therefore, in order to acquire consistent presentation of multispectral data across varying illumination conditions, a framework for illuminant-invariant representation of multispectral data needs to be developed.

For obtaining illuminant invariant representation of multispectral images, a framework is developed in this work. It is named as *multispectral constancy*. This framework use the concepts of computational color constancy and extends them for high dimensional spectral data. The core concept of multispectral constancy is the transformation of multispectral data taken under unknown illumination conditions into a representation as if the data is taken under a known (canonical) illuminant. For such representation, the acquired image data has to be transformed. The multispectral image transformation methods developed in this thesis involves incorporating the spectral characteristics of imaging sensor and the compensation for the illuminant. For a given sensor configuration, a *spectral adaptation transform* is developed for incorporating the intrinsic properties of imaging sensor. Details of this transform and demonstration the influence of sensor sensitivities of an imaging system on the obtained data are provided in Chapters 4 and 5. The second part of multispectral constancy deals with scene illumination, which can be obtained either by placing a white diffuser in the scene or by estimating it. For estimation of scene illumination in multispectral images, a study of the potential algorithms that are developed for color images is performed.

Finally, methods developed during this study are demonstrated for spectral reconstruction of material surfaces. Images of various objects under different illumina-

tions are acquired. We use the concept of multispectral constancy and demonstrate its ability to perform under uncontrolled imaging conditions.

We realize that an efficient multispectral imaging system should be able to retain the spectral and spatial properties of material surfaces being imaged. Keeping this in view, a high spatial and spectral resolution data set of hyperspectral images is created. These hyperspectral images contain various samples from textile, wood, rocks, food items and vegetation. This hyperspectral image dataset is made available for public use and will help in enabling the research community to develop algorithms for spectral and spatial analysis using this data. Examples of such analysis include spectral decomposition of the high resolution hyperspectral data and texture analysis.

To summarize, we develop this work with the aim of enabling the use of multispectral imaging systems for uncontrolled illumination conditions. To achieve this goal, we propose illuminant invariant representation of multispectral imaging. Such a representation is obtained through multispectral constancy, which involves illuminant estimation and spectral adaptation transform.

1.4 List of published articles

Throughout the span of this study, a number of articles have been published, see the list below.

Article A H. A. Khan, J. B. Thomas, J. Y. Hardeberg, and O. Laligant, “Spectral adaptation transform for multispectral constancy,” *Journal of Imaging Science and Technology*, vol. 62, no. 2, pp 020504-1-020504-12, 2018.

Article B H. A. Khan, J. B. Thomas, J. Y. Hardeberg, and O. Laligant, “Illuminant estimation in multispectral imaging,” *Journal of the Optical Society of America A*, vol. 34, no. 6, pp 1085-1098, 2017.

Article C H. A. Khan, J. B. Thomas and J. Y. Hardeberg, "Analytical survey of highlight detection in color and spectral images", in Proceedings of *Computational Color Imaging Workshop, CCIW*, Milan, Italy, Lecture Notes in Computer Science, vol 10213. Springer, Cham, 2017.

Article D H. A. Khan, J. B. Thomas and J. Y. Hardeberg, "Towards highlight based illuminant estimation in multispectral images," in Proceedings of *International Conference on Image and Signal Processing*, Cherbourg, France, Lecture Notes in Computer Science, vol 10884. Springer, Cham, 2018.

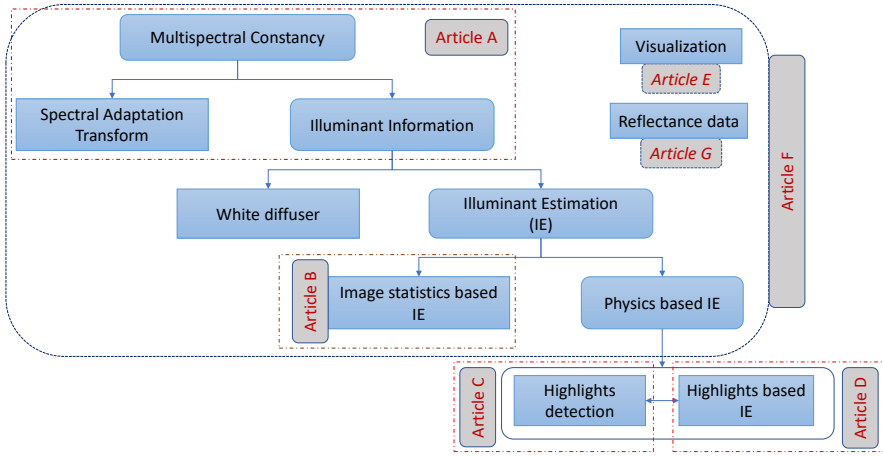


Figure 1.1: Research outline and publications. Article A provides the idea and framework for multispectral constancy that involve spectral adaptation transform and illuminant information. The later part involves the of white diffuser in a scene or illuminant estimation, which is further divided into two parts. Article B describes the image statistics based illuminant estimation methods, while Article C and D describe highlights detection methods, and illuminant estimation from those highlights, respectively. Article F provides a demonstration of the overall concept of multispectral constancy, involving multispectral image visualization (Article E) and spectral data for experiments (Article G).

Article E H. A. Khan, and P. Green, "Color characterization methods for a multispectral camera," in Proceedings of *IS&T International Symposium on Electronic Imaging: Color Imaging XXIII: Displaying, Processing, Hardcopy, and Applications*, San Francisco, USA, 2018.

Article F H. A. Khan, J. B. Thomas and J. Y. Hardeberg, "Towards the use of multispectral camera as spectrophotometer in uncontrolled illumination conditions", *draft paper*.

Article G H. A. Khan, S. Mihoubi, B. Mathon, J. B. Thomas and J. Y. Hardeberg, “HyTexiLa: High resolution visible and near infrared hyperspectral texture images,” *Sensors*, vol. 18, no. 7, 2045; 2018.

1.5 Thesis organization

This thesis is organized in the form of compilation of published articles. Fig. 1.1 gives an overview of the research outline and the publications originated from spe-

cific modules of this research. Chapter 2 provides an overview and background of the research carried out in this work. Discussion on the contributions in this thesis is present in Chapter 3. Chapter 4 consist of **Article A**. In this chapter, the concept of multispectral constancy and data transformation is provided. The preliminary results of this work is also presented in a conference paper[1]. Since the material in this conference publication is already covered in **Article A**, therefore it is not included as part of this dissertation Illuminant estimation in multispectral imaging is discussed in Chapters 5 and 7. Chapter 5 is based on the **Article B**. In this chapter, image statistics based illuminant estimation methods are discussed along with analysis of different multispectral sensors and number of channels. A review of the specular reflection based illuminant estimation methods in color images are provided in **Article D**. This article forms Chapter 7, where the potential algorithms for highlights based illuminant estimation, that can be extended for multispectral imaging are discussed. For such methods, highlight detection in images is the first step and those techniques are discussed in **Article C** and it forms Chapter 6 of this thesis.

Chapter 8 consist of **Article E** where a discussion on the use of multispectral camera for obtaining the colorimetric values of a scene is provided. This work is later used for visualization of multispectral data. A demonstration of multispectral constancy is provided in Chapter 9, which consist of **Article F**. In Chapter 10, we describe the development of a hyperspectral dataset of reflectance data which is presented in **Article G**. This dataset is created as a part of this research and is made publicly available for research and analysis. Finally, a discussion on the work presented in this thesis and the conclusion is in Chapter 11, along with the future perspectives.

[1] H.A. Khan, J. B. Thomas and J. Y. Hardeberg, "Multispectral constancy based on spectral adaptation transform", *Scandinavian Conference on Image Analysis, SCIA*, pp 459-470, Tromsø, Norway, Lecture Notes in Computer Science, vol 10270. Springer, Cham, 2017.

Chapter 2

Research Overview

2.1 Introduction

In this chapter, an overview of the research work is presented. We start with the brief description of image formation, imaging techniques and multispectral imaging. The research carried out in this thesis is focused on illuminant invariant representation of multispectral images. In color vision, the consistent representation of images is known as color constancy and is discussed in Section 2.4. After providing the fundamental knowledge, Section 2.6 starts with the introduction of multispectral constancy, which is the contribution from the work presented in this dissertation. Throughout the rest of this chapter, the necessary fundamental knowledge for understanding the imaging model and the need for consistent representation of multispectral images is provided.

2.2 Imaging techniques

Recording the interaction of electromagnetic radiations with objects, along with the spatial and volumetric sense is termed as digital imaging [1]. This interaction occurs due to absorption or reflection of electromagnetic radiation from the surface of objects. These electromagnetic radiations contain information across the wavelength spectrum and various sensors are designed to capture information in a particular region. A typical color camera records the radiation ranging from approximately 380 nm to 780 nm while the infrared imaging sensors are able to record the radiations in approximately 800 nm to 1100 nm [2]. X-ray sensors operate in the wavelength spectrum of about 0.01 nm to 10 nm while the positron emission tomography captures the information in gamma rays (about 10^{-11} meters). The focus of this study is on the electromagnetic radiation-based digital imaging in the

visible spectral range and will be developed by introducing the image formation and digital imaging model in the following.

Objects are perceived by their radiance in the visible region of the electromagnetic spectrum. The recorded radiance from a given object depends on its material properties, its shape and location in the scene. Considering non-luminous objects, the intensity, position and spectral characteristics of the scene illuminant also play a major role in image generation. Finally, the spectral sensitivity of filters is another important parameter in image creation. In a simple imaging model with three channels, the image values $\mathbf{f} = (R, G, B)^T$ are dependent on the light source $\mathbf{e}(\lambda)$, surface reflectance $\mathbf{r}(\lambda)$ and camera sensitivity functions $\mathbf{c}(\lambda) = \{r(\lambda), g(\lambda), b(\lambda)\}$, within the visible electromagnetic wavelength spectrum ω as

$$\mathbf{f} = \int_{\omega} \mathbf{e}(\lambda)\mathbf{r}(\lambda)\mathbf{c}(\lambda)d\lambda. \tag{2.1}$$

In the human visual system, there are three cone types in the eye which are sensitive to certain wavelengths [3]. The perception of color is formed when data from these three cones is combined in the brain. Color cameras work on the same principle, having three filters, named as R, G and B filters. They are sensitive to particular regions of wavelengths and their combination forms a color image. The sensor sensitivity of color cameras varies for each manufacturer and model of camera. These sensitivities are not similar to the sensitivities of cones in the human eye and are adjusted for obtaining a sharp and vivid image. Despite such variations, the output of color images is standardized and can be used across different devices. Color cameras are inexpensive and due to developments in processing power, memory, optics and sensor technology, almost every person having a mobile phone has a built-in color camera. Color images do not require lot of processing power and space. hey are widely used in numerous applications of computer vision and since they are used in standardized color spaces, the effect of sensor sensitivity ($\mathbf{c}(\lambda)$ in Eq. 2.1) is generally ignored [4]. The spatial resolution of color images can be very high but their spectral resolution is low and generally it is difficult to recover the reflectance of surfaces from color images since their sensor sensitivities cover broad areas in the visible wavelength region. Color images are affected by the change in scene illumination and need to be corrected for its effect. The issues related to change in illumination and methods for correction are discussed in Section 2.4.

2.3 Multispectral imaging

Human eyes can sense electromagnetic radiation in the visible spectrum mostly through three type of filters (cones) while multispectral imaging cameras can ac-

quire the radiation information in many more channels. Multispectral cameras can operate not only in the visible spectrum, but also in the ultraviolet and infrared regions. A multispectral camera combines the power of spectroscopy and digital imaging. Spectroscopy deals with the study of emitted or reflected radiation from materials and its variation along the wavelength spectrum. Spectrometers are used for measurement of radiance at a single point. These instruments consist of a diffraction grating or prism that splits the incoming light into several narrow and adjacent wavelength bands. The energy in each band is measured by a separate detector. Spectral measurements of very narrow bands can be performed through the use of hundreds of such detectors. Those instruments are able to acquire spectral information from only one point while the multispectral cameras can acquire not only spectral, but also the spatial information of a whole scene. The trade-off in multispectral imaging is in the form of lower spectral resolution as compared to the spectrophotometer. Each pixel in a multispectral image consists of the radiance information in a large number of channels while each channel consists of spatial information acquired at a particular wavelength. For a multispectral imaging system, the camera sensitivity function $\mathbf{c}(\lambda)$ in Eq. 2.1 consists of several filters, ranging from dozens to hundreds. Each of those filters allows a narrow region of wavelength spectrum to pass through it. According to the sensitivity of a typical silicon sensor behind an optical system, having sensitivity range ω, a multispectral system usually provides a combination of visible and/or near infrared bands, where the imaging model defined in Eq. 2.1 still holds.

$$\mathbf{f} = \int_{\omega} \mathbf{e}(\lambda)\mathbf{r}(\lambda)\mathbf{m}(\lambda)d\lambda, \tag{2.2}$$

here the camera sensitivities are represented by $\mathbf{m}(\lambda) = \{m_1(\lambda), m_2(\lambda), ..., m_N(\lambda)\}$ and N is the total number of channels.

Use of spectral imaging gained worldwide attention after the launch of *Landsat* in 1970 and since then it has been widely used in remote sensing applications. Multispectral imaging is used in many fields including satellite and air borne remote sensing systems [5], process control in food industry [6], quality control [7, 8], cultural heritage [9, 10], medical imaging [11–14] and a number of other fields [15].

For the acquisition of multispectral images, push broom cameras have been extensively used [16]. In the push broom imaging technique, line scanning of the scene is performed and the spectrogram of a particular scan is recorded on a charged couple device sensor. Another method for acquisition of multispectral images is through the use of liquid crystal tunable filter filter (LCTF) [17, 18] and MEMS Tunable Fabry-Pérot Filters [19] on top of a monochrome sensor. Use of such filters provides direct imaging ability and line scan of the scene is no longer required.

The drawback of using these filters is the low speed of spectral sampling and need for correction of acquired data due to shift in position in case of moving objects. Recently, low cost spectral imaging devices has been introduced that use off-the-shelf components for conversion of an imaging sensor into hyperspectral camera [20].

Filter-wheel camera based cameras are also popular for acquisition of multispectral images. Such cameras consist of a set of band-pass filters which are rotated over the imaging sensor and values are recorded for each filter. A combination of these images (which are gray-scale) forms the multispectral image. A recently developed method of multispectral image acquisition is through the *spectral filter array* [21], consisting of a number of filters embedded over the imaging sensor. These filters are embedded in the same way as Bayer filter mosaic but instead of three, there are more types of filters. In this way, an image consisting of responses from several filters is acquired.

Filter-wheel cameras have high spatial resolution and there is no requirement of de-mosaicing since each of the channels is acquired separately. The major shortcoming of such cameras is the requirement of static scene and camera position. The image acquisition time depends on the number of filters and the integration time being set of each of the filter. Due to the time required for image acquisition, such cameras are not feasible for capturing moving objects. A spectral filter array (SFA) camera, on the other hand, is able to take an image of the scene in a single shot [22]. The shortcoming of SFA based camera is the low spatial resolution, since the distance between same class of filters is relatively large in the filter array pattern, as compared to Bayer filter array pattern. The acquired image from SFA consists of a mosaic with each pixel representing the imaged data from only one filter. De-mosaicing is used to reconstruct each channel and this process causes loss of spatial resolution.

The advantage of multispectral imaging is the availability of high resolution spectral and spatial data of a scene. The availability of such information aids in the development of advanced spatio-spectral models for various tasks of computer vision and data processing. The use of multispectral images in object recognition can perform better than the conventional RGB color images [23]. One of the limitations of multispectral imaging is the need for known imaging conditions in order to perform any task that deals with the spectral reflectance of materials. Therefore, most of the existing multispectral imaging systems are specifically designed and require re-calibrating when the imaging conditions are changed. This limitation is a hurdle in the use of multispectral imaging in outdoor environments where the imaging conditions are not controlled. Color imaging, on the other hand, is able to address such changes as much as possible and the data is corrected from influence

of imaging conditions. One of such condition is the change in scene illumination. In this research work, our focus is to enable the use of multispectral imaging for uncontrolled imaging environment through the use of illuminant invariant representation of multispectral images.

As expressed in Eq. 2.1, the formation of an image is mainly dependent on three key factors. Among them, the reflectance of an object and the camera sensor sensitivities remain constant for a given object and camera. However, the change in illumination condition can cause different result for the same object, taken with the same camera. It is desirable to get consistent colors of objects. In the human visual system, constant color for objects is maintained through complex processes in the brain and as a result, we perceive the same color of objects most of the time, regardless of the illumination conditions. Digital camera manufacturers put a lot of effort to achieve this quality by designing the imaging sensor for optimal performance, and also through software correction of acquired data within the imaging pipeline. The consistent representation of colors is a well studied area both in the field of neuroscience and digital image processing. In the following section, we will provide a brief introduction to this field.

2.4 Color Constancy

The ability of the human visual system to perceive consistent color of objects across different illumination conditions is known as color constancy. The human visual system has the ability to address the changes in illumination and correct the color of objects. In the color constancy mechanism, the source of illumination is not necessarily known to the observer. In a complex environment with many objects, the illumination at one point can be the mixture of direct and indirect irradiation distributed over a range of incident angles and mutual reflections. Objects and surfaces are perceived by humans with consistent colors most of the time even if there is drastic change in illumination of the scene. This ability is the result of complex processes in the human visual system and the working of this mechanism is not fully understood yet [24].

Perception of colors in a constant manner is important for daily life tasks since color of objects is one of the important features in tasks involving recognition [25, 26]. For example, the color of fruits determine if they are ripe, raw or spoiled. The ability to perceive constant colors in different environments is crucial since it avoids confusion and makes life easier. Without color constancy, color of objects would not be a reliable clue for object identification. As an example, in Fig. 2.1, the color of the cup is perceived as white by human visual system while the close-up of a patch of the surface of the cup shows that the color is different. Four different illuminations cause difference in surface color but the human brain is

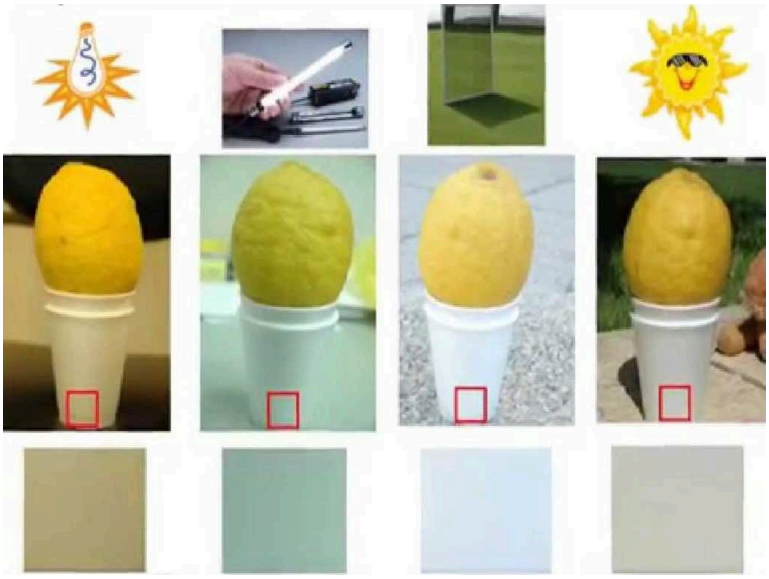


Figure 2.1: Close-up look of the color of cup under different illuminations.
Image taken from http://www.lacasamorett.com/foxgallery/color-constancy.html

able to compensate for changes in scene illumination.

Although the mechanism behind color constancy in human visual system is not fully understood yet, there are attempts to model such a system for computer vision. Such modelling is known as "computational color constancy" where the aim is to develop an illuminant-invariant representation of image data. Developing an illuminant invariant computer vision system is an open area of research and there are algorithms which are able to perform well for particular conditions and assumptions, but a universally accepted computational color constancy system does not exist.

The challenge in achieving efficient computational color constancy is that it is not easy to estimate the spectral reflectance of surface and the spectral irradiance of the incident illumination from the reflected spectral radiance into the imaging sensor. Mathematically, the recovery of reflectance from the sensor data is an ill-posed problem since the data recorded by imaging sensor is composed of surface reflectance, sensor sensitivity and the scene illumination (Eq. 2.1).

Computational color constancy plays an important role in color-based computer vision applications including object recognition, tracking and image classification [27]. Object representation and recognition from the standpoint of computer vision is discussed in detail by Dickinson [28]. For example, in the case of object

recognition, the color of the object can be used as a feature, and it should appear constant across changes in illumination [29]. So the first step in achieving a constant representation of colors is to adjust the color changes that occur due to the scene illuminant. Therefore, computational color constancy deals with the representation of a scene with the effect of the illuminant being as small as possible. There are several attempts to solve for the problem of computational color constancy by imposing constraints on the surface reflectances and illuminants [30]. There are two approaches for this. One is to compute illuminant invariant features [31, 32] while the second technique is to estimate the scene illuminant [33] and later apply a correction. In this work, we focus on the technique of estimation of scene illumination and then removing its effect from the acquired image.

From the imaging model given in Eq. 2.1, the goal of a color constancy system is to estimate the illuminant $\hat{\mathbf{e}}$, and this estimation is performed in the camera domain. Illumination in the camera domain is dependent on the camera sensor sensitivities and it defines how the illumination is recorded by the camera. For a color camera with three channels, it is represented by

$$\mathbf{e} = \int_{\omega} e(\lambda)\mathbf{c}(\lambda)d\lambda = \begin{pmatrix} R_e \\ G_e \\ B_e \end{pmatrix}. \tag{2.3}$$

The aim of illuminant estimation method is to estimate $\hat{\mathbf{e}}$ with minimum error and as close as possible to $\mathbf{e}$. Both $\mathbf{e}$ and $\hat{\mathbf{e}}$ are in the sensor domain and are composed of a set of discrete values with the dimension equal to total number of filters (N). It should be noted that $\hat{\mathbf{e}}$ is the response of each filter for the illumination (ground truth or estimated), and it is not equivalent to the spectral power distribution of the illumination itself. Some algorithms aim at recovery of spectral power distribution of illumination from $\hat{\mathbf{e}}$, while most of the techniques use the information from estimated algorithm directly for correction of the acquired image. In the following section, an overview of the illuminant estimation techniques is provided.

2.5 Illuminant estimation in color images

There are two categories of illuminant estimation algorithms; one method is to use the statistical properties of the scene, while the other is to utilize the physics-based analysis of surfaces [34]. Machine learning is also used for estimation of scene illumination. In the following sections, a brief description of those techniques is provided.

2.5.1 Statistics based illuminant estimation methods

This type of illuminant estimation method is based on the image statistics and is applied by using fixed set of parameters. One of the earlier algorithms in this category is the Land's white-patch algorithm [35] which is based on the assumption that there is at-least one pixel in the image which causes maximum reflectance and when such pixels are combined, they provide information of scene illumination. A surface with perfect reflectance is assumed to reflect back all the incident illumination on it. This assumption is alleviated by considering the color channels separately, resulting in the max-RGB algorithm [36]. This method computes the maximum response in each channel separately for each color channel of an image $\mathbf{F}$, which is taken under the unknown illuminant, as

$$\max_{\mathbf{x}} \mathbf{F}(x) = (\max_{\mathbf{x}} R(x), \max_{\mathbf{x}} G(x), \max_{\mathbf{x}} B(x)) = k\hat{\mathbf{e}} \tag{2.4}$$

where $\mathbf{x}$ is the spatial location and k is the multiplicative constant which is chosen to keep with estimated illuminant $\hat{\mathbf{e}} = (\hat{e}_R, \hat{e}_G, \hat{e}_B)$ within the unit length. A number of improvements are suggested in the max-RGB algorithm. One of the methods is to apply smoothing on the image [37] for removal of noisy pixels. Ebner [38] suggested the local space average color method which also allows pixel-wise illuminant estimation. Max-RGB algorithm is sensitive to the pre-processing strategies and the dynamic range of imaging sensor [39]. Some approaches aim at finding gray pixels and surfaces in the scene instead of white patches. Hue *et al.* [40] used such gray points in an image for estimation of illuminant temperature. In their method, RGB image is converted into YUV color space and then those pixels where $U = V = 0$ or $R = G = B = Y$ are pointed out as gray points. The values at those points is taken as the estimate for scene illuminant.

A commonly used algorithm in this category is the grey-world algorithm. It was proposed by Buchsbaum [41] and is based on the assumption that the average color of a scene is achromatic. Following this assumption, any deviation from achromaticity in the average color of a scene is caused due to the scene illumination. Therefore, by computing the average color in an image taken under unknown illumination, the color of light source can be estimated by

$$\left(\frac{\int \mathbf{F} dx}{\int dx}\right) = k\hat{\mathbf{e}} \tag{2.5}$$

The result of Grey-World algorithm was improved by Gershon *et al.* [42] by taking average reflectance of a database and assuming the average of the scene to be equal to that average reflectance. Gray-world algorithm is able to perform well when the natural scene statistics are satisfied and the image contain a large number of

surfaces, but fails in the presence of large uniformly colored surfaces. To overcome this problem, segmentation of the scene and then computing the average color of the segmented regions is suggested in [42, 43].

Finlayson and Trezzi [44] proposed that the gray-world and max-RGB algorithms are special instantiations of the more general Minkowski framework. They proposed shades of gray algorithm with Minkoeski norm, where the gray-world algorithm is the same as using the L^1 Minkowski norm while max-RGB is equivalent to using L^∞ norm. In their case, the general equation for estimation of light source becomes

$$\left(\frac{\int \mathbf{F}^p dx}{\int dx}\right)^{1/p} = k\hat{\mathbf{e}}, \tag{2.6}$$

where p is the order of the Minkowski norm. The value of p varies according to the image and dataset type and is set individually. Therefore, the optimal value of p may vary for different image datasets and even for each image within a dataset.

The assumptions in max-RGB, gray-world and shades of gray algorithm are based on the distribution of pixel values in a given image. Weijer *et al.* [45] proposed the use of higher order image statistics for estimation of scene illuminant. Their proposed algorithm is called gray-edge algorithm in which it is assumed that the average of *reflectance derivative* in a scene is achromatic. This algorithm is expressed as:

$$\left(\frac{\lfloor\int [\mathbf{F}_\sigma]^p dx}{\int dx}\right)^{1/p} = k\hat{\mathbf{e}}, \tag{2.7}$$

where F_σ is the smoothed image, after applying a Gaussian filter.

Edge-based color constancy is explored further for higher order derivatives in [46]. Celik and Tjahjadi [47] used wavelet transform to down-sample the image before applying Grey-Edge algorithm for estimation of illuminant color and for each down-sampled image, separate estimation is performed on the high pass filter's result. Decision for illuminant color is based on minimum error between the estimation in consecutive scales. There are some approaches [48, 49] which try to select the most appropriate estimation using intrinsic properties from other color constancy algorithms. Gijsenij *et al.* [50] extended the gray-edge method to incorporate a general weighting scheme for the image edges and proposed that different types of edges might contain various amounts of information. Their results suggest that the specular edges are favored for the estimation of scene illuminant. Use of specular reflection for illuminant estimation is further discussed in section 2.5.2.

Some approaches for illuminant estimation are proposed by imposing constraints on the set of plausible illuminants. Gamut mapping is one such technique where

it is assumed that *for a given illuminant, one observes only a limited number of colors*. Based on this assumption, any change in colors of the image is caused by the variation in color of the light source. The limited set of colors which can occur under a given illuminant is called the *canonical gamut* and is determined through observations of many surfaces under the known light source. Gamut mapping was introduced by Forsyth [51] and he proposed that the color of an object is its representation under a fixed canonical light, rather than as a surface reflectance function. Color-by-correlation [52] is a discrete version of Gamut mapping where the correlation matrix is used instead of canonical gamut for the considered illuminants, and is used with the image data to calculate the probability that the illumination in the test image is caused by which of the known illuminants.

Statistics based illuminant estimation methods are able to provide good results as long as the assumptions for scene statistics are fulfilled. Most of those methods require a large number of various colored surfaces in a scene. In the presence of very few colors or a large surface in the image, those methods tend to fail. For illuminant estimation in scenes with less diversity in terms of colored surfaces, dichromatic reflection model based illuminant estimation methods are used. These methods are also known as highlights based illuminant estimation algorithms. In the following section, an overview of those methods is provided.

2.5.2 Highlights based illuminant estimation methods

The imaging model of Eq. 2.1 is valid for matte surfaces with only diffuse reflection and follow the Lambertian model. In highlights based illuminant estimation methods, the information about the physical interaction between the light source and the objects in a scene is used. In practice, both specular and diffuse reflections are present in a scene. Diffuse reflection is caused when the incident light is partially absorbed and partially reflected from the surface of a material. Such diffuse reflection is omni-directional and is not dependent on the viewing angle. The color of objects and surfaces is determined by the diffuse reflection through the reflected wavelengths in a particular spectral region, while the rest of wavelengths are absorbed by that material. Specular reflections occur when all of the incident light on a material surface is reflected back. In this case, the incident light on the imaging sensor is composed of both diffuse and specular reflections from the material surface. Shafer [53] proposed to decompose the light into its diffuse $m^b(\mathbf{x})$ and specular $m^s(\mathbf{x})$ components:

$$\mathbf{f(x)} = m^b(\mathbf{x}) \int_\omega e(\lambda)s(\mathbf{x},\lambda)\mathbf{c}(\lambda)d\lambda + m^s(\mathbf{x}) \int_\omega e(\lambda)\mathbf{c}(\lambda)d\lambda. \qquad (2.8)$$

Eq. 2.8 is the basis for "dichromatic reflection model" which explains both the diffuse and specular reflections from the surface of a material. In the dichromatic

reflection model, the diffuse reflection is of low intensity and is Lambertian, while the specular reflection is generally of high intensity, directional, independent from the surface reflectance and dependent on the viewing direction.

The specular reflection part of Eq. 2.8 is the same as illuminant in the sensor domain (Eq. 2.3), along with the specular component $m^s(\mathbf{x})$. This relation suggests that solving for the diffuse and specular components can provide information of scene illumination. This information is contained in the specular reflection since the incident illumination is reflected back in that case. Eq. 2.8 shows that an image is a linear combination of the diffuse and secular components and for a camera with $\mathbf{c}(\lambda) = \{R(\lambda), G(\lambda), B(\lambda)\}$, this linear combination defines a plane in the RGB color space. The data points from the RGB color space are projected onto a space normalized by $r + g + b = 1$, where they form a line. The two points defining the line consist of the diffuse and specular points on a surface. Data points from a uniformly colored surface are distributed along this line which is called the *dichromatic line*. Assuming uniform illumination in a scene, all the dichromatic lines have one point in common, which corresponds to the scene illuminant. Therefore, if there are two or more materials being illuminated by the same light source and their dichromatic planes are identified, the vector containing illuminant information can be estimated by the intersection of these planes.

Several authors have used the two-dimensional chromaticity space by using the dichromatic reflection model for illuminant estimation. One of the earlier attempts in specular reflection based illuminant estimation are by Lee [54], and D'Zumra and Lennie [55], who proposed the use of color coordinates of specular highlights for estimation of illumination. Lehmann and Palm [56] extended Lee's red diagram [54] to the *rg-diagram* and called it "color line search". Recently, Uchimi *et al.* [57] proposed estimation of illumination from $u'v'$ chromaticity, xy chromaticity space is used in [58], and the rg color space in [59, 60]. By applying constraints on the set of possible illuminants, some algorithms [61–64] aim at finding the chromaticity of an illuminant by intersecting a dichromatic line with the Planckian locus, or a set of known illuminants. The basic idea of such methods is to calculate the projected area for a set of candidate illuminants along the Planckian locus and pick the candidate that gives the minimum error in terms of distance, as an estimate of the scene illuminant.

The inverse intensity chromaticity space (IICS) is also a physics-based illuminant estimation method introduced by Tan *et al.* [65]. It is spanned by the chromaticity of a single channel and the inverse of image intensity. A linear correlation between the image chromaticity and illumination chromaticity is formed in the IICS. This correlation allowed the estimation of scene illumination without segmenting the color beneath the highlights. Riess *et al.* [66] proposed local estimates of scene

illuminant through decomposition of the image into mini-regions. Those estimates are used for computing the dominant illuminant of the scene by using the inverse intensity chromaticity space.

One of the limitations of specular reflection based illuminant estimation methods is the requirement of pre-segmented highlight regions in an image. Detection of highlights in images is not straightforward and is still an open area of research [67]. There are some attempts to eliminate the requirement of pre-segmentation. Shi and Funt [68] proposed two Hough Transforms for computation of a histogram that represents the likelihood that a candidate intersection line is the image illumination axis. This axis is defined by the set of known illuminants and is limited for generic use. Highlights based illuminant estimation methods are limited by the dynamic range of the imaging sensor and are unable to perform well if the highlight regions in the image are saturated.

2.5.3 Machine learning based illuminant estimation

Machine Learning is also applied for illuminant estimation and is becoming a popular method. Although most of the machine learning, particularly deep learning methods aim at finding the illuminant-invariant features in images, there are attempts to use machine learning for illuminant estimation. In [69], a multi-layer neural network is trained by using the histograms of chromaticity of input images along with corresponding chromaticity of illuminant. A number of similar approaches can be found in [70–72] where the network learns the illuminant in scene through various examples. Support Vector Machine is used in [73], which is based on the higher-order structure of images. Recently, Deep Learning is also utilized in color constancy as in [74, 75]. Oh and Kim [76] used deep learning for treating illuminant estimation as a classification problem. Bianco *et al.* [77] used convolutional neural network for illuminant estimation in raw images. For generation of ground truth illumination, shades of gray, gray edge and gamut mapping is applied on the training data in their proposed method.

In the presence of a large dataset along with ground truth for training, machine learning methods are able to perform the illuminant estimation task efficiently. However, in the absence of a large training dataset, their performance can be lower than the linear methods. Machine learning, and particularly deep learning based illuminant estimation methods in color imaging have the potential to outperform conventional methods if the network is trained over a huge amount of data and the standardization of both training and test data can be ensured.

So far we have discussed the concept of color constancy and illuminant estimation for color imaging. From the next section, we will focus on multispectral ima-

ging and the work discussed in the following forms the contribution towards this dissertation.

2.6 Multispectral Constancy

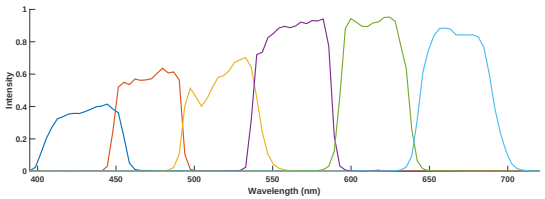
Extending the use of multispectral imaging system from heavily constrained environments to real world applications is still an open challenge. One of the major obstacles is calibration of multispectral camera according to the scene illuminant [78–82]. When the multispectral image is acquired with known (canonical) illuminant, processing of spectral data is relatively straightforward. Fig. 2.2 illustrates the effect of scene illumination on the behaviour of acquired data. An object with measured reflectance as in Fig. 2.2(c) is imaged with a filter-wheel camera with 6 sensors in the visible wavelength range (2.2(a)). Three illuminants (A, D65 and F12) are used in this example and for each of the illuminants, the sensor response for the same surface is quite different. If a surface identification system is trained to identify the response of Fig. 2.2(d) as the material corresponding to the spectral response in Fig. 2.2(c), it will fail if the camera data for the same material is obtained as in Fig. 2.2(e) or 2.2(f). In both cases, material and camera are the same but a change in illumination causes change in the final image and therefore, it is necessary to remove the effect of illumination. Once the effect of illumination is removed and the image is transformed into a canonical representation, various computer vision applications can be applied on the multispectral images.

Multispectral constancy deals with such a transformation of acquired multispectral data under uncontrolled illumination conditions, into a canonical representation. Multispectral constancy has two parts, first is the diagonal transform $\mathbf{D}^{c,ill}$ and second is the spectral adaptation transform. The elements of diagonal transform consist of illuminant in the sensor domain. These values are either extracted manually from a white patch in the scene or estimated through different methods. In the following section, multispectral data transformation is discussed.

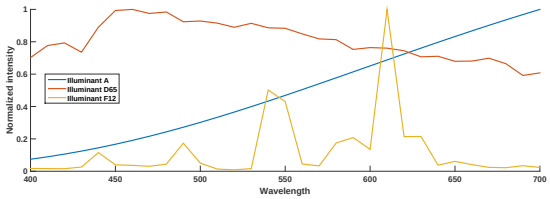
2.7 Transformation into canonical representation

The aim of multispectral constancy is to transform the acquired image into a representation as if it would have been taken under a known light source. The known light source is called as *canonical light source* and the representation of image taken under a known light source is known as *canonical representation*. This process is also expressed as "discounting the chromaticity of the illuminant" by D'Zmura and Lennie [55], and is expressed as

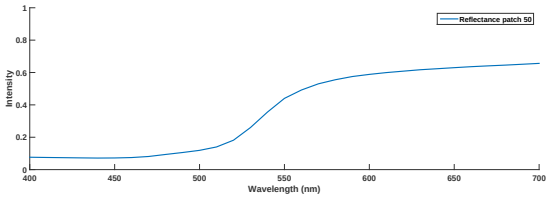
$$\mathbf{F}^{c} = \mathbf{M}^{c,ill}\mathbf{F}^{ill}, \tag{2.9}$$



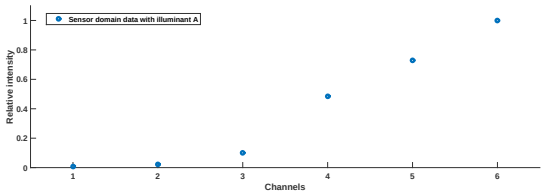
(a) 6 sensors of filter-wheel camera in visible range



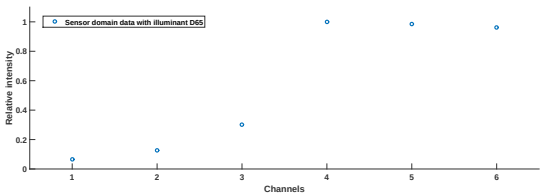
(b) Illuminants



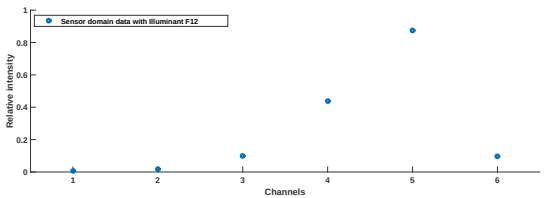
(c) Example reflectance patch



(d) Camera data with illuminant A



(e) Camera data with illuminant D65



(f) Camera data with illuminant F12

Figure 2.2: Illustration of the influence of illumination. The data obtained from camera is different in (d,e,f) varies when the scene illumination is changed for the same imaging device and surface reflectance.

where $\mathbf{F}^{ill}$ is the image taken in unknown light source and $\mathbf{F}^{c}$ is the transformed image as if taken under a canonical illuminant, while $\mathbf{M}^{c,ill}$ is the transformation matrix, which maps colors from captured image to their corresponding colors under a known illumination.

In most of the computational color constancy literature [83, 84], a diagonal transform $\mathbf{D}^{c,ill}$ is used to transform the input image taken under unknown illuminant, into a canonical representation. Such a transform requires the measurement or estimation of scene illuminant in the sensor domain. The diagonal matrix $\mathbf{D}^{c,ill}$ is a $N \times N$ matrix and the components of this matrix are the sensor responses to the illuminations $\mathbf{E}_c$ and $\mathbf{E}_{ill}$. As an alternate explanation, sensor response to the illuminants can be understood as the sensor response to a perfect white diffuser under a certain illuminant, and denoted by $\mathbf{w}_c$ and $\mathbf{w}_{ill}$ for the two illuminants. The matrix $\mathbf{D}$ is defined by

$$\mathbf{D}^{c,ill} = diag(\mathbf{w}_c / \mathbf{w}_{ill}) \tag{2.10}$$

The diagonal transform assumes that there are no inter-channel dependencies and that each of the image channels is adjusted independently for correction of image colors as

$$\begin{pmatrix} R^c \\ G^c \\ B^c \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} R^u \\ G^u \\ B^u \end{pmatrix}. \tag{2.11}$$

Eq. 2.11 describes a basic model. Thomas [85] suggested that it may be possible to extend it for multispectral images where instead of three channels, there are $N > 3$ channels.

$$\begin{pmatrix} F_1^c \\ F_2^c \\ \vdots \\ F_N^c \end{pmatrix} = \begin{pmatrix} K_1 & 0 & \dots & 0 \\ 0 & K_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & K_N \end{pmatrix} \cdot \begin{pmatrix} F_1^{ill} \\ F_2^{ill} \\ \vdots \\ F_N^{ill} \end{pmatrix} \tag{2.12}$$

where F_n^{ill} is the uncalibrated pixel of n^{th} channel, F_n^c is the calibrated pixel while K_n is the correction parameter for channel n.

The independence of color channels from each other is defined in the Retinex Model [35, 36, 86]. This assumption is closely related to the Von Kries coefficient rule. $\mathbf{D}^{c,ill}$ generally holds for color images and has been widely used in literature. However, the imaging sensors are not always independent and there are inter-channel dependencies which arise due to factors related to optics and overlapping of sensor sensitivities in the wavelength spectrum. There are efforts to address

those dependencies and several transforms are developed for color images. These transforms are generally known as "chromatic adaptation transforms" and they aim at minimizing the perceptual error in color imaging. Examples of chromatic adaptation transform are the Bradford transform [87], the sharp transform [88] and CMCCAT2000 [89]. In our work, we have introduced "spectral adaptation transform" for multispectral imaging, which is used to address the inter-channel dependencies of imaging sensor. This transform is discussed in the following section.

2.7.1 Spectral adaptation transform

The assumption of having no inter-channel dependencies for an imaging system is valid when the overlapping among channels is minimal or none. However, in the color and multispectral cameras, there is overlapping among the sensitivity regions of each filter. The diagonal transform $\mathbf{D}^{c,ill}$ does not address such inter-channel dependencies and assumes perfect isolation of each channel.

In our work on illuminant estimation in multispectral imaging, we performed simulation of multispectral sensors in three different configurations and found that sensors with more overlapping among them can help in better estimation of scene illumination (for details, please refer to Chapter 5). However, the spectral reconstruction results suggest that sensors with lower or no overlapping among them perform better. This different is caused by the spectral noise, which is introduced due to the filter's overlapping in the wavelength region. To overcome this problem in order to get optimal performance for spectral reconstruction, we have introduced the spectral adaptation transform (SAT). The idea of SAT is to incorporate the intrinsic properties of imaging sensor in the multispectral constancy pipeline for data transformation. SAT is unique for each imaging system and is calculated by using the sensor sensitivities and a set of training reflectances. The optimal SAT matrix $\mathbf{A}_{SAT}$ should minimize the error for all reflectances i and illuminants j.

$$\mathbf{F}_c^i = \mathbf{A}_{SAT}\mathbf{D}^{c,j}\mathbf{F}_j^i \tag{2.13}$$

To find the best $\mathbf{A}_{SAT}$, Eq. 2.13 is written as the explicit minimization of an error function in Eq. 2.14;

$$\min_{\mathbf{A}_{SAT}} = \sum_j |\mathbf{F}_c^i - \mathbf{A}_{SAT}\mathbf{D}^{c,j}\mathbf{f}_j^i| \tag{2.14}$$

For a given sensor configuration, a generic $\mathbf{A}_{SAT}$ is computed once and is used along with the diagonal transform as in Eq. 2.13.

The role of SAT is to address the inter-channel dependencies and optimize the diagonal transform. In Chapter 4, the role of SAT is demonstrated through the results of spectral reconstruction from multispectral images. A demonstration of

diagonal correction and SAT by using a real multispectral filter-wheel camera and material surfaces is provided in Chapter 9.

So far we have discussed SAT while assuming the information about scene illuminant in the sensor domain is measured through a white diffuser in the scene. However, in practical scenarios it is not always feasible to use diffuser while imaging. Therefore, we propose illuminant estimation in multispectral imaging, which is discussed in the following section.

2.7.2 Illuminant estimation in multispectral images

For estimation of scene illumination in multispectral imaging, we proposed the extension of four statistics based illuminant estimation methods. Those methods were developed for color images and we investigated their performance on multispectral images. Those four algorithms are;

- Gray-World Algorithm [41] $\rightarrow$ Spectral Gray-World Algorithm
- Max-RGB Algorithm [36] $\rightarrow$ Max-Spectral Algorithm
- Shades of Grey Algorithm [44] $\rightarrow$ Spectral Shades of Gray Algorithm
- Gray-Edge Algorithm [45][46] $\rightarrow$ Spectral Gray-Edge Algorithm

Illuminant is estimated for each channel individually and the final estimate is composed of a one dimensional vector with N values. Those values correspond to the illuminant is sensor domain. Details of these algorithms and their use in multispectral imaging is discussed in detail in **Article B** and **Article D** (Chapters 5 and 7). Chapter 5 provide details of statistics based illuminant estimation methods in multispectral imaging while in Chapter 7, a discussion on specular reflection based illuminant estimation methods is provided. We do not consider machine learning based illumination estimation for multispectral imaging. A machine learning based method relies on training data for creating a model that is able to predict and classify the input data. For illuminant estimation purpose, a machine learning algorithm will require standard multispectral camera data with a lot of examples that include the ground truth information about scene illumination. Machine learning and particularly, deep learning methods have been proved to outclass the classical methods for detection and classification tasks under the standardized data but in case of multispectral imaging, there is no standard way of representing the image data. Each multispectral camera has a specific number of channels and sensor sensitivities that vary among camera models and manufacturers. Some cameras cover only the visible range of wavelength spectrum, some operate in near infrared, some cameras include the ultraviolet region and there are cameras that operate in

all of those regions. For creating an efficient machine learning system for illuminant estimation, hundreds and thousands of examples using one camera need to be created for training the system and that system will not perform well for another multispectral camera with different number of channels and/or spectral range and sensor sensitivities. Due to such limitations, we do not use machine learning for illuminant estimation in multispectral imaging in our current work.

2.8 Applications of multispectral constancy

Once multispectral constancy is achieved, a multispectral imaging system can be used for obtaining the spectral reflectance of material surfaces. This spectral information along with spatial information can increase the performance of material identification and classification. The efficiency of SAT along with diagonal transform is demonstrated in Chapter 4. Once the influence of illumination is removed from the acquired multispectral data, the spectral information can be used to reconstruct the spectra of material surfaces in a scene. Use of multispectral camera for spectral reconstruction is demonstrated in Chapter 9. Multispectral imaging can also be used for acquisition of colorimetric values of a scene. In Chapter 8, multispectral data is mapped from N dimensions to 3 dimensional representation and colorimetric information is extracted. The availability of illuminant invariant multispectral imaging system will enable the use of multispectral cameras for generic computer vision applications where the spectral and spatial information will enhance the ability of surface identification and materials classification.The work presented in this dissertation will help in coming one step closer in enabling the use of multispectral cameras in the same way as color cameras and to explore their potential for numerous applications.

Bibliography

[1] K. Michael, *Electromagnetic Radiation and Interactions with Matter*. in Encyclopedia of Imaging Science and Technology. J. P. Hornak (Ed.), 2002.

[2] M. Vollmer, K.-P. Möllmann, and J. A. Shaw, "The optics and physics of near infrared imaging," in *ETOP 2015 Proceedings*, p. TPE09, Optical Society of America, 2015.

[3] A. Stockman, D. I. A. MacLeod, and N. E. Johnson, "Spectral sensitivities of the human cones," *J. Opt. Soc. Am. A*, vol. 10, pp. 2491–2521, Dec 1993.

[4] S.-B. Gao, M. Zhang, C.-Y. Li, and Y.-J. Li, "Improving color constancy by discounting the variation of camera spectral sensitivity," *J. Opt. Soc. Am. A*, vol. 34, pp. 1448–1462, Aug 2017.

[5] K. Ose, T. Corpetti, and L. Demagistri, "Multispectral satellite image processing," in *Optical Remote Sensing of Land Surface* (N. Baghdadi and M. Zribi, eds.), pp. 57 – 124, Elsevier, 2016.

[6] D. Ariana, D. E. Guyer, and B. Shrestha, "Integrating multispectral reflectance and fluorescence imaging for defect detection on apples," *Computers and Electronics in Agriculture*, vol. 50, no. 2, pp. 148 – 161, 2006.

[7] E. N. Lewis, J. Dubois, L. H. Kidder, and K. S. Haber, *Near Infrared Chemical Imaging: Beyond the Pictures*, ch. 14, pp. 335–361. Wiley-Blackwell, 2007.

[8] J. Orava, J. Parkkinen, M. Hauta-Kasari, P. Hyvönen, and A. von Wright, "Temporal clustering of minced meat by RGB and spectral imaging," *Journal of Food Engineering*, vol. 112, no. 1, pp. 112 – 116, 2012.

[9] E. Marengo, M. Manfredi, O. Zerbinati, E. Robotti, E. Mazzucco, F. Gosetti, G. Bearman, F. France, and P. Shor, "Development of a technique based on multi-spectral imaging for monitoring the conservation of cultural heritage objects," *Analytica Chimica Acta*, vol. 706, no. 2, pp. 229 – 237, 2011.

[10] G. M. A. Rahaman, J. Parkkinen, M. Hauta-Kasari, and S. H. Amirshahi, "Enhanced color visualization by spectral imaging: An application in cultural heritage," in *2017 IEEE International Conference on Imaging, Vision Pattern Recognition (icIVPR)*, pp. 1–6, Feb 2017.

[11] A. Messano and M. Singh, "Technology for multispectral scanning, detection and imaging for medical diagnosis," in *2016 10th International Symposium on Medical Information and Communication Technology (ISMICT)*, pp. 1–3, March 2016.

[12] J. R. Mansfield, "Multispectral imaging: A review of its technical aspects and applications in anatomic pathology," *Veterinary Pathology*, vol. 51, no. 1, pp. 185–210, 2014. PMID: 24129898.

[13] J. Kinnunen, J. S. Jurvelin, J. Mäkitalo, M. Hauta-Kasari, P. Vahimaa, and S. Saarakkala, "Optical spectral imaging of degeneration of articular cartilage," *Journal of Biomedical Optics*, vol. 15, no. 4, pp. 046024–1–046024–8, 2010.

[14] P. Fält, J. Hiltunen, M. Hauta-Kasari, I. Sorri, V. Kalesnykiene, J. Pietilä, and H. Uusitalo, "Spectral imaging of the human retina and computationally determined optimal illuminants for diabetic retinopathy lesion detection,"

Journal of Imaging Science and Technology, vol. 55, no. 3, pp. 30509–1–30509–10, 2011.

[15] V. C. Coffey, "Multispectral imaging moves into the mainstream," *Opt. Photon. News*, vol. 23, pp. 18–24, Apr 2012.

[16] W. L. Wolfe, *Introduction to imaging spectrometers*, vol. 25. SPIE Press, 1997.

[17] P. J. Miller, "Use of tunable liquid crystal filters to link radiometric and photometric standards," *Metrologia*, vol. 28, no. 3, p. 145, 1991.

[18] M. Hauta-Kasari, K. Miyazawa, S. Toyooka, and J. Parkkinen, "Spectral vision system for measuring color images," *J. Opt. Soc. Am. A*, vol. 16, pp. 2352–2362, Oct 1999.

[19] N. Neumann, M. Ebermann, K. Hiller, M. Seifert, M. Meinig, and S. Kurth, "MEMS Tunable Fabry-Pérot Filters for Infrared Microspectrometer Applications," in *Imaging and Applied Optics 2016*, p. AIT4B.2, Optical Society of America, 2016.

[20] F. Sigernes, M. Syrjäsuo, R. Storvold, J. Fortuna, M. E. Grøtte, and T. A. Johansen, "Do it yourself hyperspectral imager for handheld to airborne operations," *Opt. Express*, vol. 26, pp. 6021–6035, Mar 2018.

[21] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton, "Multispectral filter arrays: Recent advances and practical implementation," *Sensors*, vol. 14, no. 11, pp. 21626–21659, 2014.

[22] X. Wang, *Filter array based spectral imaging : Demosaicking and design considerations*. PhD thesis, Norwegian University of Science and Technology, 2016.

[23] Y. Manabe, K. Sato, and S. Inokuchi, "An object recognition through continuous spectral images," in *12th International Conference on Pattern Recognition*, vol. 1, pp. 858–860, Oct 1994.

[24] A. Gijsenij, T. Gevers, and J. van de Weijer, "Computational color constancy: Survey and experiments," *IEEE Transactions on Image Processing*, vol. 20, pp. 2475–2489, Sept 2011.

[25] S. Hagen, Q. C. Vuong, L. S. Scott, T. Curran, and J. W. Tanaka, "The role of color in expert object recognition," *Journal of Vision*, vol. 14, no. 9, p. 9, 2014.

[26] K. van de Sande, T. Gevers, and C. Snoek, "Evaluating color descriptors for object and scene recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 1582–1596, Sept 2010.

[27] K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms. i: Methodology and experiments with synthesized data," *IEEE Transactions on Image Processing,*, vol. 11, pp. 972–984, Sep 2002.

[28] S. J. Dickinson, "Object representation and recognition," in *What is Cognitive Science* (E. Lepore and Z. Pylyshyn, eds.), ch. 5, pp. 172–207, Basil Blackwell Publishers, 1999.

[29] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.

[30] L. T. Maloney and B. A. Wandell, "Color constancy: a method for recovering surface spectral reflectance," *J. Opt. Soc. Am. A*, vol. 3, pp. 29–33, Jan 1986.

[31] B. V. Funt and G. D. Finlayson, "Color constant color indexing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 522–529, May 1995.

[32] T. Gevers and A. W. Smeulders, "Color-based object recognition," *Pattern Recognition*, vol. 32, no. 3, pp. 453 – 464, 1999.

[33] S. D. Hordley, "Scene illuminant estimation: Past, present, and future," *Color Research and Application*, vol. 31, no. 4, pp. 303–314, 2006.

[34] H. A. Khan, J.-B. Thomas, and J. Y. Hardeberg, "Towards highlight based illuminant estimation in multispectral images," in *Proceedings of International Conference on Image and Signal Processing. Cherbourg, France*, pp. 517–525, Springer International Publishing, 2018.

[35] E. H. Land, "The retinex theory of color vision," *Scientific American*, pp. 108–128, 1977.

[36] E. H. Land and J. J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Am.*, vol. 61, pp. 1–11, Jan 1971.

[37] A. Gijsenij and T. Gevers, "Color Constancy by Local Averaging," in *14th International Conference of Image Analysis and Processing - Workshops (ICIAPW)*, pp. 171–174, Sept 2007.

[38] M. Ebner, "Color constancy based on local space average color," *Machine Vision and Applications*, vol. 20, pp. 283–301, Jul 2009.

[39] B. V. Funt and L. Shi, "The effect of exposure on MaxRGB color constancy," in *Proc. SPIE Volume 7527: Human Vision and Electronic Imaging XV*, 2010.

[40] J. Huo, Y. Chang, J. Wang, and X. Wei, "Robust automatic white balance algorithm using gray color points in images," *IEEE Transactions on Consumer Electronics*, vol. 52, pp. 541–546, May 2006.

[41] G. Buchsbaum, "A spatial processor model for object colour perception," *Journal of the Franklin Institute*, vol. 310, no. 1, pp. 1–26, 1980.

[42] R. Gershon, A. D. Jepson, and J. K. Tsotsos, "From [r,g,b] to surface reflectance: Computing color constant descriptors in images," in *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, (San Francisco, CA, USA), pp. 755–758, Morgan Kaufmann Publishers Inc., 1987.

[43] K. Barnard, L. Martin, A. Coath, and B. Funt, "A comparison of computational color constancy Algorithms. II. Experiments with image data," *IEEE Transactions on Image Processing*, vol. 11, pp. 985–996, Sep 2002.

[44] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Color and Imaging Conference*, (Scottsdale, Arizona), pp. 37–41, 2004.

[45] J. van de Weijer and T. Gevers, "Color constancy based on the grey-edge hypothesis," in *IEEE International Conference on Image Processing*, vol. 2, pp. II–722–5, Sept 2005.

[46] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Transactions on Image Processing*, vol. 16, pp. 2207–2214, Sept 2007.

[47] T. Celik and T. Tjahjadi, "Adaptive colour constancy algorithm using discrete wavelet transform," *Computer Vision and Image Understanding*, vol. 116, no. 4, pp. 561–571, 2012.

[48] A. Gijsenij and T. Gevers, "Color constancy using natural image statistics," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, June 2007.

[49] A. Gijsenij and T. Gevers, "Color constancy using natural image statistics and scene semantics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 687–698, April 2011.

[50] A. Gijsenij, T. Gevers, and J. van de Weijer, "Physics-based edge evaluation for improved color constancy," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 581–588, June 2009.

[51] D. Forsyth, "A novel algorithm for color constancy," *International Journal of Computer Vision*, vol. 5, no. 1, pp. 5–35, 1990.

[52] G. D. Finlayson, S. D. Hordley, and P. M. HubeL, "Color by correlation: a simple, unifying framework for color constancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1209–1221, Nov 2001.

[53] S. A. Shafer, "Using color to separate reflection components," *Color Research & Application*, vol. 10, no. 4, pp. 210–218, 1985.

[54] H.-C. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," *J. Opt. Soc. Am. A*, vol. 3, pp. 1694–1699, Oct 1986.

[55] M. D'Zmura and P. Lennie, "Mechanisms of color constancy," *J. Opt. Soc. Am. A*, vol. 3, pp. 1662–1672, Oct 1986.

[56] T. M. Lehmann and C. Palm, "Color line search for illuminant estimation in real-world scenes," *J. Opt. Soc. Am. A*, vol. 18, pp. 2679–2691, Nov 2001.

[57] Y. Uchimi, T. Jinno, and S. Kuriyama, "Estimation of multiple illuminant colors using color lines of single image," in *Int. Conf. on Adv. Informatics, Concepts, Theory, and Applications*, pp. 1–6, Aug 2017.

[58] J.-Y. Kim, Y.-S. Seo, and Y.-H. Ha, "Estimation of illuminant chromaticity from single color image using perceived illumination and highlight," *Journal of Imaging Science and Technology*, vol. 45, no. 3, pp. 274–282, 2001.

[59] O.-S. Kwon, Y.-H. Cho, Y.-T. Kim, and Y.-H. Ha, "Illumination estimation based on valid pixel selection in highlight region," in *IEEE International Conf. on Image Processing*, Oct 2004.

[60] E. Lakehal and D. Ziou, "Computational color constancy from maximal projections mean assumption," *Multimedia Tools and Applications*, Dec 2017.

[61] G. D. Finlayson and G. Schaefer, "Solving for colour constancy using a constrained dichromatic reflection model," *International Journal of Computer Vision*, vol. 42, pp. 127–144, May 2001.

[62] G. Schaefer, "Robust dichromatic colour constancy," in *Image Analysis and Recognition*, pp. 257–264, Springer Berlin Heidelberg, 2004.

[63] Y.-Y. Li and H.-C. Lee, "Auto white balance by surface reflection decomposition," *J. Opt. Soc. Am. A*, vol. 34, pp. 1800–1809, Oct 2017.

[64] B. Mazin, J. Delon, and Y. Gousseau, "Illuminant estimation from projections on the planckian locus," in *European Conf. on Computer Vision*, pp. 370–379, 2012.

[65] R. T. Tan, K. Nishino, and K. Ikeuchi, "Color constancy through inverse-intensity chromaticity space," *J. Opt. Soc. Am. A*, vol. 21, pp. 321–334, Mar 2004.

[66] C. Riess, E. Eibenberger, and E. Angelopoulou, "Illuminant color estimation for real-world mixed-illuminant scenes," in *International Conference on Computer Vision*, pp. 782–789, Nov 2011.

[67] H. A. Khan, J.-B. Thomas, and J. Y. Hardeberg, "Analytical survey of highlight detection in color and spectral images," in *Computational Color Imaging Workshop*, pp. 197–208, Springer, Cham, 2017.

[68] L. Shi and B. Funt, "Dichromatic illumination estimation via Hough transforms in 3D," in *Conf. on Color in Graphics, Imaging, & Vision*, pp. 259–262, 2008.

[69] V. C. Cardei, B. Funt, and K. Barnard, "Estimating the scene illumination chromaticity by using a neural network," *J. Opt. Soc. Am. A*, vol. 19, pp. 2374–2386, Dec 2002.

[70] H. R. V. Joze and M. S. Drew, "Improved machine learning for image category recognition by local color constancy," in *IEEE International Conference on Image Processing*, pp. 3881–3884, Sept 2010.

[71] J. T. Barron, "Convolutional color constancy," in *IEEE International Conference on Computer Vision*, pp. 379–387, Dec 2015.

[72] V. Agarwal, A. V. Gribok, and M. A. Abidi, "Machine learning approach to color constancy," *Neural Networks*, vol. 20, no. 5, pp. 559–563, 2007.

[73] N. Wang, D. Xu, and B. Li, "Edge-Based Color Constancy via Support Vector Regression," *IEICE Transactions on Information and Systems*, vol. 92, pp. 2279–2282, 2009.

[74] Z. Lou, T. Gevers, N. Hu, and M. Lucassen, "Color constancy by deep learning," in *British Machine Vision Conference*, 2015.

[75] W. Shi, C. C. Loy, and X. Tang, "Deep specialized network for illuminant estimation," in *European Conference on Computer Vision*, pp. 371–387, Springer, 2016.

[76] S. W. Oh and S. J. Kim, "Approaching the computational color constancy as a classification problem through deep learning," *Pattern Recognition*, vol. 61, pp. 405–416, 2017.

[77] S. Bianco, C. Cusano, and R. Schettini, "Color constancy using CNNs," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 81–89, June 2015.

[78] F. Imai and R. Berns, "Spectral estimation using trichromatic digital cameras," in *International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives*, pp. 42–49, 1999.

[79] D. Connah, S. Westland, and M. G. A. Thomson, "Recovering spectral information using digital camera systems," *Coloration Technology*, vol. 117, no. 6, pp. 309–312, 2001.

[80] E. M. Valero, J. L. Nieves, S. M. C. Nascimento, K. Amano, and D. H. Foster, "Recovering spectral data from natural scenes with an RGB digital camera and colored filters," *Color Research & Application*, vol. 32, no. 5, pp. 352–360, 2007.

[81] R. Shrestha and J. Y. Hardeberg, "Spectrogenic imaging: A novel approach to multispectral imaging in an uncontrolled environment," *Opt. Express*, vol. 22, pp. 9123–9133, Apr 2014.

[82] J. Y. Hardeberg and R. Shrestha, "Multispectral colour imaging: Time to move out of the lab?," in *Mid-term meeting of the International Colour Association (AIC)*, (Tokyo, Japan), pp. 28–32, May 2015.

[83] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *12th Color Imaging Conference*, pp. 37–41, 2004.

[84] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Transactions on Image Processing*, vol. 16, pp. 2207–2214, sep 2007.

[85] J.-B. Thomas, "Illuminant estimation from uncalibrated multispectral images," in *Colour and Visual Computing Symposium (CVCS)*, (Gjøvik, Norway), pp. 1–6, Aug 2015.

[86] E. Land, "Recent Advances in Retinex Theory and Some Implications for Cortical Computations: Color Vision and the Natural Image," *Proceedings of the National Academy of Science*, vol. 80, pp. 5163–5169, Aug. 1983.

[87] K. M. Lam, "Metamerism and colour constancy," *Ph. D. Thesis, University of Bradford*, 1985.

[88] S. E. Susstrunk, J. M. Holm, and G. D. Finlayson, "Chromatic adaptation performance of different RGB sensors," in *Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts VI*, vol. 4300, pp. 172–184, International Society for Optics and Photonics, 2000.

[89] C. Li, M. R. Luo, and B. Rigg, "Simplification of the CMCCAT97," *Color and Imaging Conference*, vol. 2000, no. 1, pp. 56–60, 2000.

Chapter 3

Contributions

There have been several research publications during this work. A summary of each of them is discussed in this chapter. In all of the articles, the literature review, methodology, experimental setup, analysis evaluation of results and manuscript writing are performed by the author of this dissertation, except for one article in which the work is performed in collaboration with another group. The contributions of author for that paper are explicitly mentioned in the description.

3.1 Article A: Spectral adaptation transform for multispectral constancy.

H. A. Khan, J. B. Thomas, J. Y. Hardeberg, and O. Laligant, "Spectral adaptation transform for multispectral constancy," *Journal of Imaging Science and Technology*, vol. 62, no. 2, pp 020504-1- 020504-12, 2018.

In this paper, we have proposed the idea of multispectral constancy for illuminant invariant representation of multispectral data. Multispectral constancy is analogous to color constancy but in higher dimensions. We propose the transformation of multispectral data, taken under uncontrolled imaging condition, into a representation that makes the data look like it was taken under a known illuminant. The known illuminant is called the *canonical illuminant* and the representation of multispectral data under such an illuminant is called *canonical representation*.

In this work, we simulated four different multispectral sensors for acquisition of multispectral data from measured reflectance values. The reflectance data used in experiments is the SFU reflectance data and Foster hyperspectral images. Multis-

pectral images are taken after conversion of reflectance data into radiance. Eight illuminants are used with SFU reflectance data while four illuminants are used to create radiance data from Foster hyperspectral images dataset. The preliminary results from simulations are also presented in a conference publication[1].

In this work, we compare the data transformation through a diagonal transform and propose spectral adaptation transform (SAT). The values of diagonal transform consist of scene illuminant in the sensor domain. This transform is used to incorporate the intrinsic properties of imaging sensor. A diagonal transform assumes that there are no inter-channel dependencies. It is true in the case of narrow band filters with almost no overlapping but when that condition is not satisfied, only the diagonal transform is not sufficient. In color imaging, chromatic adaptation transforms are applied to get the color correction. We use the same idea with multispectral imaging and develop SAT for reduction of error in image transformation. In this paper, we have presented the spectral reconstruction results with diagonal transformation of data and also by applying SAT along with diagonal transform. To evaluate the effectiveness of such transforms, we use a case when no transformation is applied to the input multispectral data. Results from spectral reconstruction show that SAT along with diagonal transform provides the best result.

For each multispectral imaging system, separate SAT needs to be calculated once and later that SAT is used every time during image transformation. Diagonal transform requires the information of scene illumination in the sensor domain and it can be acquired either by placing a white diffuser in the scene or through illuminant estimation. In the experiments, we performed estimation of scene illumination for the hyperspectral images of natural scenes. Results show a promising aspect of using SAT along with diagonal transform when evaluated in terms of root mean square error, goodness of fit and calorimetric error.

The findings from this paper are based on simulations and need to be verified through the use of multispectral camera for acquisition of data under varying illuminations. This is illustrated in Chapter 9 through the use of filter-wheel multispectral camera for acquisition of various objects and illuminations.

[1] H. A. Khan, J. B. Thomas and J. Y. Hardeberg, "Multispectral constancy based on spectral adaptation transform", *Scandinavian Conference on Image Analysis*, SCIA, pp 459-470, Tromsø, Norway, 2017.

3.2 Article B: Illuminant estimation in multispectral imaging.

H. A. Khan, J. B. Thomas, J. Y. Hardeberg, and O. Laligant, "Illuminant estimation in multispectral imaging," *Journal of the Optical Society of America A*, vol. 34, no. 6, 2017

In **Article A**, we proposed multispectral constancy which is based on spectral adaptation transform and illuminant in the sensor domain. The information of illuminant in the sensor domain can be acquired by placing a white diffuser in the scene and using the sensor values for that patch. However, it is not feasible all the time to have a white diffuser during image acquisition. In such a case, the illuminant should be estimated and then used in the diagonal transform. In this paper, we propose the methods and experiments for illuminant estimation in multispectral images.

Illuminant estimation is a well studied area in color imaging and there are many methods for this purpose. Those methods can be divided into two categories; image statistics based methods and physics based methods. We determined the feasibility of both methods for extension towards multispectral imaging and found that the image statistics based methods can be extended from three channel color images to N channels. Base on an extensive literature review, we selected four algorithms and extended them for multispectral imaging. We used the hyperspectral images of natural scenes from Foster dataset and simulated multispectral sensors. We used three different configurations for sensor sensitivities and different number of channels (3, 5, 8, 12 and 20).

The results from illuminant estimation are evaluated in terms of angular error. The performance of each of the sensor configurations is evaluated for the proposed illuminant estimation algorithms. The statistical significance of performance is assessed through Wilcoxon signed rank test. Results show that optimal performance is obtained with eight overlapping sensors and Dirac delta sensors are able to provide better results when the number of channels is increased. Among the proposed algorithms, the spectral gray-edge and max-spectral algorithms perform better and are robust to noise.

These results suggest that if the sensor sensitivities on filters are wide and overlapping, the illuminant estimation performance gets better. This is interesting due to the fact that such sensors also reduce the spectral resolution of imaged surfaces which can be observed during spectral reconstruction. Therefore, in **Article A**, we proposed spectral adaptation transform to incorporate such behavior of imaging sensor.

We also found that the bright pixels in an image provide useful information of illuminant estimation. The max-spectral algorithms used such information and provides better performance. Another useful source of information are the edges of materials in a scene which are exploited by the spectral gray-edge algorithm and gave the overall best results. These observations motivated us to explore the highlights based illuminant estimation methods in multispectral imaging. Discussion on those methods is in the following two publications.

3.3 Article C: Analytical survey of highlight detection in color and spectral images.

H. A. Khan, J. B. Thomas and J. Y. Hardeberg, "Analytical survey of highlight detection in color and spectral images", in Proceedings of *Computational Color Imaging Workshop, CCIW*, Milan, Italy, Lecture Notes in Computer Science, vol 10213. Springer, Cham, 2017.

For highlights based illuminant estimation, the first step required by most of the algorithms is to detect the specular reflection in an image. Detection of highlights is a prominent issue in computer vision, graphics and image processing. Applications which require object properties measurement or rendering are affected by specular reflection since the models assume matte diffusing surfaces most of the time. Hence, detection, and sometimes removal, of specular reflection (highlights) in an image may be critical. In this paper, we provide literature review of various highlight detection algorithms and classify them according to their techniques, assumptions and constraints. We make analysis of various categories of highlight detection and removal algorithms on the basis of the use of dichromatic reflection model, requirement of segmenting the image for highlights and the need for white balancing. We also provide an overview of the highlights detection algorithms for multispectral imaging.

A qualitative comparison of the highlights detection algorithms is presented in this paper and we pointed out that in most of the literature, results from detection and removal of highlights is presented qualitatively while comparing with other methods. Although qualitative analysis provide a general overview, quantitative measurement should also be used to provide a fair comparison. Once the highlight areas in an image are detected, the next step is to use it for estimation of scene illumination. It is described in the next paper.

3.4 Article D: Towards highlight based illuminant estimation in multispectral images.

H. A. Khan, J. B. Thomas and J. Y. Hardeberg, "Towards highlight based illuminant estimation in multispectral images," in Proceedings of *International Conference on Image and Signal Processing*, Cherbourg, France, Lecture Notes in Computer Science, vol 10884. Springer, Cham, 2018.

In this paper, we provide a review of the physics based illuminant estimation methods. In **Article C**, image statistics based illuminant estimation methods are discussed and we identified that bright pixels provide valuable information. In physics based illuminant estimation approaches, the analysis of interaction of light with surfaces is performed for estimating the scene illuminant. The optically inhomogenous objects (e.g. plastics, ceramics, paints, etc) create a neutral interface reflection condition, where the spectral power distribution of the specular reflection is the same as the incident illumination; and this property is used to estimate the scene illuminant. The basis of illuminant estimation from specular highlights in images is the dichromatic reflection model, which describes the reflected light from an object as a combination of the diffuse and specular reflection. The specular reflection is also termed as highlights in an image, and therefore, the problem of illuminant estimation is treated as the analysis of highlights from surfaces.

In this work, we reviewed the illuminant estimation methods which are based on highlights in images and identified the potential methods which can be extended to multispectral imaging. Based on this qualitative analysis, the dichromatic reflection model based methods that work in two dimensional chromaticity space can be computationally complex and unstable for extension to multispectral imaging. We identified some techniques that can be extended for multispectral imaging. Quantitative analysis of such an extension is also required before reaching to a conclusion. By combining the current work with state of the art highlight detection algorithms described in the earlier paper, the reader can pick the algorithms which best suit the type of images and data to be processed for illuminant estimation.

One of the limitations of this work is that we did not explicitly use the highlights based illuminant estimation methods further. The reason for it is the presence of saturated pixels in multispectral data that is acquired and presented in Chapter 9. However, we used the spectral gray-edge based illuminant estimation method and obtained good performance. The spectral gray-edge algorithm used the edge information from a material's surface. The incident light on a material forms sheen

reflection at the edges and provides the information about the illumination. Therefore, in a way we can express the spectral gray-edge algorithm as a method based on both image statistics and highlights based algorithm. Further investigation of physics based illuminant estimation methods for multispectral imaging is one of the future aspects from the work presented in this dissertation.

3.5 Article E: Color characterization methods for a multispectral camera.

H. A. Khan and P. Green, "Color characterization methods for a multispectral camera," in Proceedings of *IS&T International Symposium on Electronic Imaging: Color Imaging XXIII: Displaying, Processing, Hardcopy, and Applications*, San Francisco, USA, 2018

In this work, we evaluate the performance of multispectral camera as a colorimeter. We measure the camera characterization performance by two methods, linear mapping and through spectral reconstruction. Linear mapping is used in 3-channel camera characterization and we use the same method for a multispectral camera. We also investigate if instead of linear mapping, spectral reconstruction from the camera data improves the performance of color reproduction or not. The recovery of reflectance spectra is an ill-posed problem and certain assumptions are required for obtaining a unique solution. Linear methods are generally used for spectral reconstruction from the camera data and are based on training on known spectra. These methods can perform well when the test data consists of a subset of the training spectra, however, their performance is significantly reduced when the test data is different. In this paper, we also investigate the role of training spectra for the camera characterization. Five different spectral reflectance datasets are used for training the camera characterization models.We investigated the reflectance models created with reflectance spectra from ISO 17321, GretagMacbeth ColorChecker, Skin reflectance, Munsell chips and SFU dataset. The SFU dataset consists of reflectance spectra from a number of objects including paints, natural objects, color patches and dyes. The generic nature of this dataset makes it a good training data for the spectral reflectance reconstruction system. We found that by using linear method, efficient colorimetric information can be obtained and the results outperform non-linear method being tested.

We did not use illuminant estimation and data transformation methods in this study. However, this work demonstrates the use of multispectral camera as colorimeter and it can be anticipated from the promising results of this work that a multispectral camera can be used in uncontrolled imaging environment for obtaining colorimetric data of a scene. This work is also used for visualization of multispectral images

under a canonical illuminant in our work presented in the following.

3.6 Article F: Towards the use of multispectral camera as spectrophotometer in uncontrolled illumination conditions.

After defining the multispectral constancy based of spectral adaptation transform and diagonal correction for illumination, and getting promising results in simulation, we made experiments with a filter-wheel multispectral camera.Three images with varying illuminations are captured and the spectra of material surfaces is reconstructed. The scenes consisted of various objects and some of those objects are imaged through the hyperspectral camera. The reflectance data from hyperspectral images is used as ground truth and is used to compare the spectral reconstruction results.

The acquired multispectral data is transformed into canonical representation through manual selection of white patch of ColorChecker in the image and by estimating the illuminant. Spectral reconstruction is performed by using Wiener estimation method and results are evaluated in terms of root mean square error, cosine distance and spectral angle mapper. Results show a promising aspect of multispectral imaging as they can be used as a spectrophotometer for getting spectral information of a whole scene. The proposed concept of multispectral constancy is valid for both filter-wheel and snapshot type of multispectral cameras. Having such promising results, we are one step more close towards enabling the use of multispectral imaging for computer vision applications.

One of the shortcoming that we noticed while trying to perform object detection through spectral information is that not only the spectral information is required for an image but the spatial information of texture is also an important feature. In order to investigate the spectral and spatial features of objects and material surfaces, a dataset is required. We joined another team from University of Lille, France, who are experts in texture analysis. Through the collaboration of both teams, we created a hyperspectral reflectance dataset for various material surfaces. Details are that dataset are as follows;

3.7 Article G: HyTexiLa: High resolution visible and near infrared hyperspectral texture images.

> H. A. Khan, S. Mihoubi, B. Mathon, J. B. Thomas and J. Y. Hardeberg, “HyTexiLa: High resolution visible and near infrared hyperspectral texture images,” Sensors, vol. 18, no. 7, 2045; 2018.

Hyperspectral images contain more spectral information compared to color im-

ages. The high dimensionality of hyperspectral image data is an open challenge, and a trade-off between high information content and practical handling is often required. To evaluate this compromise and assess its performance, high quality data with high spatial and high spectral resolutions are required. Creating such a dataset takes time and effort, but it is necessary to the research community. Considering these needs, we present a hyperspectral image dataset of 112 textured objects falling into five different categories. Images are reflectance data that span the visible and near infrared (NIR) parts of the electromagnetic spectrum. The data is referred to as HyTexiLa (Hyperspectral Texture images acquired in Laboratory). The idea behind creating this dataset is to provide a platform for the benchmark analysis of various applications and processing. The areas where this data could be used are in the fields of image processing, computational imaging and computer vision, such as surface identification, spatio-spectral analysis of textured surfaces, image sensor simulation, color reproduction, image relighting and so on. The availability of a high spectral and spatial resolution dataset will provide easy access for the evaluation of different techniques and will allow the results to be compared. In this paper, we present that dataset and analyze its spatial and spectral properties. We define the image acquisition protocol, the distortions that occur during the acquisition of the objects, the method for the correction of such distortions and the effect of such corrections. Our focus in this paper is to present a hyperspectral dataset to the community and to provide an analysis on our dataset as a benchmark for further research. We provide the methodology for pre-processing the raw hyperspectral data before using it for the intended applications.

In this work we used principal component analysis for determining the effective dimension of hyperspectral reflectance data and used non-negative matrix factorization (NMF) for decomposition of spectral data into the corresponding components [2]. Spectral decomposition of imaged surfaces provide valuable information about their physical and chemical behavior. The spectral decomposition of a leaf provided us the information of chlorophyll and areas causing green color of the leaf. Similarly, the spectral decomposition of textile sample provided us information about different materials in it. Such a study is valuable in acquisition of spectral signatures for determination of various properties from the imaged surface.

This dataset is available for scientific use and simulations, and it provides a benchmark to test various computer vision algorithms that are related to object classification and material identification. This dataset will also help in designing optimal

[2] Author of this dissertation have partial contribution in post processing of raw hyperspectral data and no contribution in texture analysis. Details of contributions to work in this paper can also be found at the end of the published article

spectral sensors, computational imaging systems and spectral reconstruction algorithms.

3.8 Summary of contributions

In this dissertation, the concept of multispectral constancy is introduced. Multispectral constancy is analogous to color constancy but in higher spectral dimension. It is achieved through spectral adaptation transform and diagonal correction. The spectral adaptation transform is used to incorporate the intrinsic properties of imaging sensor while the diagonal correction is used for correction of scene illuminant. The information of scene illuminant in the sensor domain can be found through the use of a white diffuser in the image or through illuminant estimation. In this work, we have presented methods for illuminant estimation in the sensor domain. The demonstration of multispectral constancy is provided in two experiments where the multispectral data is used for spectral reconstruction and for obtaining colorimetric values from a scene.

Figure 1.1 gives the summary of research modules, along with the publications generated by working on specific modules or sections of this research. The true potential of multispectral imaging can be realized after those systems are widely used imaging in uncontrolled imaging environments. The work presented in this dissertation makes it one step closer towards the use of multispectral images for computer vision applications.

Chapter 4

Spectral adaptation transform for multispectral constancy

Article A

This chapter is a reformatted reprint of the publication

H. A. Khan, J. B. Thomas, J. Y. Hardeberg, and O. Laligant, "Spectral adaptation transform for multispectral constancy," *Journal of Imaging Science and Technology*, vol. 62, no. 2, pp 020504-1-020504-12, 2018.

The spectral reflectance of an object surface provides valuable information of its characteristics. Reflectance reconstruction from multispectral image data is typically based on certain assumptions. One of these common assumptions is that the same illumination is used for system calibration and image acquisition. We propose the concept of multispectral constancy which transforms the captured sensor data into an illuminant-independent representation, analogously to the concept of computational color constancy. We propose to transform the multispectral image data to a canonical representation through spectral adaptation transform. The performance of such a transform is tested on measured reflectance spectra and hyperspectral reflectance images. We also investigate the robustness of the transform to the inaccuracy of illuminant estimation in natural scenes. Results of reflectance reconstruction show that the proposed spectral adaptation transform is efficient and is robust to error in illuminant estimation.

4.1 Introduction

The formation of an image depends on the spectral reflectance of the surface being viewed, the spectral power distribution and intensity of the illumination, and the spectral sensitivities of the sensors. In the human visual system, the three cone types act as sensors. In case of a camera with 3 channels, the RGB filters together with the imaging sensor, play a similar role. The human visual system has the natural ability to perceive constant color of surfaces despite the change in spectral composition of the illuminant, and this ability to discard illumination effects is called *Color Constancy* [1]. Creation of such models for illuminant invariant representation of scenes in computer vision is called computational color constancy [2]. An illuminant invariant representation is important for computer vision applications including object recognition, tracking and image classification [3]. There are two major techniques for achieving computational color constancy. One method is to compute illuminant invariant features, and a second method is to estimate the illuminant and later apply a correction [4]. In this paper, we use the former method of illuminant estimation.

The advancement in sensor technology has developed the use of multispectral imaging for indoor and close range imaging. The ability of multispectral imaging in acquisition of more spectral information is useful for object and material classification and identification by means of spectral reconstruction [5–9] of surfaces in a scene. The need for spectral reconstruction of surfaces was recognized in 1980s [10–14]. Since then, many methods are developed to provide spectral reconstruction from the camera data. Most of these methods rely on the use of training data to learn the mapping between the camera data and the desired spectra. This process is called *calibration* of the system and is performed through a training set of measured reflectances and radiance data with a given illuminant. To maintain a reasonable accuracy, the same illuminant is required for scene acquisition. This limitation of having the same illuminant for calibration and image acquisition is a major shortcoming for generic use of multispectral imaging [15].

To overcome this limitation, we propose to transform the acquired multispectral data under any unknown illumination into its canonical representation. This transformation requires the estimation of scene illuminant. For multispectral images, estimation of illuminant in the sensor domain by using image statistics based methods is proposed by Thomas [16] and Khan *et al.* [17]. In their work, the illuminant estimation methods for color images are extended from three channels to K channels. The estimated illuminant is used for the diagonal correction [18] to transform the input sensor data into a canonical representation. We call this consistent representation of multispectral data as *multispectral constancy*. A preliminary proposal

of this concept is provided in [19].

To achieve multispectral constancy, we use a diagonal transform and also introduce a *spectral adaptation transform* (SAT). The concept of SAT being proposed is closely related to the data based sensor sharpening transform by Finlayson *et al.* [20]. The main difference between these strategies is that instead of finding a sharpening transform and then diagonalizing it, we want to optimize the result after applying the diagonal transform, where the elements of diagonal transform are obtained from illuminant estimation in the sensor domain. The proposed idea of multispectral constancy and optimization of diagonal transform is tested with a simulated multispectral camera on the reflectance dataset and hyperspectral images of real scenes from the Foster dataset [21]. The advantage of having a canonical representation is that the spectral reconstruction system can be calibrated with a canonical illuminant and hence, the condition of having same illuminants for training under which the scene is acquired, is no longer needed. Results show a promising aspect of the use of multispectral imaging for outdoor scene acquisition under uncontrolled illumination condition.

This paper is organized as follows. In Section 4.2, we define multispectral constancy after formalizing a model for the multispectral image acquisition system. System calibration and spectral reconstruction are defined as linear problem in the section. Section 4.3 defines our experimental protocol based on simulations. Results are analyzed in Section 4.4 before we conclude.

4.2 Multispectral constancy

4.2.1 Definition

In a simplified noiseless imaging model, formation of an image depends on the spectral sensitivity of imaging sensor $c(\lambda)$, spectral reflectance of the surface $r(\lambda)$ and the spectral power distribution of illuminant $e(\lambda)$. This formation for the visible wavelength spectrum ω is defined as

$$f = \int_{\omega} r(\lambda)e(\lambda)c(\lambda)d\lambda \tag{4.1}$$

In practice, we can formulate an extended and discrete version of Eq. 4.1 as;

$$\mathbf{F} = \mathbf{REC} \tag{4.2}$$

Considering the spectral sampling (N) of 10nm within the wavelength range of 400nm to 700nm and K number of spectral channels, $\mathbf{F}$ is $S \times K$ matrix (S is the number of spectral samples), $\mathbf{R}$ is $S \times N$ matrix of surface reflectance, $\mathbf{E}$ is

the diagonal matrix ($N \times N$) of the scene illuminant and $\mathbf{C}$ is $N \times K$ matrix, consisting of spectral sensitivities of the channels.

Here we consider two cases of image acquisition with the same imaging system. One image is acquired with a canonical illumination $\mathbf{E}_c$ and another image with an unknown illuminant $\mathbf{E}_{ill}$. We present both cases in parallel in Eq. 4.3.

$$\mathbf{F}_{ill} = \mathbf{R}\mathbf{E}_{ill}\mathbf{C} \quad ; \quad \mathbf{F}_c = \mathbf{R}\mathbf{E}_c\mathbf{C} \tag{4.3}$$

To perform the spectral reflectance estimation $\hat{\mathbf{R}}$ from the imaged data in both of the above mentioned cases, a generalized inverse, denoted by $^+$, can be applied as;

$$\hat{\mathbf{R}} = \mathbf{F}_{ill}\mathbf{C}^+\mathbf{E}_{ill}^+ \quad ; \quad \hat{\mathbf{R}} = \mathbf{F}_c\mathbf{C}^+\mathbf{E}_c^+ \tag{4.4}$$

However, $\mathbf{C}$ is not necessarily a square matrix and therefore, it is not trivial to compute the inverse. Computational procedures are applied to achieve the task of spectral reconstruction. There are several works in literature where the spectral reconstruction is performed by linear transform through a calibration matrix. Linear methods are popular for learning the mapping between camera data and desired output (spectral reflectance). The idea behind using linear methods is that when the reflectance spectra is continuous and band-limited [10, 13], the statistical analysis of measured reflectances of standard color samples are enough to calibrate the spectral reconstruction system [14, 22].

By using the calibration matrix $\mathbf{W}$, the equations for spectral reflectance reconstruction become $\hat{\mathbf{R}} = \mathbf{F}_{ill}\mathbf{W}_{ill}$ and $\hat{\mathbf{R}} = \mathbf{F}_c\mathbf{W}_c$ for both cases, respectively. This calibration is specific for a given illumination $\mathbf{E}_c$. The scene illuminant plays a direct role in the spectral reconstruction [7, 23], because of its effect in the camera image formation, as in Eq. 4.1.

With $\mathbf{F}_c$, the spectral reconstruction system can work efficiently, since the calibration is already performed with the same canonical illuminant, but with $\mathbf{F}_{ill}$, the calibration matrix needs to be recomputed with $\mathbf{E}_{ill}$. Measurement of scene illuminant and calibration matrix for each change in imaging environment is a difficult task and is not a practical solution for a situation where the illumination is not constant all the time. To avoid this problem, we propose to transform the acquired multispectral data $\mathbf{F}_{ill}$ into a canonical representation $\mathbf{F}_c$. Thus, $\mathbf{W}_c$ can be used for the spectral reconstruction from the multispectral data, being taken under any illumination. We denote such a transform as $\mathbf{M}^{c,ill}$, which maps the camera data $\mathbf{F}_{ill}$, taken under unknown illuminant $\mathbf{E}_{ill}$, into its canonical representation $\mathbf{F}_c$.

$$\mathbf{F}_c = \mathbf{M}^{c,ill}\mathbf{F}_{ill}. \tag{4.5}$$

Once such a transform is available, then the problem of spectral reconstruction is limited to finding the transform $\mathbf{M}^{c,ill}$. The spectral reconstruction in this case is mathematically represented as;

$$\hat{\mathbf{R}} = \mathbf{W}_c \mathbf{M}^{c,ill} \mathbf{F}_{ill} \tag{4.6}$$

We call this concept of illuminant invariant representation of multispectral data as *multispectral constancy*. By achieving multispectral constancy, the requirement of having the same illumination for calibration and scene is no longer required.

In this work, we propose the use of illuminant estimation for achieving multispectral constancy. In [19], a diagonal transform is used as $\mathbf{M}^{c,ill}$ and the preliminary results for spectral reconstruction from the sensor data were provided. Such a transform requires the estimation of scene illuminant in the sensor domain and the sensor data is corrected from effects of illumination through a diagonal transform ($\mathbf{D}$). The diagonal matrix $\mathbf{D}^{c,ill}$ is a $K \times K$ matrix and the components of this matrix are the sensor responses to the illuminations $\mathbf{E}_c$ and $\mathbf{E}_{ill}$. As an alternate explanation, sensor response to the illuminants can be understood as the sensor response to a perfect white diffuser under a certain illuminant, and denoted by $\mathbf{w}_c$ and $\mathbf{w}_{ill}$ for the two illuminants. The matrix $\mathbf{D}$ is defined as;

$$\mathbf{D}^{c,ill} = diag(\mathbf{w}_c / \mathbf{w}_{ill}) \tag{4.7}$$

With the use of diagonal transform as in Eq. 4.7, the problem of color constancy (for 3-channel images) and multispectral constancy (for multispectral data) is reduced to the estimation of K parameters of the diagonal transform. $\mathbf{D}$ is applied on each channel independently and is used in many color constancy algorithms [24, 25].

With a perfect white diffuser, Eq. 4.5 holds and the input illuminant is transformed into the desired canonical illuminant with a diagonal transform. However, when the surface reflectance is not constant across the wavelength spectrum, then this transformation generates errors. The diagonal transform also works well when the bandwidth of spectral sensitives of filters are within 100 to 150 nanometers [26, 27]. However, in case of large band filters, the diagonal transform may not be sufficient. Therefore, we investigate the performance of diagonal transform and propose to minimize the error during diagonal transformation through a spectral adaptation transform $\mathbf{A}_{SAT}$. The idea of SAT is to incorporate the sensors response for efficient transformation, along with the diagonal transform. The concept of spectral adaptation transform is analogous to chromatic adaptation transform [28], while here we consider multispectral data, rather than the three-channel color images.

4.2.2 Related state of the art

To improve the performance of $\mathbf{M}$ for color images, Finlayson *et al.* [20] proposed *sensor sharpening* which aims at finding a linear combination of the spectral sensitives of a camera, with respect to which a diagonal transform for illuminant transform works the best. The idea in this technique is to transform the sensor responses so that they appear to be *sharper* than the original ones. Finlayson *et al.* [20] proposed three methods for finding the sharpening transform $\mathbf{T}$. The first method is called *sensor based sharpening*, and finds $\mathbf{T}$ by optimizing the intuitive notion of sharpness for each filter individually. This optimization requires the knowledge of sharpening interval in the wavelength spectrum and the Lagrange multiplier [20]. The second method is called *perfect sharpening*, in which it is assumed that the surface reflectances can be fitted in a three dimensional linear model while the illuminants can be fitted in a two dimensional linear model. The third method is *data based sharpening*, in which RGBs are generated from a set of known spectral reflectances, using a specific camera and two illuminations, one as canonical illumination $\mathbf{E}_c$ while the other as test illumination $\mathbf{E}_{ill}$. The sharpening transform $\mathbf{T}$ is added on both sides of Eq. 4.5 to reduce the error in mapping. It is expressed as;

$$\mathbf{TF}_c = \mathbf{MTF}_{ill}. \tag{4.8}$$

According to Eq. 4.8, $\mathbf{M}$ can be optimal in the least-square sense if it can be defined by the Moore-Penrose inverse ($^+$);

$$\mathbf{M} = \mathbf{TF}_c[\mathbf{TF}_{ill}]^+ \tag{4.9}$$

Since $\mathbf{T}$ is still unknown in Eq. 4.9, the unknown terms are shifted to the left hand side;

$$\mathbf{T}^{-1}\mathbf{MT} = \mathbf{F}_c[\mathbf{F}_{ill}]^+ \tag{4.10}$$

The sharpening transform T is found through eigenvector decomposition of the expression on left side of Eq. 4.10. In their work, the expression $[\mathbf{T}]^{-1}\mathbf{MT}$ is diagonalized so that $\mathbf{M}$ is replaced by a diagonal matrix $\mathbf{D}$. Finlayson *et al.* [20] found similar $\mathbf{T}$ for sensor based sharpening and data based sharpening.

The sharpening transform can sometimes have negative values and there are some investigations to reduce this problem [29]. Data based sharpening was further improved by Barnard *et al.* [30] by using an average of measured illuminants, and introduced a parameter for prioritizing the positivity. Drew and Finlayson [31] proposed data driven positivity by adding constraints to ensure that all the values in sharpening transform are positive. Chong *et al.* [32] proposed the measurement

tensor technique for finding $\mathbf{T}$. Spherical sampling [33] is also introduced as a technique for spectral sharpening where the discretely sampled points on a sphere are found and related to the original sensors. Sharpening through filter chart calibration was proposed by Abdellatif [34]. An overview of spectral sharpening methods is provided by Corral and Bertalmío [29].

In case of multispectral imaging, the concept of spectral sharpening is not straightforward due to the higher dimensionality of sensors. In [33], spectral sharpening for six multispectral sensors is discussed. Their method is computationally expensive since there is need to generate a sphere of the sensor dimension which is not trivial. For a dimension above 3, the spherical sampling [33] has to be extended to hyper-sphere computation. Due to these complications, we do not use the sensor based sharpening method. We do not use the data based sharpening method either, because we are not interested in diagonalizing the result of Eq. 4.10. Instead, we are using $\mathbf{D}$ as the diagonal matrix containing the illuminant in the sensor domain, as in Eq. 4.7, and to improve the efficiency of such a transform in the same way as in spectral sharpening. It can be argued that instead of improving the diagonal transform, why not find any other linear transform which can serve the same purpose more efficiently? The reason for preferring the diagonal transform (and its improvement) in our work is the fact that knowledge of scene illuminant is of major importance in computer vision applications. Therefore, either it can be measured for a specific scene or can be estimated, and this information can be used by our proposed method. Another reason is that most of the color constancy algorithms are defined in terms of a diagonal transform, where the elements of diagonal transform are found through illuminant estimation in the sensor domain.

For achieving illuminant invariant representation in multispectral imaging, Abrardo *et al.* [35] proposed the concept of color constancy for multispectral imaging by linearly transforming the sensor response under an unknown illuminant. The transformation matrix $\mathbf{M}$ is determined through the least-square solution as;

$$\mathbf{M} = \mathbf{F}_c \mathbf{F}_{ill}^{+} = \mathbf{F}_c \mathbf{F}_{ill}^{t} (\mathbf{F}_{ill} \mathbf{F}_{ill}^{t})^{-1} \tag{4.11}$$

The transformation $\mathbf{M}$ is applied to the acquired multispectral data and then spectral reflectance is estimated. It is interesting to note that the right side of Eq. 4.11 resembles with the concept of data based sharpening (Eq. 4.10). This method works well when the canonical representation ($\mathbf{F}_c$) for the same scene is available. In [35], authors claim that 97% accuracy in spectral reconstruction is achieved by their method. The problem with this technique is that in the absence of $\mathbf{F}_c$, it is not possible to use Eq. 4.11. In fact, our aim is to transform the input multispectral data into its canonical representation, while in Eq. 4.11, the availability of canonical representation of same scene is assumed, which is not feasible for every image

being captured.

4.2.3 Proposal for computation of M

It can be observed that Eq. 4.10 and 4.11 are originally formulated from Eq. 4.5 and the goal is to find the transform $\mathbf{M}$. In spectral sharpening, $\mathbf{M}$ is found by eigenvector decomposition of Eq. 4.10 and diagonalizing it, while the product of $\mathbf{F}_c$ and $\mathbf{F}_{ill}$ is used directly in Eq. 4.11. We propose decomposing $\mathbf{M}$ into two elements; $\mathbf{D}^{c,ill}$ as in Eq. 4.7 and $\mathbf{A}_{SAT}$. In this proposed method, the diagonal transform performs the transformation of multispectral data taken under unknown illuminant $\mathbf{E}_{ill}$, into its canonical representation under the illuminant $\mathbf{E}_c$. The role of $\mathbf{A}_{SAT}$ is to incorporate the spectral response of imaging sensors and improve the transformation of camera data into its canonical representation. The optimal $\mathbf{A}_{SAT}$ should minimize the error for all reflectances i and illuminants j.

$$\mathbf{F}_c^i = \mathbf{A}_{SAT}\mathbf{D}^{c,j}\mathbf{F}_j^i \tag{4.12}$$

To find the best $\mathbf{A}_{SAT}$, Eq. 4.12 is written as the explicit minimization of an error function in Eq. 4.13;

$$\min_{\mathbf{A}_{SAT}} = \sum_j |\mathbf{F}_c^i - \mathbf{A}_{SAT}\mathbf{D}^{c,j}\mathbf{f}_j^i| \tag{4.13}$$

For a given sensor configuration, a generic $\mathbf{A}_{SAT}$ is created by minimizing the error for all the reflectances and illuminations in the training dataset.

The advantage of our proposed technique is that only the knowledge of sensor sensitives of camera is required to compute $\mathbf{A}_{SAT}$. This specific $\mathbf{A}_{SAT}$ for a camera can be used to transform the multispectral data captured under any illuminant, into its canonical representation through the diagonal correction for illumination. The only parameter left in the proposed system is the performance of illuminant estimation algorithm, since the elements of diagonal transform are obtained through the illuminant estimation.

The multispectral data being acquired under any illumination is transformed into a canonical representation after estimating the scene illuminant, and then the precalibrated $\mathbf{W}_c$ is used. The advantage of our proposed idea of multispectral constancy is that it is no more necessary to acquire the multispectral data in a controlled environment.

Once the transform $\mathbf{M}^{c,ill} = \mathbf{A}_{SAT}\mathbf{D}^{c,ill}$ is defined, spectral reconstruction can be mathematically represented as;

$$\hat{\mathbf{R}} = \mathbf{W}_c\mathbf{A}_{SAT}\mathbf{D}^{c,ill}\mathbf{F}_{ill} \tag{4.14}$$

Preliminary results by using the diagonal transform $\mathbf{D}^{c,ill}$ are provided in [19]. In the current work, we are using $\mathbf{D}^{c,ill}$ along with $\mathbf{A}_{SAT}$. The experimental protocol for testing the idea of multispectral constancy through SAT is provided in the following section.

4.3 Experiments

To implement and validate the idea of multispectral constancy through SAT, we use measured reflectance data of 1995 surfaces from the SFU reflectance dataset [36]. These surfaces include the 1269 Munsell chips, 24 patches of the X-rite ColorChecker, 170 natural objects [37], 120 Dupont paint chips [37], 350 surfaces from the Krinov dataset and 57 surfaces measured by Barnard *et al.* [36]. This dataset is used as the training data for calibration of $\mathbf{W}_c$ with D65 as the canonical illuminant. The reflectance data is in the wavelength range of 400 to 700 nm with 10 nm sampling.

For testing the proposed idea of multispectral constancy, we use 24 patches from the X-rite ColorChecker and 1296 Munsell chips as the test data and acquire them using the simulated multispectral sensors. Spectral reconstruction is done from the acquired multispectral data as in Eq. 4.14. To validate the usefulness of proposed idea for natural outdoor images, we use the reflectance data from Foster dataset of hyperspectral images [21]. These hyperspectral images are within wavelength range of 400 nm to 720 nm, but we use these images within wavelength range of 400 nm to 700 nm, since the training dataset [36] is within this range. In the following sections, the details of experimental setup are provided.

4.3.1 Sensor

Diagonal transform works well when the sensor sensitives are narrow-band (ideally Dirac delta functions). Following this, it seems as the use of narrow band filters is the optimal choice to be used in imaging systems. The term $\mathbf{F}_c\mathbf{F}_{ill}^t$ in Eq. 4.11 resembles the *spectral sharpening*, introduced by Finlayson *et al.* [20]. They proposed a linear combination to transform the original sensor responses into narrower bands. However, from a practical point of view, one reason for not employing such narrow band systems is that the acquisition time, complexity and cost of such system is high as compared to wide band systems [38]. Other reasons are the fact that the narrow band systems are not an optimal choice for illuminant estimation [16, 17] and demosaicing (when multispectral filter array is used [39]). Wang *et al.* [40] studied the influence of increase in number of bands and found that increasing the number of spectral bands cause reduction in performance of spectral reconstruction. Also, the efficiency of illuminant estimation algorithms decreases when the number of spectral filters is increased [17]. In this work, we are limit-

ing the experiments to linear systems and our aim is to investigate the concept of multispectral constancy for a generalized multispectral imaging system. We use 5, 8 and 12 spectral bands for testing the proposed framework of spectral constancy. In the results of illuminant estimation in multispectral images by Khan *et al.* [17], the accuracy of illuminant estimation is reduced when the number of spectral bands is increased beyond 8. On the other hand, having more than 8 bands still increase the spectral estimation. For this reason, we limit the number of bands within this range for our experiments.

In our experiments, we use a Gaussian model of sensor sensitivities. Such a model has been extensively used in the literature to simulate sensors or to approximate Fabry-Pérot filter transmittance [41]. This configuration is called the equi-Gaussian [16]. This sensor configuration is tested for illuminant estimation [17] and spectral estimation [19]. The Full Width at Half Maximum (FWHM) of the sensor sensitivities decrease with increase in number of bands and the overlap between adjacent bands remain approximately the same. By increasing the number of bands (K) in this configuration, we are gradually shifting from typical multispectral sensors towards hyperspectral sensors.

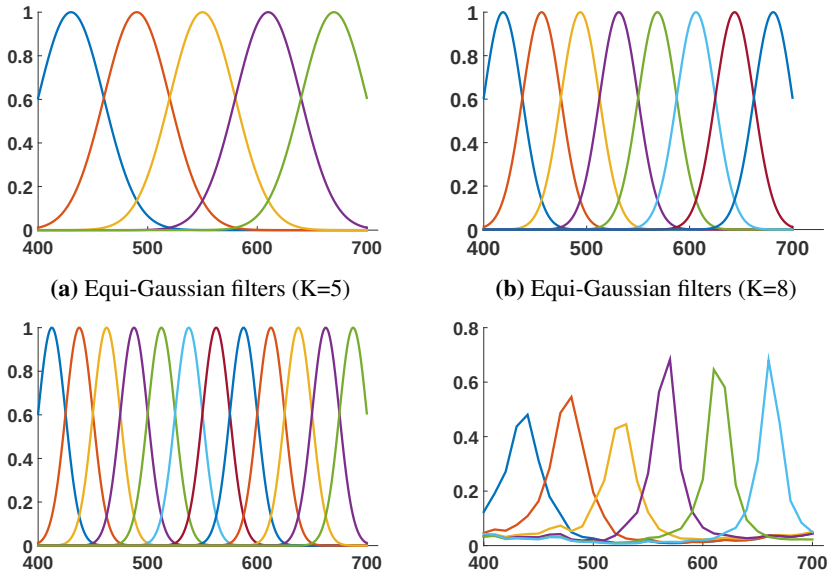
For testing the proposed method of SAT for spectral reconstruction, we also use measured sensitivities of a real implementation of spectral filter array (SFA) camera [42]. There are eight filters in this SFA camera. The first six filters are used in our experiments as the available spectral reflectance data for training (SFU dataset) is within the wavelength range of 400 nm to 700 nm. Figure 4.1 shows the spectral sensitivities of each filter being used in the experiments.

4.3.2 $\mathbf{A}_{SAT}$ computation

For a given sensor configuration, $\mathbf{A}_{SAT}$ is computed by using the set of 102 illuminants and surface reflectance of 1995 surfaces from the SFU data [36]. Radiance data is generated using an illuminant j (from the set of 102 illuminants) and is acquired as multispectral data. For each illuminant, a diagonal transform is created by using Eq. 4.7 and the multispectral data under that illuminant is transformed by using Eq. 4.5. The desired output after applying $\mathbf{M}^{c,ill}$ is the multispectral data under the canonical illuminant (D65). As discussed in Section 4.2.3, we propose to decompose $\mathbf{M}^{c,ill}$ into diagonal transform $\mathbf{D}^{c,ill}$, and $\mathbf{A}_{SAT}$. This $\mathbf{A}_{SAT}$ is unique for a given sensor configuration and is computed by using Eq. 4.13. Once $\mathbf{A}_{SAT}$ is computed, Eq. 4.12 is used to transform $\mathbf{F}_{ill}$ into $\mathbf{F}_c$.

4.3.3 Illuminant estimation

For testing and validating the proposed concept of spectral adaptation transform, we keep $\mathbf{D}^{c,ill}$ fixed for the spectral reflectance data. For obtaining the elements



(a) Equi-Gaussian filters (K=5)

(b) Equi-Gaussian filters (K=8)

(c) Equi-Gaussian filters (K=12)

(d) 6 sensors from the SFA

Figure 4.1: Spectral sensitivities of the filters being used

of $\mathbf{D}^{c,ill}$, we use the test and canonical illuminants in the sensor domain. Initially we want to validate the efficiency of $\mathbf{A}_{SAT}$ while testing the reflectance dataset, therefore we assume that the illuminant is estimated in the sensor domain without error.

In practical cases, the information about scene illumination may not be available all the time. This is where the illumination estimation algorithms come into play. By using an illuminant estimation algorithm, the elements of diagonal transform are obtained and then the pre-calculated SAT is used.

For testing the proposed concept for natural scenes, we perform illuminant estimation in the multispectral data of natural scenes and use the Max-Spectral Algorithm, which is the extension of Max-RGB algorithm [43]. The extension of this algorithm from color to spectral is proposed and analyzed in detail in [16, 17]. The estimated values of illuminant, which are defined in the sensor domain, are used in the diagonal transform $\mathbf{D}^{c,ill}$. For each image, $\mathbf{D}^{c,ill}$ is estimated individually and this estimation may consist of error as well. We report the error in illuminant estimation in form of angular error (ΔA).

4.3.4 Spectral reflectance reconstruction

As explained in Section 2, a calibration matrix $\mathbf{W}$ is required for the spectral reconstruction from camera data. It is obtained by using measured reflectance spectra $\mathbf{R}_t$ and the camera sensor sensitivities ($\mathbf{M}$). For reducing the error between original spectra $\mathbf{R}$ and the estimated spectra $\hat{\mathbf{R}}$, a covariance matrix of a set of measured reflectance samples can be used. Those measured reflectance samples provide the a-priori statistical information about the surfaces in a scene [44]. If the a-priori information is well chosen, error in the spectral reconstruction can be minimal.

There are several methods being proposed for the spectral reconstruction in literature [45]. We use a linear method for clarity, namely the Wiener estimation [46] because of its robustness to noise. It is defined as

$$\mathbf{W} = \mathbf{R}_t\mathbf{R}_t^T(\mathbf{CE})^T((\mathbf{CE})\mathbf{R}_t\mathbf{R}_t^T(\mathbf{CE})^T + \mathbf{G})^{-1}. \tag{4.15}$$

Here, $\mathbf{R}_t\mathbf{R}_t^T$ and $\mathbf{G}$ are the autocorrelation matrices of the training spectra and additive noise, respectively. $\mathbf{G}$ is in the form of a diagonal matrix consisting of the variance of noise σ^2. Training for obtaining the matrix $\mathbf{W}_c$ is performed with CIE illuminant D65 as the canonical illuminant $\mathbf{E}_c$. The obtained calibration matrix is used for spectral reconstruction in Eq. 4.14.

4.3.5 Evaluation

For testing the proposed idea of multispectral constancy, radiance data is simulated from the measured test spectra with CIE illuminants A, D50, D55, D75, F5, F7 and F12. We also use the LED (Philips SlimStyle: 2700K) as an illuminant in the experiments.

The proposed idea is also tested on hyperspectral images of scenes consisting of vegetation and urban areas. We create radiance data from these images using illuminants A, D50, D55 and D75. The same procedure of reflectance reconstruction for each pixel of simulated multispectral image is performed and evaluated.

To measure the performance of the spectral reconstruction, we compare the reconstruction $\hat{\mathbf{r}}$ for each patch of the reconstructed reflectance with the corresponding measured reflectance $\mathbf{r}$, through root mean square error (RMSE) as

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(r_j - \hat{r}_j)^2} \tag{4.16}$$

We also use goodness of fit coefficient (GFC) [47] for evaluation of spectral recon-

struction results. For each reconstructed reflectance $\hat{\mathbf{r}}$, GFC is calculated as

$$\text{GFC} = \frac{\mathbf{r}^T\hat{\mathbf{r}}}{\sqrt{(\mathbf{r}^T\mathbf{r})(\hat{\mathbf{r}}^T\hat{\mathbf{r}})}}. \tag{4.17}$$

In case of the multispectral images from Foster dataset, RMSE is calculated for each pixel of the estimated reflectance, with the original reflectance in the hyperspectral image. We also compute the colorimetric error by a linear mapping of the reflectance data into its corresponding CIEXYZ. The CIEXYZ is converted into CIELab by taking white point of D65. Error between CIELab from the original reflectance and reconstructed reflectance is calculated in terms of CIEDE2000 [48]. For each evaluation metric, we include three methods. First method is by doing nothing to the input multispectral data and using Eq. 4.15 for spectral reconstruction. We call this method as *do nothing*. In second method, simple diagonal transform is applied to the input data before using it for spectral reconstruction, while in the third experiment, the input data is first transformed by using the proposed SAT. Results are discussed in Section 4.4.

4.4 Results

The results of spectral reconstruction of Munsell reflectance data, by using eight different illuminations and four different sensor configurations, are shown in form of graphs in Fig. 4.2, consisting of mean RMSE, mean GFC and mean CIEDE2000. For each test illuminant, we compare the spectral reconstruction results after applying diagonal transform to the input multispectral data, SAT and *do nothing*. The RMSE results show that for illuminant A, the diagonal transform reduces the error as comparing to *do nothing*, but for illuminants D50, D55 and D75, applying only a simple diagonal transform significantly increases the error in spectral reconstruction as compared to when the SAT is applied. There is slight increase in RMSE with the use of diagonal transform for illuminants F5, F7 and F12, while there is no change in error when LED is used as illuminant source. These results show that in terms of RMSE, applying diagonal transform to input multispectral data increases the error in comparison to *do nothing*. This trend is consistent among the 5, 8 and 12 channel cameras being used. Results obtained from simulated 6 channel SFA camera also show similar results. When our proposed SAT is used along with diagonal transform on the input multispectral data, then RMSE is reduced significantly. By increasing the number of channels, there is further reduction in RMSE for all illuminants except F12, where RMSE is minimum when 8 channels are used. Overall, the significant reduction in RMSE shows the efficiency of our proposed SAT for spectral reconstruction.

When the spectral reconstruction is evaluated in terms of GFC, diagonal transform and SAT perform closely. *Do nothing* performs lower for illuminants A, F5, F12 and LED while the performance difference is less significant for other test images. There is no change in performance by increasing the number of channels, except for illuminant F12, where the performance of *do nothing* goes down by increasing the number of channels. The increase in error for *do nothing* is because the bands become more sensitive to illumination changes when they get narrower.

Colorimetric evaluation of the spectral reconstruction reveals the interesting fact that there is no difference in *do nothing* and applying the diagonal transform. Error in terms of CIEDE2000 for both diagonal transform and *do nothing*, is increased slightly when the number of channels are increased from 5 to 12. However, applying SAT significantly reduces the colorimetric error, as can be seen in Fig. 4.2. The performance of SAT behaves opposite to other two methods and error is slightly reduced with increase in number of channels.

For testing the proposed idea of multispectral constancy on images containing real scenes, we use the Foster hyperspectral reflectance data [21]. We use four different sensor configurations as simulation of multispectral camera for acquiring the data. Sensor configuration for these cameras is shown in Fig. 4.1. Before acquiring the multispectral data, each hyperspectral reflectance image is converted into radiance image by applying an illuminant. We use illuminants A, D50, D55 and D75.

Illuminant estimation in each multispectral image is performed by using the Max-spectral algorithm [17]. Diagonal transform is applied to each multispectral image with the estimated illuminant. That estimated illuminant contain error which is evaluated in terms of angular error (ΔA). Fig. 4.3 show the performance of spectral reconstruction in terms of mean RMSE, mean GFC and mean CIEDE2000 of the eight images being tested. Detailed results along with ΔA are provided in the supplementary data file.

With the Foster dataset, the error in spectral reconstruction is larger than the results of Munsell dataset. The main reason is the error in illuminant estimation. As the elements of diagonal transform are an estimation of the illuminant, therefore the error in illuminant estimation is intensified in the spectral reconstruction as well. Fig. 4.3 show RMSE results for 5, 8, 12, and 6 (SFA) channels, respectively. *Do nothing* and simple diagonal transform produces almost similar results except for multispectral data acquired with 8 channels. In that case, simple diagonal transform performs better as compared to *do nothing*. RMSE is reduced significantly when SAT is used, which shows that our proposed method performs effectively in reducing RMSE.

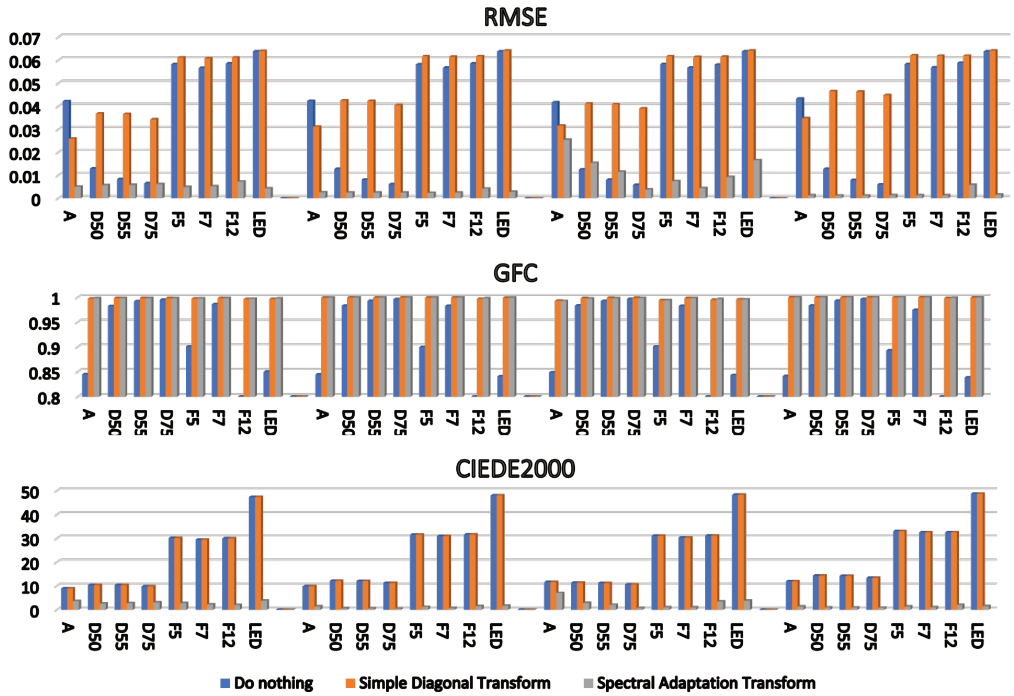


Figure 4.2: Evaluation of spectral reconstruction results of Munsell spectral reflectance. The column on left show results obtained from 5 channel camera, 2[nd] column consist of results obtained from 8 channel camera, column 3 show results obtained from 12 channel multispectral data while the last column show results obtained from 6 channel SFA camera. These results show that by using the proposed $\mathbf{A}_{SAT}$, error in spectral reconstruction is reduced in terms of RMSE, GFC and CIEDE2000

In terms of GFC, *do nothing* performs better than simple diagonal transform, except for illuminant A, where the *do nothing* performs significantly lower. Same trend can be seen across all four sensor configurations being tested. SAT is able to perform slightly better than diagonal transform and *do nothing*. Evaluation in terms of CIEDE2000 shows that *do nothing* and simple diagonal transform performs same except for 5 channels, where error is increased when diagonal transform is used. By using the proposed SAT, there is decrease in CIEDE2000 which shows that the proposed idea is valid with images containing real scenes. Although the colorimetric error is still large, but it should be kept in mind that inaccuracy in illuminant estimation also play its role in the overall error in spectral reconstruction. The average RMSE with 5 and 12 channels are almost equal and the same result in obtained with SFA, while RMSE is comparatively larger when 8 channels are used. In terms of colorimetry, the error is reduced gradually by increasing the number of channels. When simple diagonal transform is used then CIEDE2000 remains almost equal for 8 and 12 channels while the error is reduced for all sensor configurations when SAT is used. Therefore, by increasing the number of filters, there is slight improvement in the spectral reconstruction results. However, the performance may become more sensitive to the imaging noise. Those results are based on simulations and are still to be validated for the experimentally captured multispectral data. Detailed results from the experiments, including the angular error in illuminant estimation for each image, mean and maximum errors of RMSE, GFC and CIEDE2000, are provided in the supplementary data file.

In most of the illuminant estimation algorithms proposed in literature, the efficiency of algorithm is evaluated in the terms of angular error ΔA and a diagonal transform is applied to the input images. In most of the cases, the transformed images appear to be taken under a canonical (usually white) illuminant and it is assumed that the effect of scene illuminant is removed through the diagonal transform. Results from our experiments show that a simple diagonal transform is not sufficient. For illustration, we show an example of radiance image from Foster dataset and the corresponding transformations (Fig. 4.8. In this example, the color rendering of radiance spectral image with test illumination and reconstructed spectral images (after applying diagonal transform and SAT), are shown. D55 is used for creating the radiance scene and multispectral image is acquired with 12 equi-Gaussian channels Illuminant estimation gives ΔA of 0.1013. As can be seen in Fig. 4.8, the effect due to D55 is removed after the transformations and the output images appear to be almost the same visually but the RMSE, GFC and CIEDE2000 evaluation shows that there is difference among these 2 images in terms of spectra and colorimetry. This suggests that the effect of such transformations has to be verified in computer vision applications.

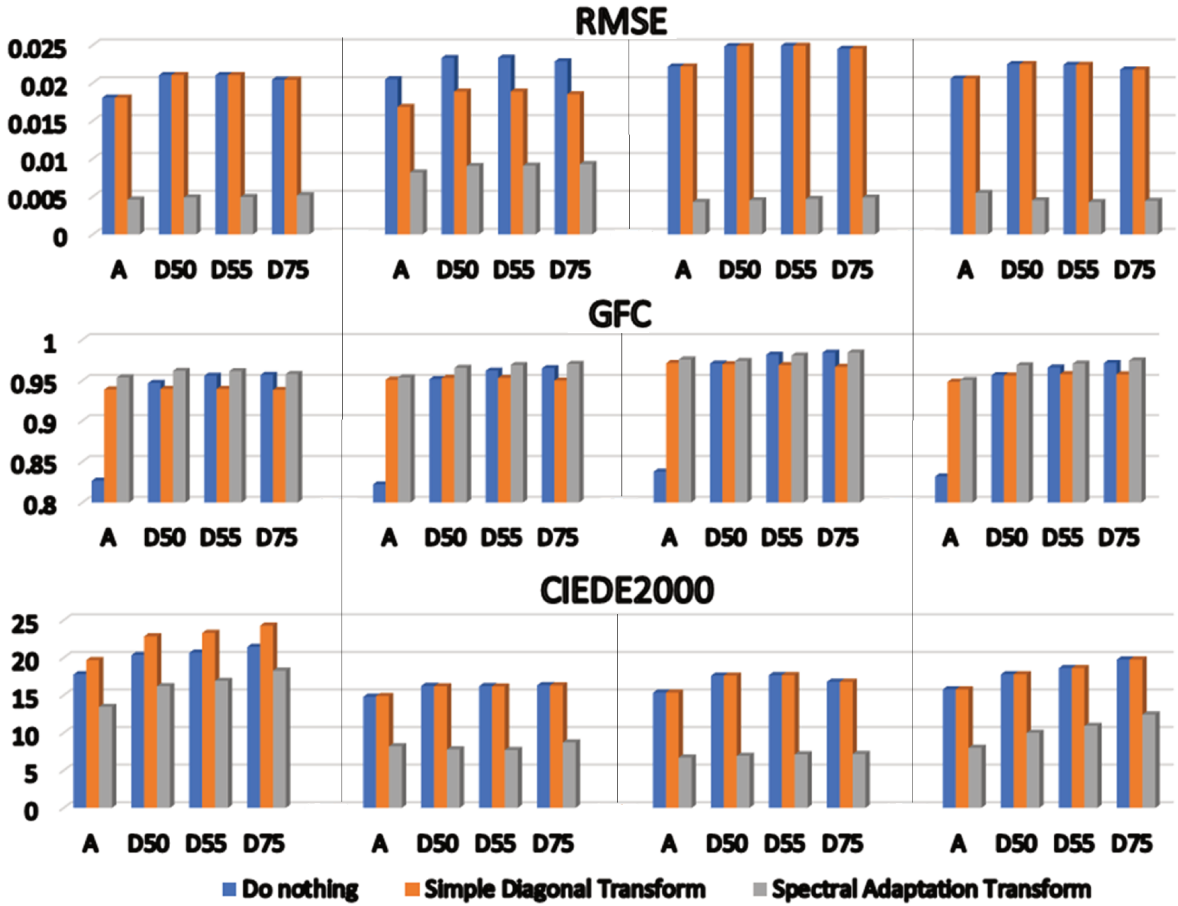


Figure 4.3: Evaluation results from the Foster dataset. Among the results, column 1 show results obtained with 5 channel multispectral camera, column 2 consist of results with 8 channel multispectral camera, column 3 show results obtained from 12 channel multispectral data while the last column show results obtained from 6 channel SFA camera.

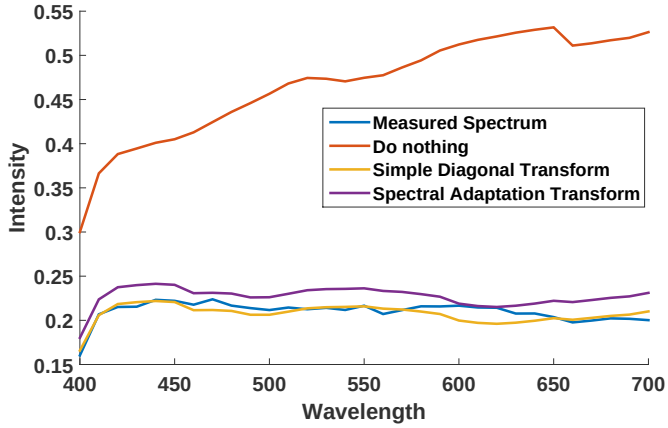
To visualize the reconstructed spectra, eight examples are provided in Fig. 4.4, 4.5, 4.6 & 4.7. In each Fig. two sample spectra for one image of Foster dataset are shown. These samples are selected on the basis of best and worst GFC values. Each Fig. provides the comparison between the measured spectral reflectance and the spectral reconstruction results with *do nothing*, simple diagonal transform and SAT. For each figure, the measured values of ΔA, GFC and RMSE are provided in the captions.

The reconstruction results using simple diagonal transform and SAT match closely, while the *do nothing* results are significantly lower, as they are not close to the measured spectra in terms of shape and scale. This observation shows that with illuminant estimation and then applying the transformation into canonical representation has the clear advantage. Upon a careful comparison, it can be seen that by applying SAT, the scale of reconstructed spectra is brought closer to the original spectra. This explains the reason that why our proposed SAT is able to reduce the error in spectral reconstruction. The visualization of spectra shows that with the use of SAT, there is no significant difference in the shape of reconstructed spectra. The main difference is in scale of intensity, which explains that why GFC remains same while RMSE and CIEDE2000 metrics provide different results. One of the drawbacks of using linear method for spectral reconstruction is its limitation in reconstruction of spectra containing spikes. Linear spectral reconstruction assumes smooth spectra and provides a smoothed approximation to the input spiky spectra. This effect can be seen in Fig. 4.4(b), 4.5(b), 4.6(b) & 4.7(b), where the linear reconstruction is unable to address the spikes in the measured spectra.

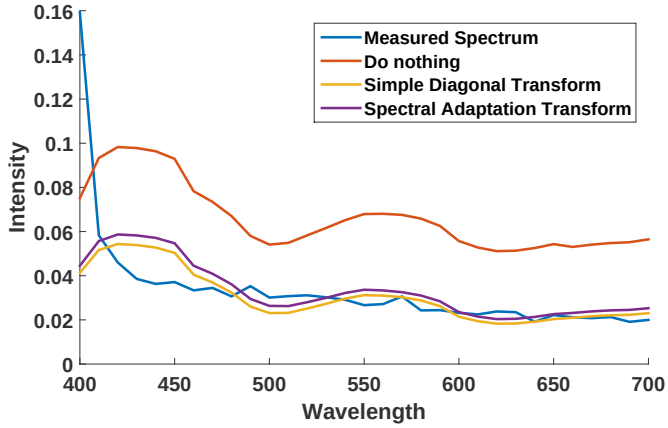
These observations lead to the question that what is the purpose of color and multispectral constancy. If it has to be applied for visual correction only, then a simple diagonal transform may be efficient, as seen in Fig. 4.8. If the purpose is to reconstruct the spectral reflectance, then a simple diagonal transform is not helpful while GFC metric indicates that there is slight improvement when simple diagonal transform is used. Evaluation in terms of CIEDE2000 indicates that the results with *do nothing* and diagonal transform are the same. However, all the evaluation metrics show that the proposed SAT is able to reduce the error significantly. Having said this, the next goal is to investigate that what is the role of improving the currently used evaluation metrics in computer vision applications, such as object detection and material identification.

4.5 Conclusion and perspectives

In this work, we propose the concept of multispectral constancy, which defines the illuminant invariant representation of multispectral images. Multispectral constancy is achieved through a spectral adaptation transform (SAT), which trans-

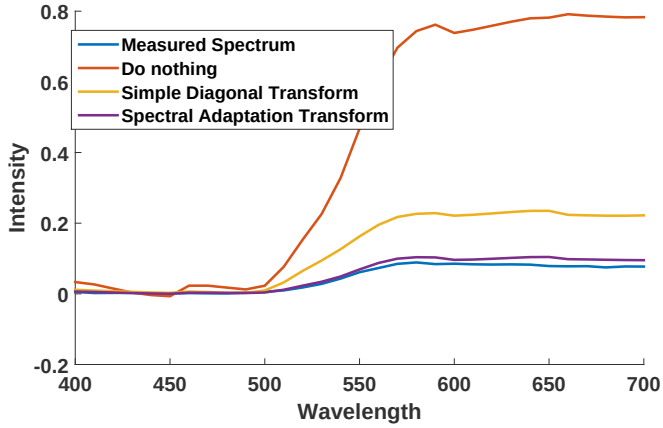


(a) GFC with simple diagonal transform: 0.9994, GFC with SAT: 0.9994.
RMSE with simple diagonal transform: 0.0489, RMSE with SAT: 0.0033

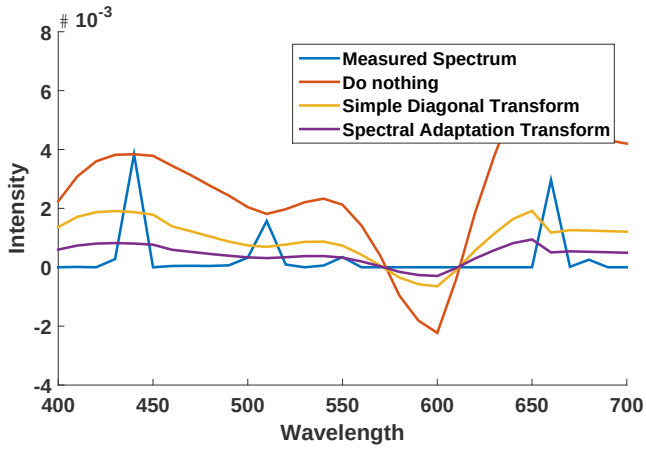


(b) GFC with simple diagonal transform: 0.8463. GFC with SAT: 0.8466
RMSE with simple diagonal transform: 0.0124 RMSE with SAT 0.0039

Figure 4.4: Spectral reconstruction of 2 sample spectra from Foster dataset (image 7). Illuminant D55 is used to create radiance data, simulated multispectral camera with 5 equi-Gaussian filters is used and illuminant estimation error ΔA is 0.0162

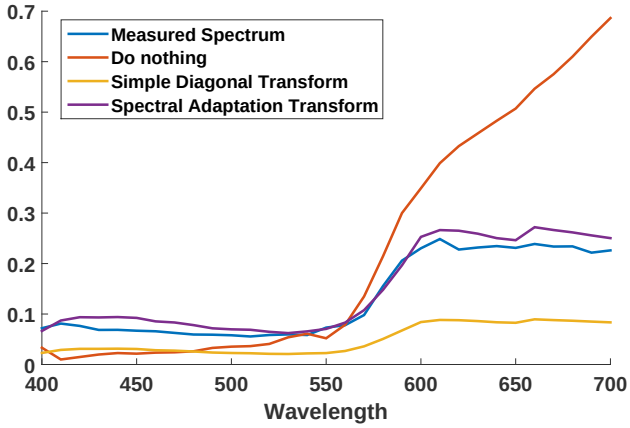


(a) GFC with simple diagonal transform: 0.9985, GFC with SAT: 0.9990.
RMSE with simple diagonal transform: 0.0562, RMSE with SAT: 0.0023

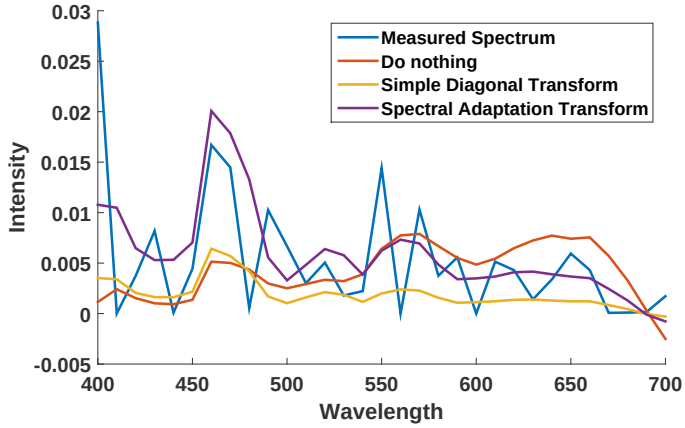


(b) GFC with simple diagonal transform: 0.3848, GFC with SAT: 0.3956.
RMSE with simple diagonal transform: 0.0013, RMSE with SAT: 0.0001

Figure 4.5: Spectral reconstruction of 2 sample spectra from Foster dataset (image 4). Illuminant D50 is used to create radiance data, simulated multispectral camera with 8 equi-Gaussian filters is used and illuminant estimation error ΔA is 0.0373

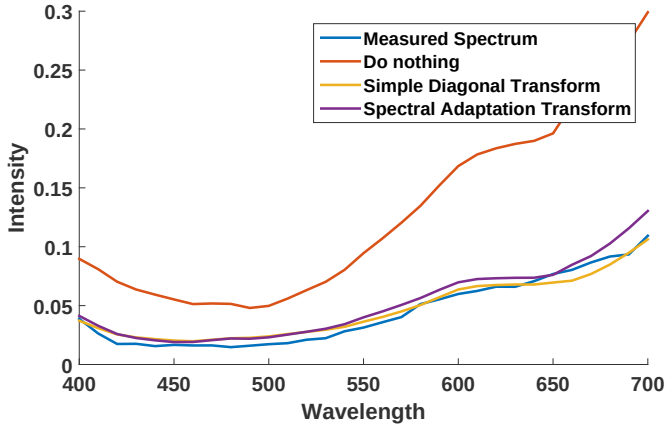


(a) GFC with simple diagonal transform: 0.9977, GFC with SAT: 0.9981.
RMSE with simple diagonal transform: 0.0202, RMSE with SAT: 0.0036

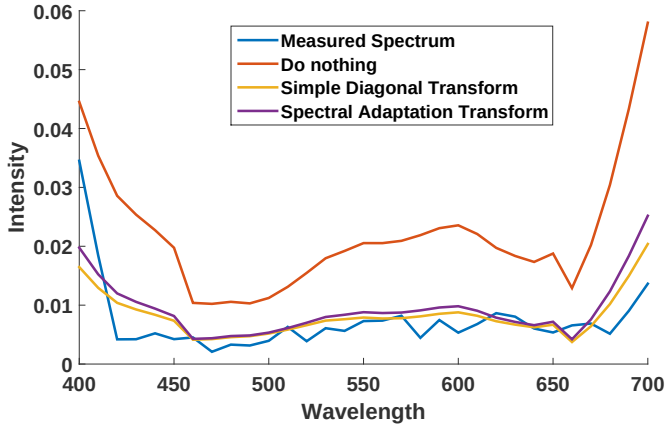


(b) GFC with simple diagonal transform: 0.7652, GFC with SAT: 0.7666.
RMSE with simple diagonal transform: 0.0036, RMSE with SAT: 0.0013

Figure 4.6: Spectral reconstruction of 2 sample spectra from Foster dataset (image 6). Illuminant A is used to create radiance data, simulated multispectral camera with 12 equi-Gaussian filters is used and illuminant estimation error ΔA is 0.0545



(a) GFC with simple diagonal transform: 0.9946, GFC with SAT: 0.9971.
RMSE with simple diagonal transform: 0.0183, RMSE with SAT: 0.0015



(b) GFC with simple diagonal transform: 0.8832, GFC with SAT: 0.8840.
RMSE with simple diagonal transform: 0.0036, RMSE with SAT: 0.0008

Figure 4.7: Spectral reconstruction of 2 sample spectra from Foster dataset (image 8). Illuminant D75 is used to create radiance data, simulated multispectral camera with 6 SFA channels is used and illuminant estimation error ΔA is 0.0884

Figure 4.8: Color rendering of multispectral image acquired with 12 channels. Images from left to right: Rendering under illuminant D55, Diagonal transform, SAT. ΔA is 0.1013

forms the data representation from an unknown illuminant, to a canonical one. This SAT is obtained by training the system. Advantage of the proposed technique is that the only information required is the sensor sensitivities of imaging system.

With the idea of multispectral constancy being developed, what needs further attention is the establishment of efficient illuminant estimation methods for multispectral imaging. The proposed framework for spectral reconstruction is sensitive to accuracy in illuminant estimation. This dependency is seen when spectral reconstruction is performed on the images from Foster dataset. The performance of SAT is affected with error in illuminant estimation and this dependency may be addressed through development of efficient illuminant estimation algorithms. Once such an efficient illuminant estimation algorithm in the sensor domain is available, our proposed SAT is able to show efficient results as demonstrated in the simulations.

The colorimetric difference also needs to be investigated so that an acceptable level of accuracy for color based systems can be achieved. Finally, it has yet to be investigated that what level of accuracy is required in the spectral reconstruction so that the estimation can be used for spectral information based applications. Such an application can be object detection and identification on the basis of its spectral properties. Hence. our proposed idea of multispectral constancy is opening research questions which needs to be investigated, in order to enable the generic use of multispectral imaging for real-world applications. Improvement in results can be done by carefully selecting the training data which can represent the objects in the scene. In this way, the SAT and spectral reconstruction system can be calibrated for a specific scenario. We did not work on improvement in selection of training data in this work as our main aim is to study the effectiveness of proposed idea of multispectral constancy and SAT under general conditions.

Simulation results show that for unimodal sensitivities and linear sensor, the proposed SAT is an adequate transform and permits to achieve efficient reflectance reconstruction. However, when the SAT is evaluated based on an estimate of illumination, error in illuminant estimates makes the performance to decrease. It is still to be investigated that what accuracy is required for illuminant estimation to make this concept beneficial. Further work shall investigate these directions and define the limits of using this approach.

Bibliography

[1] D. H. Foster, "Color constancy," *Vision Research*, vol. 51, no. 7, pp. 674–700, 2011. Vision Research 50th Anniversary Issue: Part 1.

[2] A. Gijsenij, T. Gevers, and J. Van De Weijer, "Computational color con-

stancy: Survey and experiments," *IEEE Transactions on Image Processing*, vol. 20, pp. 2475–2489, sep 2011.

[3] K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms - Part I: Methodology and experiments with synthesized data," *IEEE Transactions on Image Processing*, vol. 11, pp. 972–984, sep 2002.

[4] S. D. Hordley, "Scene illuminant estimation: Past, present, and future," *Color Research & Application*, vol. 31, no. 4, pp. 303–314, 2006.

[5] F. H. Imai and R. S. Berns, "Spectral estimation using trichromatic digital cameras," in *Proceedings of the International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives*, pp. 1–8, Chiba University Chiba, Japan, 1999.

[6] J. Y. Hardeberg, *Acquisition and Reproduction of Color Images: Colorimetric and Multispectral Approaches*. Universal Publishers, 2001.

[7] D. Connah, S. Westland, and M. G. A. Thomson, "Recovering spectral information using digital camera systems," *Coloration Technology*, vol. 117, pp. 309–312, nov 2001.

[8] E. M. Valero, J. L. Nieves, S. M. C. Nascimento, K. Amano, and D. H. Foster, "Recovering spectral data from natural scenes with an RGB digital camera and colored filters," *Color Research & Application*, vol. 32, no. 5, pp. 352–360, 2007.

[9] J. Y. Hardeberg and R. Shrestha, "Multispectral colour imaging: Time to move out of the lab?," in *Mid-term meeting of the International Colour Association (AIC)*, (Tokyo, Japan), pp. 28–32, May 2015.

[10] L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," *J. Opt. Soc. Am. A*, vol. 3, pp. 1673–1683, oct 1986.

[11] L. T. Maloney and B. A. Wandell, "Color constancy: a method for recovering surface spectral reflectance," *J. Opt. Soc. Am. A*, vol. 3, pp. 29–33, Jan 1986.

[12] R. Hall, *Illumination and color in computer generated imagery*. Monographs in Visual Communication, New York, NY: Springer New York, 1989.

[13] J. P. S. Parkkinen, J. Hallikainen, and T. Jaaskelainen, "Characteristic spectra of Munsell colors," *J. Opt. Soc. Am. A*, vol. 6, pp. 318–322, feb 1989.

[14] T. Jaaskelainen, J. Parkkinen, and S. Toyooka, "Vector-subspace model for color representation," *J. Opt. Soc. Am. A*, vol. 7, pp. 725–730, apr 1990.

[15] R. Shrestha and J. Y. Hardeberg, "Spectrogenic imaging: a novel approach to multispectral imaging in an uncontrolled environment.," *Optics Express*, vol. 22, pp. 9123–33, apr 2014.

[16] J.-B. Thomas, "Illuminant estimation from uncalibrated multispectral images," in *Colour and Visual Computing Symposium (CVCS)*, (Gjøvik, Norway), pp. 1–6, Aug 2015.

[17] H. A. Khan, J. B. Thomas, J. Y. Hardeberg, and O. Laligant, "Illuminant estimation in multispectral imaging," *J. Opt. Soc. Am. A*, vol. 34, pp. 1085–1098, Jul 2017.

[18] G. D. Finlayson, M. S. Drew, and B. V. Funt, "Color constancy: generalized diagonal transforms suffice," *J. Opt. Soc. Am. A*, vol. 11, pp. 3011–3019, Nov 1994.

[19] H. A. Khan, J. B. Thomas, and J. Y. Hardeberg, "Multispectral constancy based on spectral adaptation transform," in *20th Scandinavian Conference on Image Analysis, SCIA, Tromsø, Norway, June 12–14, 2017, Proceedings, Part II* (P. Sharma and F. M. Bianchi, eds.), pp. 459–470, Springer International Publishing, 2017.

[20] G. D. Finlayson, M. S. Drew, and B. V. Funt, "Spectral sharpening: sensor transformations for improved color constancy," *J. Opt. Soc. Am. A*, vol. 11, pp. 1553–1563, may 1994.

[21] D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster, "Frequency of metamerism in natural scenes," *J. Opt. Soc. Am. A*, vol. 23, p. 2359, oct 2006.

[22] Y. Zhao and R. S. Berns, "Image-based spectral reflectance reconstruction using the matrix R method," *Color Research & Application*, vol. 32, pp. 343–351, oct 2007.

[23] N. Eslahi, S. H. Amirshahi, and F. Agahian, "Recovery of spectral data using weighted canonical correlation regression," *Optical Review*, vol. 16, pp. 296–303, may 2009.

[24] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," *Proceedings of the Twelfth Color Imaging Conference*, pp. 37–41, 2004.

[25] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Transactions on Image Processing*, vol. 16, pp. 2207–2214, sep 2007.

[26] J. A. Worthey and M. H. Brill, "Heuristic analysis of Von Kries color constancy," *J. Opt. Soc. Am. A*, vol. 3, pp. 1708–1712, Oct 1986.

[27] D. A. Forsyth, "A novel algorithm for color constancy," *International Journal of Computer Vision*, vol. 5, pp. 5–35, aug 1990.

[28] R. W. G. Hunt, C. Li, and M. R. Luo, "Chromatic adaptation transforms," *Color Research & Application*, vol. 30, no. 1, pp. 69–71, 2005.

[29] J. Vazquez-Corral and M. Bertalmío, "Spectral sharpening of color sensors: Diagonal color constancy and beyond," *Sensors*, vol. 14, pp. 3965–3985, feb 2014.

[30] K. Barnard, F. Ciurea, and B. Funt, "Sensor sharpening for computational color constancy.," *J. Opt. Soc. Am. A*, vol. 18, pp. 2728–43, nov 2001.

[31] M. Drew and G. Finlayson, "Spectral sharpening with positivity," *J. Opt. Soc. Am. A*, vol. 17, no. 8, pp. 1361–70, 2000.

[32] H. Y. Chong, S. J. Gortler, and T. Zickler, "The von Kries hypothesis and a basis for color constancy," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1–8, IEEE, 2007.

[33] G. D. Finlayson, J. Vazquez-Corral, S. Süsstrunk, and M. Vanrell, "Spectral sharpening by spherical sampling," *J. Opt. Soc. Am. A*, vol. 29, no. 7, p. 1199, 2012.

[34] M. Abdellatif, "Physics-based spectral sharpening through filter-chart calibration," *Color Research & Application*, vol. 40, pp. 564–576, dec 2015.

[35] A. Abrardo, L. Alparone, I. Cappellini, and A. Prosperi, "Color constancy from multispectral images," *International Conference on Image Processing*, vol. 3, pp. 570–574, 1999.

[36] K. Barnard, L. Martin, B. Funt, and A. Coath, "A data set for color research," *Color Research & Application*, vol. 27, no. 3, pp. 147–151, 2002.

[37] M. J. Vrhel, R. Gershon, and L. S. Iwan, "Measurement and analysis of object reflectance spectra," *Color Research & Application*, vol. 19, no. 1, pp. 4–9, 1994.

[38] R. Shrestha and J. Y. Hardeberg, "Evaluation and comparison of multispectral imaging systems," in *Color and Imaging Conference*, pp. 107–112, 2014.

[39] P. J. Lapray, X. Wang, J. B. Thomas, and P. Gouton, "Multispectral filter arrays: Recent advances and practical implementation," *Sensors*, vol. 14, pp. 21626–21659, nov 2014.

[40] X. Wang, J.-B. Thomas, J. Y. Hardeberg, and P. Gouton, "Multispectral imaging: narrow or wide band filters?," *Journal of the International Colour Association*, vol. 12, pp. 44–51, 2014.

[41] P. J. Lapray, J. B. Thomas, P. Gouton, and Y. Ruichek, "Energy balance in spectral filter array camera design," *Journal of the European Optical Society*, vol. 13, p. 1, dec 2017.

[42] J. B. Thomas, P. J. Lapray, P. Gouton, and C. Clerc, "Spectral characterization of a prototype SFA camera for joint visible and NIR acquisition," *Sensors*, vol. 16, p. 993, jun 2016.

[43] E. H. Land and J. J. McCann, "Lightness and Retinex Theory," *Journal of the Optical Society of America*, vol. 61, pp. 1–11, jan 1971.

[44] J. Conde, H. Haneishi, M. Yamaguchi, N. Ohyama, and J. Baez, "Spectral reflectance estimation of ancient Mexican codices, multispectral images approach," *Revista Mexicana de Fisica*, vol. 50, no. 5, pp. 484–489, 2004.

[45] D. Connah, J. Y. Hardeberg, and S. Westland, "Comparison of linear spectral reconstruction methods for multispectral imaging," in *International Conference on Image Processing, ICIP*, vol. 3, pp. 1497–1500, 2004.

[46] H.-L. Shen, P.-Q. Cai, S.-J. Shao, and J. H. Xin, "Reflectance reconstruction for multispectral imaging by adaptive Wiener estimation," *Optics Express*, vol. 15, no. 23, pp. 15545–15554, 2007.

[47] J. Hernández-Andrés, J. Romero, and R. L. Lee, "Colorimetric and spectroradiometric characteristics of narrow-field-of-view clear skylight in Granada, Spain," *J. Opt. Soc. Am. A*, vol. 18, pp. 412–420, Feb 2001.

[48] CIE, *Improvement to industrial colour-difference evaluation.* CIE technical report 142-2001, Vienna: Central Bureau of the CIE, 2001.

Chapter 5

Illuminant estimation in multispectral imaging

Article B

This chapter is a reformatted reprint of the publication

H. A. Khan, J. B. Thomas, J. Y. Hardeberg, and O. Laligant, "Illuminant estimation in multispectral imaging," *Journal of the Optical Society of America A*, vol. 34, no. 6, pp 1085-1098, 2017.

With the advancement in sensor technology, use of multispectral imaging is gaining wide popularity for computer vision applications. Multispectral imaging is used to achieve better discrimination between the radiance spectra, as compared to the color images. However, it is still sensitive to illumination changes. This study evaluates the potential evolution of illuminant estimation models from color to multispectral imaging. We first present a state of the art on computational color constancy and then extend a set of algorithms to use them in multispectral imaging. We investigate the influence of camera spectral sensitivities and number of channels. Experiments are performed on simulations over hyperspectral data. The outcomes indicate that extension of computational color constancy algorithms from color to spectral, give promising results and may have the potential to lead towards efficient and stable representation across illuminants. However, it is highly dependent on spectral sensitivities and noise. We believe that the development of illuminant invariant multispectral imaging systems will be a key enabler for further use of this technology.

5.1 Introduction

Objects are perceived by their radiance in the visible region of the electromagnetic spectrum and for a given object, the radiance depends on its material properties, its shape and location in the scene. Intensity, position and spectral characteristics of the illuminant also play a major role in image generation. The spectral sensitivity of filters is another important parameter in image creation. In a simple imaging model with three channels, the image values $\mathbf{f} = (R, G, B)^T$ are dependent on the light source $e(\lambda)$, surface reflectance $r(\lambda)$ and camera sensitivity functions $\mathbf{c}(\lambda) = \{r(\lambda), g(\lambda), b(\lambda)\}$, as

$$\mathbf{f} = \int_{\omega} e(\lambda) r(\lambda) \mathbf{c}(\lambda) d\lambda. \tag{5.1}$$

In the human visual system, the three cone types are sensitive to certain wavelengths in photopic vision [1]. In the case of a camera with 3 channels, the color filters play a similar role. Multispectral imaging is being used to capture more spectral details in a scene as compared to conventional color images. Recently emerging technologies, such as the spectral filter arrays [2–4], enable a broader range of usage domains for multispectral imaging. The use of multispectral images in object recognition can perform better than the conventional RGB color images [5]. An example of multispectral imaging system to determine quality attributes and ripeness stage in strawberries was proposed by Liu *et al.* [6]. In that work, the imaging system is first radiometrically calibrated using both a diffuse white and dark target. Similarly, most existing multispectral imaging systems are specifically designed and needs to be re-calibrated when the imaging conditions are changed. Extending the use of multispectral imaging system from heavily constrained environments to real world applications is still an open challenge. One of the major obstacles is calibration of multispectral camera according to the scene illuminant [7–11]. In this work, we investigate the use of illuminant estimation algorithms for multispectral imaging systems.

We propose to extend the illuminant estimation algorithms from 3-channels to N-channels. Recently, Thomas [12] investigated the physical validity of these illuminant estimation algorithms by applying them on uncalibrated multispectral images (MSI) with 3, 5, 12 and 20 bands. That work showed that there is a huge variability due to scene contents, and suggests that the number and potential configuration of bands has an important influence on the results. In this work, we extend those preliminary results to a more general and exhaustive investigation through an experimental framework where we simulate a multispectral imaging system using different number of sensors and configurations. In [12], only equi-Gaussian filters are used in simulations and evaluation is provided in form of an-

gular error and goodness of fit coefficient. In this work, we use equi-Gaussian filters, Dirac delta form of filters and overlapping equi-energy filters for the evaluation of the effect of the filter configuration on illuminant estimation. We use the extension of specific illuminant estimation algorithms, which contain simple assumptions, provide efficient performance with natural scenes, and are robust to illumination changes since they do not require any training. We evaluate the results in form of angular error. We also map the illuminant in sensor domain into xy chromaticity space and then evaluate the xy chromaticity error. In this way, we are able to compare the performance of illuminant estimation algorithms and configurations between varying number of filters by reducing data into a common dimensionality. The experimental framework presented here can be extended for more sophisticated illuminant estimation algorithms as well, in order to develop optimal illuminant estimation system for multispectral imaging.

This paper is organized as follows. In Section 2, we briefly discuss computational color constancy and previous research on illuminant estimation in color images. In Section 3, we discuss previous work done on illuminant estimation in MSI and define the methodology for extension of illuminant estimation algorithms to higher dimensions. In Section 4, we present the experimental setup, the 5^{th} section consists of results and discussion, while Section 6 concludes the paper.

5.2 Computational Color Constancy Review

The captured color of objects generally changes when observed under different light sources, since the creation of an image is dependent not only on the spectral reflectance property of the object's surface and the camera's sensor sensitivity, but also on the incident illuminant on the object, as in Eq. 5.1. The human visual system has the natural ability to perceive constant color of surfaces despite the change in spectral composition of the illuminant [13] and this ability to discard illumination effects is called "Color Constancy" [14]. Color constancy is usually defined in the context of natural scenes along with flat matte and diffuse materials by a so-called "equivalent illumination model" [15, 16]. Creating such a model for color constancy in computer vision is called computational color constancy (CCC). Developing an illuminant invariant computer vision system is an open area of research and there are algorithms which are able to perform well for particular conditions and assumptions, but still a universally accepted CCC system does not exist.

CCC plays an important role in color-based computer vision applications including object recognition, tracking and image classification [17]. Object representation and recognition from the standpoint of computer vision is discussed in detail in [18]. For example, in case of object recognition, the color of the object can be used

as a feature, and it should appear constant across changes in illumination [19]. So the first step in achieving a constant representation of colors is to adjust the color changes due to the illuminant. CCC therefore deals with the representation of a scene with the effect of the illuminant being as small as possible. There are basically two approaches for this. One is to compute illuminant invariant features [20, 21] and a second to estimate the illuminant [22] and later apply a correction. Our work focuses on illuminant estimation in a scene.

The problem of developing an efficient and generic CCC algorithm obviously depends strongly on the illuminant estimation in a given scene, which indeed is not a straightforward task. The core challenge for CCC is that the data acquired is a combination of three unknown factors; surface reflectance properties, color of illuminant and sensor sensitivities. Maloney and Wandell [23] showed that color constancy is indeed impossible without applying restrictions on spectral reflectance and illuminations.

From the imaging model given in Eq. 5.1, the goal of a color constancy system is to estimate the illuminant $\hat{\mathbf{e}}$, and this estimation is performed in the camera domain:

$$\mathbf{e} = \begin{pmatrix} R_e \\ G_e \\ B_e \end{pmatrix} = \int_{\omega} e(\lambda)\mathbf{c}(\lambda)d\lambda. \tag{5.2}$$

In Eq. 5.2, $\mathbf{e}$ corresponds to the illuminant's projection over filters (IPF), which is a set of discrete values with the dimension equal to total number of filters (N). It should be noted that IPF is the response of each filter for the illumination (ground truth or estimated), and it is not equivalent to the spectral power distribution of the illumination itself.

Since the sensor's response $\mathbf{f}$ is a combination of three unknown factors, therefore the estimation of scene illuminant $\hat{\mathbf{e}}$ is an ill-posed problem [24] and certain assumptions have to be made in order to estimate the scene illuminant. Once the illuminant is estimated within the sensor domain, then correction is applied to the acquired image in order to represent it as it would have been taken under a known light source. This process is also expressed as "discounting the chromaticity of the illuminant" by D'Zmura and Lennie [25]. This transformation is performed as

$$\mathbf{F}^c = \mathbf{D}^{u,c}\mathbf{F}^u, \tag{5.3}$$

where $\mathbf{F}^u$ is the image taken in unknown light source and $\mathbf{F}^c$ is the transformed image as if taken under a canonical illuminant, while $\mathbf{D}^{u,c}$ is the spectral adaptation transform matrix, which maps colors from captured image to their corresponding colors under a known illumination. The independence of color channels from each

other is defined in the Retinex Model [26–28]. This assumption is closely related to the Von Kries coefficient rule [29, 30]. Land's White-Patch Algorithm [28], proposes that there is at least one pixel in each color channel which cause maximum reflection of the illuminant and when such maximum responses are combined, they form the color of illuminant. This assumption is alleviated by considering the color channels separately, resulting in the max-RGB algorithm [27].

The Grey-World Algorithm was proposed by Buchsbaum [31] and is based on the assumption that the average color of a scene is achromatic. The result of Grey-World algorithm was improved by Gershon *et al.* [32] by taking average reflectance of a database and assuming the average of the scene to be equal to that average reflectance.

The Shades of Gray Algorithm was introduced by Finlayson and Trezzi [33]. This is a general form of max-RGB and Grey-World Algorithms where it is shown that Grey-World algorithm is the same as using the L^1 Minkowski norm while max-RGB is equivalent to using L^∞ norm. In their case, the general equation for estimation of light source becomes

$$\left(\frac{\int F^p dx}{\int dx}\right)^{1/p} = k\mathbf{e}, \tag{5.4}$$

where k is a constant is p is the order of the Minkowski norm.

The Gray-edge Algorithm proposed by Weijer *et al.* [34], assumes that average of reflectance *derivative* in a scene is achromatic. This algorithm is expressed as:

$$\left(\frac{\int [F_\sigma]^p dx}{\int dx}\right)^{1/p} = k\mathbf{e}, \tag{5.5}$$

where F_σ is the smoothed image, after applying a Gaussian filter.

Edge-based CCC is explored further for higher order derivatives in [35]. Celik and Tjahjadi [36] used wavelet transform to down-sample the image before applying Grey-Edge algorithm for estimation of illuminant color, and for each down-sampled image, separate estimation is performed on the high pass filter's result. Decision for illuminant color is based on minimum error between the estimation in consecutive scales. CCC based on spatio-temporal statistics in a scene was proposed by Chakrabarti *et al.* [37] where the spatial features of object surfaces are also accounted for in the determination of the illuminant. That work is improved in [38] by using only the edge information for achieving computational efficiency. There are some approaches which try to select the most appropriate estimation using intrinsic properties from other color constancy algorithms [39].

Gamut mapping is also used in CCC. It was introduced by Forsyth [40]. He proposed that the color of an object is its representation under a fixed canonical light, rather than as a surface reflectance function. It is based on the assumption that *for a given illuminant, one observes only a limited number of colors*. Based on this assumption, any change in colors of the image is caused by the variation in color of the light source. The limited set of colors which can occur under a given illuminant is called the *canonical gamut* and is determined through observations of many surfaces under the known light source. Gijsenij *et al.* [41] proposed gamut estimation for illuminant by using higher order statistics. Their results show that for lower number of surfaces, pixel-based gamut mapping performs well but with large number of surfaces, the efficiency of edge-based gamut mapping increases. Color-by-correlation [42] is a discrete version of Gamut mapping where the correlation matrix is used instead of canonical gamut for the considered illuminants, and is used with the image data to calculate the probability that the illumination in the test image is caused by which of the known illuminants.

Hue *et al.* [43], proposed an automatic white balancing algorithm by using gray points in an image for estimation of illuminant temperature. In their method, RGB image is converted into YUV color space and then those pixels where $U = V = 0$ or $R = G = B = Y$ are pointed out as Gray Points. A feedback system is used to estimate those points and then remaining pixels are corrected by adjusting the gain of R or B channel according to the illuminant color being detected. Yoon *et al.* [44] proposed dichromatic line space where a dichromatic slope is formed within dichromatic line space. Illuminant chromaticity is estimated through intersection of those lines. Ratnasingam and Collins [45] proposed two features that are described to represent chromaticity and are independent of intensity and correlated color temperature of illuminant in a scene. Sapiro [46] presented probabilistic Hough transform approach where a surface is selected according to the defined distribution and is used to recover the illuminant while using it along with sensor response. Bayesian formulation for solving CCC is used by Brainard and Freeman [47] where each surface and light is represented by basis functions for which probability distribution is defined. Xiong and Funt [48] used stereo images for extraction of 3-D information as an additional source for illuminant estimation. Use of 6 channels is proposed by Finlayson *et al.* [49] in the chromagenic algorithm. The additional three channels are acquired by using chromagenic filter being placed in front of the sensor. The information from these channels is used to estimate the scene's illuminant from a set of known illuminants. Modification in chromagenic algorithm is proposed by Fredembach and Finlayson in the bright-chromagenic algorithm [50], by using only the brightest pixels in the two images.

Assuming that the subspace of reflectances of all surfaces is linear and in a small dimension then the number of sensors, Maloney-Wandell algorithm [51] propose that the sensor responses for the surfaces under one illuminant fall within a linear subspace of the same dimensionality. Estimation of surface colors under two illuminants using Retinex Theory is proposed by Barnard *et al.* [52] and Finlayson *et al.* [53]. Nieves *et al.* [54] proposed a linear pseudo-inverse method for recovery of spectral power distribution of the illuminants using a learning-based procedure. Their algorithm is based on the detection of naturally occurring bright areas in natural images, acquired through the color camera.

Machine Learning is also applied for illuminant estimation. In [55], multilayer neural network is trained using histograms of chromaticity of input images along with corresponding chromaticity of illuminant. A number of similar approaches can be found in [56, 57] [58]. Support Vector Machine is used in [59], which is based on the higher-order structure of images. Recently, Deep Learning is also utilized in color constancy as in [60][61]. Bianco *et al.* [62] used convolutional neural network for illuminant estimation in raw images. For generation of ground truth illumination, shades of gray, gray edge and gamut mapping is applied on the training data in their proposed method. Oh and Kim [63] treat this as illuminant classification problem by using deep learning.

We consider multispectral images taken in outdoor environment that can be generated by any mixture of illuminants. We are also interested in investigating the effect of number of filters and their configurations for illuminant estimation. We propose to select a set of illuminant estimation algorithms which can handle any type of illuminant without requiring prior training and provide straightforward extension to N dimensions. We also require the estimated illuminant to be in sensor dimension and not in the xy chromaticity space so that it can be used for spectral adaptation transform ($\mathbf{D}^{u,c}$ in Eq. 5.3). Following our review, we chose to investigate on the extension of gray-world, max-RGB, shades of gray and gray-Edge algorithms. Another reason for selection of these algorithms is the diversity of spectral imaging systems in term of spectral sensitivities and number of channels in our experiments. Initially we do not select the learning based algorithms as we are interested in generic illuminant estimation framework without the need of prior training. Although use of classification methods shows improvement in performance of illuminant estimation, the major problem with such techniques is availability of training data and the limited set of illuminations being considered. This is not a major problem in case of color images but may be troublesome in spectral images. Another constraint is the diversity of spectral imaging systems in term of spectral sensitivities and the number of channels. Therefore, we limit our investigations to "equivalent illumination models".

5.3 Illuminant estimation from multispectral images

In this section, we will first discuss the previous work done for illuminant estimation in multispectral images and then define our proposed idea for extension of existing illuminant estimation algorithms from color to multispectral images.

5.3.1 Related work

In this section, we define the formation of a multispectral image and then review the literature on illuminant estimation on these images. Spectral imaging can be defined as an array of N channels representing several spectral components at each spatial location. Use of spectral imaging gained worldwide attention after the launch of Landsat in 1970 and since then it has been widely used in remote sensing applications. With the development in sensor technology, use of spectral imaging in short-range imaging is also expanded. A survey on hyperspectral and multispectral imaging technologies is provided by Vagni [64]. In this work, we are considering only *multispectral* images acquired through short-range imaging techniques, where N, the number of spectral filters, is typically in the range of 5 to 20 [65].

According to the sensitivity of a typical silicon sensor behind an optical system, having sensitivity range from 400 nm to 1100 nm, a multispectral system usually provides a combination of visible and/or near infrared bands, where the imaging model defined in Eq. 5.1 still holds.

$$\mathbf{f} = \int_{\omega} e(\lambda) r(\lambda) \mathbf{m}(\lambda) d\lambda, \tag{5.6}$$

where we now represent the camera sensitivities as $\mathbf{m}(\lambda) = \{m_1(\lambda), m_2(\lambda), ..., m_N(\lambda)\}$.

Mosny and Funt [66] investigated the role of additional information acquired through multispectral imaging in order to improve the performance of already existing color constancy algorithms for illuminant chromaticity estimation. They used chromagenic algorithm [49], Maloney-Wandell algorithm [51], Grey-World algorithm [31] and Max-RGB [27]. Multispectral images were synthesized for their experiments by using the spectral sensitivity of Sony DXC-930 camera. For additional bands acquisition simulation, the sensitivity curves were shifted by $\pm$ 16*nm*. They used 3, 6 and 9 bands for image acquisition along with 1995 surfaces and 287 illuminants. For representation of results, median angular error in the sensor domain and median angular error for illuminants estimates converted to RGB space, were used. According to their evaluation, there is slight improvement with 6 channels but overall there is no significant improvement in illuminant chromaticity estimation by increasing the number of bands. Such experiments are performed on real world data in [67] where authors have used 28 scenes being photographed

with 10 different illuminations. For image acquisition, cool and warm filters were used with camera. Their evaluation methods show the same results that additional spectral bands does not contribute significantly towards illuminant chromaticity estimation.

Shrestha and Hardeberg [10] proposed a spectrogenic imaging system where two images are acquired from a scene; One normal RGB image and one filtered-RGB image. Illuminant estimation of the scene using these two images is performed using the Chromagenic Algorithm [49] and its modification proposed by Fredembach and Finlayson [50]. 87 illuminants were used for training the system and an illuminant with minimum fitting error was selected as the potential illuminant for the scene.

It is worth noting that the purpose of Mosny and Funt [66, 67] was to investigate if there is any improvement in illuminant estimation achieved by increasing the number of filters, while in our work we want to investigate the extension of illuminant estimation into multispectral domain. The system proposed by Shrestha and Hardeberg [10] is limited in term of bands and illuminants. We are interested in development of an illuminant estimation framework for multispectral imaging with any number of bands and with any mixture of illuminants so that it can be used for outdoor image acquisition without requiring calibration.

5.3.2 Proposed multispectral illuminant estimation algorithms

In this work, we propose four algorithms for investigation, which are instantiations of a class of models referred to as "equivalent illumination models" and they assume a "flat-matte-diffuse" condition. These algorithms are computational attempts to implement the model of human visual system for color constancy using natural image statistics. We evaluate the performance of those algorithms with multispectral data by extending those techniques to N-dimensions and get the estimate of illuminant in sensor domain. We rename those algorithms so that the confusion between color information and spectral information is eliminated.

- Gray-World Algorithm [31] $\rightarrow$ Spectral Gray-World Algorithm
- Max-RGB Algorithm [27] $\rightarrow$ Max-Spectral Algorithm
- Shades of Grey Algorithm [33] $\rightarrow$ Spectral Shades of Gray Algorithm
- Gray-Edge Algorithm [34][35] $\rightarrow$ Spectral Gray-Edge Algorithm

In the Gray-World algorithm, it is assumed that the average reflectance of a scene is gray or achromatic. We extend this definition for the case of multispectral images

by assuming that the average reflectance in an N-dimensional image is constant;

$$\left(\frac{\int r(\lambda)dx}{\int dx}\right) = k. \tag{5.7}$$

Using Eq. 5.4 with $p = 1$, the illuminant can be estimated by computing the average pixel values for each channel.

$$\frac{\int \mathbf{f}(\lambda)dx}{\int dx} = \frac{1}{\int dx}\int\int_{\omega} e(\lambda)r(\lambda)\mathbf{m}(\lambda)d\lambda dx$$
$$= k\int_{\omega} e(\lambda)\mathbf{m}(\lambda)d\lambda = k\hat{\mathbf{e}} \tag{5.8}$$

The term $\hat{\mathbf{e}}$ is the estimate of illuminant in sensor domain. The same technique is used for spectral gray-edge algorithm where each channel is treated according to Eq. 5.5 after smoothing through Gaussian filter with $\sigma = 2$ and extraction of edges through derivative in both spatial axis. In case of spectral shades of gray algorithm, Eq. 5.4 is used with value of p higher than 1, while for Max-Spectral algorithm, we treat each spectral band separately to get the pixels with maximum response and use them for estimating the illuminant according to the originally presented hypothesis where authors used color images.

Our implementation strategy for extension of these algorithms is slightly different than [12] as we consider each channel of a multispectral image separately. It is worth mentioning that both shades of gray and gray-edge algorithm use Minkowski norm p and in [33], authors declare that with $p = 6$, best results are obtained. In our experiments, we keep the same value of p as proposed by authors, however we perform experiment to obtain optimized value for this parameter and discuss it in the results section.

5.4 Experimental Setup

5.4.1 Data Preparation

We use hyperspectral images from the Foster Dataset 2004 [68], which are acquired in the wavelength range of 400-720 nm. This dataset contains reflectance data from natural scenes and is adequate for our purposes because of its natural image statistics, which are fundamental to the proposed methods (Fig. 5.1). In order to prepare radiance data, we use D65 and F11 illuminants. We also test the framework using a combination of D65 and F11 illuminations to simulate a scene having mix D65-F11 illuminants (Fig. 5.2). D65 is used as standard daylight illuminant while F11 resembles the spectral response of Sodium-vapor lamp [69], which would typically represent an example of outdoor lighting, e.g. road or ski

tracks. Illuminant F5 is also used in the experiments and we found similar results as being obtained with F11 illuminant. In this paper, we present the results obtained from the multispectral data generated through F11 illuminant.

We also consider noise in the multispectral imaging system. Typically, the main sources of noise are photon shot noise, dark current noise, read noise and quantization noise [70]. We do not consider photon shot noise and dark current noise since the Foster Dataset 2004 is already corrected for these type of noise. We do not consider quantization noise either since the data is already quantized at 12-bit. We simulate the additive read noise in our experiments as normally distributed Gaussian noise with zero mean and 2% variance [71].

5.4.2 Sensor Configuration

The performance of the proposed algorithms would be affected by the spectral sensitivities of the sensors that capture the radiance [72]. In our experiments, we use a Gaussian model of sensor sensitivities. Such model has been extensively used in the literature to simulate sensors or to approximate Fabry-Pérot filter transmittance [73]. For our experiments, three sensor configurations S^g, S^d and S^{50} are investigated. Within the visible range, we define S^g as equi-Gaussian [12]. Full Width at Half Maximum (FWHM) of the sensor sensitivities decrease with increase in number of bands and the overlap between adjacent bands remain approximately same. By increasing the number of bands in this configuration, we are gradually shifting from typical multispectral sensors towards hyperspectral sensors. S^d configuration is a simulation of Dirac delta function where only the band corresponding to mean of the Gaussian filter is selected while rest of bands are discarded. It is of interest to test if such a configuration will provide any help in estimating the illuminants with spiky behaviour (e.g. F illuminants). Configuration S^{50} consists of equi-energy filters, having a fixed FWHM and $\sigma = 50nm$, which is different from S^g where the FWHM of filters get changed with change in number of bands. Using this configuration, we evaluate the effect of overlapping of filters for illuminant estimation.

In addition to the above explained filter configurations, we also consider different number of bands. 3 bands for simulating an instantiation of RGB cameras. 5 and 8 bands are used for simulating a typical multispectral camera [7]. 12 bands are used to get the best spectral reconstruction [74] while 20 bands are deployed to approach the properties of a hyperspectral sensor. Figure 5.3 show the three different configurations with 8 spectral filters.

5.4.3 Evaluation

We consider images with different number of bands, therefore the quantitative evaluation is not straightforward, especially when comparing results obtained with different number of bands. We consider different quality evaluation metrics, which include evaluation on the basis of angular error, goodness-of-fit coefficient (GFC) [75] and Normalized Mean Square Error (NRMSE). These 3 evaluation metrics are used only when the dimension of filters is the same and therefore results obtained from different number of filters cannot be compared. The estimated illuminants and ground truth illuminant are normalized by dividing each value from the maximum so that the range is within [0-1] and relative errors are evaluated. The three indicators are very similar in the way they evaluate the similarity between data. We determined the correlation among the computed metrics and found that the correlation between angular error and GFC is -0.987, while the correlation between angular error and NRMSE is 0.975 on our data. Therefore, we decide to discuss and analyse the results in terms of angular error in this paper.

Calculation of angular error (ΔA) between the original illuminant $\mathbf{e}$ and the estimated illuminant $\hat{\mathbf{e}}$ is computed in radians as in Eq. 5.9. This is commonly used in CCC literature.

$$\Delta A = \arccos \frac{\mathbf{e}^T \hat{\mathbf{e}}}{\sqrt{(\mathbf{e}^T \mathbf{e})(\hat{\mathbf{e}}^T \hat{\mathbf{e}})}} \tag{5.9}$$

where $\mathbf{e}$ is the ground truth illuminant directly acquired in sensor domain while $\hat{\mathbf{e}}$ is the estimated illuminant in sensor domain. When the estimated and measured illuminant are the same then it returns zero error.

The comparison of performance is done among five different number of spectral filters (3, 5, 8, 12 & 20), three different filter configurations (equi-Gaussian, Dirac delta & equi-energy filters), and 4 different algorithms (spectral gray-world, max-spectral, spectral shades of gray & spectral gray-edge). The estimated illuminant for all these configurations is compared with the ground-truth illuminant in the sensor domain.

To be able to compare results obtained from different number of filters, we project the data into the chromaticity space, where they could be compared at the expense of an error on the projection definition. We call this evaluation metric "xy chromaticity difference" where we perform a camera linear colorimetric calibration based on mean square error fitting on the reflectance of X-Rite ColorChecker, similar to the work of [76] where authors used that technique for color reproduction of MSI. We get the CIEXYZ of both estimated and the ground truth illuminants using this method. xy values are computed from these values and the chromatic

distance between them is observed in terms of Euclidean distance. This method enables us to compare the results obtained from different number of filters with each other. To verify the validity of this technique, we compared the ground truth illuminants in sensor domain with the chromaticity value of D65 and found that the Euclidean distance between them varies between 0.000934 and 0.00523 in the xy chromaticity space, which is very small and therefore, we can neglect the chromaticity error introduced during mapping of illuminant from sensor domain to the xy chromaticity space.

We present the results in form of mean angular error, and in order to compare the statistical significance of results, Wilcoxon signed rank test (WST) is applied. The use of WST is recommended by Hordley and Finlayson [77] and is used for evaluation of illuminant estimation performance [50, 78, 79]. We investigate the statistical significance among results at 95% confidence level and provide the WST scores in term of the sum of positive scores in the same way as being provided by Bianco *et al.* [78]. Higher score means that one particular algorithm along with sensors configuration is able to perform well as compared to others. A lower WST score means that the performance is significantly low in comparison with the rest. To illustrate the visual difference among the ground truth illuminant and the estimated illuminant, we have included examples in form of plots. In each figure, IPF for $\mathbf{e}$ and $\hat{\mathbf{e}}$ can be compared when the number of filters are same.

5.5 Results and Discussion

We have provided the results in Tables 5.1-5.6. Table 5.1 shows that in the noiseless case with 3 filters, Spectral Gray Edge S_3^g performs the best, followed by S_3^{50} and then Max-Spectral S_3^g. S_3^d configuration performs the worst for all 4 algorithms. Illuminant estimation from noisy data shows also show same results. There is a slight improvement in mean error in some cases when noise data is used but this slight change is not statistically significant and the overall results are robust with noise. With 5 bands (Table 5.2), Spectral Gray Edge S_5^{50} is the best followed by Max-Spectral S_5^{50} for $D65$ and mix D65-F11 illuminants. $F11$ shows different behaviour as Max-Spectral S_5^g performs best and Spectral Edge S_5^{50} follows. With noisy data, Spectral Gray Edge S_5^{50} gives consistent performance in terms of WST ranking while performance of Max-Spectral S_5^{50} is significantly reduced in case of F11 illuminant. Table 5.3 shows that with 8 filters the trend for best performance shifts from Spectral Edge to Max-Spectral as S_8^{50} performs best for both illuminants. However, in case of $F11$, it is interesting to note that Spectral Shades of Gray S_8^d performs the second best. This behavior is explained by the spikes in $F11$ illuminant and S^d configuration is able to detect those spikes more efficiently. However, with noisy data, Shades of Gray S_8^d is unable to perform

anymore and Spectral Gray Edge S_8^d gets the second best ranking while the rest of trend remains almost the same. For 12 bands, Max-Spectral S_{12}^{50} achieves the best estimate, followed by Spectral Gray-Edge S_{12}^{50} as seen in Table 5.4. Performance of those algorithms remains similar in presence of noise. In Table 5.5, results from using 20 filters show that Max-Spectral S_{20}^{50} and Spectral Gray Edge S_{20}^{50} performs almost the same in both conditions.

(a) I1

(b) I2

(c) I3

(d) I4

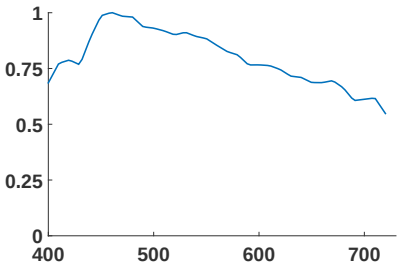
(e) I5

(f) I6

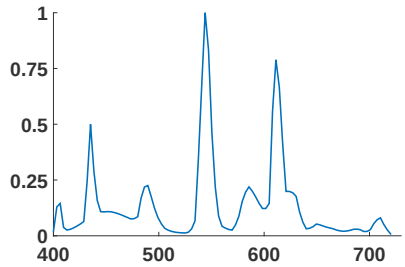
(g) I7

(h) I8

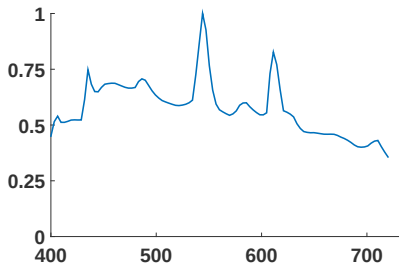
Figure 5.1: Rendering of hyperspectral images from Foster Dataset 2004 into RGB with D65 illuminant. The hyperspectral images are acquired within the wavelength range of 400nm to 720nm with 10nm sampling. Each hyperspectral image consists of 33 spectral bands.



(a) D65 Illuminant

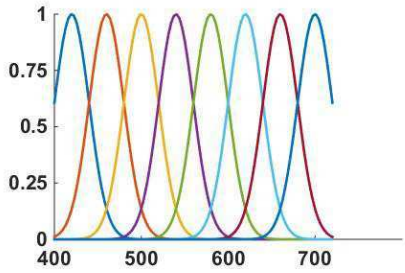


(b) F11 Illuminant

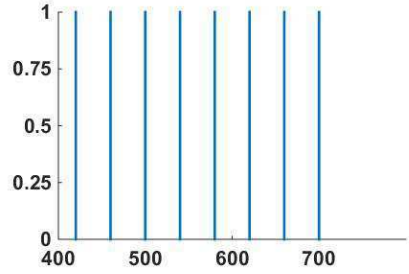


(c) Mix D65-F11 Illuminant

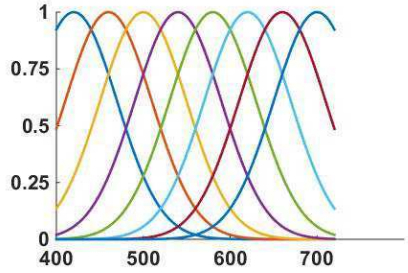
Figure 5.2: Illuminants used for creating radiance data from hyperspectral reflectance data are shown. Normalization is performed by diving each value by the maximum of that illuminant so that all values are within range of [0-1]. In (c), the mix D65-F11 illuminant consists of 50% D65 and 50% F11.



(a) S_8^g. Equi-Gaussian filters



(b) S_8^d. Dirac delta filters



(c) S_8^{50}. Equi-energy filters

Figure 5.3: Filter Configurations. For a configuration denoted S_x^y, x is the number of filters and y represent the configuration where g=equi-Gaussian, d=Dirac delta, f=filter with constant FWHM. Here we show examples with 8 filters

Table 5.1: Performance of illuminant estimation algorithms and filter configurations for 3 bands. Green cells show the best performance and yellow cells show the 2[nd] best performance, for a particular algorithm and filters configuration. Spectral Gray-Edge with S_3^g configuration shows the best result.

Algorithm	Filter. Confg.	D65				F11				Mixed illuminant (D65 and F11)			
		Without Noise		With Noise		Without Noise		With Noise		Without Noise		With Noise	
		Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank
Spectral Gray World	S_3^g	0.1881	2	0.1761	2	0.1318	2	0.1125	2	0.1812	2	0.1691	2
	S_3^d	0.2923	0	0.2686	0	0.3184	0	0.2926	0	0.2954	0	0.2723	0
	S_3^{50}	0.1948	1	0.1822	1	0.1388	1	0.1182	1	0.1881	1	0.1753	1
Max Spectral	S_3^g	0.1148	5	0.1164	3	0.0729	4	0.0735	3	0.1103	5	0.1083	4
	S_3^d	0.1583	0	0.1516	0	0.1485	0	0.1612	0	0.1616	0	0.1544	0
	S_3^{50}	0.1196	4	0.1108	4	0.0749	4	0.0743	3	0.1151	4	0.1117	3
Spectral Shades of Gray	S_3^g	0.1229	3	0.1256	3	0.0846	3	0.0861	3	0.1186	3	0.1202	3
	S_3^d	0.1568	1	0.1574	1	0.1545	1	0.1718	1	0.1573	1	0.158	1
	S_3^{50}	0.1264	2	0.1283	2	0.0869	2	0.0887	2	0.122	2	0.1242	2
Spectral Gray Edge	S_3^g	0.1003	7	0.1011	6	0.0723	7	0.072	7	0.0968	7	0.0971	7
	S_3^d	0.1364	1	0.1364	1	0.1096	2	0.1093	3	0.1366	1	0.1363	1
	S_3^{50}	0.1037	6	0.1035	5	0.0742	5	0.0741	4	0.1002	6	0.1	6

Table 5.2: Performance of illuminant estimation algorithms and filter configurations for 5 bands. Spectral Gray-Edge with S_5^{50} configuration shows best result for $D65$ and mix D65-F11 illuminants but is placed second in case of $F11$ illuminant where Max-Spectral S_5^{50} performs best with noise-free data.

Algorithm	Filter. Confg.	D65				F11				Mixed illuminant (D65 and F11)			
		Without Noise		With Noise		Without Noise		With Noise		Without Noise		With Noise	
		Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank
Spectral Gray World	S_5^g	0.2323	1	0.217	1	0.1675	1	0.1413	1	0.2244	1	0.2092	1
	S_5^d	0.2843	0	0.2602	0	0.2803	0	0.2638	0	0.2882	0	0.2657	0
	S_5^{50}	0.1893	2	0.1772	2	0.1264	2	0.1095	2	0.1816	2	0.1697	2
Max Spectral	S_5^g	0.1192	3	0.1288	2	0.0876	4	0.0878	4	0.1125	4	0.1159	4
	S_5^d	0.1482	1	0.1465	1	0.1224	0	0.1305	0	0.1546	1	0.1537	1
	S_5^{50}	0.1018	5	0.1017	7	0.0655	9	0.0776	6	0.0981	7	0.0978	8
Spectral Shades of Gray	S_5^g	0.1454	2	0.146	2	0.1036	1	0.1077	2	0.1408	2	0.1416	2
	S_5^d	0.165	1	0.1643	1	0.136	0	0.1533	0	0.1662	1	0.1663	1
	S_5^{50}	0.1224	3	0.1239	3	0.083	6	0.0848	6	0.1178	3	0.1194	3
Spectral Gray Edge	S_5^g	0.1224	3	0.1218	3	0.0949	3	0.0952	3	0.1193	4	0.1191	3
	S_5^d	0.141	1	0.1411	1	0.119	1	0.1199	1	0.1413	2	0.1408	2
	S_5^{50}	0.1016	6	0.1012	7	0.0737	7	0.0743	7	0.0984	8	0.0983	8

Table 5.3: Performance of illuminant estimation algorithms and filter configurations for 8 bands. Max-Spectral with S_8^{50} performs best.

Algorithm	Filter. Confg.	D65				F11				Mixed illuminant (D65 and F11)			
		Without Noise		With Noise		Without Noise		With Noise		Without Noise		With Noise	
		Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank
Spectral Gray World	S_8^g	0.2748	1	0.257	1	0.1848	0	0.159	0	0.2654	1	0.2476	1
	S_8^d	0.3292	0	0.3059	0	0.0976	1	0.1326	0	0.3007	0	0.2779	0
	S_8^{50}	0.1884	2	0.1765	2	0.1237	1	0.1088	2	0.1807	2	0.1689	2
Max Spectral	S_8^g	0.1378	3	0.1334	4	0.0895	1	0.0942	2	0.1287	4	0.1241	5
	S_8^d	0.1552	2	0.151	2	0.0993	0	0.1673	0	0.1624	2	0.1661	1
	S_8^{50}	0.0986	9	0.1022	8	0.0631	7	0.0728	7	0.0942	8	0.1005	6
Spectral Shades of Gray	S_8^g	0.1623	1	0.1625	2	0.1114	1	0.1195	0	0.1564	2	0.1569	2
	S_8^d	0.1846	0	0.185	1	0.0653	5	0.1537	0	0.1704	1	0.1714	1
	S_8^{50}	0.1213	4	0.123	4	0.0823	2	0.085	4	0.1167	4	0.118	4
Spectral Gray Edge	S_8^g	0.1355	4	0.1362	2	0.1024	1	0.1027	1	0.1317	2	0.131	2
	S_8^d	0.1498	2	0.1502	3	0.0656	3	0.07	6	0.1408	2	0.1411	3
	S_8^{50}	0.1008	7	0.1019	8	0.0729	3	0.0723	5	0.0975	8	0.0985	8

Table 5.4: Performance of illuminant estimation algorithms and filter configurations for 12 bands. Max-Spectral with S_{12}^{50} performs best.

Algorithm	Filter. Confg.	D65				F11				Mixed illuminant (D65 and F11)			
		Without Noise		With Noise		Without Noise		With Noise		Without Noise		With Noise	
		Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank
Spectral Gray World	S_{12}^g	0.3007	1	0.2808	1	0.2022	0	0.1737	0	0.2905	1	0.2707	1
	S_{12}^d	0.344	0	0.3204	0	0.1806	1	0.215	0	0.3276	0	0.3027	0
	S_{12}^{50}	0.1881	2	0.1762	2	0.1251	2	0.1087	4	0.1804	2	0.1688	2
Max Spectral	S_{12}^g	0.1439	3	0.1447	4	0.107	2	0.1217	3	0.1386	4	0.139	2
	S_{12}^d	0.1574	2	0.158	2	0.0975	2	0.2122	0	0.1561	2	0.1545	2
	S_{12}^{50}	0.0996	9	0.1041	9	0.0642	7	0.0727	7	0.0966	8	0.101	8
Spectral Shades of Gray	S_{12}^g	0.1746	2	0.175	3	0.1152	2	0.134	4	0.1679	3	0.1681	3
	S_{12}^d	0.1964	0	0.1959	2	0.0912	3	0.2195	0	0.1851	2	0.1855	2
	S_{12}^{50}	0.1214	4	0.1232	4	0.0821	3	0.085	6	0.1168	4	0.1194	4
Spectral Gray Edge	S_{12}^g	0.1456	3	0.1455	4	0.1066	1	0.1078	3	0.1414	3	0.1414	3
	S_{12}^d	0.161	2	0.1618	2	0.0804	5	0.0931	6	0.153	2	0.1536	2
	S_{12}^{50}	0.1007	8	0.1009	9	0.0728	5	0.0737	7	0.0973	8	0.0976	9

Table 5.5: Performance of illuminant estimation algorithms and filter configurations for 20 bands. Max-Spectral with S_{20}^{50} configuration performs best for D65 and F11, while Spectral Gray Edge S_{20}^{50} performs best for mixed illuminants.

Algorithm	Filter. Confg.	D65				F11				Mixed illuminant (D65 and F11)			
		Without Noise		With Noise		Without Noise		With Noise		Without Noise		With Noise	
		Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank	Mean ΔA	Rank
Spectral Gray World	S_{20}^{g}	0.3212	1	0.3002	1	0.1994	1	0.1886	2	0.3099	1	0.2886	1
	S_{20}^{d}	0.3391	0	0.3158	0	0.2251	0	0.2535	0	0.3309	0	0.3061	0
	S_{20}^{50}	0.188	2	0.1762	2	0.1232	2	0.1087	5	0.1803	2	0.1687	2
Max Spectral	S_{20}^{g}	0.1521	2	0.1499	4	0.0956	2	0.1253	3	0.1445	4	0.1443	3
	S_{20}^{d}	0.1621	2	0.1576	2	0.1036	3	0.1851	0	0.1503	3	0.1493	3
	S_{20}^{50}	0.0996	9	0.1047	8	0.0626	9	0.0688	9	0.097	8	0.1023	8
Spectral Shades of Gray	S_{20}^{g}	0.1843	1	0.1848	3	0.1168	3	0.1572	2	0.1768	3	0.1772	3
	S_{20}^{d}	0.1944	0	0.1941	2	0.1317	2	0.2307	0	0.1885	2	0.1876	2
	S_{20}^{50}	0.1214	4	0.1232	4	0.0821	5	0.0854	7	0.1166	4	0.119	4
Spectral Gray Edge	S_{20}^{g}	0.1531	3	0.1536	4	0.1087	2	0.1089	5	0.1485	4	0.1489	4
	S_{20}^{d}	0.163	2	0.1633	3	0.1209	2	0.1284	3	0.16	2	0.16	2
	S_{20}^{50}	0.1006	8	0.1009	9	0.0728	7	0.0732	8	0.0972	9	0.0973	9

We also compare performance on individual multispectral images to determine the effect of scene content on illuminant estimation. In the following, analysis is provided on the data being generated with D65 illuminant. With 3 channels, images I1, I2, I4, I6 and I8 show good performance with Spectral Gray Edge S_3^g while with images I3, I5 and I7, Max-Spectral S_3^g performs the best. To illustrate the difference in projection of ground truth illuminant and the estimated illuminant, some examples are shown in Figures 5.4-5.8. In each figure, the x-axis represents each filter among the N filters and configuration, while the y-axis represent values of $\mathbf{e}$ and $\hat{\mathbf{e}}$, corresponding to the IFP. The points in figures are joined through straight lines so that the overall behavior can be observed easily. It is worth noting that the results for different number of filters are not comparable across Figures 5.4-5.8, since the dimension of filters is change in each of them. Figure 5.4 show the estimated illuminants in sensor domain for I3 and I4 when Spectral Gray Edge S_3^g is used. For 5 filters, I3 and I7 performs best with Max-Spectral S_5^{50} while other images show good results with Spectral Gray Edge S_5^{50}. I6 performs worst with Max-Spectral S_5^{50} which is the reason that this algorithm and configurations gets the 2$^{\text{nd}}$ best rank while Spectral Gray Edge S_5^{50} gets highest score for 5 channels. Figure 5.5 shows the estimated illuminant in sensor domain for I5 and I3. Figure 5.5 shows poor performance of illuminant estimation for I3 and I5. At this stage, the trend of improvement in Max-Spectral can already be observed which becomes clear with 8 channels as Max-Spectral S_8^{50} performs best for all images except I6 which works well with Spectral Gray Edge S_8^{50}. The performance of Max-Spectral S_8^{50} for images I3 and I5 is shown in Figure 5.6. Same behavior is shown by individual images with 12 and 20 channels as well. Figures 5.7 and 5.8 show performance of Max-Spectral S^{50} for I6 and I7 when number of channels are 12 and 20 respectively. In other images, there is close tie between Max-Spectral S_8^{50} and Spectral Gray Edge S_8^{50} but images I3 and I6 do not perform well with Spectral Gray Edge S_8^{50} and thus causing it to get overall 2$^{\text{nd}}$ rank. Angular errors for all the algorithms, number of filters, filter configurations and illuminants being used are provided in the supplementary material. We have also provided the error in terms of xy chromaticity for each of the individual image along with the other parameters being tested, in the supplementary material.

Overall, the configuration S^{50} performs the best among tested filter configurations. Max-Spectral and Spectral Gray-Edge attain good results while Spectral Gray World shows the worst result for all cases. S^d shows slightly better performance with $F11$ illuminant but otherwise it also performs worst. It is interesting to note that Spectral Gray-Edge performs better with 3 bands but by increasing the number of bands, Max-Spectral algorithms starts performing the best among the tested algorithms. We investigate that trend by altering the value of Minkowski norm p as in Eq. 5.4 and 5.5. When the value of p parameter is increased, more

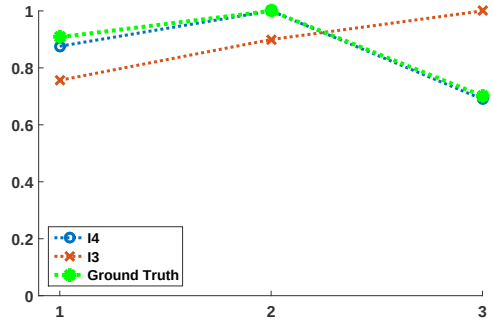


Figure 5.4: Illuminant's Projection over Filters (IPF) of D65 with N=3 channels and estimated illuminants with Spectral Gray Edge S_3^g for images I3 and I4 with ΔA of 0.228 and 0.0158, respectively. I4 gives good result while I3 performs worst.

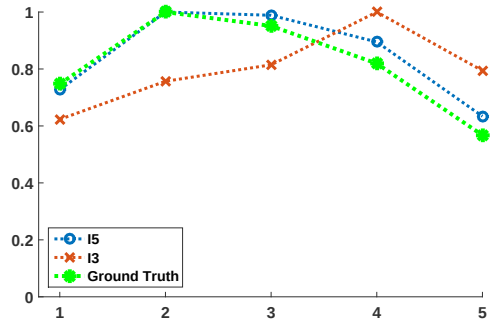


Figure 5.5: Illuminant's Projection over Filters (IPF) of D65 with N=5 channels and estimated illuminants with Spectral Gray Edge S_5^{50} in sensor domain for images I3 and I5 with ΔA of 0.2284 and 0.0457, respectively. I5 gives good result while I3 performs worst.

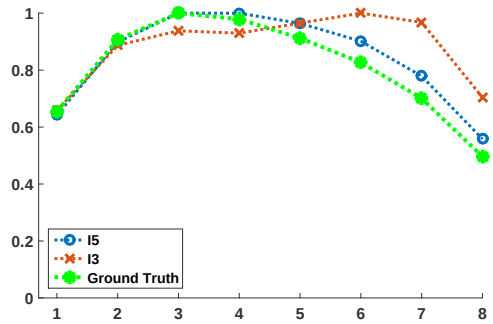


Figure 5.6: Illuminant's Projection over Filters (IPF) of D65 with N=8 channels and estimated illuminants with Max-Spectral S_8^{50} in sensor domain for images I3 and I5 with ΔA of 0.1142 and 0.0446, respectively. I5 gives good result while I3 performs worst.

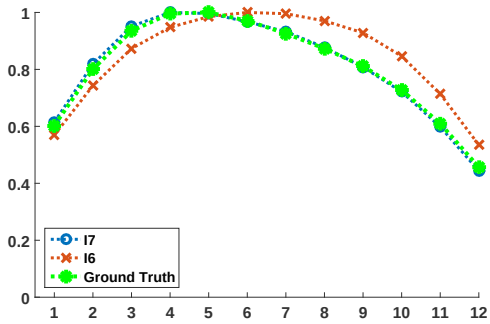


Figure 5.7: Illuminant's Projection over Filters (IPF) of D65 with N=12 channels and estimated illuminants with Max-Spectral S_{12}^{50} in sensor domain for images I6 and I7 with ΔA of 0.0838 and 0.0117, respectively. I7 gives good result.

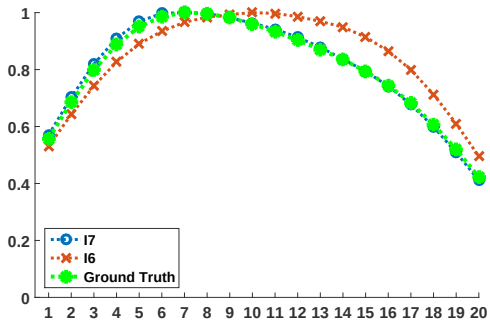


Figure 5.8: Illuminant's Projection over Filters (IPF) of D65 with N=20 channels and estimated illuminants with Max-Spectral S_{20}^{50} in sensor domain for images I6 and I7 with ΔA of 0.0839 and 0.0117, respectively. I7 gives good result.

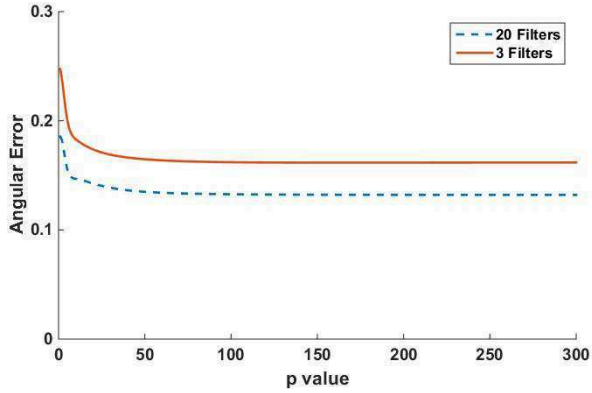


Figure 5.9: Change in Angular Error with variation in p. We did experiments for $p = 1 - 1000$ but show results only up to 300 because there is no change in error value as the value of p is increased beyond 100.

weight is given to bright pixels in an image and this ultimately lead towards Max-Spectral algorithm. We performed tests with values of p varying from 1 to 1000. The results show very interesting observation that as more weight is given to bright pixels in a scene, the illuminant estimation gets better. Thus it explains why Max-Spectral algorithm perform well specially with increase in number of bands. Figure 5.9 shows the change in angular error with variation in value of p.

Tables 5.1-5.5 provide analysis of the performance of the proposed algorithms along with a given sensor configuration, in terms of ΔA. However, these results cannot be compared across the different number of filters because ΔA can be compared between two vectors only if they have the same dimension (in our case, the ground-truth and estimated illuminants are in the sensor dimension).

Mosny and Funt [66, 67] performed their evaluation in rg chromaticity space. In their method, RGB of estimated illuminant is obtained after identifying which illuminant from a database of known illuminants it is most similar to, and using that illuminant's rgb, as the conversion value. Based on this evaluation, they concluded that there is minor improvement in increasing number of bands from 3 to 6 for illuminant estimation but further increase to 9 bands does not provide any improvement. For evaluating the effect of number of bands, we perform the evaluation based on chromaticity error in Table 5.6 but with a different approach as defined in Section 5.4.3. The comparison is performed among 5 different number of spectral filters (3, 5, 8, 12 & 20), 3 sensor configurations (equi-Gaussian, Dirac delta & equi-energy filters), and 4 algorithms (Spectral gray-world, Max-Spectral, Spectral shades of gray & Spectral gray-edge algorithm). We have used

Table 5.6: Ranking based on xy chromaticity error in terms of Euclidean distance (ED). Green cells show the best performance, yellow cells show the 2nd best and blue cells represent the 3rd best performance, for a particular algorithm and filters configuration. Spectral Edge S_8^{50} gives the best performance for $D65$ illuminant. With $F11$ illuminant, Max-Spectral S_{20}^{d} performs the best and is followed by Max-Spectral S_8^{g}. In case of mix D65-F11 illuminant, Max-Spectral S_{20}^{d} gives the best result and Spectral Edge S_8^{50} comes after it. Overall, Max-Spectral algorithm gives consistent performance and illuminant is best estimated with 8 filters.

No. of filters	Algorithm	Filter. Config.	D65				F11				Mixed illuminant (D65 and F11)			
			Without Noise		With Noise		Without Noise		With Noise		Without Noise		With Noise	
			Mean ED	Rank	Mean ED	Rank	Mean ED	Rank	Mean ED	Rank	Mean ED	Rank	Mean ED	Rank
3	Spectral Gray World	S_3^g	0.0702	3	0.0621	3	0.0575	4	0.0469	10	0.0692	2	0.0608	6
		S_3^d	0.0741	1	0.0671	2	0.0592	2	0.0577	6	0.0727	0	0.0664	3
		S_3^{50}	0.0712	2	0.0633	2	0.0583	3	0.0478	9	0.0702	1	0.0619	4
	Max Spectral	S_3^g	0.0274	6	0.0268	16	0.028	10	0.0272	15	0.0274	7	0.0251	27
		S_3^d	0.0317	2	0.0296	8	0.034	2	0.0362	8	0.0322	2	0.0306	14
		S_3^{50}	0.0283	3	0.0236	28	0.0278	10	0.0223	29	0.0283	3	0.0256	24
	Spectral Shades of Gray	S_3^g	0.034	3	0.0345	6	0.0336	2	0.0332	11	0.0343	3	0.0342	6
		S_{20}^d	0.0346	4	0.0351	7	0.0344	8	0.0379	11	0.0345	4	0.0349	7
		S_3^{50}	0.0342	3	0.0345	6	0.0334	3	0.0333	12	0.0345	3	0.0347	6
	Spectral Gray Edge	S_3^g	0.0281	10	0.0284	13	0.029	10	0.0287	13	0.0283	14	0.0283	12
		S_3^d	0.0285	12	0.0286	15	0.0264	22	0.0262	22	0.0285	15	0.0284	17
		S_3^{50}	0.0282	10	0.028	15	0.0286	11	0.0285	13	0.0283	13	0.0283	21
5	Spectral Gray World	S_5^g	0.0736	2	0.065	3	0.0629	1	0.0516	7	0.0728	1	0.0642	4
		S_5^d	0.0819	1	0.072	2	0.067	1	0.0796	4	0.0804	0	0.0722	3
		S_5^{50}	0.0747	2	0.0648	3	0.0627	2	0.0499	7	0.0738	1	0.0637	4
	Max Spectral	S_5^g	0.0263	15	0.0276	15	0.0262	17	0.0212	30	0.0232	29	0.0218	34
		S_5^d	0.028	14	0.028	15	0.0337	3	0.0425	8	0.0319	5	0.0303	9
		S_5^{50}	0.0249	16	0.0202	38	0.029	17	0.0309	15	0.0241	28	0.0219	26
	Spectral Shades of Gray	S_5^g	0.0355	2	0.0354	6	0.0328	3	0.0305	14	0.0354	2	0.0351	7
		S_5^{50}	0.036	3	0.0358	6	0.0331	3	0.0484	7	0.0363	3	0.0363	7
		S_5^d	0.0352	2	0.035	6	0.033	3	0.0306	13	0.0351	2	0.0347	7
	Spectral Gray Edge	S_5^g	0.0313	3	0.0312	6	0.0299	9	0.0301	13	0.0313	3	0.0311	7
		S_5^d	0.0302	5	0.0302	10	0.0288	12	0.0291	13	0.0302	6	0.03	10
		S_5^{50}	0.0311	4	0.0309	6	0.0304	9	0.0308	13	0.0313	4	0.0314	7
8	Spectral Gray World	S_8^g	0.0702	8	0.0628	4	0.06	7	0.0473	8	0.0693	6	0.0619	8
		S_8^d	0.0725	2	0.0645	2	0.0579	9	0.0376	8	0.0712	1	0.0626	5
		S_8^{50}	0.0701	7	0.0687	2	0.0605	7	0.0575	6	0.0693	6	0.0676	3
	Max Spectral	S_8^g	0.0267	15	0.0255	21	0.0217	32	0.0169	51	0.0229	25	0.0212	35
		S_8^d	0.0274	10	0.0273	9	0.0296	4	0.0496	7	0.027	5	0.0256	13
		S_8^{50}	0.0546	2	0.0925	1	0.0292	15	0.082	2	0.0331	2	0.1005	2
	Spectral Shades of Gray	S_8^g	0.0311	3	0.0312	7	0.0287	13	0.0271	20	0.0308	4	0.0309	8
		S_8^d	0.0316	4	0.032	7	0.0285	8	0.0493	7	0.0311	4	0.0313	8
		S_8^{50}	0.0239	19	0.0262	12	0.023	26	0.0251	23	0.0234	19	0.0258	7
	Spectral Gray Edge	S_8^g	0.0267	20	0.0269	20	0.0257	20	0.0255	21	0.0267	19	0.0265	21
		S_8^d	0.0263	24	0.0265	26	0.0256	22	0.0239	30	0.0262	25	0.0265	26
		S_8^{50}	0.0232	33	0.0159	43	0.0232	24	0.0258	17	0.0222	30	0.0168	44
12	Spectral Gray World	S_{12}^g	0.0728	2	0.0651	2	0.0621	2	0.0478	8	0.0718	1	0.0643	3
		S_{12}^d	0.0754	1	0.0669	2	0.0729	0	0.1015	2	0.0753	0	0.0667	3
		S_{12}^{50}	0.0729	1	0.0482	4	0.0619	2	0.05	9	0.0719	1	0.0592	5
	Max Spectral	S_{12}^g	0.0277	14	0.0275	9	0.0241	17	0.0205	33	0.0231	20	0.0218	26
		S_{12}^d	0.0292	12	0.0274	15	0.0259	14	0.0782	3	0.0245	15	0.0227	23
		S_{12}^{50}	0.1152	0	0.103	2	0.3765	0	1.4405	0	0.1457	0	0.3697	0
	Spectral Shades of Gray	S_{12}^g	0.0316	5	0.0319	8	0.0295	13	0.0277	20	0.0313	4	0.0315	8
		S_{12}^d	0.0322	4	0.0324	7	0.0322	2	0.1069	2	0.0321	4	0.0327	8
		S_{12}^{50}	0.0298	3	0.0823	1	0.0367	3	0.0657	3	0.0293	4	0.0354	4
	Spectral Gray Edge	S_{12}^g	0.0268	23	0.0268	25	0.0259	20	0.0262	22	0.0267	21	0.0268	24
		S_{12}^d	0.0268	24	0.027	24	0.0271	13	0.0443	7	0.0269	23	0.0272	23
		S_{12}^{50}	0.0339	13	0.0459	15	0.0396	3	0.3079	0	0.0337	14	0.118	3
20	Spectral Gray World	S_{20}^g	0.0726	2	0.0649	2	0.0619	2	0.0443	10	0.0717	1	0.064	3
		S_{20}^d	0.0702	2	0.062	3	0.0652	1	0.0387	10	0.0704	1	0.0609	4
		S_{20}^{50}	0.0729	2	0.0723	1	0.0619	2	0.0393	11	0.0719	1	0.0441	6
	Max Spectral	S_{20}^g	0.0282	14	0.027	17	0.0225	28	0.0182	37	0.023	28	0.0221	31
		S_{20}^d	0.0256	15	0.0388	6	0.0199	45	0.0398	8	0.0219	39	0.0188	47
		S_{20}^{50}	0.4377	0	0.3142	0	0.2082	1	0.2405	0	0.195	0	0.4725	0
	Spectral Shades of Gray	S_{20}^g	0.032	4	0.0324	7	0.0296	13	0.0273	22	0.0316	4	0.0318	8
		S_{20}^d	0.0328	3	0.0331	6	0.0301	10	0.0434	7	0.0327	4	0.0322	8
		S_{20}^{50}	0.1013	0	0.2115	0	0.034	14	0.0497	7	0.0914	0	0.2327	1
	Spectral Gray Edge	S_{20}^g	0.0274	19	0.0273	22	0.0261	20	0.0263	23	0.0273	22	0.0274	22
		S_{20}^d	0.0289	17	0.0291	17	0.0257	27	0.0286	17	0.0289	20	0.0291	19
		S_{20}^{50}	0.1146	0	0.1693	0	0.0347	13	0.1177	2	0.1409	0	0.5641	0

Euclidean distance for evaluation of xy chromaticity error since we are assuming that our evaluation is in terms of physical measurement. Using the xy chromaticity space allows us to retain our assumption and enables the comparison among the ground-truth illuminant and the estimated illuminant.

Evaluation based on xy chromaticity error for D65 shows that best result is obtained from Spectral Gray Edge S_8^{50} and 2nd best are the Spectral Gray Edge S_8^d and S_{12}^d. However, there is a significant statistical difference among S^{50} and S^d for this illuminant which becomes more prominent in the case of noisy data.

With F11 illuminant, Max-Spectral S_{20}^d performs the best and is followed by Max-Spectral S_8^g. This behavior of is explained from the spectral power distribution of F11, as shown in Fig. 5.2b. The spiky character of this illuminant can be best acquired with the ideal Dirac delta type of filters. However, in presence of noise, performance of Max-Spectral S_{20}^d is significantly reduced. Max-Spectral S_8^g performs best in case of noisy data and is followed by Max-Spectral S_{20}^g.

In case of mix D65-F11 illuminant, Max-Spectral S_{20}^d performs the best while Spectral Gray Edge S_8^{50} and Max-Spectral S_5^g perform 2nd and 3rd best respectively. Since the behavior of mix D65-F11 illuminant is influenced by peaks of $F11$ illuminant, therefore, S_{20}^d performs best in this case. Same trend continues in case of noisy data where the statistically significant difference among results is more prominent in light of WST rankings.

It is interesting to note that by increasing the number of channels beyond 8, there is a reduction in performance of illuminant estimation algorithms. It suggests that spectral resolution should also be maintained in a multispectral imaging system. As noticed from Table 5.6, S_{20}^{50} configuration performs the worst because of huge overlapping among filter sensitivities. This leads to the conclusion that by increasing the number of bands, more noise is introduced during image acquisition and therefore, performance of illuminant estimation algorithm is degraded. To validate this, we performed an additional illuminant estimation experiment using the native spectral resolution of the data, which is equivalent to Dirac delta configuration with 33 filters (S_{33}^d). There is no improvement in results when compared with the already obtained results from 20 channels and it performs the worst when noise is added to the system. This fact is also observed by Wang *et al.* [74] where the spectral reconstruction results start degrading after increasing the number of filters beyond 12.

Although the results and ranks are based only on 8 images of similar contents and may not lead to a strong conclusion, our investigation suggests several general behaviors. First, overlapping equi-energy filters may be most suitable for

natural or smooth illuminations. Although there may be loss of spectral resolution in case of using large overlapping sensors, since natural illuminations behave smoothly throughout the visible spectrum, overlapping equi-energy filters are able to perform well. We observe the same trend after noise is added to the images before illuminant estimation. Second, the Max-Spectral and Spectral Gray Edge algorithms provide better results than the other tested algorithms in general. Result is rather dependent on image content also, and in some of the images, better estimate of illuminant is achieved (data seems to follow the illumination), in other the results are quite noisy. Third, we found contradictory results as compared to Mosny and Funt [66] and our results suggest that illuminant chromaticity can be better retrieved when we increase the number of bands. However, the impact on color rendering is yet to be investigated. The optimum number of bands seems to be around 8. Finally, we still cannot provide clear indications on how good illuminant estimation is in term of usability. In practice, the indicator used only provides relative ranking and objective indications on quality. Further analysis is required to understand what accuracy should be achieved for acquiring illuminant invariant representation of multispectral images.

5.6 Conclusion and Future Work

In this work, we proposed to extend illuminant estimation from color to multispectral imaging. Based on an extensive review of state of the art algorithms for computational color constancy, we selected 4 algorithms that belong to the class of *equivalent illumination models*, and extended them from 3-channels to N-channels. We named those extended algorithms *spectral gray world*, *max-spectral*, *spectral shades of gray*, and *spectral gray edge* algorithms. Results show that both spectral gray edge and max-spectral algorithms perform well in illuminant estimation. Comparison among three different sensor sensitivities is also performed and the overlapping equi-energy filters are able to estimate the illuminant more accurately as compared to equi-Gaussian or Dirac delta functions for limited number of channels. The same results are obtained when noise is added to the image data which shows that the proposed extended algorithms for illuminant estimation are robust to noise.

The illuminant estimation results obtained from simulated multispectral sensors show promising aspects of application of the proposed framework. Based on these results, future work could be derived in three directions. First, development of new algorithms or further extension of more sophisticated illuminant estimation algorithms from color to spectral may be performed. Second, the validity of the proposed framework may be evaluated for real data acquired with a multispectral camera. The evaluation can also be performed in terms of color difference and

the spectral reconstruction error. Finally further development in evaluation and usability of this framework may be performed, for instance by evaluating surface classification under different illuminations.

Bibliography

[1] A. Stockman, D. I. A. MacLeod, and N. E. Johnson, "Spectral sensitivities of the human cones," *J. Opt. Soc. Am. A*, vol. 10, pp. 2491–2521, Dec 1993.

[2] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton, "Multispectral filter arrays: Recent advances and practical implementation," *Sensors*, vol. 14, no. 11, pp. 21626–21659, 2014.

[3] J.-B. Thomas, P.-J. Lapray, P. Gouton, and C. Clerc, "Spectral Characterization of a Prototype SFA Camera for Joint Visible and NIR Acquisition," *Sensors*, vol. 16, no. 7, p. 993, 2016.

[4] R. Shrestha, J. Y. Hardeberg, and R. Khan, "Spatial arrangement of color filter array for multispectral image acquisition," in *Proc. SPIE, Sensors, Cameras, and Systems for Industrial, Scientific, and Consumer Applications XII*, vol. 7875, pp. 787503–787503–9, Feb. 2011.

[5] Y. Manabe, K. Sato, and S. Inokuchi, "An object recognition through continuous spectral images," in *12th International Conference on Pattern Recognition*, vol. 1, pp. 858–860, Oct 1994.

[6] C. Liu, W. Liu, X. Lu, F. Ma, W. Chen, J. Yang, and L. Zheng, "Application of multispectral imaging to determine quality attributes and ripeness stage in strawberry fruit," *PLOS ONE*, vol. 9, pp. 1–8, 02 2014.

[7] F. Imai and R. Berns, "Spectral estimation using trichromatic digital cameras," in *International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives*, pp. 42–49, 1999.

[8] D. Connah, S. Westland, and M. G. A. Thomson, "Recovering spectral information using digital camera systems," *Coloration Technology*, vol. 117, no. 6, pp. 309–312, 2001.

[9] E. M. Valero, J. L. Nieves, S. M. C. Nascimento, K. Amano, and D. H. Foster, "Recovering spectral data from natural scenes with an RGB digital camera and colored filters," *Color Research & Application*, vol. 32, no. 5, pp. 352–360, 2007.

[10] R. Shrestha and J. Y. Hardeberg, "Spectrogenic imaging: A novel approach to multispectral imaging in an uncontrolled environment," *Opt. Express*, vol. 22, pp. 9123–9133, Apr 2014.

[11] J. Y. Hardeberg and R. Shrestha, "Multispectral colour imaging: Time to move out of the lab?," in *Mid-term meeting of the International Colour Association (AIC)*, (Tokyo, Japan), pp. 28–32, May 2015.

[12] J.-B. Thomas, "Illuminant estimation from uncalibrated multispectral images," in *Colour and Visual Computing Symposium (CVCS)*, (Gjøvik, Norway), pp. 1–6, Aug 2015.

[13] O. Bertr and C. Tallon-Baudry, "Oscillatory gamma activity in humans: a possible role for object representation," *Trends in Cognitive Sciences*, vol. 3, pp. 151–162, 1999.

[14] M. Ebner, *Color Constancy*. Wiley Publishing, 1st ed., 2007.

[15] D. H. Brainard and L. T. Maloney, "Surface color perception and equivalent illumination models," *Journal of Vision*, vol. 11, no. 5, p. 1, 2011.

[16] D. H. Brainard, W. A. Brunt, and J. M. Speigle, "Color constancy in the nearly natural image. 1. Asymmetric matches," *J. Opt. Soc. Am. A*, vol. 14, pp. 2091–2110, Sep 1997.

[17] K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms. i: Methodology and experiments with synthesized data," *IEEE Transactions on Image Processing,*, vol. 11, pp. 972–984, Sep 2002.

[18] S. J. Dickinson, "Object representation and recognition," in *What is Cognitive Science* (E. Lepore and Z. Pylyshyn, eds.), ch. 5, pp. 172–207, Basil Blackwell Publishers, 1999.

[19] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.

[20] B. V. Funt and G. D. Finlayson, "Color constant color indexing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 522–529, May 1995.

[21] T. Gevers and A. W. Smeulders, "Color-based object recognition," *Pattern Recognition*, vol. 32, no. 3, pp. 453 – 464, 1999.

[22] S. D. Hordley, "Scene illuminant estimation: Past, present, and future," *Color Research and Application*, vol. 31, no. 4, pp. 303–314, 2006.

[23] L. T. Maloney and B. A. Wandell, "Color constancy: a method for recovering surface spectral reflectance," *J. Opt. Soc. Am. A*, vol. 3, pp. 29–33, Jan 1986.

[24] D. Cheng, B. Price, S. Cohen, and M. S. Brown, "Effective learning-based illuminant estimation using simple features," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1000–1008, June 2015.

[25] M. D'Zmura and P. Lennie, "Mechanisms of color constancy," *J. Opt. Soc. Am. A*, vol. 3, pp. 1662–1672, Oct 1986.

[26] E. Land, "Recent Advances in Retinex Theory and Some Implications for Cortical Computations: Color Vision and the Natural Image," *Proceedings of the National Academy of Science*, vol. 80, pp. 5163–5169, Aug. 1983.

[27] E. H. Land and J. J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Am.*, vol. 61, pp. 1–11, Jan 1971.

[28] E. H. Land, "The retinex theory of color vision," *Scientific American*, pp. 108–128, 1977.

[29] J. von Kries, "Influence of adaptation on the effects produced by luminous stimuli," *D.L. MacAdam (Ed.), Sources of color science*, pp. 109–119, 1970.

[30] von Kries J., "Theoretische Studien über die Umstimmung des Sehorgans," *Handbuch der Physiologie des Menschen*, pp. 211–212, 1905.

[31] G. Buchsbaum, "A spatial processor model for object colour perception," *Journal of the Franklin Institute*, vol. 310, no. 1, pp. 1–26, 1980.

[32] R. Gershon, A. D. Jepson, and J. K. Tsotsos, "From [r,g,b] to surface reflectance: Computing color constant descriptors in images," in *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, (San Francisco, CA, USA), pp. 755–758, Morgan Kaufmann Publishers Inc., 1987.

[33] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Color and Imaging Conference*, (Scottsdale, Arizona), pp. 37–41, 2004.

[34] J. van de Weijer and T. Gevers, "Color constancy based on the grey-edge hypothesis," in *IEEE International Conference on Image Processing*, vol. 2, pp. II–722–5, Sept 2005.

[35] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Transactions on Image Processing*, vol. 16, pp. 2207–2214, Sept 2007.

[36] T. Celik and T. Tjahjadi, "Adaptive colour constancy algorithm using discrete wavelet transform," *Computer Vision and Image Understanding*, vol. 116, no. 4, pp. 561–571, 2012.

[37] A. Chakrabarti, K. Hirakawa, and T. Zickler, "Color constancy with spatio-spectral statistics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, pp. 1509–1519, Aug 2012.

[38] M. Rezagholizadeh and J. Clark, "Edge-based and efficient chromaticity spatio-spectral models for color constancy," in *International Conference on Computer and Robot Vision*, pp. 188–195, May 2013.

[39] A. Gijsenij and T. Gevers, "Color constancy using natural image statistics and scene semantics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 687–698, April 2011.

[40] D. Forsyth, "A novel algorithm for color constancy," *International Journal of Computer Vision*, vol. 5, no. 1, pp. 5–35, 1990.

[41] A. Gijsenij, T. Gevers, and J. van de Weijer, "Generalized gamut mapping using image derivative structures for color constancy," *International Journal of Computer Vision*, vol. 86, no. 2, pp. 127–139, 2010.

[42] G. D. Finlayson, S. D. Hordley, and P. M. HubeL, "Color by correlation: a simple, unifying framework for color constancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1209–1221, Nov 2001.

[43] J. Huo, Y. Chang, J. Wang, and X. Wei, "Robust automatic white balance algorithm using gray color points in images," *IEEE Transactions on Consumer Electronics*, vol. 52, pp. 541–546, May 2006.

[44] K.-J. Yoon, Y. J. Chofi, and I.-S. Kweon, "Dichromatic-based color constancy using dichromatic slope and dichromatic line space," in *IEEE International Conference on Image Processing*, vol. 3, pp. III–960–3, Sept 2005.

[45] S. Ratnasingam and S. Collins, "Study of the photodetector characteristics of a camera for color constancy in natural scenes," *J. Opt. Soc. Am. A*, vol. 27, pp. 286–294, Feb 2010.

[46] G. Sapiro, "Color and illuminant voting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, pp. 1210–1215, Nov 1999.

[47] D. H. Brainard and W. T. Freeman, "Bayesian color constancy," *J. Opt. Soc. Am. A*, vol. 14, pp. 1393–1411, Jul 1997.

[48] W. Xiong and B. Funt, "Stereo Retinex," in *The 3rd Canadian Conference on Computer and Robot Vision*, pp. 15–15, June 2006.

[49] G. D. Finlayson, S. D. Hordley, and P. Morovic, "Chromagenic colour constancy," in *10th Congress of the International Colour Association*, 2005.

[50] C. Fredembach and G. Finlayson, "The bright-chromagenic algorithm for illuminant estimation," *Journal of Imaging Science and Technology*, vol. 54, no. 4, p. 040906–1–040908–11, 2008.

[51] B. A. Wandell, "The synthesis and analysis of color images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 2–13, Jan 1987.

[52] K. Barnard, G. Finlayson, and B. Funt, "Color constancy for scenes with varying illumination," *Computer Vision and Image Understanding*, vol. 65, no. 2, pp. 311–321, 1997.

[53] G. D. Finlayson, B. V. Funt, and K. Barnard, "Color constancy under varying illumination," in *5th International Conference on Computer Vision*, pp. 720–725, Jun 1995.

[54] J. L. Nieves, C. Plata, E. M. Valero, and J. Romero, "Unsupervised illuminant estimation from natural scenes: an rgb digital camera suffices," *Appl. Opt.*, vol. 47, pp. 3574–3584, Jul 2008.

[55] V. C. Cardei, B. Funt, and K. Barnard, "Estimating the scene illumination chromaticity by using a neural network," *J. Opt. Soc. Am. A*, vol. 19, pp. 2374–2386, Dec 2002.

[56] H. R. V. Joze and M. S. Drew, "Improved machine learning for image category recognition by local color constancy," in *IEEE International Conference on Image Processing*, pp. 3881–3884, Sept 2010.

[57] J. T. Barron, "Convolutional color constancy," in *IEEE International Conference on Computer Vision*, pp. 379–387, Dec 2015.

[58] V. Agarwal, A. V. Gribok, and M. A. Abidi, "Machine learning approach to color constancy," *Neural Networks*, vol. 20, no. 5, pp. 559–563, 2007.

[59] N. Wang, D. Xu, and B. Li, "Edge-Based Color Constancy via Support Vector Regression," *IEICE Transactions on Information and Systems*, vol. 92, pp. 2279–2282, 2009.

[60] Z. Lou, T. Gevers, N. Hu, and M. Lucassen, "Color constancy by deep learning," in *British Machine Vision Conference*, 2015.

[61] W. Shi, C. C. Loy, and X. Tang, "Deep specialized network for illuminant estimation," in *European Conference on Computer Vision*, pp. 371–387, Springer, 2016.

[62] S. Bianco, C. Cusano, and R. Schettini, "Color constancy using CNNs," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 81–89, June 2015.

[63] S. W. Oh and S. J. Kim, "Approaching the computational color constancy as a classification problem through deep learning," *Pattern Recognition*, vol. 61, pp. 405–416, 2017.

[64] F. Vagni, *Survey of Hyperspectral and Multispectral Imaging Technologies*. Technichnical Report. North Atlantic Treaty Organization, Sensors and Electronics Technology Panel, 2007.

[65] K.-S. Lee, W. B. Cohen, R. E. Kennedy, T. K. Maiersperger, and S. T. Gower, "Hyperspectral versus multispectral data for estimating leaf area index in four different biomes," *Remote Sensing of Environment*, vol. 91, no. 3–4, pp. 508–520, 2004.

[66] M. Mosny and B. Funt, "Multispectral colour constancy," *Color and Imaging Conference*, no. 1, pp. 309–313, 2006.

[67] M. Mosny and B. Funt, "Multispectral color constancy: real image tests," in *SPIE, Human Vision and Electronic Imaging XII*, vol. 6492, pp. 64920S–6, 2007.

[68] D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster, "Frequency of metamerism in natural scenes," *J. Opt. Soc. Am. A*, vol. 23, pp. 2359–2372, Oct 2006.

[69] M. DiCosola, "Understanding illuminants," tech. rep., X-Rite, 1995.

[70] M. Rezagholizadeh and J. J. Clark, "Image sensor modeling: Color measurement at low light levels," *Journal of Imaging Science and Technology*, vol. 58, no. 3, pp. 30401–1–30401–11, 2014.

[71] K. Barnard and B. Funt, "Camera characterization for color research," *Color Research & Application*, vol. 27, pp. 152–163, 6 2002.

[72] G. D. Finlayson, M. S. Drew, and B. V. Funt, "Spectral sharpening: sensor transformations for improved color constancy," *J. Opt. Soc. Am. A*, vol. 11, pp. 1553–1563, May 1994.

[73] P.-J. Lapray, J.-B. Thomas, P. Gouton, and Y. Ruichek, "Energy balance in spectral filter array camera design," *Journal of the European Optical Society-Rapid Publications*, vol. 13, no. 1, p. 1, 2017.

[74] X. Wang, J.-B. Thomas, J. Y. Hardeberg, and P. Gouton, "Multispectral imaging: narrow or wide band filters?," *Journal of the International Colour Association*, vol. 12, pp. 44–51, July 2014. http://jaic.jsitservices.co.uk/index.php/JAIC/article/view/149.

[75] J. Hernández-Andrés, J. Romero, J. L. Nieves, and R. L. Lee, "Color and spectral analysis of daylight in southern europe," *J. Opt. Soc. Am. A*, vol. 18, pp. 1325–1335, Jun 2001.

[76] J. Conde, H. Haneishi, M. Yamaguchi, N. Ohyama, and J. Baez, "Cie-xyz fitting by multispectral images and mean square error minimization with a linear interpolation function," *Revista Mexicana de Física*, vol. 6, p. 601–607, 2004.

[77] S. D. Hordley and G. D. Finlayson, "Reevaluation of color constancy algorithm performance," *J. Opt. Soc. Am. A*, vol. 23, pp. 1008–1020, May 2006.

[78] S. Bianco, F. Gasparini, and R. Schettini, "Consensus-based framework for illuminant chromaticity estimation," *Journal of Electronic Imaging*, vol. 17, p. 023013, 2008.

[79] G. D. Finlayson, R. Zakizadeh, and A. Gijsenij, "The reproduction angular error for evaluating the performance of illuminant estimation algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PP, no. 99, pp. 1–1, 2016.

Chapter 6

Analytical survey of highlight detection in color and spectral images

Article C

This chapter is a reformatted reprint of the publication

H. A. Khan, J. B. Thomas and J. Y. Hardeberg, "Analytical survey of highlight detection in color and spectral images", *Computational Color Imaging Workshop, CCIW*, Milan, Italy, Lecture Notes in Computer Science, vol 10213. Springer, Cham, 2017.

Detection of highlights is a prominent issue in computer vision, graphics and image processing. Applications which require object properties measurement or rendering are affected by specular reflection since the models assume matte diffusing surfaces most of the time. Hence, detection, and sometimes removal, of specular reflection (highlights) in an image may be critical. Several methods are proposed for addressing this issue. In this paper, we present a review and analysis of these techniques in color and spectral images.

6.1 Introduction

The process of extracting information from an image, and its transformation into a useful representation, enables the description of intrinsic characteristics of objects in the scene. Barrow *et al.* [1] introduced the term "intrinsic images" and suggested that the function of the human visual system at its earlier stage is to determine the orientation of the illumination and the surface being observed. The authors describe such details of scenes in term of shading and reflectance images, which are collectively denoted as intrinsic images. There are other intrinsic properties as well, including shading, reflectance, diffuse reflection components and specular reflection components. Diffuse reflection is caused by scattering of light in all directions after hitting the surface, while specular reflection occurs when incident light is reflected in a single direction. Lee *et al.* [2] presented a neutral interface reflection model by examining the light reflection problem through the use of bi-directional spectral-reflectance distribution function (BSRDF) for specifying both incident and reflected beam geometry. They proposed that specular reflection is identical to scene illuminant in color while diffuse reflection contains the intrinsic properties of the surface. There can be significant variations in appearance of a surface in presence of specular reflection, and they appear as an additional surface property which is not intrinsic [3, 4].

We are interested in the detection of specular reflection in images. Several computer vision applications such as 3D reconstruction, object detection, recognition, target tracking, and dichromatic editing use the intrinsic information and are mostly based on the assumption of the surfaces having perfect diffusion. In most of such algorithms, specular reflections and highlights are termed as outliers [5]. However, the presence of specular reflection is unavoidable in most of the real world scenes since the materials and surfaces not only possess diffuse reflection but also show specular reflections, which is explained in the Dichromatic Illumination Model (DIM) [6]. Hence, the assumption of absence of specular reflection for such algorithms introduce constraints and reduce their robustness. There is a loss of details in case of specular reflection, for example texture and color of surface being observed. Highlight removal is often considered as an inpainting problem [7–10]. This methodology is implemented after identification of highlights. In this paper, we discuss the methods for highlight detection. Discussion about removal of specular highlights is out of scope of this paper.

As an example, Figure 6.1 shows an image of a printed circuit board. The information about connectivity of the circuit is totally washed out in the area which is under direct specular reflection. In addition, there are highlight spots on the metallic surfaces of diodes and resistors. Such spots follow the DIM and create

ambiguity concerning the intrinsic characteristics of the object surface. It may thus be desirable to remove the effect of highlights from the image for recovery of information.

Highlight detection algorithms can be classified into various categories on the basis of data being used for input. For color images, we propose two major categories, namely the *single-image* and *multi-images* based techniques, presented in Section 2 and 3, respectively. Furthermore, a relatively recent development in imaging technology is spectral imaging. The problem of specular highlights occurs also in spectral images, therefore we have included it in this review as another category (Section 4). However, this review focuses only on close range spectral imaging and does not address remote sensing images. In Section 5, we discuss and compare key features of the presented algorithms, before concluding.

6.2 Single color image based techniques

We have defined two major categories of algorithms to detect and remove specularities in a single color image. Those categories use either the color space analysis, or spatial information analysis.

The techniques based on color space analysis treat an image pixel by pixel. Klinker *et al.* [11] classified color pixels in the categories of *diffuse*, *highlight* and *saturated* pixels. A diffuse pixel is defined as a pixel containing only the body reflectance (although the color of body is influenced by scene illuminant), a highlight pixel contain both body and specular reflections while a saturated pixel is created

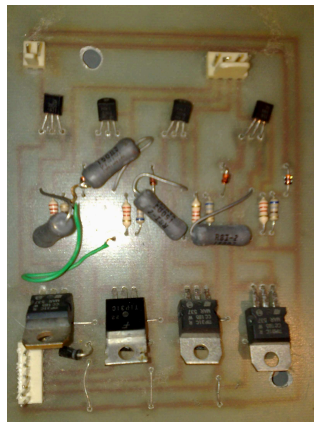


Figure 6.1: Image of printed circuit board with transistors, resistors and diodes. Specular highlights can be observed in the highly saturated areas, where it washes out the information. In addition, specular spots are also observed on the metallic parts and resistors.

when a highlight pixel exceeds the maximum measurable light intensity of camera sensor. Klinker *et al.* [11] analyzed the color histogram and observed that the specular and diffuse components from a uniform surface form a skewed T shape. To separate these components, convex polygon fitting technique is used in their work. Linking color space with DIM [6] is also proposed [12, 13]. This color information is used to separate reflection components by fitting it into a dichromatic plane.

Transformation into other color spaces for detection of specularities is also a technique where the characteristics of an adequate color space are exploited. Schlüns and Teschner [14, 15] transformed the image from RGB to Y'U'V' color space. Bajcsy *et al.* [16] proposed S-space for analysis of variation in color of objects. There are three orthogonal basis functions in S space named S_0, S_1 and S_2. The S_0 basis function corresponds to specular reflection in S space. A data-driven color space called SUV color space was introduced by Mallick *et al.* [17]. Yang *et al.* [18] proposed Ch-CV space. This color space is spanned by maximum chromaticity (Ch) and the coefficient of variation (CV) in RGB. Yang *et al.* [19] proposed separation of specular and diffuse components in HSI color space as further improvement in their already proposed Ch-CV space. Recently, Akashia and Okatani [20] proposed an optimization technique for sparse non-negative matrix factorization for the identification of specular reflections in an image.

In spatial information based techniques, detection of specularities is performed through the use of local information in an image. Tan *et al.* [21] introduced Maximum Chromaticity-Intensity Space to differentiate between the maximum intensity and maximum chromaticity in an image. A pseudo-diffuse component image is created, which is later utilized for separation of specular reflection from the image. The ratio of intensities and colors amongst neighbouring pixels is preserved in the pseudo-diffuse image and is called the specular free image. The specular free image is obtained by Yoon *et al.* [22] through subtracting the minimum value for each channel from the input image. Shen and Cai [23] introduced a modified specular free image by adding an offset to the subtraction method provided in [22]. This offset can either be constant for the whole image [24] or can be varying for each pixel [23]. In [4], intensity logarithmic differentiation is used on both the specular-free and the input image for determination of diffuse pixels. Using this method, performance of the technique of creation of specularity free image is improved for highly textured surfaces. However, the position of highlights should be known for applying this method. Liu *et al.* [25] proposed the preservation of surface color saturation by initially producing a specular free image and then increasing the achromatic component of diffuse chromaticity. It is important to note that in all the specularity free image creation techniques, the input image is nor-

malized for illuminant. Hence, the illuminant should either be known or should be estimated first before applying those techniques.

Yang *et al.* [26, 27] treat the specular pixels as noise and use a bilateral filter for smoothing the maximum fraction of color components. In this way, the noise caused by specular pixels is eliminated. Kim *et al.* [28] observed that a diffuse pixel has low intensity in, at least, one channel. They called it the *"dark channel"* and proposed that the dark channel of an image contains no specular reflection. Their technique uses a maximum a-posteriori formulation that helps in the recovery of specular reflection and chromaticity. An *et al.* [29] proposed the pure diffuse pixel distribution model. This model is built on the assumption that there is at-least one purely diffuse pixel for each material in the scene. Shen and Zheng [30] assumed that some area of a surface contain only the diffuse reflection. They define the range value for an image as maximum minus minimum intensity for a surface and observed that the intensity ratio between maximum value of a pixel and range value is independent of the geometry of the surface. With the assumption for certain pixels to be diffuse, specular components from remaining pixels are computed through their proposed model.

Highlight detection techniques based on a single color image are practical as they do not require any additional hardware and data. However a problem associated with such algorithms is that they rely on image statistics and are based on strong prior assumptions. Therefore, such methods are not robust for change in imaging environment but works reasonably well when the required conditions are fulfilled.

6.3 Multiple images based techniques

The use of multiple images for separation of reflectance components from the scene is proposed in a number of studies. Since highlights are not intrinsic properties of an image, they can occur at any point and are dependent on the viewing angle. The direction of illumination also has its impact on the location of highlight spots. A surface area of an image which is affected by highlight in one image can appear as a diffuse surface if the viewing angle, or the illumination direction is changed. Based on this phenomenon, multiple images based highlight removal techniques are proposed in the literature, which require a number of images, captured using different imaging conditions.

Lee and Bajcsy [31] proposed the use of Lambertian consistency, which states that the Lambertian reflection does not vary in brightness and spectral contents with change in the viewing angle while the specular reflection changes the behaviour. They defined the spectral distance as euclidean distance between two colour points in a three-dimensional space. Minimum spectral distance is calculated to detect the

inconsistency in color among two images of same scene but with different viewing angle. Sato and Ikeuchi [32] introduced temporal-color space analysis by using a moving light source. Lin and Shum [33] used different illuminations for the same scene, and then proposed linear basis functions for separating diffuse and specular components. Lin *et al.* [34] used stereo images for the detection of specularities. Weiss [35] acquired an image sequence with varying specularities and used maximum likelihood estimation by assuming that the change in illumination gives rise to sparse filter outputs.

Feris *et al.* [36] used flash images taken with same point of view but different positions of flash for recovering the diffuse component. Agrawal *et al.* [37] proposed a method for image enhancement by using two images of a scene. One is taken with flash and one without flash. Reflection from the flash image is removed by using a gradient projection scheme.

Chen *et al.* [38] reconstructed the specular field by using histograms of the same image but with different intensities. Yang *et al.* [39] proposed statistical methods for removal of specularities from stereo images. They assume non-overlapping highlight regions in their method. Wang *et al.* [40] used three cameras for taking images of transparent plastic package containing tablets. They normalized the acquired images and then generated an image consisting of average intensities of corresponding pixels. In this way, pills are retained with higher intensity while varying regions of specular reflection are removed. Generation of specularity map from video sequence is proposed by Prinet *et al.* [41]. Recently, Wang *et al.* [42] proposed the use of light field imaging technology for capturing multiple views of a scene and then used that information for detection of specularities in the scene.

Nayar *et al.* [43] noticed that the specular reflection is highly polarized and proposed the use of polarization filter to separate the diffuse and specular reflections. They used a polarization filter by placing it in front of a camera and observed that by rotating the polarization filter, the brightness of diffuse materials is not changed, while the specular reflection is changed, since it is highly polarized, and varies following a cosine function. Polarization based methods are also proposed by Wolff [44, 45], Kim *et al.* [46], Atkinson and Hancock [47, 48], Müller [49], Umeyama and Godin [50], Lamond *et al.* [51] and Zhang *et al.* [52].

Although the above mentioned specularity detection techniques are able to show good performance, the major limitation associated with them is the availability of multiple images of the same scene with varying illumination direction or viewing angle. This limitation causes those techniques to become less practical compared to the approaches which are based on a single image.

6.4 Spectral image based techniques

Hyperspectral and multispectral imaging has been used extensively for remote sensing. Recently, with advancement in sensor technology, spectral imaging is widely available for imaging of objects at shorter distance. Such imaging comes with the problem of highlights, the same as in the case of color images. However, since spectral images contain more data compared to a color image, therefore highlight detection techniques are not exactly the same.

Bochko and Parkkinen [53] proposed probabilistic principal component analysis for detection of spectral and diffuse parts in images. Fu *et al.* [54] proposed orthogonal subspace projection (OSP) and dichromatic model for specularity-free representation of hyperspectral images. OSP results in projecting the radiance and illumination spectrum being orthogonal to each other in the subspace. In this way, the illuminant spectra is removed from mixed spectra and a spectral image without highlights is obtained. It is important to note that prior information about illuminant is required in this method. Koirala *et al.* [55] proposed spectral unmixing of end-members for separation of specular and diffuse components in spectral images. Spectral end-members consist of pure spectra corresponding to objects in a scene. Spectral unmixing method is widely used in hyperspectral remote sensing where the end-members correspond of pure spectra of land cover classes. In the method by Koirala *et al.* [55], Automated target generation program (ATGP) is employed for selection of end-members. When the scene illuminant is known, then the initial target detected by ATGP is the illuminant spectra. In case of unknown illuminant, a pixel with maximum value along the whole spectra is considered as the initial target. Using constrained energy minimization, the diffuse part of the image is identified.

6.5 Analysis of various specularity detection methods

In Table 6.1, we compare the characteristics of various specularity detection techniques. In these tables, different techniques are divided in general categories. The concept behind the technique is briefly defined along with the general assumptions being made. Use of DIM is common in the specularity detection techniques but there are some algorithms that do not use this model, so it is also given in the table. White balancing (WB) and image segmentation (Seg.) is also a constraint in such algorithms. We make analysis of various categories of highlight detection algorithms on the basis of the above expressed factors. Finally, general remarks about strength of those techniques is provided.

Table 6.1: Comparison of characteristics of highlight removal techniques

Category	Images	Concept	Assumptions	DIM use	WB Req.	Seg. Req.	Strength
Color Histogram Analysis [11][12][13]	Single	Skewed T-shape formation between specular and diffuse components	Lambertian body reflection, Interface reflection is function with a sharp peak around the angle of perfect mirror reflection.	Yes	Yes	Yes	Works well for dielectric materials but not for metals
Dichromatic Illumination based Model [6]	Single	Description of reflected light from a dielectric object as linear combination of object color and highlight	Single Illumination, existence of matte cluster	Yes	Yes	Yes	Works well for dielectric materials but not for metals
Color Space Transformation [14][15][16] [17][18][19]	Single	Segmentation of specular regions for obtaining the max. diffuse chromaticity in each segmented region	Single Illumination, each segmented cluster has uniform diffuse chromaticity	Yes	Yes	Yes	Iterative process can be time consuming,Colors are distorted in some algorithms. Overall good result on dielectric materials with single illumination.
Saturation preservation model [25]	Single	Creation of over-saturated diffuse reflectance and then putting the achromatic regions back to the diffuse reflection.	Pure white illuminant	Yes	Yes	No	Color saturation of surface is preserved in this method.
Sparse matrix factorization [20]	Single	Sparse non-negative matrix factorization for separation of specular components.	Single illuminant, presence of diffuse component for every surface	Yes	Yes	No	No assumption about spatial priors.
Pseudo diffuse image [3][4][21]	Single	Dark channel prior.	Pure white illuminant	Yes	Yes	Yes	Color ratio among neighbouring pixels is preserved.
Inpainting technique [7–10]	Single	Removal of highlight part through use of neighbouring pixels info. and inpainting	Single illuminant, presence of diffuse component for every surface	No	Yes	Yes	Preservation of color for dielectric surfaces.

Table 6.1 – *Continued from previous page*

Category	Images	Concept	Assumptions	DIM use	WB Req.	Seg. Req.	Strength
Multiple images acquisition [31–42]	Multiple	Highlights behave different when viewing angle or illumination is changed	Points which show specular reflection in one image can behave purely diffuse in another image when viewing condition is changed	Yes	___	___	Efficient detection of highlights when the required conditions and no. of images are available.
Polarization [43–52]	Multiple	Use of polarization filter during image acquisition.	Specular highlights are polarized while diffuse reflections are unpolarized	Yes	No	No	With use of additional hardware (polarization filter), highlights can be detected efficiently.
Orthogonal Subspace Projection [54]	Spectral	Separation of radiance and illumination spectrum through orthogonal subspace projection	SPD of Illuminant is known.	Yes	Yes	No	With knowledge of a particular illuminant, highlight caused by it can be removed.
PCA [53]	Spectral	Use of probabilistic PCA for separation of specular and diffuse components	Colored objects already segmented, Constant Hue.	Yes	No	Yes	Performs efficient highlight removal if segmentation is already performed.
Spectral unmixing [55]	Spectral	Spectral unmixing method	End-member of highlight part is SPD of illuminant.	Yes	No	No	With the use of a proper spectral segmentation method, highlights can be removed .

6.6 Conclusion

In this paper, we provide a survey of highlight detection algorithms in color and spectral images. In most of the literature, results from detection and removal of highlights is presented qualitatively while comparing with other methods. Although qualitative analysis provide a general overview, quantitative measurement should also be used to provide a fair comparison.

Generally, strong assumptions and priors are used for highlight detection. However, the highlight detection is still able to provide much of the useful information and is therefore a critical issue in image analysis. We also review state-of-the-art on highlight detection in spectral images. There is less work done on it and much is yet to be explored. The amount of spectral information being acquired in spectral imaging can certainly provide benefit for highlight detection and is still an open area of research.

Bibliography

[1] H. Barrow and J. Tanenbaum, "Recovering intrinsic scene characteristic from images," *Computer Vision System*, pp. 3–26, 1978.

[2] H. C. Lee, E. J. Breneman, and C. P. Schulte, "Modeling light reflection for computer color vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 402–409, Apr 1990.

[3] P. Tan, L. Quan, and S. Lin, "Separation of highlight reflections on textured surfaces," in *proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2 of *CVPR '06*, (Washington, DC, USA), pp. 1855–1860, IEEE Computer Society, 2006.

[4] R. T. Tan and K. Ikeuchi, "Separating reflection components of textured surfaces using a single image," in *Digitally Archiving Cultural Objects*, (Boston, MA), pp. 353–384, Springer US, 2008.

[5] A. Artusi, F. Banterle, and D. Chetverikov, "A survey of specularity removal methods," *Computer Graphics Forum*, vol. 30, no. 8, pp. 2208–2230, 2011.

[6] S. A. Shafer, "Using color to separate reflection components," *Color Research & Application*, vol. 10, no. 4, pp. 210–218, 1985.

[7] J. W. Park and K. H. Lee, "Inpainting highlights using color line projection," *IEICE-Transactions on Information and Systems*, vol. E90-D, pp. 250–257, Jan. 2007.

[8] Budianto and D. P. K. Lun, "Inpainting for fringe projection profilometry based on geometrically guided iterative regularization," *IEEE Transactions on Image Processing*, vol. 24, pp. 5531–5542, Dec 2015.

[9] F. Ortiz and F. Torres, "A new inpainting method for highlights elimination by colour morphology," in *proceedings of 3rd International Conference on Advances in Pattern Recognition*, pp. 368–376, Springer Berlin Heidelberg, 2005.

[10] P. Tan, S. Lin, L. Quan, and H.-Y. Shum, "Highlight removal by illumination-constrained inpainting," in *proceedings of 9th IEEE International Conference on Computer Vision*, vol. 1, pp. 164–169, Oct 2003.

[11] G. Klinker, S. Shafer, and T. Kanade, "Using a color reflection model to separate highlights from object color," in *proceedings of 1st International Conference on Computer Vision*, pp. 145–150, 1991.

[12] G. J. Klinker, S. A. Shafer, and T. Kanade, "The measurement of highlights in color images," *International Journal of Computer Vision*, vol. 2, no. 1, pp. 7–32, 1988.

[13] G. J. Klinker, S. A. Shafer, and T. Kanade, "A physical approach to color image understanding," *International Journal of Computer Vision*, vol. 4, no. 1, pp. 7–38, 1990.

[14] K. Schlüns and M. Teschner, "Analysis of 2D color spaces for highlight elimination in 3d shape reconstruction," in *proceedings of the Asian Conference on Computer Vision II*, pp. 801–805, 1995.

[15] K. Schlüns and M. Teschner, "Fast separation of reflection components and its application in 3D shape recovery," *Color and Imaging Conference*, no. 1, pp. 48–51, 1995.

[16] R. Bajcsy, S. W. Lee, and A. Leonardis, "Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation," *International Journal of Computer Vision*, vol. 17, no. 3, pp. 241–272, 1996.

[17] S. P. Mallick, T. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Specularity removal in images and videos: A PDE approach," in *proceedings of 9th European Conference on Computer Vision*, (Graz, Austria), pp. 550–563, Springer Berlin Heidelberg, 2006.

[18] J. Yang, Z. Cai, L. Wen, Z. Lei, G. Guo, and S. Z. Li, "A new projection space for separation of specular-diffuse reflection components in color images," in *proceedings of 11th Asian Conference on Computer Vision*, (Daejeon, Korea), pp. 418–429, Springer Berlin Heidelberg, 2013.

[19] J. Yang, L. Liu, and S. Z. Li, "Separating specular and diffuse reflection components in the HSI color space," in *IEEE International Conference on Computer Vision Workshops*, pp. 891–898, Dec 2013.

[20] Y. Akashi and T. Okatani, "Separation of reflection components by sparse non-negative matrix factorization," *Computer Vision and Image Understanding*, vol. 146, pp. 77–85, 2016.

[21] R. T. Tan and K. Ikeuchi, "Separating reflection components of textured surfaces using a single image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 178–193, Feb 2005.

[22] K. j. Yoon, Y. Choi, and I. S. Kweon, "Fast separation of reflection components using a specularity-invariant image representation," in *International Conference on Image Processing*, pp. 973–976, Oct 2006.

[23] H.-L. Shen and Q.-Y. Cai, "Simple and efficient method for specularity removal in an image," *Applied Optics*, vol. 48, pp. 2711–2719, May 2009.

[24] H.-L. Shen, H.-G. Zhang, S.-J. Shao, and J. H. Xin, "Chromaticity-based separation of reflection components in a single image," *Pattern Recognition*, vol. 41, no. 8, pp. 2461–2469, 2008.

[25] Y. Liu, Z. Yuan, N. Zheng, and Y. Wu, "Saturation-preserving specular reflection separation," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3725–3733, June 2015.

[26] Q. Yang, S. Wang, and N. Ahuja, "Real-time specular highlight removal using bilateral filtering," in *11th European Conference on Computer Vision*, (Heraklion, Greece), pp. 87–100, Springer Berlin Heidelberg, Sept. 2010.

[27] Q. Yang, J. Tang, and N. Ahuja, "Efficient and robust specular highlight removal," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, pp. 1304–1311, June 2015.

[28] H. Kim, H. Jin, S. Hadap, and I. Kweon, "Specular reflection separation using dark channel prior," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1460–1467, June 2013.

[29] J. Suo, D. An, X. Ji, H. Wang, and Q. Dai, "Fast and high quality highlight removal from a single image," *IEEE Transactions on Image Processing*, vol. 25, pp. 5441–5454, Nov 2016.

[30] H.-L. Shen and Z.-H. Zheng, "Real-time highlight removal using intensity ratio," *Applied Optics*, vol. 52, pp. 4483–4493, Jul 2013.

[31] S. W. Lee and R. Bajcsy, "Detection of specularity using colour and multiple views," *Image and Vision Computing*, vol. 10, no. 10, pp. 643–653, 1992.

[32] Y. Sato and K. Ikeuchi, "Temporal-color space analysis of reflection," *Journal of the Optical Society of America A*, vol. 11, pp. 2990–3002, Nov 1994.

[33] S. Lin and H.-Y. Shum, "Separation of diffuse and specular reflection in color images," in *proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*, vol. 1, pp. 341–346, 2001.

[34] S. Lin, Y. Li, S. B. Kang, X. Tong, and H.-Y. Shum, "Diffuse-specular separation and depth recovery from image sequences," in *proceedings of 7th European Conference on Computer Vision*, (Copenhagen, Denmark), pp. 210–224, Springer Berlin Heidelberg, May 2002.

[35] Y. Weiss, "Deriving intrinsic images from image sequences," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV*, vol. 2, pp. 68–75, 2001.

[36] R. Feris, R. Raskar, K.-H. Tan, and M. Turk, "Specular highlights detection and reduction with multi-flash photography," *Journal of the Brazilian Computer Society*, vol. 12, no. 1, pp. 35–42, 2006.

[37] A. Agrawal, R. Raskar, S. K. Nayar, and Y. Li, "Removing photography artifacts using gradient projection and flash-exposure sampling," *ACM Transactions on Graphics*, vol. 24, pp. 828–835, July 2005.

[38] T. Chen, M. Goesele, and H. P. Seidel, "Mesostructure from specularity," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 1825–1832, 2006.

[39] Q. Yang, S. Wang, N. Ahuja, and R. Yang, "A uniform framework for estimating illumination chromaticity, correspondence, and specular reflection," *IEEE Transactions on Image Processing*, vol. 20, pp. 53–63, Jan 2011.

[40] C. Wang, S. I. Kamata, and L. Ma, "A fast multi-view based specular removal approach for pill extraction," in *IEEE International Conference on Image Processing*, pp. 4126–4130, Sept 2013.

[41] V. Prinet, M. Werman, and D. Lischinski, "Specular highlight enhancement from video sequences," in *IEEE International Conference on Image Processing*, pp. 558–562, Sept 2013.

[42] H. Wang, C. Xu, X. Wang, Y. Zhang, and B. Peng, "Light field imaging based accurate image specular highlight removal," *PLOS ONE*, vol. 11, pp. 1–17, 06 2016.

[43] S. K. Nayar, X.-S. Fang, and T. Boult, "Separation of reflection components using color and polarization," *International Journal of Computer Vision*, vol. 21, no. 3, pp. 163–186, 1997.

[44] L. B. Wolff, "Classification of material surfaces using the polarization of specular highlights," in *proceedings of SPIE - The International Society for Optical Engineering*, vol. 1005, pp. 206–213, Jan. 1988.

[45] L. B. Wolff, "Polarization-based material classification from specular reflection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 1059–1071, Nov 1990.

[46] D. W. Kim, S. Lin, K.-S. Hong, and H. Shum, "Variational specular separation using color and polarization," in *proceedings of the IAPR Workshop on Machine Vision Applications*, 2002.

[47] G. A. Atkinson and E. R. Hancock, "Recovery of surface orientation from diffuse polarization," *IEEE Transactions on Image Processing*, vol. 15, pp. 1653–1664, June 2006.

[48] G. A. Atkinson and E. R. Hancock, "Two-dimensional BRDF estimation from polarisation," *Computer Vision and Image Understanding*, vol. 111, no. 2, pp. 126–141, 2008.

[49] V. Müller, "Polarization-based separation of diffuse and specular surface-reflection," in *Verstehen akustischer und visueller Informationen*, pp. 202–209, Springer Berlin Heidelberg, 1995.

[50] S. Umeyama and G. Godin, "Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 639–647, May 2004.

[51] B. Lamond, P. Peers, and P. Debevec, "Fast image-based separation of diffuse and specular reflections," in *ACM SIGGRAPH Sketches*, SIGGRAPH '07, (New York, NY, USA), ACM, 2007.

[52] L. Zhang, E. R. Hancock, and G. A. Atkinson, "Reflection component separation using statistical analysis and polarisation," in *proceedings of 5th Iberian Conference, IbPRIA*, (Las Palmas de Gran Canaria, Spain), pp. 476–483, Springer Berlin Heidelberg, 2011.

[53] V. Bochko and J. Parkkinen, "Highlight analysis using a mixture model of probabilistic PCA," in *proceedings of the 4th WSEAS International Conference on Signal Processing, Robotics and Automation*, ISPRA'05, (Stevens Point, Wisconsin, USA), pp. 15:1–15:5, World Scientific and Engineering Academy and Society (WSEAS), 2005.

[54] Z. Fu, R. T. Tan, and T. Caelli, "Specular free spectral imaging using orthogonal subspace projection," in *18th International Conference on Pattern Recognition (ICPR)*, vol. 1, pp. 812–815, 2006.

[55] P. Koirala, P. Pant, M. Hauta-Kasari, and J. Parkkinen, "Highlight detection and removal from spectral image," *Journal of the Optical Society of America A*, vol. 28, pp. 2284–2291, Nov 2011.

Chapter 7

Towards highlight based illuminant estimation in multispectral images

Article D

This chapter is a reformatted reprint of the publication

H. A. Khan, J. B. Thomas and J. Y. Hardeberg, "Towards highlight based illuminant estimation in multispectral images," to appear in proceedings of *International Conference on Image and Signal Processing*, Cherbourg, France, 2018.

We review the physics based illuminant estimation methods, which extract information from highlights in images. Such highlights are caused by specular reflection from the surface of dielectric materials, and according to the dichromatic reflection model, provide cues about the illumination. This paper analyzes different categories of highlight based illuminant estimation techniques for color images from the point of view of their extension to multispectral imaging. We find that the use of chromaticity space for multispectral imaging is not straightforward and imposing constraints on illuminants in the multispectral imaging domain may not be efficient either. We identify some methods that are feasible for extension to multispectral imaging, and discuss the advantage of using highlight information for illuminant estimation.

7.1 Introduction

Color constancy is the ability of the human visual system to perceive constant color of surfaces despite the change in spectral composition of the illuminant. Creating such capability for a computer vision system is called "computational color constancy", and it can be achieved through two major techniques. One method is to use the statistical properties of the scene, while the other is to utilize the physics-based analysis of surfaces. Statistics-based methods can provide fairly accurate results when the assumptions are satisfied. One such common assumption is the presence of a variety of surfaces and objects in a scene. Generally, these methods work poorly in images with low color complexity [1].

In physics based illuminant estimation approaches, the analysis of interaction of light with surfaces is performed for estimating the scene illuminant. The optically inhomogenous objects (e.g. plastics, ceramics, paints, etc) create a neutral interface reflection [2] condition, where the spectral power distribution of the specular reflection is the same as the incident illumination; and this property is used to estimate the scene illuminant. The basis of illuminant estimation from specular highlights in images is the dichromatic reflection model [3], which describes the reflected light from an object as a combination of the diffuse and specular reflection. The specular reflection is also termed as highlights in an image, and therefore, the problem of illuminant estimation is treated as the analysis of highlights from surfaces [4, 5].

With advancement in sensor technology, use of multispectral and hyperspectral imaging is increasing for close range imaging applications [6]. Those images also contain diffuse and specular reflections which can be helpful for estimation of scene illuminant. The information of scene illumination is helpful in achieving multispectral constancy [7, 8]. In this paper, we discuss different categories of highlight based illuminant estimation methods for color images and their possible extension to spectral images. This paper is organized as follows; In Section 7.2, dichromatic reflection model and its use for illuminant estimation in images is discussed. In Section 7.3, different categories of work on illuminant estimation from highlights in color images are analyzed from the point of view of their possible extension towards multispectral imaging. In Section 7.4, use of highlights in multispectral imaging and the existing works dedicated to spectral imaging is discussed, followed by conclusion in Section 7.5.

7.2 Dichromatic reflection model

In a simplified color imaging model, the image values **f(x)** (where **x** represents spatial coordinates) are dependent on the light source $e(\lambda)$, the surface albedo

$s(\mathbf{x}, \lambda)$ and the camera sensitivity function $\mathbf{c}(\lambda)$. It is represented as:

$$\mathbf{f(x)} = \int_{\omega} e(\lambda)s(\mathbf{x}, \lambda)\mathbf{c}(\lambda)d\lambda. \tag{7.1}$$

Shafer [3] proposed to decompose the light into its diffuse $m^b(\mathbf{x})$ and specular $m^s(\mathbf{x})$ components:

$$\mathbf{f(x)} = m^b(\mathbf{x}) \int_{\omega} e(\lambda)s(\mathbf{x}, \lambda)\mathbf{c}(\lambda)d\lambda + m^s(\mathbf{x}) \int_{\omega} e(\lambda)\mathbf{c}(\lambda)d\lambda \tag{7.2}$$

In the dichromatic reflection model, the diffuse reflection is of low intensity and is Lambertian, while the specular reflection is generally of high intensity, directional, independent from the surface albedo and dependent on the viewing direction. The illuminant in the sensor domain $\mathbf{e}$ is defined as:

$$\mathbf{e} = \int_{\omega} e(\lambda)\mathbf{c}(\lambda)d\lambda \tag{7.3}$$

Eq. 7.3 is the second part of Eq. 7.2. This relation states that the specular reflection part of Eq. 7.2 provides information about the illuminant. For a camera with $\mathbf{c}(\lambda) = \{R(\lambda), G(\lambda), B(\lambda)\}$, this linear combination defines a plane in the RGB color space. The data points from the RGB color space are projected onto a space normalized by $r + g + b = 1$, where they form a line. The two points defining the line consist of the diffuse and specular points on a surface. Data points from a uniformly colored surface are distributed along this line which is called the *dichromatic line*. Assuming uniform illumination in a scene, all the dichromatic lines have one point in common, which corresponds to the scene illuminant. Therefore, if there are two or more materials being illuminated by the same light source and their dichromatic planes are identified, the vector containing illuminant information can be estimated by the intersection of these planes. There are some methods that do not explicitly use the dichromatic reflection model, but still use highlights for illuminant estimation. In the following section, we discuss different categories of highlight based illuminant estimation models along with an analysis of their potential extension towards multispectral imaging.

7.3 Inversion of dichromatic reflection model

In this section, we analyze the methods based on the dichromatic reflection for illuminant estimation and present empirical observations for their extension towards multispectral imaging.

7.3.1 Illuminant estimation in two-dimensional chromaticity space

Using two-dimensional chromaticity space allows the removal of intensity component and simplifies the illuminant estimation problem in color images. One of

the earlier attempts in highlight based illuminant estimation are by Lee [9], and D'Zmura and Lennie [10], who proposed the use of color coordinates of specular highlights for estimation of illumination. Lehmann and Palm [11] extended Lee's red diagram [9] to the *rg-diagram* and called it "color line search". Recently, Uchimi *et al.* [12] proposed estimation of illumination from $u'v'$ chromaticity, xy chromaticity space is used in [13], and the rg color space in [14, 15].

One way to think about extension of dichromatic reflection model for multispectral imaging is to extend the results obtained in two dimensional chromaticity space. For example, if estimation is performed in rg space, the third component can be calculated directly as $b = 1 - (r + g)$. However, this approach may not lead to an accurate three dimensional estimation because the errors of estimation in each component are accumulated in the third component [15]. If we imagine the same procedure for multispectral images, then there are two major issues. The first is the choice of base components or channels for creation of dichromatic lines. It is not intuitive to call any combination of two channels from multispectral channels as a chromaticity space. The second problem is that a space satisfying the condition $c_1 + c_2 + ... + c_N = 1$, where c_i is the i^{th} channel, leads to very small values. This causes instability in any numerical computation.

Lakehal and Ziou [15] proposed that to minimize the error in illuminant estimation while extending the two dimensional estimate into three dimensions, the chromatic coordinates in the rg and gb sub-spaces can be used. For each sub-space, the intersection of dichromatic lines is computed and the error for each coordinate of both sub-spaces are combined to form an objective function. They proposed to minimize this function individually for each R, G and B channel. That method shows promising results for color images, but if we want to extend it to multispectral imaging then we have to create at least $(N - 1)$ sub-spaces, where N is the total number of channels in a multispectral camera. It might be possible to create a reduced number of sub-spaces but it is not obvious how the spectral correlation may permit the reduction of number of planes. Even if somehow the dichromatic lines and their intersections are computed for each subspace, projecting them to N dimensions will be computationally expensive and unstable because of being very sensitive to errors in estimation in any sub-space. In the presence of such issues, extending these methods is not expected to bring promising results.

7.3.2 Constraint on possible illuminants

For reducing the estimation error in a sub-space, some authors propose to impose constraints on the set of possible illuminants. Some algorithms [16–19] aim at finding the chromaticity of an illuminant by intersecting a dichromatic line with the Planckian locus, or a set of known illuminants. The basic idea of such meth-

ods is to calculate the projected area for a set of candidate illuminants along the Planckian locus and pick the candidate that gives the minimum error in terms of distance, as an estimate of the scene illuminant. Toro and Funt [20] assumed the presence of a fixed number of different materials in an image patch. This method requires a discrete set of candidate illuminants. Each of the illuminants from the given set is tested sequentially, and the illumination is estimated by solving a set of simultaneous linear equations using a veronese projection of multilinear constraints. Toro [21] proposed optimizing a cost function on the color of illuminant. The scene illuminant is found by fitting the color of candidate illuminants to the image colors. Shi and Funt [22] proposed two Hough Transforms for computation of a histogram that represents the likelihood that a candidate intersection line is the image illumination axis. In these methods, illumination estimation is obtained from a set of candidate illuminations.

Here we discuss the feasibility of extending the algorithms based on finding the intersection of dichromatic lines along the Planckian locus. Using this technique for color images, the possibilities are to find either none, one or two points of intersection. In case of no intersection, the closest point between the line and the curve of Planckian locus is identified as the scene illuminant. If there are two intersections, then one of them is selected as the correct intersection on the basis of some heuristics. Therefore, in theory it is possible that the constraint of approximating the illuminants to the curve of the black-body radiator may help in determination of scene illuminant from a single surface in color images. However, it was found in [23] that approximating the illuminants by the curve of black body radiator is not advantageous for illuminant estimation in color images. Also, the Planckian locus needs to be defined for the multispectral imaging domain. Constraining the set of possible illuminants also have the same issues of computational complexity and instability as discussed in Section 7.3.1, and therefore, it does not guarantee better results. Exhaustive quantitative investigations would need to be performed for evaluation of the complexity, assumptions and results for the extension of such algorithms as their extension to multispectral imaging is not straightforward.

7.3.3 Inverse intensity chromaticity space

The inverse intensity chromaticity space (IICS) was introduced by Tan *et al.* [24]. It is spanned by the chromaticity of a single channel and the inverse of image intensity. A linear correlation between the image chromaticity and illumination chromaticity is formed in the IICS. This correlation allowed the estimation of scene illumination without segmenting the color beneath the highlights. This method has the potential of extension for multispectral imaging, since it deals with each channel separately. One of the limitations of this algorithm is the detection of highlight regions, which can be improved by selecting a suitable method for multispectral

imaging [25]. Riess *et al.* [26] proposed local estimates of scene illuminant by decomposing the image into mini-regions and then using those estimates for computing the dominant illuminant of the scene by using the inverse intensity chromaticity space. The proposed method is an extension of [24] and also has the potential for extension to higher dimensional multispectral data.

7.3.4 Use of polarizing filter

The use of polarizing filter for obtaining specular only image is proposed in [27, 28]. In [27], knowledge of geometry of surfaces is required and its extension for multispectral imaging is not obvious. However, Badawi *et al.* [28] proposed the use of spatially constrained independent component analysis (ICA) and the mutual information-based ICA for estimation of scene illumination. This method has the potential to be extended for multispectral imaging. It is also worth noting that the proposed method in [28] is tested on hyperspectral images after thresholding the bright pixels and promising results are obtained.

7.4 Illuminant information from multispectral images

7.4.1 Statistics based methods

The statistics based methods for illuminant estimation in color images are extended by Khan *et al.* [29] for multispectral images and they found that algorithms based on bright pixels (Max-RGB) and edge information (Gray-Edge) provide good results. Use of highlights may further improve those methods. We identified the method by Gijsenij *et al.* [30] who proposed the use of specular and shadow edges in an image for estimation of scene illuminant. In their approach, initially the illuminant estimation is performed in the same way as in the gray-edge algorithm [31]. The obtained estimate is improved by iteratively removing the color cast and assigning weights to the edges. The illuminant is estimated again by using the updated weights until a stable estimate of scene illuminant is found. Use of edge information suggests that there is some sheen reflection around the edges and this can be a valuable cue for the estimation of scene illuminant. It can be anticipated that extension of [30] to multispectral imaging can also bring promising results.

Drew *et al.* [32] proposed to estimate the scene illuminant by taking the mean of specular pixels or pixels from near-white materials. In their method, a planar constraint is imposed which assumes that the log-relative-chromaticity values for near-specular pixels are orthogonal to the illuminant chromaticity. They defined zeta-image which is formed by taking dot product of the log-relative-chromaticity and the light direction. Their method works well when there are strong specularities in the scene but the performance is significantly reduced when the specular

reflection from colored surfaces is low. However, when their proposed planar constraint is used along with gray-world and gray-edge algorithms, the performance of illuminant estimation is increased. It will be interesting to evaluate the effect of imposing planar constraint and creating the so-called zeta images from multispectral data, for estimation of scene illumination.

7.4.2 Low-rank matrix factorization of multispectral data

There are some methods for highlights based illuminant estimation in multispectral imaging that use the factorization property of data matrix. An *et al.* [33] proposed illuminant estimation from highlight in multispectral images by assuming that the chromaticity of specular components is uniform over all the pixels. They also assumed that the surfaces with specular reflection have their diffuse counterparts in the specular free regions. With this assumption, it is suggested that the diffuse component of a specular pixel can be reconstructed from another diffuse pixel containing the same material and the correlation between the spectra of specular and diffuse reflection is low. By assuming a single light source, or multiple light sources with normalized spectrum, a low-rank matrix is formed by the specular pixels and is used to estimate the scene illuminant. The proposed method in [33] requires more than 4 channels for its implementation. Low-rank matrix factorization is further analyzed by Zheng *et al.* [34] for signal separation. This method works well even with non-uniform illuminations but the performance is degraded in presence of multiple illuminants or shadows. Their work is improved by Chen *et al.* [35] by introducing a conditional random field model for handling scenes with multiple illuminations and shadows.

7.4.3 Use of Singular Value Decomposition

Tominaga and Wandell [36] observed that color planes from different objects intersect in a certain line, which gives the information of scene illumination. They computed the color planes of reflected light through singular value decomposition (SVD). This method is used for illuminant estimation in multispectral imaging by Huynh and Robles-Kelly [37]. They assumed that a scene is uniformly illuminated and used the uniform-albedo patches for illuminant estimation. Using this technique, satisfying performance is obtained for monochromatic patches. However, the pre-processing of images before applying this method is the key factor which can significantly degrade the performance if not done properly. Imai *et al.* [38] proposed that the surface containing specular reflection is two dimensional. They projected the segmented high dimensional spectral data from the specular reflection, onto a two-dimensional space which is spanned by the first two principal components. Within this space, the pixel distribution is divided into two clusters and it is assumed that the specular highlight cluster is coincident to the

directional vector of the illuminant. Correct identification of highlights is crucial for the performance of their proposed method. This method is extended to multiple illuminants in [39, 40].

The SVD and principal components based illuminant estimation methods are already demonstrated to perform well for multispectral imaging but they are dependent on the correct segmentation of highlights region in image. A combination of highlight detection algorithm in multispectral imaging and the proposed algorithms in this section can provide promising results for illuminant estimation in multispectral imaging.

7.5 Conclusion

In this paper, we reviewed the illuminant estimation methods which are based on highlights in images and identified the potential methods which can be extended to multispectral imaging. Based on this qualitative analysis, the dichromatic reflection model based methods that work in two dimensional chromaticity space can be computationally complex and unstable for extension to multispectral imaging. We identified some techniques that can be extended for multispectral imaging. Quantitative analysis of such an extension is also required before reaching to a final conclusion. By combining the current work with state of the art highlight detection algorithms [25], the reader can pick the algorithms which best suit the type of images and data to be processed for illuminant estimation.

Bibliography

[1] G. D. Finlayson, S. D. Hordley, and P. M. Hubel, "Color by correlation: a simple, unifying framework for color constancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1209–1221, Nov 2001.

[2] H. C. Lee, E. J. Breneman, and C. P. Schulte, "Modeling light reflection for computer color vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 402–409, Apr 1990.

[3] S. A. Shafer, "Using color to separate reflection components," *Color Research & Application*, vol. 10, no. 4, pp. 210–218, 1985.

[4] G. J. Klinker, S. A. Shafer, and T. Kanade, "The measurement of highlights in color images," *International Journal of Computer Vision*, vol. 2, pp. 7–32, Jun 1988.

[5] S. Lin and H.-Y. Shum, "Separation of diffuse and specular reflection in color images," in *Conf. on Computer Vision and Pattern Recognition*, pp. 341–346, 2001.

[6] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton, "Multispectral filter arrays: Recent advances and practical implementation," *Sensors*, vol. 14, no. 11, pp. 21626–21659, 2014.

[7] H. A. Khan, J. B. Thomas, and J. Y. Hardeberg, "Multispectral constancy based on spectral adaptation transform," in *20th Scandinavian Conf. on Image Analysis*, pp. 459–470, June 2017.

[8] H. A. Khan, J.-B. Thomas, J. Y. Hardeberg, and O. Laligant, "Spectral adaptation transform for multispectral constancy," *Journal of Imaging Science and Technology*, vol. 62, no. 2, pp. 1020504–1– 020504–12, 2018.

[9] H.-C. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," *J. Opt. Soc. Am. A*, vol. 3, pp. 1694–1699, Oct 1986.

[10] M. D'Zmura and P. Lennie, "Mechanisms of color constancy," *J. Opt. Soc. Am. A*, vol. 3, pp. 1662–1672, Oct 1986.

[11] T. M. Lehmann and C. Palm, "Color line search for illuminant estimation in real-world scenes," *J. Opt. Soc. Am. A*, vol. 18, pp. 2679–2691, Nov 2001.

[12] Y. Uchimi, T. Jinno, and S. Kuriyama, "Estimation of multiple illuminant colors using color lines of single image," in *Int. Conf. on Adv. Informatics, Concepts, Theory, and Applications*, pp. 1–6, Aug 2017.

[13] J.-Y. Kim, Y.-S. Seo, and Y.-H. Ha, "Estimation of illuminant chromaticity from single color image using perceived illumination and highlight," *Journal of Imaging Science and Technology*, vol. 45, no. 3, pp. 274–282, 2001.

[14] O.-S. Kwon, Y.-H. Cho, Y.-T. Kim, and Y.-H. Ha, "Illumination estimation based on valid pixel selection in highlight region," in *IEEE International Conf. on Image Processing*, Oct 2004.

[15] E. Lakehal and D. Ziou, "Computational color constancy from maximal projections mean assumption," *Multimedia Tools and Applications*, Dec 2017.

[16] G. D. Finlayson and G. Schaefer, "Solving for colour constancy using a constrained dichromatic reflection model," *International Journal of Computer Vision*, vol. 42, pp. 127–144, May 2001.

[17] G. Schaefer, "Robust dichromatic colour constancy," in *Image Analysis and Recognition*, pp. 257–264, Springer Berlin Heidelberg, 2004.

[18] Y.-Y. Li and H.-C. Lee, "Auto white balance by surface reflection decomposition," *J. Opt. Soc. Am. A*, vol. 34, pp. 1800–1809, Oct 2017.

[19] B. Mazin, J. Delon, and Y. Gousseau, "Illuminant estimation from projections on the planckian locus," in *European Conf. on Computer Vision*, pp. 370–379, 2012.

[20] J. Toro and B. Funt, "A multilinear constraint on dichromatic planes for illumination estimation," *IEEE Transactions on Image Processing*, vol. 16, pp. 92–97, Jan 2007.

[21] J. Toro, "Dichromatic illumination estimation without pre-segmentation," *Pattern Recognition Letters*, vol. 29, no. 7, pp. 871 – 877, 2008.

[22] L. Shi and B. Funt, "Dichromatic illumination estimation via Hough transforms in 3D," in *Conf. on Color in Graphics, Imaging, & Vision*, pp. 259–262, 2008.

[23] M. Ebner and C. Herrmann, "On determining the color of the illuminant using the dichromatic reflection model," in *27th Pattern Recognition Symposium DAGM*, pp. 1–8, Springer Berlin Heidelberg, 2005.

[24] R. T. Tan, K. Nishino, and K. Ikeuchi, "Color constancy through inverse-intensity chromaticity space," *J. Opt. Soc. Am. A*, vol. 21, pp. 321–334, Mar 2004.

[25] H. A. Khan, J.-B. Thomas, and J. Y. Hardeberg, "Analytical survey of highlight detection in color and spectral images," in *Computational Color Imaging Workshop*, pp. 197–208, 2017.

[26] C. Riess, E. Eibenberger, and E. Angelopoulou, "Illuminant color estimation for real-world mixed-illuminant scenes," in *International Conference on Computer Vision*, pp. 782–789, Nov 2011.

[27] K. Hara and K. Nishino, "Variational estimation of inhomogeneous specular reflectance and illumination from a single view," *J. Opt. Soc. Am. A*, vol. 28, pp. 136–146, Feb 2011.

[28] W. K. M. Badawi, C. C. Chibelushi, M. N. Patwary, and M. Moniri, "Specular-based illumination estimation using blind signal separation techniques," *IET Image Processing*, vol. 6, pp. 1181–1191, Nov 2012.

[29] H. A. Khan, J.-B. Thomas, J. Y. Hardeberg, and O. Laligant, "Illuminant estimation in multispectral imaging," *J. Opt. Soc. Am. A*, vol. 34, pp. 1085–1098, 2017.

[30] A. Gijsenij, T. Gevers, and J. van de Weijer, "Improving color constancy by photometric edge weighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, pp. 918–929, May 2012.

[31] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Transactions on Image Processing*, vol. 16, pp. 2207–2214, Sept 2007.

[32] M. S. Drew, H. R. V. Joze, and G. D. Finlayson, "The zeta-image, illuminant estimation, and specularity manipulation," *Computer Vision and Image Understanding*, vol. 127, pp. 1 – 13, 2014.

[33] D. An, J. Suo, H. Wang, and Q. Dai, "Illumination estimation from specular highlight in a multispectral image," *Optics Express*, vol. 23, pp. 17008–17023, Jun 2015.

[34] Y. Zheng, I. Sato, and Y. Sato, "Illumination and reflectance spectra separation of a hyperspectral image meets low-rank matrix factorization," in *Conf. on Computer Vision and Pattern Recognition*, pp. 1779–1787, June 2015.

[35] X. Chen, M. S. Drew, and Z.-N. Li, "Illumination and reflectance spectra separation of hyperspectral image data under multiple illumination conditions," *Electronic Imaging*, no. 18, pp. 194–199, 2017.

[36] S. Tominaga and B. A. Wandell, "Standard surface-reflectance model and illuminant estimation," *J. Opt. Soc. Am. A*, vol. 6, pp. 576–584, Apr 1989.

[37] C. P. Huynh and A. Robles-Kelly, "A solution of the dichromatic model for multispectral photometric invariance," *International Journal of Computer Vision*, vol. 90, pp. 1–27, Oct 2010.

[38] Y. Imai, Y. Kato, H. Kadoi, T. Horiuchi, and S. Tominaga, "Estimation of multiple illuminants based on specular highlight detection," in *Computational Color Imaging Workshop*, pp. 85–98, 2011.

[39] N. T. D. Hang, T. Horiuchi, K. Hirai, and S. Tominaga, "Estimation of two illuminant spectral power distributions from highlights of overlapping illuminants," in *Signal-Image Technology & Internet-Based Systems*, pp. 434–440, Dec 2013.

[40] Y. Kato, T. Horiuchi, and S. Tominaga, "Estimation of multiple light sources from specular highlights," in *International Conference on Pattern Recognition*, pp. 2033–2086, Nov 2012.

Chapter 8

Color characterization methods for a multispectral camera

Article E

This chapter is a reformatted reprint of the publication

H. A. Khan, and P. Green, "Color characterization methods for a multispectral camera," in *Proceedings of IS&T International Symposium on Electronic Imaging: Color Imaging XXIII: Displaying, Processing, Hardcopy, and Applications*, San Francisco, USA, 2018.

With advances in sensor technology, the availability of multispectral cameras and their use are increasing. Having more information compared to a three-channel camera has its advantages but the data must be handled appropriately. In this work, we are interested in multispectral camera characterization. We measure the camera characterization performance by two methods, by linear mapping and through spectral reconstruction. Linear mapping is used in 3-channel camera characterization and we use the same method for a multispectral camera. We also investigate whether instead of linear mapping, spectral reconstruction from the camera data improves the performance of color reproduction. The recovery of reflectance spectra is an underdetermined problem and certain assumptions are required for obtaining a unique solution. Linear methods are generally used for spectral reconstruction from the camera data and are based on training on known spectra. These methods can perform well when the test data

consists of a subset of the training spectra, however, their performance is significantly reduced when the test data is different. In this paper, we also investigate the role of training spectra for camera characterization. Five different spectral reflectance datasets are used for training the camera characterization models. Finally we provide a comparison between the linear mapping and spectral reconstruction methods for multispectral camera characterization and also test the camera characterization framework on hyperspectral images of natural scenes.

8.1 Introduction

A common procedure for obtaining device independent color information from a digital camera is to characterize the camera in terms of CIE tristimulus values [1–6]. Ideally, the spectral sensitivities of a digital camera with three color channels would satisfy the Luther condition, in which it is assumed that they are linear transformation of the CIE color matching function [7]. However, real cameras do not follow the Luther condition because of manufacturing constraints and limitations in design. Because of this issue, device metamerism may occur, where two surfaces with different spectral behavior produce identical camera response when imaged under the same illumination and hence provide the same XYZ values. Such metamerism limits the colorimetric application of the imaging device.

With advances in sensor technology, the use of multispectral imaging for indoor scene acquisition under controlled conditions have also increased. The advantage of multispectral imaging is the ability to acquire more spectral information from a scene, which can be used for spectral reflectance reconstruction [8, 9] of the object's surfaces. Most of the natural objects and surfaces have smooth spectral properties which are helpful for multispectral imaging [10]. With an effective spectral reconstruction from the multispectral data, the device can be used as imaging spectrophotometer and can enable the measurement of device-independent and illuminant-independent image data [11]. Use of multispectral imaging techniques have been proposed for the characterization of imaging devices [5, 12, 13]. The use of digital three channel color cameras as colorimetric devices has been investigated extensively in last decades [14], while multispectral cameras are mostly used for extraction of spectral information. We are interested to also use the multispectral camera as a colorimetric device and obtain the colorimetric information through camera characterization.

Camera characterization involves finding a transform between the camera input space and CIE XYZ color space. Ideally, this transform should be linear and invertible. Such a transform can be found by imaging the surfaces with known

spectral reflectance and colorimetric values.

One method for device characterization is to recover the spectral properties of the surfaces being imaged and then compute the tristimulus values from the reconstructed spectral reflectances [5]. There are many methods proposed in the literature for spectral reconstruction. Low dimensional linear models are used in spectral reconstruction from multispectral data with the assumption of smoothness of spectra [8]. The efficiency of a spectral reconstruction system is dependent on the training data and performs well if the test data consist of similar spectral properties. In this work, we investigate the role of training dataset for spectral reconstruction by using non-linear methods. The rest of paper is organized as follows; In Section 2, we provide a brief review of the multispectral imaging system and spectral reconstruction. In Section 3, we explain the experimental framework, discussion of results in Section 4, and followed by conclusions and future aspects.

8.2 Multispectral imaging system

In a simplified noiseless imaging model, formation of an image depends on the spectral sensitivity of the imaging sensor $m(\lambda)$, the spectral reflectance of the surface $r(\lambda)$ and the spectral power distribution of the illuminant $e(\lambda)$. This formation for the visible wavelength spectrum ω is defined as

$$f = \int_{\omega} r(\lambda)e(\lambda)m(\lambda)d\lambda \tag{8.1}$$

In practice, we can formulate an extended and discrete version of Eq. 8.1 as $\mathbf{F} = \mathbf{REM}$. Considering the spectral sampling of 10nm within the wavelength spectrum of 400nm to 720nm and K number of spectral channels, $\mathbf{F}$ is $S \times N$ matrix (S is the number of spectral samples being acquired across N wavelengths), $\mathbf{R}$ is $S \times N$ matrix of surface reflectance, $\mathbf{E}$ is the diagonal matrix ($N \times N$) of the scene illuminant and $\mathbf{M}$ is $N \times K$ matrix, consisting of the spectral sensitivities of the channels.

Spectral reconstruction aims at estimating $\mathbf{R}$ from $\mathbf{F}$. If the term $\mathbf{EM}$ is known and invertible, then the solution is a simple inverse as $\hat{\mathbf{R}} = \mathbf{WF}$, where $\mathbf{W} = \mathbf{E}^{-1}\mathbf{M}^{-1}$ and $\hat{\mathbf{R}}$ is the estimated spectra. Such a transformation does not produce efficient results for spectral reconstruction and the calibration matrix $\mathbf{W}$ is created through a set of measured training spectra. There are many methods for creating $\mathbf{W}$ by using linear methods. One of the methods is through the pseudo inverse, proposed by Maloney and Wandell [10], as;

$$\mathbf{W} = \mathbf{R}_t\mathbf{F}_t^T(\mathbf{F}_t\mathbf{F}_t^T)^{-1} \tag{8.2}$$

Here the subscript t represents training data, while the superscript T represents the transpose of a matrix.

In most of the spectral reconstruction work, a linear mapping is considered sufficient. There are many methods for determining a linear transform. The idea for using such a transformation is that a few basis functions are helpful in representing the spectral reflectance [10, 15]. To determine the basis functions, Principal Component Analysis (PCA) and Karhunen-Loeve transformation is mostly used over a set of measured spectral samples. Imai and Berns [16] proposed the use of principal component analysis over the training data, before using it in Eq. 8.2. In the pseudo inverse methods proposed in [10, 16], the dimensionality of the reflectance model is equal to or less than the number of camera channels [17]. Shi and Healey [13] proposed the use of the first three basis functions separately from the remaining basis functions and trained the calibration matrix $\mathbf{W}$. Wiener estimation [18] is also an efficient linear method for spectral reconstruction. It is defined as;

$$\mathbf{W} = \mathbf{R}_t\mathbf{R}_t^T(\mathbf{ME})^T((\mathbf{ME})\mathbf{R}_t\mathbf{R}_t^T(\mathbf{ME})^T + \mathbf{G})^{-1} \tag{8.3}$$

Here, $\mathbf{R}_t\mathbf{R}_t^T$ and $\mathbf{G}$ are the autocorrelation matrices of training spectra and additive noise, respectively. $\mathbf{G}$ is in the form of a diagonal matrix consisting of the variance of noise σ^2.

Linear mapping method is computationally efficient and has been used successfully for color camera characterization. As in the analysis provided at [19], three sets of training data were used separately in the training phase and the characterization error is measured after transformation of input data. The three spectral datasets being used for training were ISO-17321-2 [20], Image Engineering dataset and GretagMacbeth ColorChecker. The ISO 17321 data was used to test the characterization. As expected, the error is smallest when the representative data for the test scene was used in the training phase. This suggests that domain specific training data should be used for specific applications. We use the same concept for characterization of a multispectral camera and train the system with different reflectance datasets for comparison of performance and effect of training dataset. If $\mathbf{X}$ is the CIE XYZ representation of a known spectra $\mathbf{R}$, then the image $\mathbf{F}$ can be transformed into its corresponding colorimetric values through at transformation matrix $\mathbf{W}_x$ as;

$$\mathbf{X} = \mathbf{W}_x\mathbf{F} \tag{8.4}$$

By using known reflectance spectra and their corresponding colorimetric values, Eq. 8.4 can be used to find the transformation matrix $\mathbf{W}_x$ through the pseudo inverse of $\mathbf{F}$ as;

$$\mathbf{W}_x = \mathbf{X}\mathbf{F}^+ \tag{8.5}$$

In this work, we also use non-linear mapping for modelling the mapping between the reflectance spectra and the camera's response. Non-linear methods have been used for spectral reconstruction in previous literature. Ribés and Schmitt [21] used

mixture density networks and genetic algorithms for the estimation of spectral reflectance reconstruction of fine art paintings. Artificial neural network is used by Mansouri *et al.* [22], where the authors formulated the problem of spectral reflectance reconstruction using hetro-associative memories. Yang and Stark [23] proposed the projection onto convex sets, which is an iterative algorithm that projects onto a sequence of constraint sets and eventually produces a solution that satisfies the a-priori constraints. Gómez *et al.* [24] used a feed-forward back propagation neural network for estimation of spectra in oil paintings. Chane *et al.* [25] used a neural network, which is trained by using the 15-band multispectral luminescence acquisitions and the corresponding spot spectroscopy luminescence data. Kandi [26] used a neural network to estimate the spectral information of printed samples from camera response data under two illuminants and to investigate the colorimetric characterization of camera. Recently, Hajipour and Nateri [27] used tristimulus values as input to a neural network for recovery of spectral reconstruction. Radial basis functions (RBF) are also used for spectral reconstruction [28, 29]. RBF is used to model the spectral response of camera for the known spectra. In the training phase, M number of clusters are formed by RBF which aims at minimizing the distance between the similar spectra through the linear least square method [30] and weights ω_i is assigned to each center c_i. The distance being used is Euclidean distance and denoted by $||.||$. A dataset of measured spectra and the same samples being captured by the given camera are used to train the model M as;

$$M = \omega_o + \sum_{i=1}^{M} \omega_i \phi(||\mathbf{f} - c_i||) \tag{8.6}$$

Once the model is trained, it is used to estimate the spectral reflectance of a surface $\hat{\mathbf{r}}$ from the camera input $\mathbf{f}$ as;

$$\hat{\mathbf{r}} = M(\mathbf{f}) \tag{8.7}$$

In ISO 17321-2, a procedure for characterization of digital cameras is defined which is achieved through spectral measurements. Sensor spectral sensitivity of the camera and a training set, comprising of measured spectral radiances are used to compute a linear transform. This transform is used to map the input data at the testing stage. In our work, we use a linear method, and also RBF to create a reflection model for a given camera and the training spectra. The reflectance model is used to estimate the reflectance of input camera data and this reflectance is utilized in the measurement of colorimetric characteristics of the camera. Details of the procedure are explained in the following Section.

8.3 Experiments

In the experiments, we use the following spectral datasets;

- ISO 17321 [20] – spectral radiance measurements of 14 natural objects.
- SFU reflectance data [31] – Spectral reflectance of 1995 surfaces.
- Skin colour reflectance database [32] – Spectral measurement of 4392 skin samples taken from 960 people.
- Munsell color chips [33] – Spectral measurement of 1269 surfaces.
- GretagMacbeth ColorChecker – Spectral measurement of 24 patches.

The data being used is in the range of 400 nm to 700 nm, with 10 nm spectral sampling. For camera data acquisition, we use spectral sensitivities of two multispectral cameras. One camera is the spectral filter array (SFA) camera [34] and the other is filter wheel multispectral camera from pixelteq [35]. Originally both SFA and pixelteq multispectral cameras are of 8 channels, but since the reflectance data that we have is limited in the spectral range of 400 nm to 700 nm, we omit the filters in ultra-violet and infra-red wavelength regions. The sensor sensitivities of the simulated filters are shown in Fig. 8.1. Illuminant D65 is used for creation of radiance data for both training and testing.

In the training phase, we use the simulated multispectral cameras to acquire data from a particular reflectance dataset, which is to be used as training data for camera characterization. As mentioned in Section 2, we use two methods for camera characterization; one is through a linear transform (Eq. 8.5) and other method is through non-linear method (RBF). RBF is used to create the reflectance model for a given set of training spectra and the camera response over that spectra. In the training phase, the weights of hidden layers are adjusted according to the output, until the minimum mean square error is reached. The maximum iterations (epochs) are set to 300 in our experiments. Theoretically, the RBF network create clusters for the input data and the training data by adjusting the weights of the hidden layers. When the test data is given as the input to already trained reflectance model, the output nodes are assigned to the closest matching cluster. There are 31 output nodes which collectively form the output as reflectance spectra from 400nm to 700nm, with 10nm sampling. The structure of RBF is shown in Fig. 8.2.

To measure the performance of camera characterization, estimated colorimetric values from camera data are compared with the ground truth data. In the linear

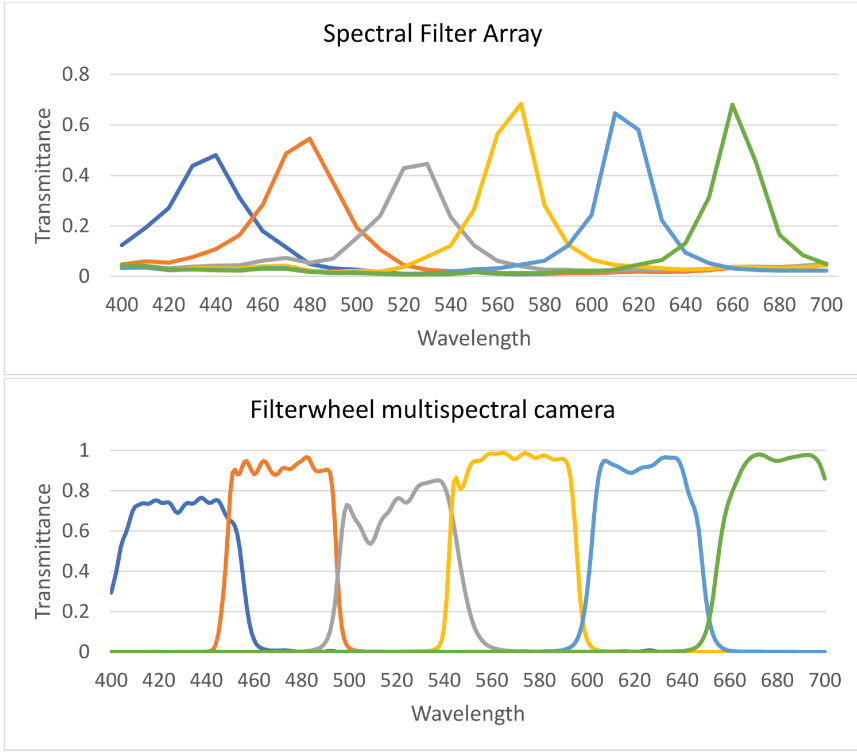


Figure 8.1: Spectral sensitivities of SFA and filter wheel multispectral cameras used in simulations.

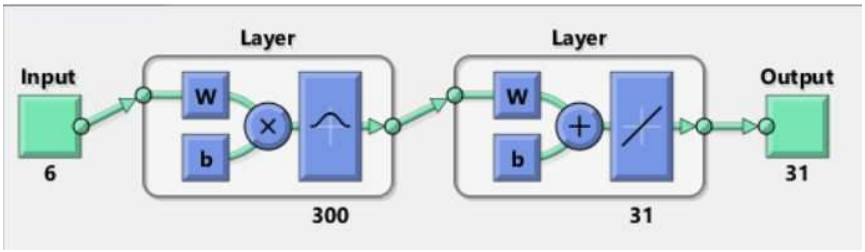


Figure 8.2: The structure of RBF based network. Input consists of a vector of sensor response (6 in this Fig.) and output is the reconstructed spectrum, consisting of 31 elements. The weights of the hidden layer at input side are adjusted in the training phase till the minimum mean square error is achieved, or maximum number of epochs (300 in our experiments) is reached. The hidden layer at output side consists of 31 nodes and provide the estimation of spectrum for the input camera data.

transform, the output is CIE XYZ values while for the output from RBF, the reconstructed spectra is transformed into XYZ by using the CIE 1931 2° Standard Observer and then the XYZ values are converted into CIELAB color space values. For measurement of color difference, we use the CIEDE2000 [36] error metric (also denoted by ΔE_{00}). We use the white point of illuminant D65 in the calculation of XYZ and CIELAB values from the spectra. For each result of ΔE_{00}, the mean, median, maximum and 95$^{\text{th}}$ percentile of error is provided. Results are discussed in the following Section.

8.4 Results and Discussion

Tables 8.1-8.10 provide the colorimetric error in terms of ΔE_{00}. Each Table consist of results obtained by using 2 simulated multispectral cameras and a particular reflectance data for training. For each training reflectance dataset, two methods of camera characterization are used; one is the linear method and other is RFB. The odd numbered tables show results from the linear method while even numbered tables contain results obtained from the RBF method.

The ISO 17321 reflectance data consists of 14 spectral reflectances, and the result of using this spectra as a training set is shown in Tables 8.1 and 8.2. When the linear method is used, the mean colorimetric error is below one for the skin reflectance dataset and the ISO 17321 reflectance itself, but the performance with SFU reflectance is significantly reduced. The performance of SFA is better than the filter wheel camera in terms of ΔE_{00}. In the case of RBF, Table 8.2 shows that when the same reflectance data is used as test data, the error is almost zero, however, when the test data being used is different, then the error values are large. The high values of ΔE_{00} shows that RBF is not a good choice for camera characterization

when the ISO 17321 reflectance data is used for training.

In Tables 8.3 and 8.4, SFU reflectance data is used for training. By using the linear method, the mean of ΔE_{00} is below 2 for both simulated multispectral cameras. However, it can be seen that the maximum value of error is significantly higher when the same SFU dataset is used as test data. Nonetheless, system trained with the SFU dataset is able to keep the ΔE_{00} fairly low for all the reflectance datasets being used for testing. Using RBF, ΔE_{00} for SFU data as test is reduced compared to the linear method and mean error values for Munsell data as test are also below 2. The reason for the lower error values for Munsell dataset while using SFU reflectances as training is the presence of Munsnell reflectances within the SFU dataset. This can also be seen with the lower error values from ColorChecker reflectances as they are also present in the SFU dataset. However, the results are significantly degraded when ISO 17321 data is used for testing and error values are higher with the RBF technique. The same trend can be seen when skin reflectance data is used for training. Tables 8.5 and 8.6 show that with the linear method, there is a lower error of the test spectra while ΔE_{00} is significantly reduced when the same skin reflectance data is used for training and testing in RBF. However, the error is high in the RBF technique when the test data is not related to training data.

Tables 8.7-8.8 consist of ΔE_{00} obtained when Munsell reflectance data is used for training. These results also follow the same trend and the linear method provides smaller errors compared to RBF. The result from RBF is the best only when the same data is used for training and testing. Tables 8.9 and 8.10 consist of results obtained by using ColorChecker reflectance data for training. Once again we see that RBF does not perform better than the linear method, unless the same data is used for training and testing. This trend indicates that when the training data is small, the ΔE_{00} is near to zero when adequate training data is used. However, this is a case of over-fitting and does not apply to generic real world scenarios. Also, such small datasets perform the worst when tested with different spectra. This indicates that the training data should have more samples. However, the training samples should have enough relevance to the test spectra and simply increasing the number of training samples do not provide an advantage.

The best result in terms of ΔE_{00} was obtained when the SFU reflectance dataset was used for training. However, it should be noted that the SFU dataset already consists of the 24 Macbeth color checker patches and 1269 Munsell chips. Additionally, it include reflectance spectra of 120 Dupont paint chips , 170 natural objects , the 350 surfaces in Krinov data set and 57 additional measured surface reflectances. Such a generalized reflectance data is expected to perform well, compared to other datasets. The model trained with SFU dataset is also able to get good results for the ISO 17321 and skin reflectances. This low error indicates that the

linear method along with SFU dataset being used for training provides efficient performance, when tested with a simulated multispectral camera.

The results being obtained by using different training spectra provide interesting information about the behavior of a reflectance estimation system. When the training data is highly correlated with the surfaces being tested, the colorimetric characterization can be performed efficiently. This suggests that we can create reflectance data with as many samples as are available. This can avoid the problem of overfitting and underfitting. It is yet to be investigated how much accuracy is required for computer vision applications, which rely on the combination of spectral and colorimetric information.

Table 8.1: ΔE_{00} results from simulated SFA and filter wheel cameras on 5 different reflectance datasets. Camera characterization is done through linear transformation of input values and reflectance data from ISO 17321 is used for training.

ΔE_{00}	Spectral Filter Array					Filterwheel multispectral camera				
	Test data									
	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker
Mean	0.7194	3.5423	0.946	2.6786	3.3893	0.9883	4.196	2.0603	3.7516	5.0281
Median	0.3213	2.7606	0.822	2.3474	2.6463	0.4337	3.7645	1.9961	3.7541	3.5608
Max	2.4943	20.0891	3.2796	15.0928	8.5842	6.1738	30.8002	5.922	27.993	15.1954
95%	2.4502	10.9307	1.8843	6.132	8.4703	5.3738	10.0045	3.5481	7.7155	13.3755

Table 8.2: ΔE_{00} results from simulated SFA and filter wheel cameras on 5 different reflectance datasets. Colorimetric values from camera input are obtained after spectral reconstruction through RBF, and reflectance data from ISO 17321 is used for training.

ΔE_{00}	Spectral Filter Array					Filterwheel multispectral camera				
	Test data									
	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker
Mean	0.0001	25.2965	16.4351	29.7926	23.0418	0.0001	36.6186	63.2236	40.3671	37.0267
Median	0.0001	19.7864	10.8621	23.6507	18.108	0.0001	32.8391	70.8413	37.4879	33.6008
Max	0.0001	92.9274	58.0097	99.7072	50.2504	0.0001	141.0209	105.3386	149.0721	93.1658
95%	0.0001	59.4169	38.551	72.0706	44.9412	0.0001	79.3924	86.5554	84.9848	71.3004

Table 8.3: ΔE_{00} results from simulated SFA and filter wheel cameras on 5 different reflectance datasets. Camera characterization is done through linear transformation of input values and SFU reflectance data is used for training.

ΔE_{00}	Spectral Filter Array					Filterwheel multispectral camera				
	Test data									
	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker
Mean	1.7584	1.1919	0.9249	0.8282	0.9957	1.7431	1.2836	0.4329	0.9254	0.958
Median	1.6313	0.832	0.9177	0.7382	0.9537	1.8697	0.9392	0.3338	0.8432	0.8008
Max	3.8625	11.3672	2.0604	4.028	2.6986	2.9034	14.8269	2.7687	7.079	2.7164
95%	3.6899	3.8758	1.3144	1.69	2.4889	2.8989	3.6974	1.0873	1.8486	2.2451

Table 8.4: ΔE_{00} results from simulated SFA and filter wheel cameras on 5 different reflectance datasets. Colorimetric values from camera input are obtained after spectral reconstruction through RBF, and SFU reflectance data is used for training.

ΔE_{00}	Spectral Filter Array					Filterwheel multispectral camera				
	Test data									
	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker
Mean	20.3457	0.4943	5.8507	1.4383	2.224	16.2154	0.585	6.904	1.8177	1.6987
Median	12.6682	0.2839	4.7344	1.3454	1.2189	6.8432	0.3148	5.0181	1.755	1.4933
Max	72.5969	10.3802	38.6426	7.9147	17.0861	111.5009	13.781	44.2994	4.5246	4.8049
95%	69.8636	1.5687	11.3124	1.9207	9.4372	95.7351	2.1196	17.7355	2.2121	3.8698

Table 8.5: ΔE_{00} results from simulated SFA and filter wheel cameras on 5 different reflectance datasets. Camera characterization is done through linear transformation of input values and skin reflectance dataset is used for training.

ΔE_{00}	Spectral Filter Array					Filterwheel multispectral camera				
	Test data									
	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker
Mean	2.1332	1.8638	0.1564	1.1973	1.4295	2.189	1.9801	0.2592	1.3363	1.7026
Median	1.9484	1.2062	0.1298	1.0235	1.103	2.3253	1.3586	0.2128	1.1818	1.4986
Max	7.2685	13.5945	1.3601	5.1716	4.8582	5.2003	21.4672	2.26	8.4396	3.9487
95%	6.4475	7.1785	0.3684	2.5389	4.1924	5.0175	6.2132	0.6132	2.8947	3.6536

Table 8.6: ΔE_{00} results from simulated SFA and filter wheel cameras on 5 different reflectance datasets. Colorimetric values from camera input are obtained after spectral reconstruction through RBF, and skin reflectance dataset is used for training.

ΔE_{00}	Spectral Filter Array					Filterwheel multispectral camera				
	Test data									
	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker
Mean	19.1457	10.6788	0.1257	11.4708	17.4727	14.8275	19.4485	0.2013	19.8317	23.2347
Median	11.6453	7.096	0.1033	7.7014	14.4252	5.7423	13.0797	0.1617	13.2395	16.1735
Max	72.6046	112.8374	0.7225	69.1411	33.0798	77.0516	171.6711	1.0909	168.5628	103.2295
95%	66.613	29.3613	0.3024	31.4903	32.5279	67.9785	54.0677	0.4997	55.8913	89.5829

Table 8.7: ΔE_{00} results from simulated SFA and filter wheel cameras on 5 different reflectance datasets. Camera characterization is done through linear transformation of input values and munsell reflectance dataset is used for training.

ΔE_{00}	Spectral Filter Array					Filterwheel multispectral camera				
	Test data									
	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker
Mean	2.145	1.3519	2.0956	0.7224	0.9969	1.8357	1.5966	0.6957	0.8778	1.0108
Median	2.0369	0.7722	2.0932	0.6355	0.6841	1.7431	0.8409	0.6622	0.7276	0.6461
Max	4.1234	12.0243	3.9997	3.7152	2.7537	4.1763	25.4742	2.9312	11.4406	3.1553
95%	3.9508	5.2581	2.7588	1.5458	2.5626	3.8847	6.2594	1.0421	1.9906	2.5951

Table 8.8: ΔE_{00} results from simulated SFA and filter wheel cameras on 5 different reflectance datasets. Colorimetric values from camera input are obtained after spectral reconstruction through RBF, and munsell reflectance dataset is used for training.

ΔE_{00}	Spectral Filter Array					Filterwheel multispectral camera				
	Test data									
	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker
Mean	24.0039	2.9209	9.3015	0.1884	2.1101	15.5738	5.1576	11.7369	0.2483	1.9564
Median	14.5737	1.3599	8.448	0.1334	1.2064	8.6756	1.7684	10.8623	0.1651	1.2751
Max	110.9224	76.2363	34.1916	2.9005	8.2254	80.9067	99.7878	41.0444	1.9762	8.6787
95%	98.0849	8.8584	15.3718	0.5308	7.9414	70.9825	24.3242	19.2322	0.7876	6.7308

Table 8.9: ΔE_{00} results from simulated SFA and filter wheel cameras on 5 different reflectance datasets. Camera characterization is done through linear transformation of input values, and reflectance data from 24 patches of ColorChecker is used for training.

ΔE_{00}	Spectral Filter Array					Filterwheel multispectral camera				
	Test data									
	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker
Mean	1.7404	1.6828	1.0393	1.2164	1.0956	2.204	2.197	0.6699	1.7178	1.1093
Median	1.7835	1.142	1.0493	1.0405	0.7157	2.2295	1.5898	0.5744	1.4805	0.9898
Max	3.3151	12.749	2.4795	4.2034	3.2838	3.8507	19.9727	3.8322	11.5692	3.0324
95%	3.2561	5.7123	1.5598	2.6905	2.4682	3.7146	6.5214	1.5944	3.74	2.6195

Table 8.10: ΔE_{00} results from simulated SFA and filter wheel cameras on 5 different reflectance datasets. Colorimetric values from camera input are obtained after spectral reconstruction through RBF, and reflectance data from 24 patches of ColorChecker is used for training.

ΔE_{00}	**Spectral Filter Array**					**Filterwheel multispectral camera**				
	Test data									
	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker	ISO 17321	SFU	Skin Refl.	Munsell	ColorChecker
Mean	23.0606	6.7614	10.5938	5.0789	0.0001	16.5937	6.7808	8.7854	3.7128	0.0001
Median	14.3294	4.6457	8.821	4.0598	0.0001	8.0794	3.0843	7.7112	2.5198	0.0001
Max	79.9477	65.128	37.6771	33.2546	0.0001	67.2223	122.6417	33.1485	31.1151	0.0001
95%	75.5436	24.2184	21.3308	12.9415	0.0001	62.5552	23.1108	15.4279	10.8399	0.0001

Table 8.11: ΔE_{00} results from simulated SFA on the Foster dataset of hyperspectral reflectance images. Camera characterization is done through linear transformation of input values and SFU reflectance data is used for training.

ΔE_{00}	**Reflectance images from Foster hyperspectral data**							
	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8
Mean	7.3972	1.7079	3.7541	2.1279	1.3803	3.2125	3.5741	4.8736
Median	6.1865	1.2901	1.9380	1.2586	1.1031	3.0211	3.2730	2.8670
Max	34.2251	39.2696	82.2016	39.8810	20.9365	28.1000	16.6525	35.2226
95%	16.1055	4.0502	13.6360	7.1016	3.3286	6.2330	6.9585	17.9389

In order to investigate the camera characterization results on natural scenes, we use SFA for simulation of multispectral image capture on the Foster hyperspectral dataset of natural scenes [37]. These hyperspectral images consist of eight natural scenes from 400 to 720 nm, but we use the data from 400 to 700 nm since reflectance data for training is available in this region. The reason for using the SFA multispectral camera in simulations is that it is able to provide better results for ΔE_{00} compared to the filter wheel camera, except in a few cases. Also, we note that since the SFU reflectance data provide better results compared to the other five datasets used for training, we use the SFU data for training the linear model. This model is used to obtain colorimetric information from the multispectral images taken with simulated SFA camera. D65 illumination is used in the simulations. Table 8.11 show ΔE_{00} for each image in the Foster dataset. Although mean ΔE_{00} for these images is relatively low, except for image 1, the maximum and 95 percentile error values are high. The reason for this error is the absence of relevant spectra in the training dataset. Therefore, if the training spectra consists of a representative spectra and as close to the test images as possible, the error in ΔE_{00} can be reduced significantly. Nonetheless, the results from Table 8.11 shows that using linear methods, a multispectral camera can be efficiently characterized for obtaining colorimetric information.

One of the limitations of the proposed spectral reconstruction framework is that the camera data should be taken in the same illumination under which the system is trained. This condition can be fulfilled when the data acquisition is performed in controlled conditions and the illumination is measured during the experiments. However, measurement of illumination by a telespectroradiometer in outdoor scenes is not always feasible. In the case of a change in illumination, the system should be trained again with the new illumination, and it must be made sure that the same illumination was present during the test image acquisition as well. To overcome this limitation, the concept of *multispectral constancy* was recently proposed [38]. In this method, the estimation of illumination in multispectral images [39] is performed after the acquisition, and the effect of illumination is removed from the multispectral data. The spectral reconstruction system is trained/calibrated with the training data acquired with the same camera under a canonical illuminant. The acquired image is also transformed into a canonical representation and then is provided as input to the spectral reconstruction system. With an efficient multispectral constancy and spectral reconstruction system, use of the multispectral camera can be increased widely for imaging in uncontrolled environments.

8.5 Conclusion

In this paper, we analyse the role of training data for obtaining colorimetric information from two simulated multispectral cameras. We use a linear method for direct conversion from multispectral data into colorimetric data, and a spectral reconstruction method whereby the spectral estimation is performed first and this information is subsequently used to obtain colorimetric data. For spectral reconstruction, a Radial basis function based network is trained to create a model of reflectance and the input data from the multispectral camera is used to obtain the spectral reconstruction. The training data plays an important role in defining the efficiency of the camera characterization. We investigated the reflectance models created with reflectance spectra from ISO 17321, GretagMacbeth ColorChecker, Skin reflectance, Munsell chips and SFU dataset. The SFU dataset consists of reflectance spectra from a number of objects including paints, natural objects, color patches and dyes. The generic nature of this dataset makes it a good training data for the spectral reflectance reconstruction system. We found that by using a linear method, efficient colorimetric information can be obtained and the results outperform the non-linear method being tested. The results obtained are promising and now further work is required to validate these findings by using the real camera and imaging environment.

Bibliography

[1] B. A. Wandell and J. E. Farrell, "Water into wine: converting scanner RGB to tristimulus xyz," *Proc. SPIE, , Device-Independent Color Imaging and Imaging Systems Integration*, vol. 1909, pp. 92–101, 1993.

[2] J. Farrell, D. Sherman, and B. W, "How to turn your scanner into a colorimeter," in *Proc. of IS&T 10th Int. Congress on Adv. in Non-Impact Printing Technologies*, pp. 579–581, 1994.

[3] T. Johnson, "Methods for characterizing colour scanners and digital cameras," *Displays*, vol. 16, no. 4, pp. 183 – 191, 1996.

[4] W. Wu, J. P. Allebach, and M. Analoui, "Imaging colorimetry using a digital camera," *Journal of Imaging Science and Technology*, vol. 44, pp. 267–279, 2000.

[5] G. D. Finlayson and P. M. Morovic, "Metamer constrained color correction," *Journal of Imaging Science and Technology*, vol. 44, no. 4, pp. 295–300, 2000.

[6] G. Hong, M. R. Luo, and P. A. Rhodes, "A study of digital camera colori-

metric characterization based on polynomial modeling," *Color Research & Application*, vol. 26, no. 1, pp. 76–84, 2001.

[7] F. H. Imai, S. Quan, M. R. Rosen, and R. S. Berns, "Digital camera filter design for colorimetric and spectral accuracy," in *Proc. of third international conference on multispectral color science*, pp. 13–16, University of Joensuu, Finland, 2001.

[8] J. Y. Hardeberg, *Acquisition and Reproduction of Color Images: Colorimetric and Multispectral Approaches*. Parkland, FL: Dissertation.com, 2001.

[9] D. Connah, J. Y. Hardeberg, and S. Westland, "Comparison of linear spectral reconstruction methods for multispectral imaging," in *International Conference on Image Processing, ICIP*, vol. 3, pp. 1497–1500, 2004.

[10] L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," *J. Opt. Soc. Am. A*, vol. 3, pp. 1673–1683, Oct 1986.

[11] V. Cheung, C. Li, J. Hardeberg, D. Connah, and S. Westland, "Characterization of trichromatic color cameras by using a new multispectral imaging technique," *J. Opt. Soc. Am. A*, vol. 22, pp. 1231–1240, Jul 2005.

[12] R. S. Berns and M. J. Shyu, "Colorimetric characterization of a desktop drum scanner using a spectral model," *Journal of Electronic Imaging*, vol. 4, no. 4, pp. 360–372, 1995.

[13] M. Shi and G. Healey, "Using reflectance models for color scanner calibration," *J. Opt. Soc. Am. A*, vol. 19, pp. 645–656, Apr 2002.

[14] V. Cheung, S. Westland, D. Connah, and C. Ripamonti, "A comparative study of the characterisation of colour cameras by means of neural networks and polynomial transforms," *Coloration Technology*, vol. 120, no. 1, pp. 19–25, 2004.

[15] T. Jaaskelainen, J. Parkkinen, and S. Toyooka, "Vector-subspace model for color representation," *J. Opt. Soc. Am. A*, vol. 7, pp. 725–730, Apr 1990.

[16] F. H. Imai and R. S. Berns, "Spectral estimation using trichromatic digital cameras," in *Proceedings of the International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives*, vol. 42, pp. 1–8, Chiba University Chiba, Japan, 1999.

[17] D. Connah, S. Westland, and M. G. A. Thomson, "Recovering spectral information using digital camera systems," *Coloration Technology*, vol. 117, no. 6, pp. 309–312, 2001.

[18] H.-L. Shen, P.-Q. Cai, S.-J. Shao, and J. H. Xin, "Reflectance reconstruction for multispectral imaging by adaptive wiener estimation," *Opt. Express*, vol. 15, pp. 15545–15554, Nov 2007.

[19] International Color Consortium, "*Digital camera spectral calibration*," 2017. http://www.color.org/resources/photography/spectral_calibration.xalter.

[20] ISO/TR 17321-2:2012, "*Graphic technology and photography – Colour characterization of digital still cameras (DSCs)*." https://www.iso.org/standard/58003.html.

[21] A. Ribés and F. Schmitt, "A fully automatic method for the reconstruction of spectral reflectance curves by using mixture density networks," *Pattern Recognition Letters*, vol. 24, no. 11, pp. 1691 – 1701, 2003. Colour Image Processing and Analysis. First European Conference on Colour in Graphics, Imaging, and Vision.

[22] A. Mansouri, F. S. Marzani, and P. Gouton, "Neural networks in two cascade algorithms for spectral reflectance reconstruction," in *IEEE International Conference on Image Processing*, vol. 2, pp. II–718–21, Sept 2005.

[23] Y. Yang and H. Stark, "Solutions of several color-matching problems using projection theory," *J. Opt. Soc. Am. A*, vol. 11, pp. 89–96, Jan 1994.

[24] C. Osorio-Gómez, E. Mejía-Ospino, and J. Guerrero-Bermúdez, "Spectral reflectance curves for multispectral imaging, combining different techniques and a neural network," *Revista mexicana de física*, vol. 55, no. 2, pp. 120–124, 2009.

[25] C. S. Chane, M. Thoury, A. Tournié, and J.-P. Echard, "Implementation of a neural network for multispectral luminescence imaging of lake pigment paints," *Applied Spectroscopy*, vol. 69, no. 4, pp. 430–441, 2015.

[26] S. G. Kandi, "Estimating spectral and colorimetric data of printed samples from digital camera responses under two illuminants by neural network," *Journal of Printing Science and Technology*, vol. 47, no. 6, pp. 392–400, 2010.

[27] A. Hajipour and A. Shams-Nateri, "Effect of classification by competitive neural network on reconstruction of reflectance spectra using principal component analysis," *Color Research & Application*, vol. 42, no. 2, pp. 182–188, 2017.

[28] M. D. Buhmann, *Radial basis functions: theory and implementations*, vol. 12. Cambridge university press, 2003.

[29] R. M. H. Nguyen, D. K. Prasad, and M. S. Brown, *Training-Based Spectral Reconstruction from a Single RGB Image*, pp. 186–201. Cham: Springer International Publishing, 2014.

[30] S. Chen, C. F. N. Cowan, and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Transactions on Neural Networks*, vol. 2, pp. 302–309, Mar 1991.

[31] K. Barnard, L. Martin, B. Funt, and A. Coath, "A data set for color research," *Color Research & Application*, vol. 27, no. 3, pp. 147–151, 2002.

[32] K. Xiao, J. M. Yates, F. Zardawi, S. Sueeprasan, N. Liao, L. Gill, C. Li, and S. Wuerger, "Characterising the variations in ethnic skin colours: a new calibrated data base for human skin," *Skin Research and Technology*, vol. 23, no. 1, pp. 21–29, 2017.

[33] University of Joensuu, "*Database - Munsell Colors Matt (Spec).*" http://cs.joensuu.fi/spectral/databases/download/munsell_spec_matt.htm.

[34] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton, "Multispectral filter arrays: Recent advances and practical implementation," *Sensors*, vol. 14, no. 11, pp. 21626–21659, 2014.

[35] "SpectroCam™ Multispectral Wheel Cameras." https://pixelteq.com/spectrocam/.

[36] CIE, *Improvement to industrial colour-difference evaluation.* CIE technical report 142-2001, Vienna: Central Bureau of the CIE, 2001.

[37] D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster, "Frequency of metamerism in natural scenes," *J. Opt. Soc. Am. A*, vol. 23, p. 2359, oct 2006.

[38] H. A. Khan, J. B. Thomas, and J. Y. Hardeberg, "Multispectral constancy based on spectral adaptation transform," in *20th Scandinavian Conference*

on Image Analysis, Tromsø, Norway (P. Sharma and F. M. Bianchi, eds.), pp. 459–470, Springer International Publishing, June 12–14, 2017.

[39] H. A. Khan, J. B. Thomas, J. Y. Hardeberg, and O. Laligant, "Illuminant estimation in multispectral imaging," *J. Opt. Soc. Am. A*, vol. 34, no. 6, 2017.

Chapter 9

Towards the use of multispectral camera as spectrophotometer in uncontrolled illumination conditions

Article F

This chapter is a tentative journal article that is in process of being written and not submitted yet.

H. A. Khan, , J. B. Thomas and J. Y. Hardeberg, "Towards the use of multispectral camera as spectrophotometer in uncontrolled illumination conditions".

Multispectral constancy is recently proposed for enabling illuminant invariant representation of multispectral data. This chapter consists of the experimental case study of multispectral constancy explained in the previous chapters. Multispectral constancy is evaluated through the use of multispectral camera as a spectrophotometer for reconstruction of surface reflectance. Three images with varying illuminations are captured and the spectra of material surfaces is reconstructed. Results show that use of multispectral constancy is beneficial for both filter-wheel and snapshot multispectral cameras. The work presented in this paper will make us one step closer in enabling the use of multispectral cameras in uncontrolled imaging conditions.

9.1 Introduction

With advancements in sensor technology, use of multispectral and hyperspectral imaging is increasing for close range imaging applications [1]. Multispectral imaging is used to capture more spectral details in a scene as compared to conventional color images. Recently emerging technologies, such as the spectral filter arrays [2, 3], enable a broader range of usage domains for multispectral imaging. The ability of multispectral imaging in acquisition of more spectral information is useful for material classification and identification by means of spectral reconstruction [4–8] of surfaces in a scene. The need for spectral reconstruction of surfaces was recognized in 1980s [9, 10]. Since then, many methods are developed to provide spectral reconstruction from the camera data. Most of these methods rely on the use of training data to learn the mapping between camera data and the desired spectra. This process is called *calibration* of the system and is performed through a training set of measured reflectances and radiance data with a given illuminant. To maintain a reasonable accuracy, the same illuminant is required during the scene acquisition. This limitation of having the same illuminant for calibration and image acquisition is a major shortcoming for generic use of multispectral imaging [11].

Acquisition of reflectance spectrum of a surface can be performed with a spectrometer as well but it provides spectral information of a single point. Hyperspectral camera is another option for obtaining spectral information of a scene, it is not a practical solution as the bulky hyperspectral imaging system has be moved frequently. Multispectral imaging provides the intermediate solution as such a camera can be easily be used as a hand held device. The trade-off in this solution is lower number of spectral channels as compared to a hyperspectral image and thus, less spectral information.

In this paper, we are interested in using a multispectral camera for spectral reconstruction of surfaces in a scene. This will enable the use of a multispectral camera as a spectrophotometer that can be dynamically calibrated. So far, the use of multispectral cameras is limited to remote sensing and indoor laboratory conditions. One of the reasons for their limited use is the requirement of known imaging conditions. Scene illumination is one of the key factor in such conditions and generally it is either assumed to be constant (as in case of remote sensing) or measured separately, before processing the acquired multispectral data. Extending the use of multispectral imaging system from heavily constrained environments to real world applications is still an open challenge. One of the major obstacles is calibration of multispectral camera according to the scene illuminant [7, 8, 12–14]. To address this issue, the idea of multispectral constancy is recently proposed by Khan *et al.*

[15, 16]. The aim of this work is to present a practical demonstration of multispectral constancy and evaluate the results in terms of spectral reconstruction. Three scenes are created in a viewing booth and are acquired through a multispectral camera under five different illuminations. Estimation of scene illuminant is performed and the multispectral data is transformed into a canonical representation. Linear method is used for spectral reconstruction by using a calibration matrix which is formed with the canonical illuminant. This method allows the use of multispectral camera as a spectrophotometer and its use in the uncontrolled imaging environment.

9.2 Multispectral constancy

In [16] the concept of multispectral constancy is proposed. Multispectral constancy enables an illuminant invariant representation of the multispectral data. It involves the transformation of multispectral data taken under an unknown illuminant $\mathbf{F}_{ill}$, into a canonical representation $\mathbf{F}_c$. Such a transform is represented by $\mathbf{M}^{c,ill}$ and consists of two parts. One is the diagonal transform $\mathbf{D}^{c,ill}$. The components of this transformation matrix are the sensor responses to the illuminations $\mathbf{E}_c$ (canonical illuminant) and $\mathbf{E}_{ill}$ (scene illuminant). It is represented by

$$\mathbf{D}^{c,ill} = diag(\mathbf{E}_c/\mathbf{E}_{ill}). \tag{9.1}$$

Illuminant in the sensor domain can be measured by using a white diffuser in the scene. It is a commonly used practice during calibration of multispectral and hyperspectral cameras in laboratory conditions. However, it is not always feasible if the camera is used outdoors. In the absence of a white diffuser, the scene illuminant has to be estimated. Discussion of illuminant estimation is provided in Section 9.3.4. The second component of $\mathbf{M}^{c,ill}$ is the spectral adaptation transform (SAT). It is represented by $\mathbf{A}_{SAT}$ and its purpose is to incorporate the intrinsic properties of imaging sensors. The details of SAT are discussed in [16].

Once the components of $\mathbf{M}^{c,ill}$ are computed, the input multispectral data is transformed into its canonical representation as

$$\mathbf{F}_c = \mathbf{M}^{c,ill}\mathbf{F}_{ill} = \mathbf{A}_{SAT}\mathbf{D}^{c,ill}\mathbf{F}_{ill} \tag{9.2}$$

For analysis of this concept on real multispectral data, we take multispectral images, perform illuminant estimation on them, transform them into canonical representation and then evaluate the result through spectral reconstruction of material surfaces in the scene. The details of the experimental setup are discussed in the following section.

9.3 Experimental setup

9.3.1 Objects and surfaces

For the demonstration of multispectral constancy, we selected objects of various categories and acquired their reflectance data through hyperspectral camera. The detail of those objects, image acquisition and reflectance computation is provided in [17]. The reflectance data of those objects is used as ground truth data for matching the performance of spectral reconstruction of the corresponding surfaces from multispectral images.

9.3.2 Scenes setup

To create a complex scene, the objects are placed in a viewing booth. Three scenes created for capturing the various materials. They consist of a scene of kitchen containing utensils, spices and pieces of clothes, and two scenes with different combination of textiles. The scenes are set in a viewing booth (GretagMacbeth Spectralight III) and each image is captured with illuminant A, D65, horizon light, coolwhite and TL84 illuminant. In Fig. 9.1, color illustration of those scenes is shown by using the method described in [18]. These images are shown under simulation of illumination D65.

9.3.3 Image acquisition

After setting up the scenes in viewing booth, a multispectral filter wheel camera (Pixelteq Spectrocam^{TM}) [19] is used to acquire multispectral images. For our experiments, we use the channels in the visible range. Sensitivities of the filters from the multispectral camera are shown in Fig. 9.2. During acquisition of multispectral image from a filter-wheel camera, the exposure time for each channel is set individually. The reason for doing so is to acquire the maximum information in terms of contrast and brightness. In our experiments, the exposure time for each channel is set empirically by observing the histogram of a particular channel and avoiding over-exposure and under-exposure as much as possible. The exposure time for each channel is channel with the change in illumination. By doing so, the compensation for illumination is already incorporated into the resultant multispectral image. However, this is not the case for snapshot cameras, such as the multispectral filter array camera [1] since the sensor integration time is same for all of the channels. In the experiments, we compare results of manually adjusted integration time with a simulation of snapshot camera to analyze the difference in both imaging techniques.

Another issue is the “chromatic aberration” effect along channels which results in blur. Snapshot multispectral cameras also face the same problem [20]. In filter-

(a) Kitchen scene

(b) Textile_1 scene

(c) Textile_2 scene

Figure 9.1: Color rendering of the scenes created in viewing booth and rendered under D65 illuminant.

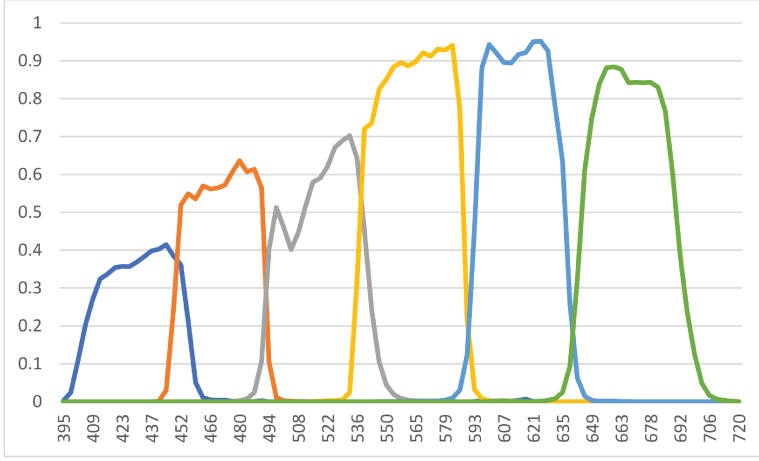


Figure 9.2: Camera sensitivity in the visible range.

wheel multispectral camera, this issue is resolved by focusing the lens for each channel individually and then combining the resulting images to ensure that the images for each channel are as sharp as possible. The unfavorable effect of this process is the slight translation distortion. This distortion can be corrected through affine transformation. However, in the experiments we do not perform it since our method for illuminant estimation, image transformation and spectral reconstruction of surfaces (except the edges) is robust to this distortion.

9.3.4 Illuminant estimation and image transformation

In the work on illuminant estimation in multispectral imaging by Khan *et al.* [21], four algorithms based on the statistics based illuminant estimation method were extended for multispectral imaging. They found that spectral gray-edge algorithm, which is the extension of gray-edge algorithm [22], is robust and performs better than the other three algorithms tested in their work. Based on their observations, we use the spectral gray-edge algorithm for estimation of scene illuminant. The spectral gray-edge algorithm assumes that the average of reflectance *derivative* of each channel of the multispectral image is achromatic. This algorithm is expressed as:

$$\left(\frac{[\int[F_\sigma]^p dx}{\int dx}\right)^{1/p} = \hat{e}, \tag{9.3}$$

where F_σ is the smoothed image, after applying a Gaussian filter and p is the order of the Minkowski norm. In our experiments, we use $p = 5$ and $\sigma = 2$ and estimate the scene illuminant in the sensor domain after masking out the ColorChecker from the scene.

9.3.5 Spectral reflectance reconstruction

For spectral reconstruction of material surfaces from the camera data through linear method, a calibration matrix $\mathbf{W}$ is required. It is obtained by using measured reflectance spectra $\mathbf{R}_t$ and the camera sensor sensitivities ($\mathbf{M}$). For reducing the error between original spectra $\mathbf{R}$ and the estimated spectra $\hat{\mathbf{R}}$, a covariance matrix of a set of measured reflectance samples can be used. Those measured reflectance samples provide the a-priori statistical information about the surfaces in a scene [23]. If the a-priori information is well chosen, error in the spectral reconstruction can be minimal.

There are several methods being proposed for the spectral reconstruction in literature [24]. We use a linear method, namely the Wiener estimation [25] because of its robustness to noise. It is defined as

$$\mathbf{W} = \mathbf{R}_t\mathbf{R}_t^T(\mathbf{CE})^T((\mathbf{CE})\mathbf{R}_t\mathbf{R}_t^T(\mathbf{CE})^T + \mathbf{G})^{-1}. \tag{9.4}$$

Here, $\mathbf{R}_t\mathbf{R}_t^T$ and $\mathbf{G}$ are the autocorrelation matrices of the training spectra and additive noise, respectively. $\mathbf{G}$ is in the form of a diagonal matrix consisting of the variance of noise σ^2. Training for obtaining the matrix $\mathbf{W}_c$ is performed with CIE illuminant E as the canonical illuminant $\mathbf{E}_c$ and the SFU spectral reflectance data [26]. The obtained calibration matrix is used for spectral reconstruction,

$$\hat{\mathbf{R}} = \mathbf{W}_c\mathbf{A}_{SAT}\mathbf{D}^{c,ill}\mathbf{F}_{ill}. \tag{9.5}$$

9.3.6 Evaluation

To measure the performance of the spectral reconstruction, we compare the reconstruction $\hat{\mathbf{r}}$ for each selected material surface with the corresponding measured reflectance $\mathbf{r}$, through root mean square error (RMSE) as

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\mathbf{r}_j - \hat{\mathbf{r}}_j)^2} \tag{9.6}$$

For further evaluation of the reconstructed spectra, cosine distance and spectral angle mapper (SAM) for are also used to compare the results with original reflectance spectra. The equations for cosine distance and SAM are given in Eq. 9.7 and 9.8, respectively.

$$\text{cosine distance} = 1 - \frac{\mathbf{r}.\hat{\mathbf{r}}}{\sqrt{\mathbf{r}.\mathbf{r^t}}\sqrt{\hat{\mathbf{r}}.\hat{\mathbf{r}}^{\mathbf{t}}}} \tag{9.7}$$

$$\text{SAM} = \arccos\frac{\sum_{j=1}^{N}\mathbf{r}_j.\hat{\mathbf{r}}_j}{\sqrt{\sum_{j=1}^{N}\mathbf{r}_j^2}\sqrt{\sum_{j=1}^{N}\hat{\mathbf{r}}_j^2}} \tag{9.8}$$

We also report the error in illuminant estimation in form of angular error (ΔA). It is computed between the values obtained from white patch of ColorChecker $\mathbf{e}$ and the estimated illuminant $\hat{\mathbf{e}}$ as;

$$\Delta A = \arccos\frac{\mathbf{e}^T\hat{\mathbf{e}}}{\sqrt{(\mathbf{e}^T\mathbf{e})(\hat{\mathbf{e}}^T\hat{\mathbf{e}})}} \tag{9.9}$$

Since we do not have ground truth reflectance values for all of the surfaces in the scene, a binary mask is applied on each image for selection of those surfaces whose reflectance data is acquired by the hyperspectral camera. The binary mask is created for each of the three images individually and sensor values for the selected surfaces are obtained. Those values are used further for spectral reconstruction and comparison with ground truth reflectance of the same material.

9.4 Experimental results

For the demonstration of spectral reconstruction results, we selected 16 material surfaces from both"Kitchen" and "Textile 1" images, while 20 material surfaces from the "Textile 2" image. The reflectance data for these surfaces is acquired from hyperspectral camera and is used as ground truth for comparison of results. We present results from two imaging techniques in this paper; one is through the filter-wheel camera where the integration time for each channel is adjusted manually while the other imaging technique is snapshot camera where all channels are acquired with the same integration time. We get the snapshot like effect by normalizing the acquired image through filter-wheel camera with respect to sensor integration time for each channel. In this way, all of the channels appear to be taken with the same integration time. For each test illuminant, results of spectral reconstruction are shown for five experiments which are;

- "Do nothing", when the multispectral data is used for spectral reconstruction, without any processing.
- "Patch select", the white patch of ColorChecker is selected manually and those values are used as the scene illuminant. Data transformation is performed by two methods,
  - Using the extracted values in a diagonal transform, $\mathbf{D}^{c,ill}$
  - Using SAT along with the diagonal transform, $\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$
- "SG-E", illuminant estimation with spectral gray-edge algorithm and then transforming the data using;
  - $\mathbf{D}^{c,ill}$ with values obtained from illuminant estimation.
  - Using SAT along with the diagonal transform, $\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$

Analysis of the results is provided in the following.

9.4.1 Influence of the imaging technique

We used filter-wheel camera and manually set the integration time for each channel. This is a standard practice during image acquisition with such camera and the aim is to acquire each channel with minimum saturated and dark pixels. Generally the integration time is set through observation of image and the histogram. The aim of this practice is to enhance the contrast of the acquired image. On the other hand, a snapshot camera captures the image data with same integration time for all of the channels. Tables 9.1, 9.3 & 9.5 show results with manually set integration

time for each channel and this time varies for each illumination. Tables 9.2, 9.4 & 9.6 show results obtained after the integration time is normalized for all channels. Based on the results, it can be seen that while the results with *do nothing* varies with the change in imaging technique, the proposed multispectral constancy is robust to change in imaging technique. Although there are variations in the results of both techniques and it is not possible to conclude which imaging technique is better, the spectral reconstruction error is reduced when multispectral constancy is used as compared with *do nothing*. This observation illustrates the importance of removing the effect of scene illuminant before using the image data for further analysis.

9.4.2 Manual selection of white patch vs. illuminant estimation

The results from spectral reconstruction provide an interesting observation that the error in spectral reconstruction is slightly reduced when correction is applied after illuminant estimation instead of manual selection of white patch in the ColorChecker. The reason for this effect is due to the presence of saturated pixels in the channels which results in erroneous transformation towards the canonical representation. During illuminant estimation, the pixels with peak value (255 for 8-bit image) are ignored by the spectral gray-edge algorithm and the resultant estimate of scene illuminant is free from saturated values. In a few cases, the error in reconstructed spectra is slightly increased when image is transformed through the estimated illuminant. The reason for it is the error in illuminant estimation itself. Although the improvement in spectral reconstruction result is not significant when data is transformed by using the estimated illuminant, it is worth noting that in practical situations, it is not feasible to use ColorChecker with every image acquisition. The results from illuminant estimation show the strength and robustness of spectral gray-edge algorithm for illuminant estimation in multispectral imaging and this will be a key factor in enabling the use of multispectral camera in uncontrolled imaging environment.

9.4.3 Performance of spectral adaptation transform

During simulation in [16], SAT showed significant improvement in spectral reconstruction results. However, the improvement due to SAT in results from our experiments is marginal. The purpose of SAT is to incorporate the intrinsic characteristics of imaging sensor and address the inter-channel overlapping. In [16], SAT is able to improve results of simulated overlapping sensors. The overlapping of filters in the filter-wheel camera used in our experiments is small as can be seen in Fig. 9.2. The diagonal transform assumes that there are no inter-channel dependencies in the camera and each of the channels is treated individually. SAT accounts for those dependencies and if they are small (as for the filter-wheel cam-

era in our experiments), the influence of SAT is also low. When we observed the acquired SAT from Eq. 9.2, we found that the diagonal entries of this matrix consist of large numbers while the rest of the entries are very small. This explains why SAT does not change the results significantly in comparison with a diagonal transform. However, it does not mean that SAT is of no use since it is able to address the inter-channel dependencies, particularly in the case of snap-shot spectral filter arrays.

9.4.4 Discussion on results

To provide an overview of the results in Tables 9.1-9.6, we show the RMSE results in Fig. 9.3. They show promising results in terms of spectral reconstruction from the multispectral images. Use of multispectral constancy is able to reduce the error significantly as compared to *do nothing*. These results show that a multispectral camera can be used as a spectrophotometer for outdoor imaging by incorporating the idea of multispectral constancy in the imaging pipeline. Although we can see marked improvement in spectral reconstruction results in terms of RMSE, cosine distance and SAM, it is yet to be investigated that how much accuracy is required in order to enable the use of multispectral cameras for computer vision applications. Use of multispectral constancy for extraction of spectral information along with rest of features being used by computer vision applications will definitely improve the accuracy and open new paths for research by using multispectral imaging.

Table 9.1: Spectral reconstruction result of selected surfaces from the "Kitchen" image. Each channel is acquired with manually adjusted integration time. Results are presented in form of mean and 95 percentile of error metric. Results of manually selecting the white patch of ColorChecker is shown with "Patch select", while "SG-E" show results of Spectral gray-edge algo. with angular error (ΔA), after ColorChecker is masked out from image.

Test Illuminant	**Image Transformation**		**RMSE**		**Cosine Distance**		**SAM**	
			Mean	95 %	Mean	95 %	Mean	95 %
A	Do nothing		1.7007	3.7506	0.0867	0.1571	0.403	0.568
	Patch select	$\mathbf{D}^{c,ill}$	1.3796	2.8477	0.0386	0.0773	0.2701	0.3949
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2673	2.7666	0.0347	0.067	0.2572	0.3679
	SG-E ΔA 0.0579	$\mathbf{D}^{c,ill}$	1.2607	2.7954	0.0328	0.0671	0.2497	0.3679
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.1656	2.8097	0.0297	0.0579	0.2386	0.342
Coolwhite	Do nothing		1.91	4.4173	0.1278	0.2297	0.4925	0.6913
	Patch select	$\mathbf{D}^{c,ill}$	1.0331	3.8032	0.0225	0.0495	0.2058	0.3155
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.0084	3.8485	0.0212	0.0455	0.1995	0.3015
	SG-E ΔA 0.0976	$\mathbf{D}^{c,ill}$	1.1715	3.7569	0.0337	0.0715	0.2507	0.38
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.1277	3.7981	0.0312	0.0678	0.2421	0.3699
D65	Do nothing		1.907	4.3645	0.0816	0.1524	0.3887	0.5594
	Patch select	$\mathbf{D}^{c,ill}$	1.1652	3.1008	0.0245	0.0572	0.2123	0.3397
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.0904	3.0948	0.0229	0.0608	0.2044	0.3503
	SG-E ΔA 0.093	$\mathbf{D}^{c,ill}$	1.16	3.1658	0.0244	0.046	0.2145	0.3046
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.092	3.1741	0.0231	0.0464	0.2083	0.3058
Horizon	Do nothing		2.1	4.0221	0.0996	0.1743	0.4271	0.5993
	Patch select	$\mathbf{D}^{c,ill}$	1.3095	2.9973	0.0356	0.0629	0.2609	0.3563
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2163	3.05	0.0325	0.0543	0.2499	0.3309
	SG-E ΔA 0.0904	$\mathbf{D}^{c,ill}$	1.2389	3.1152	0.0355	0.0594	0.2601	0.3463
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.1661	3.1667	0.0333	0.056	0.2524	0.3362
TL84	Do nothing		1.7456	4.0384	0.1073	0.1858	0.4532	0.6194
	Patch select	$\mathbf{D}^{c,ill}$	1.0044	3.6175	0.0255	0.0574	0.2149	0.3401
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	0.9857	3.6702	0.0245	0.06	0.2089	0.3474
	SG-E ΔA 0.0934	$\mathbf{D}^{c,ill}$	1.0769	3.5472	0.0296	0.0625	0.2349	0.3551
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.0397	3.5985	0.0274	0.0534	0.2262	0.3282

Table 9.2: Spectral reconstruction result of selected surfaces from the "Kitchen" image, taken with simulation of snapshot camera. Results are presented in form of mean and 95 percentile of error metric. Results of manually selecting the white patch of ColorChecker is shown with "Patch select", while "SG-E" show results of Spectral gray-edge algo. with angular error (ΔA), after ColorChecker is masked out from image.

Test Illuminant	**Image Transformation**		**RMSE**		**Cosine Distance**		**SAM**	
			Mean	95 %	Mean	95 %	Mean	95 %
A	Do nothing		1.5142	3.0859	0.06	0.3041	0.2848	0.8002
	Patch select	$\mathbf{D}^{c,ill}$	1.2582	2.7207	0.0318	0.0662	0.2455	0.364
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.156	2.7262	0.0285	0.0564	0.2333	0.3365
	SG-E ΔA 0.0548	$\mathbf{D}^{c,ill}$	1.2876	2.7812	0.0368	0.0688	0.2639	0.373
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.1945	2.8154	0.0335	0.0599	0.2531	0.3478
Coolwhite	Do nothing		2.0593	4.5579	0.3201	0.6067	0.7886	1.1665
	Patch select	$\mathbf{D}^{c,ill}$	1.0326	3.8029	0.0224	0.0493	0.2054	0.3147
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.0079	3.8483	0.0211	0.0453	0.1991	0.3007
	SG-E ΔA 0.1427	$\mathbf{D}^{c,ill}$	1.1662	3.8015	0.0342	0.0771	0.252	0.3946
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.1299	3.8398	0.0321	0.0738	0.2449	0.3858
D65	Do nothing		1.3305	3.1833	0.0301	0.0777	0.2329	0.3968
	Patch select	$\mathbf{D}^{c,ill}$	1.1643	3.1003	0.0244	0.0572	0.2119	0.3397
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.0896	3.0948	0.0229	0.0608	0.204	0.3503
	SG-E ΔA 0.0692	$\mathbf{D}^{c,ill}$	1.2494	3.197	0.0288	0.0538	0.2337	0.3293
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.1678	3.1265	0.0266	0.047	0.225	0.3078
Horizon	Do nothing		1.4976	3.5967	0.0946	0.4329	0.3689	0.9651
	Patch select	$\mathbf{D}^{c,ill}$	1.4139	2.9828	0.0368	0.0621	0.265	0.354
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.3066	2.994	0.0338	0.0557	0.2542	0.3353
	SG-E ΔA 0.0504	$\mathbf{D}^{c,ill}$	1.4165	3.0549	0.0425	0.0719	0.2827	0.3815
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.3228	3.093	0.0398	0.0707	0.274	0.3781
TL84	Do nothing		1.8658	4.2665	0.3205	0.6623	0.7936	1.2259
	Patch select	$\mathbf{D}^{c,ill}$	1.0054	3.625	0.0259	0.0608	0.2157	0.3499
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	0.9882	3.6774	0.025	0.064	0.21	0.3582
	SG-E ΔA 0.0675	$\mathbf{D}^{c,ill}$	1.1141	3.5222	0.0326	0.0713	0.2468	0.3789
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.0685	3.5738	0.0299	0.0611	0.2363	0.3509

Table 9.3: Spectral reconstruction result of selected surfaces from the "Textile 1" image. Each channel is acquired with manually adjusted integration time. Results are presented in form of mean and 95 percentile of error metric. Results of manually selecting the white patch of ColorChecker is shown with "Patch select", while "SG-E" show results of Spectral gray-edge algo. with angular error (ΔA), after ColorChecker is masked out from image.

Test Illuminant	**Image Transformation**		**RMSE**		**Cosine Distance**		**SAM**	
			Mean	95 %	Mean	95 %	Mean	95 %
Coolwhite	Do nothing		1.8657	4.1576	0.323	0.4823	0.818	1.0266
	Patch select	$\mathbf{D}^{c,ill}$	0.6197	2.4179	0.0323	0.1563	0.22	0.5553
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	0.623	2.5232	0.0327	0.1555	0.2236	0.5535
	SG-E ΔA 0.0422	$\mathbf{D}^{c,ill}$	0.6489	2.1213	0.038	0.1929	0.235	0.616
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	0.6342	2.2421	0.0365	0.189	0.2285	0.6093
D65	Do nothing		0.9675	1.975	0.0499	0.1951	0.2808	0.6244
	Patch select	$\mathbf{D}^{c,ill}$	0.5934	1.9496	0.0362	0.1783	0.2266	0.5929
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	0.5876	2.0273	0.0371	0.178	0.2316	0.5928
	SG-E ΔA 0.0926	$\mathbf{D}^{c,ill}$	0.7395	2.0679	0.0388	0.1909	0.2426	0.6089
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	0.6848	2.0023	0.0363	0.1863	0.2321	0.6014
Horizon	Do nothing		0.9473	2.3188	0.1344	0.306	0.4923	0.8016
	Patch select	$\mathbf{D}^{c,ill}$	0.8052	2.0146	0.0343	0.1411	0.2413	0.5218
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	0.7694	1.9766	0.034	0.1404	0.2391	0.5203
	SG-E ΔA 0.034	$\mathbf{D}^{c,ill}$	0.8032	1.9973	0.0364	0.153	0.242	0.5519
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	0.7731	1.961	0.0367	0.1532	0.2422	0.5533
TL84	Do nothing		1.5226	3.9476	0.4034	0.6191	0.9197	1.1798
	Patch select	$\mathbf{D}^{c,ill}$	0.6703	2.4296	0.0402	0.1766	0.2436	0.5938
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	0.6785	2.5263	0.0424	0.1771	0.2543	0.595
	SG-E ΔA 0.1021	$\mathbf{D}^{c,ill}$	0.663	2.3945	0.0359	0.1738	0.225	0.5868
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	0.6662	2.4911	0.0375	0.1736	0.2334	0.5868

Table 9.4: Spectral reconstruction result of selected surfaces from the "Textile 1" image, taken with simulation of snapshot camera. Results are presented in form of mean and 95 percentile of error metric. Results of manually selecting the white patch of ColorChecker is shown with "Patch select", while "SG-E" show results of Spectral gray-edge algo. with angular error (ΔA), after ColorChecker is masked out from image.

Test Illuminant	**Image Transformation**		**RMSE**		**Cosine Distance**		**SAM**	
			Mean	95 %	Mean	95 %	Mean	95 %
Coolwhite	Do nothing		2.2736	4.4972	0.1197	0.2531	0.4724	0.7272
	Patch select	$\mathbf{D}^{c,ill}$	1.251	2.6715	0.0317	0.1155	0.2303	0.4807
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2268	2.7705	0.0315	0.1142	0.2301	0.4735
	SG-E ΔA 0.0976	$\mathbf{D}^{c,ill}$	1.2093	2.7357	0.0334	0.1313	0.2345	0.5139
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.1979	2.828	0.0333	0.1295	0.2343	0.506
D65	Do nothing		1.7803	4.2289	0.0769	0.2045	0.3629	0.6488
	Patch select	$\mathbf{D}^{c,ill}$	1.2	3.4305	0.0361	0.1567	0.239	0.558
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.212	3.5152	0.0356	0.1549	0.2358	0.5494
	SG-E ΔA 0.093	$\mathbf{D}^{c,ill}$	1.2399	3.2981	0.0355	0.1577	0.2333	0.5672
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2245	3.3836	0.0345	0.1543	0.231	0.5577
Horizon	Do nothing		1.944	4.3626	0.104	0.2038	0.427	0.6497
	Patch select	$\mathbf{D}^{c,ill}$	2.1323	5.0341	0.1204	0.25	0.4849	0.7138
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	2.149	5.0757	0.1216	0.2548	0.4872	0.7206
	SG-E ΔA 0.0904	$\mathbf{D}^{c,ill}$	1.7996	4.541	0.0705	0.1978	0.3609	0.6248
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.8272	4.6056	0.0717	0.2033	0.3641	0.6332
TL84	Do nothing		2.0818	4.4509	0.1153	0.2812	0.4578	0.7688
	Patch select	$\mathbf{D}^{c,ill}$	1.2714	2.9514	0.0438	0.1634	0.2664	0.5654
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2676	3.0732	0.0446	0.1666	0.2681	0.568
	SG-E ΔA 0.0934	$\mathbf{D}^{c,ill}$	1.2377	2.8104	0.0409	0.1599	0.2563	0.5652
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.221	2.9062	0.0409	0.1584	0.2573	0.5572

Table 9.5: Spectral reconstruction result of selected surfaces from the "Textile 2" image. Each channel is acquired with manually adjusted integration time. Results are presented in form of mean and 95 percentile of error metric. Results of manually selecting the white patch of ColorChecker is shown with "Patch select", while "SG-E" show results of Spectral gray-edge algo. with angular error (ΔA), after ColorChecker is masked out from image.

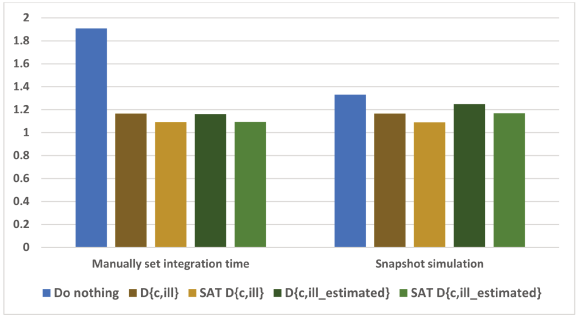
Test Illuminant	**Image Transformation**		**RMSE**		**Cosine Distance**		**SAM**	
			Mean	95 %	Mean	95 %	Mean	95 %
A	Do nothing		1.8455	4.1913	0.0837	0.1966	0.3882	0.6377
	Patch select	$\mathbf{D}^{c,ill}$	1.2922	2.8812	0.0396	0.1381	0.2476	0.5317
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2435	2.9499	0.0376	0.1338	0.2417	0.5223
	SG-E	$\mathbf{D}^{c,ill}$	1.2274	2.852	0.0366	0.1546	0.2377	0.5625
	ΔA 0.0579	$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2041	2.9587	0.0355	0.151	0.2361	0.5532
Coolwhite	Do nothing		2.2736	4.4972	0.1197	0.2531	0.4724	0.7272
	Patch select	$\mathbf{D}^{c,ill}$	1.251	2.6715	0.0317	0.1155	0.2303	0.4807
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2268	2.7705	0.0315	0.1142	0.2301	0.4735
	SG-E	$\mathbf{D}^{c,ill}$	1.2093	2.7357	0.0334	0.1313	0.2345	0.5139
	ΔA 0.0976	$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.1979	2.828	0.0333	0.1295	0.2343	0.506
D65	Do nothing		1.7803	4.2289	0.0769	0.2045	0.3629	0.6488
	Patch select	$\mathbf{D}^{c,ill}$	1.2	3.4305	0.0361	0.1567	0.239	0.558
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.212	3.5152	0.0356	0.1549	0.2358	0.5494
	SG-E	$\mathbf{D}^{c,ill}$	1.2399	3.2981	0.0355	0.1577	0.2333	0.5672
	ΔA 0.093	$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2245	3.3836	0.0345	0.1543	0.231	0.5577
Horizon	Do nothing		1.944	4.3626	0.104	0.2038	0.427	0.6497
	Patch select	$\mathbf{D}^{c,ill}$	2.1323	5.0341	0.1204	0.25	0.4849	0.7138
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	2.149	5.0757	0.1216	0.2548	0.4872	0.7206
	SG-E	$\mathbf{D}^{c,ill}$	1.7996	4.541	0.0705	0.1978	0.3609	0.6248
	ΔA 0.0904	$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.8272	4.6056	0.0717	0.2033	0.3641	0.6332
TL84	Do nothing		2.0818	4.4509	0.1153	0.2812	0.4578	0.7688
	Patch select	$\mathbf{D}^{c,ill}$	1.2714	2.9514	0.0438	0.1634	0.2664	0.5654
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2676	3.0732	0.0446	0.1666	0.2681	0.568
	SG-E	$\mathbf{D}^{c,ill}$	1.2377	2.8104	0.0409	0.1599	0.2563	0.5652
	ΔA 0.0934	$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.221	2.9062	0.0409	0.1584	0.2573	0.5572

Table 9.6: Spectral reconstruction result of selected surfaces from the "Textile_2" image, taken with simulation of snapshot camera. Results are presented in form of mean and 95 percentile of error metric. Results of manually selecting the white patch of ColorChecker is shown with "Patch select", while "SG-E" show results of Spectral gray-edge algo. with angular error (ΔA), after ColorChecker is masked out from image.

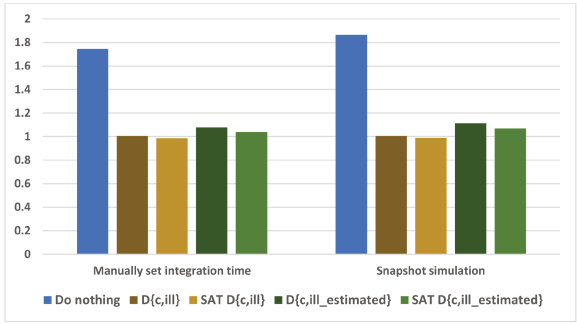
Test Illuminant	**Image Transformation**		**RMSE**		**Cosine Distance**		**SAM**	
			Mean	95 %	Mean	95 %	Mean	95 %
A	Do nothing		1.6211	4.1602	0.1145	0.3778	0.4207	0.8976
	Patch select	$\mathbf{D}^{c,ill}$	1.2962	2.854	0.0397	0.1382	0.2473	0.5321
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2429	2.9196	0.0376	0.1339	0.2406	0.5227
	SG-E	$\mathbf{D}^{c,ill}$	1.2822	2.8042	0.0381	0.1527	0.2404	0.5599
	ΔA 0.0533	$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.238	2.8956	0.0365	0.1484	0.2373	0.5505
Coolwhite	Do nothing		2.3063	5.1156	0.2552	0.6308	0.6832	1.1913
	Patch select	$\mathbf{D}^{c,ill}$	1.2374	2.6408	0.0315	0.1185	0.229	0.4866
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2151	2.744	0.0315	0.1173	0.2291	0.4793
	SG-E	$\mathbf{D}^{c,ill}$	1.2124	2.69	0.0329	0.1305	0.233	0.5121
	ΔA 0.0422	$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.1984	2.7875	0.0329	0.1288	0.2329	0.5044
D65	Do nothing		1.3168	3.0995	0.0442	0.1434	0.2778	0.5391
	Patch select	$\mathbf{D}^{c,ill}$	1.1992	3.4287	0.0361	0.1566	0.2389	0.5579
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2111	3.5136	0.0355	0.1548	0.2357	0.5493
	SG-E	$\mathbf{D}^{c,ill}$	1.3004	3.1749	0.0378	0.1594	0.2387	0.5722
	ΔA 0.0926	$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2553	3.2501	0.036	0.1548	0.2343	0.5623
Horizon	Do nothing		2.4914	4.868	0.1667	0.4982	0.51	1.0451
	Patch select	$\mathbf{D}^{c,ill}$	2.1986	5.1516	0.128	0.2643	0.5006	0.7365
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	2.2144	5.19	0.1292	0.2702	0.5027	0.7452
	SG-E	$\mathbf{D}^{c,ill}$	1.7334	4.4552	0.0628	0.1771	0.3367	0.5848
	ΔA 0.034	$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.7607	4.5202	0.0633	0.1797	0.3388	0.5867
TL84	Do nothing		2.2905	4.7938	0.2626	0.6477	0.705	1.2093
	Patch select	$\mathbf{D}^{c,ill}$	1.2874	2.8718	0.0433	0.1598	0.2642	0.5613
		$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2695	2.9824	0.0437	0.1593	0.2652	0.5549
	SG-E	$\mathbf{D}^{c,ill}$	1.2472	2.7941	0.0408	0.1594	0.2558	0.5648
	ΔA 0.1021	$\mathbf{A}_{SAT}\mathbf{D}^{c,ill}$	1.2263	2.889	0.0409	0.1578	0.2566	0.5567

9.5 Conclusion

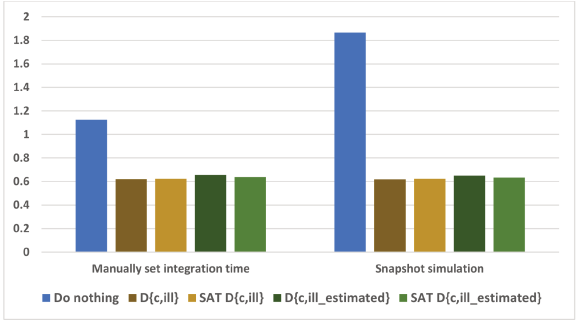
In this paper, we provide practical demonstration and evaluation of the concept of multispectral constancy. Three different scenes are created in viewing booth and multispectral images are captured while using different illuminants. The acquired multispectral data is transformed into canonical representation through manual selection of white patch of ColorChecker in the image and by estimating the illuminant. Spectral reconstruction is performed by using Wiener estimation method and results are evaluated in terms of RMSE, cosine distance and spectral angle mapper. Results show a promising aspect of multispectral imaging as they can be used as a spectrophotometer for getting spectral information of a whole scene. The proposed concept of multispectral constancy is valid for both filter-wheel and snapshot type of multispectral cameras. Having such promising results, we are one step closer towards enabling the use of multispectral imaging for computer vision applications.



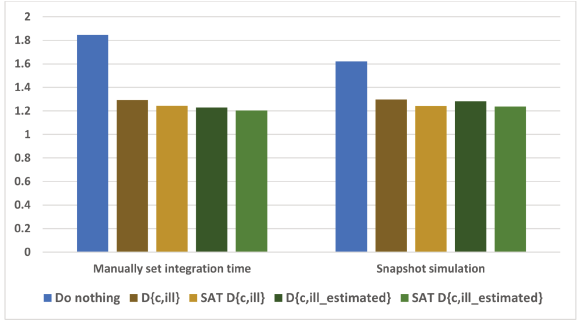
(a) RMSE results from *Kitchen* image taken under simulation of D65 illuminant.



(b) RMSE results from *Kitchen* image taken under simulation of TL84 illuminant.



(c) RMSE results from *Textile_1* image taken under simulation of coolwhite illuminant.



(d) RMSE results from *Textile_2* image taken under simulation of A illuminant.

Figure 9.3: RMSE from spectral reconstruction. Each fig. provides comparison of *multispectral constancy* with *do nothing*. The comparison of imaging techniques (manually set integration time and snapshot) is also provided in the same figure. It can be observed that use of multispectral constancy is robust to imaging technique and reduces the error in comparison to *do nothing*. $D(c, ill)$ represents diagonal transform with illuminant values taken from white patch of ColorChecker while $D(c, ill_estimated)$ is the diagonal transform with estimated illuminant values.

Bibliography

[1] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton, "Multispectral filter arrays: Recent advances and practical implementation," *Sensors*, vol. 14, no. 11, pp. 21626–21659, 2014.

[2] J.-B. Thomas, P.-J. Lapray, P. Gouton, and C. Clerc, "Spectral Characterization of a Prototype SFA Camera for Joint Visible and NIR Acquisition," *Sensors*, vol. 16, no. 7, p. 993, 2016.

[3] R. Shrestha, J. Y. Hardeberg, and R. Khan, "Spatial arrangement of color filter array for multispectral image acquisition," in *Proc. SPIE, Sensors, Cameras, and Systems for Industrial, Scientific, and Consumer Applications XII*, vol. 7875, pp. 787503–787503–9, Feb. 2011.

[4] F. H. Imai and R. S. Berns, "Spectral estimation using trichromatic digital cameras," in *Proceedings of the International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives*, pp. 1–8, Chiba University Chiba, Japan, 1999.

[5] J. Y. Hardeberg, *Acquisition and Reproduction of Color Images: Colorimetric and Multispectral Approaches*. Universal Publishers, 2001.

[6] D. Connah, S. Westland, and M. G. A. Thomson, "Recovering spectral information using digital camera systems," *Coloration Technology*, vol. 117, pp. 309–312, nov 2001.

[7] E. M. Valero, J. L. Nieves, S. M. C. Nascimento, K. Amano, and D. H. Foster, "Recovering spectral data from natural scenes with an RGB digital camera and colored filters," *Color Research & Application*, vol. 32, no. 5, pp. 352–360, 2007.

[8] J. Y. Hardeberg and R. Shrestha, "Multispectral colour imaging: Time to move out of the lab?," in *Mid-term meeting of the International Colour Association (AIC)*, (Tokyo, Japan), pp. 28–32, May 2015.

[9] L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," *J. Opt. Soc. Am. A*, vol. 3, pp. 1673–1683, oct 1986.

[10] J. P. S. Parkkinen, J. Hallikainen, and T. Jaaskelainen, "Characteristic spectra of Munsell colors," *J. Opt. Soc. Am. A*, vol. 6, pp. 318–322, feb 1989.

[11] R. Shrestha and J. Y. Hardeberg, "Spectrogenic imaging: a novel approach to multispectral imaging in an uncontrolled environment.," *Optics Express*, vol. 22, pp. 9123–33, apr 2014.

[12] F. Imai and R. Berns, "Spectral estimation using trichromatic digital cameras," in *International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives*, pp. 42–49, 1999.

[13] D. Connah, S. Westland, and M. G. A. Thomson, "Recovering spectral information using digital camera systems," *Coloration Technology*, vol. 117, no. 6, pp. 309–312, 2001.

[14] R. Shrestha and J. Y. Hardeberg, "Spectrogenic imaging: A novel approach to multispectral imaging in an uncontrolled environment," *Opt. Express*, vol. 22, pp. 9123–9133, Apr 2014.

[15] H. A. Khan, J. B. Thomas, and J. Y. Hardeberg, "Multispectral constancy based on spectral adaptation transform," in *20th Scandinavian Conf. on Image Analysis*, pp. 459–470, June 2017.

[16] H. A. Khan, J.-B. Thomas, J. Y. Hardeberg, and O. Laligant, "Spectral adaptation transform for multispectral constancy," *Journal of Imaging Science and Technology*, vol. 62, no. 2, pp. 1020504–1– 020504–12, 2018.

[17] H. A. Khan, S. Mihoubi, B. Mathon, J.-B. Thomas, and J. Y. Hardeberg, "HyTexiLa: High Resolution Visible and Near Infrared Hyperspectral Texture Images," *Sensors*, vol. 18, no. 7, p. 2045, 2018.

[18] H. A. Khan and P. Green, "Color characterization methods for a multispectral camera," in *International Symposium on Electronic Imaging 2018: Color Imaging XXIII: Displaying, Processing, Hardcopy, and Applications*, (San Francisco, United States), IS&T , Jan. 2018.

[19] "SpectroCam Multispectral Wheel Cameras." https://pixelteq.com/spectrocam/. Accessed: 07-06-2018.

[20] C. Ni, J. Jia, M. Howard, K. Hirakawa, and A. Sarangan, "Single-shot multispectral imager using spatially multiplexed fourier spectral filters," *J. Opt. Soc. Am. B*, vol. 35, pp. 1072–1079, May 2018.

[21] H. A. Khan, J.-B. Thomas, J. Y. Hardeberg, and O. Laligant, "Illuminant estimation in multispectral imaging," *J. Opt. Soc. Am. A*, vol. 34, pp. 1085–1098, 2017.

[22] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Transactions on Image Processing*, vol. 16, pp. 2207–2214, Sept 2007.

[23] J. Conde, H. Haneishi, M. Yamaguchi, N. Ohyama, and J. Baez, "Spectral reflectance estimation of ancient Mexican codices, multispectral images approach," *Revista Mexicana de Fisica*, vol. 50, no. 5, pp. 484–489, 2004.

[24] D. Connah, J. Y. Hardeberg, and S. Westland, "Comparison of linear spectral reconstruction methods for multispectral imaging," in *International Conference on Image Processing, ICIP*, vol. 3, pp. 1497–1500, 2004.

[25] H.-L. Shen, P.-Q. Cai, S.-J. Shao, and J. H. Xin, "Reflectance reconstruction for multispectral imaging by adaptive Wiener estimation," *Optics Express*, vol. 15, no. 23, pp. 15545–15554, 2007.

[26] K. Barnard, L. Martin, B. Funt, and A. Coath, "A data set for color research," *Color Research & Application*, vol. 27, no. 3, pp. 147–151, 2002.

Chapter 10

HyTexiLa: High resolution visible and near infrared hyperspectral texture images

Article G

This chapter is a reformatted reprint of the publication

H. A. Khan, S. Mihoubi, B. Mathon, J. B. Thomas and J. Y. Hardeberg, "HyTexiLa: High resolution visible and near infrared hyperspectral texture images," *Sensors*, 18(7), 2045; 2018.

We present a dataset of close range hyperspectral images of materials that span visible and near infrared: HyTexiLa (Hyperspectral Texture images acquired in Laboratory). The data is intended to provide high spectral and spatial resolution reflectance images of 112 materials to study spatial and spectral textures. In this paper we discuss the calibration of data and the method for addressing the distortions during image acquisition. We provide a spectral analysis based on non-negative matrix factorization to quantify the spectral complexity of the samples and extend local binary pattern operators to hyperspectral texture analysis. Results demonstrate that although the spectral complexity of each of the textures is generally low, increasing the number of bands permits better texture classification, with the opponent band local binary pattern feature giving the best performance.[1]

[1] Author of this dissertation have major contributions in Section 10.1, 10.2, 10.4, and partial contribution in Section 10.3.

10.1 Introduction

When the electromagnetic waves are incident on a material's surface, some part of it is absorbed and some of it is reflected back. The reflected electromagnetic waves from a surface depends on its material composition, physical and chemical properties, surface roughness and the geometry of incident and reflected waves. The behavior of reflected waves from a surface in a particular wavelength region is known as the spectral signature of that material. To obtain the spectral signature from a surface, devices like spectrophotometer and hyperspectral cameras are used. The advantage of hyperspectral devices is the acquisition of high resolution spectral and spatial information of a material or scene, whereas the spectrophotometer is able to acquire spectral information from one point of the selected surface.

For the acquisition of hyperspectral images, push broom cameras have been extensively used [1]. In push broom imaging technique, line scanning of the scene is performed and the spectrogram of a particular scan is recorded on charged couple device sensor. Another method for acquisition of hyperspectral images is through the use of liquid crystal tunable filter (LCTF) [2] and MEMS Tunable Fabry-Pérot Filters [3] on top of a monochrome sensor. Use of such filters provide direct imaging ability and line scan of the scene is no longer required. The drawback of using such filters is low speed of spectral sampling and lower number of channels compared to push broom cameras. Recently, low cost hyperspectral imaging devices are introduced that use off-the-shelf components for conversion of an imaging sensor into hyperspectral camera [4].

Analysis of hyperspectral images is performed for image data exploration, classification and quantification tasks. There are various multivariate analysis techniques that are applied to hyperspectral images for data decomposition, pre-processing and performing regression or classification analyses. Some of the methods for hyperspectral data analysis include principal component analysis [5], linear discriminant analysis [6], partial least squares discriminant analysis [7], K-nearest neighbor algorithm [8], support vector machine [9], least-squares support vector machine [10], artificial neural network [11], deep recurrent neural network [12] and deep convolutional neural networks [13]. Hyperspectral imaging is widely used in a number of applications, e.g. remote sensing [14], medical diagnosis [15], cultural heritage [16, 17], face recognition [18], food quality control [19], color management and printing [20, 21]. A number of applications are developed through multivariate analysis of hyperspectral images which include dried blood spots investigations [22], measurement of oil and protein contents in peanut varieties [23], classification of muscle foods [24], detection of defects on peaches [25] and numerous other applications.

Hyperspectral images contain more spectral information as compared to color images. The high dimensionality of hyperspectral image data is an open challenge, and a trade-off between high information content and practical handling is often required. To evaluate this compromise and assess its performance, high quality data with high spatial and high spectral resolutions are required. Creating such a dataset takes time and effort but it is necessary to the research community.

Considering these needs, we present a hyperspectral image dataset of 112 textured objects falling into five different categories. Images are reflectance data that spans visible and near infrared (NIR) parts of the electromagnetic spectrum. The data is referred to as HyTexiLa: Hyperspectral Texture images acquired in Laboratory. The idea behind creating this dataset is to provide a platform for benchmark analysis of various applications and processing. The areas where this data could be used are in the fields of image processing, computational imaging and computer vision, such as surface identification, spatio-spectral analysis of textured surfaces, image sensor simulation, color reproduction, image relighting and so on. The availability of a high spectral and spatial resolution dataset will provide an easy access for evaluation of different techniques and compare the results. In this paper, we present that dataset and analyze its spatial and spectral properties. We have defined the image acquisition protocol, the distortions that occur during the acquisition of the objects, the method for the correction of such distortions and the effect of such corrections. Our focus in this paper is to present a hyperspectral dataset to the community and provide analysis on our dataset as benchmark for further research. We provide the methodology for pre-processing the raw hyperspectral data before using it for the intended applications. This paper is organized as follows. First we analyze the existing hyperspectral datasets and then in Section 10.3, we discuss the hyperspectral image acquisition protocol and the processing of spectral data for addressing the distortions that occur during the acquisition. In Section 10.4, spectral analysis of the data is provided. Section 10.5 consists of the texture analysis of the object surfaces in our dataset, followed by the conclusion.

10.2 Comparison of HyTexiLa with existing hyperspectral datasets

Several hyperspectral image datasets are publicly available. In this section, we provide an overview of the existing hyperspectral image datasets that consist of close range images, including natural scenes and indoor images of objects. In this Section, we provide a brief review of existing hyperspectral image datasets and compare them with the properties of our proposed dataset. We do not include the remote sensing datasets in this Section because remote sensing does not assume the same imaging model. We also exclude reflectance datasets obtained from point-

wise spectrometers because they will not contain any spatial relationship between spectra, and those multispectral image datasets which contain less than 10 channels.

Table 10.1 gives an overview of the existing hyperspectral image datasets. The datasets in [26–34] consist of outdoor scenes which include vegetation and buildings. The wavelength range in [26, 30, 34–38] is from 400 nm to 700 nm while the datasets [27, 30, 31] have spectral range from 400 nm to 720 nm. The spectral sampling in [26, 27, 29–33, 35, 36, 39] is 10 nm. The dataset in [28, 32, 40, 41] have spectral range of 400 nm to 1000 nm. The datasets in [35, 36] consist of images of objects taken indoor. There are some datasets which consist of a specific category of objects like paintings [42], leaves from apple tree [43], textiles [39], wood samples [44], and honey samples [41]. Most of the available datasets are sampled or re-sampled at 10 nm intervals while some of them have different sampling intervals, e.g. [28] has 1.25 nm sampling step while datasets in [43] and [41] have the spectral sampling of 6 nm and 4.7 nm, respectively.

In this paper, we provide a new dataset of hyperspectral images that is made publicly available. This dataset consists of samples from five different categories of materials. The spectral range of this dataset is from 405.37 nm to 995.83 nm. As identified in Section 10.2, there are only few datasets that span both the visible and NIR region. The information in NIR is valuable for material classification [46, 47], identification of textile fibers [48] and minerals [49]. Although some of the hyperspectral datasets provide information in the NIR region, they are either for specific samples as in [41–43], or consist of outdoor scenes with many objects as in [28, 32]. Each image in our dataset consists of one specific, rather flat object. The spectral resolution of data is also an important aspect to be considered. In most of the hyperspectral image datasets mentioned in Tab. 10.1, the spectral sampling is 10 nm, which is considered enough for representation of spectra [50]. A spectral sampling of 5 nm is recommended for Munsell spectra in [50] but it is also worth noting that more closely sampled data may be required when spectra are less smooth. One of such cases is the presence of spiky illumination spectral power distributions, where the use of more accurate spectral intervals is required. The availability of high resolution spectral and spatial data will allow the development and testing of advanced resampling methods. There is an existing dataset for textile samples [39] but it contain spectral information in the visible region and the sampling is of 10 nm between the bands. Some datasets consist of food items [41, 45] but do not provide high spatial resolution. Similarly, the datasets containing images of urban and natural scenes do not provide high spatial resolution for each material. The CAVE dataset [35] contain detailed images of various materials that are captured from close range but they are in the visible region and the spatial

Table 10.1: Summary of various characteristics of existing and our proposed Hyperspectral image dataset (N/A means information Not Available from the dataset description).

Dataset	Nature of images	Camera	No. of images	Spatial Resolution	Spectral range [nm]	No. of channels
Bristol (1994) [26]	Outdoor scenes of vegetation	Pasecon integrating camera tube	29	256x256	400 - 700	31
Natural scenes (2002) [29]	Urban and rural scenes	Pulnix TM-1010 with VariSpec tunable birefringent filter	8	1024x1024	410 - 710	31
Natural scenes (2004) [30]	Reflectance image of natural scenes	Hamamatsu C4742-95-12ER with VariSpec liquid crystal tunable filter	8	1344x1024	400 - 720	33
Natural scenes (2015) [31]	Radiance image of natural scenes	Hamamatsu C4742-95-12ER with VariSpec liquid crystal tunable filter	30	1344x1024	400 - 720	33
Time-Lapse (2015) [27]	Images of 5 natural scenes taken at different times	Hamamatsu C4742-95-12ER with VariSpec liquid crystal tunable filter	33	1344x1024	400 - 720	33
ICVL (2016) [28]	Urban and rural scenes	Specim PS Kappa DX4	201	1392x1300	400 - 1000	519
Harvard (2011) [33]	Indoor and outdoor images	Nuance FX, CRI Inc.	50	1392x1040	420 - 720	31
UGR (2015) [32]	Outdoor scenes	Photon V-EOS	14	1392x1040	400 - 1000	61
CAVE (2008) [35]	Materials and objects	Apogee Alta U260 with VariSpec liquid crystal tunable filter	32	512x512	400 - 700	31
East Anglia (2004) [36]	Everyday objects placed in viewing booth	Applied Spectral Imaging Spectracube camera	22	Various resolutions	400 - 700	31
SIDQ (2015) [40]	Pseudo-flat objects	HySpex-VNIR-1600	9	500x500	400 - 1000	160
Brainard (1998) [37]	Indoor scenes	Kodak KA4200 CCD with Optical Thin Films filter	9	2000x2000	400 - 700	31
Nordic sawn timbers (2014) [44]	Wood samples	N/A	107	320x800	300-2500	440
Scien (2012) [45]	Various objects, scenes and faces	N/A	106	Various resolutions	Various range	Various channels
Paintings (2017) [42]	Paintings	IRIS II filter wheel camera	23	2560x2048	360 - 1150	23
Ancient manuscripts (2012) [38]	Printed documents	Hamamatsu C4742-95-12ER with VariSpec liquid crystal tunable filter	3	1344x1024	400 - 700	33
Apple tree leaves (2018) [43]	Near infrared images of healthy & infected leaves	HySpex SWIR-320m-e	N/A	Various resolutions	960 - 2490	256
SpecTex (2017) [39]	Textiles	ImSpector V8	60	640x640	400 - 780	39
Honey (2017) [41]	Honey samples	Surface Optic Corporation SOC710-VP	32	520x696	400 - 1000	126
Singapore (2014) [34]	Outdoor images of natural objects, man made objects, buildings	Specim PFD-CL-65-V10E	66	Various resolutions	400 - 700	31
HyTexiLa Our proposed dataset	Textured materials of 5 different categories	HySpex VNIR-1800	112	1024x1024	400 - 1000	186

resolution is low.

Our dataset is spectrally sampled at about 3.19 nm and can be used with simulation of most of the fluorescent tube and light emitting diode based illumination sources. The advantage of our dataset when compared with the already available databases is the availability of high resolution spectral and spatial data. The spectral range in our dataset covers both the visible and NIR region. Each image in the dataset contains a unique material and it can be used as ground-truth spectral data for surface identification applications. Our dataset is available for public use in the form of reflectance data in which each image is cropped to a uniform size. The raw data consists of radiance image captured along with the white diffuser and gray reflectance patch. Since the raw data is of large size (about 442 GB), we do not provide it publicly but if someone is interested in processing of the raw hyperspectral data, it can also be provided upon specific request to the authors.

10.3 Image acquisition and processing

In this Section, we present and discuss the objects being imaged and the hyperspectral image acquisition setup. The issues faced during image acquisition and protocols to correct for those issues are also part of this Section. We describe how we correct for distortions inherent to the system and how we compute reflectance images from the input radiance data.

10.3.1 Notations

We first list the notations and conventions used in this article. Functions are usually denoted in roman font, sets in calligraphic font, scalar variables in italic font. All indexes start from 0 and $[A]$ denotes the set of integers $\{0, 1, \ldots, A-1\}$. $\mathcal{B}$ is the set of wavelengths $\{\lambda^k\}_{k\in[K]}$ which ranges from 405.37 nm to 995.83 nm at 3.19 nm intervals. Hyperspectral images are denoted in bold font and spectral channels in italic font. $\mathbf{D}$, $\mathbf{I}$, $\mathbf{R}$ respectively denote the uncorrected (distorted) radiance image, the radiance image, the reflectance image. These images are made of K channels of $M \times N$ pixels associated to K spectral bands centered at wavelengths $\lambda_k \in \mathcal{B}$. D^k, I^k and R^k denote the k-th channel of respective hyperspectral images. A pixel coordinate p of an image is denoted by a couple of bounded integer values $p = (x, y) \in [M] \times [N] \subset \mathbb{Z}_+^2$. $\mathbf{D}_p$ (resp. $\mathbf{I}_p$, $\mathbf{R}_p$) and $\mathbf{D}(x, y)$ (resp. $\mathbf{I}(x, y)$, $\mathbf{R}(x, y)$) denote the vector (spectral) value of $\mathbf{D}$ (resp. $\mathbf{I}$, $\mathbf{R}$) at pixel $p = (x, y)$. Similarly, D_p^k (resp. I_p^k, R_p^k) and $D^k(x, y)$ (resp. $I^k(x, y)$, $R^k(x, y)$) denote the value of channel D^k (resp I^k, R^k) at pixel $p = (x, y)$. $\mathbf{S} = \{\mathrm{S}(\lambda^k)\}_{\lambda^k\in\mathcal{B}}$ is a vector of values of the spectral power distribution of the scene illumination S, sampled at each $\lambda^k \in \mathcal{B}$. Finally, $\lfloor\cdot\rceil$ denotes the nearest integer function, $\lceil\cdot\rceil$ denotes the ceiling function, $\langle\cdot,\cdot\rangle$ denotes the inner

product and $||\cdot||$ denotes the Euclidean norm.

10.3.2 Objects in the dataset

The different objects chosen for the dataset are divided into five classes: textile, wood, vegetation, food, stone. These samples are a set of textured object surfaces that exhibit different properties while remaining fairly flat. The imaged objects are reasonably flat so that the blurring effect due to distance variations from camera is minimal during the scanning process. After the completion of image acquisition, we analyzed each image thoroughly and discarded a few samples which were too blurry. We also excluded materials that exhibits high gloss in the considered viewing conditions to keep up with a simple imaging model as in Eq. (10.1), which allows an easy computation of reflectance values:

$$\forall k \in [K],\ \forall (x,y) \in [M] \times [N],\ I^k(x,y) = R^k(x,y) \cdot \mathrm{S}(\lambda^k). \tag{10.1}$$

Finally, for this dataset we provide 65 textile samples of various types and colors. They were selected among textile samples provided by the humanitarian association *Don De Soie* located in Templeuve in France. In the wood category, there are 18 different samples of wood that we could find in the area around Mjøsa lake in Norway. There are 15 images for the vegetation category, which consists of leaves, flowers and one mushroom. The food category consists of 10 samples of tea, coffee and spices, and there are 4 samples in the stone category. Image labels include the material category, the acquisition sequence number and a specific name if the material has been precisely identified. In case of textile, wood, and vegetation samples, labels also include the color of the material (suffix $\_colorname$ if there are at least two samples of the same type but with different colors) and/or the acquisition side (suffix $\_back$ if the back of the material has been acquired together with the front). We provide the sRGB rendering of each image along with the dataset for natural visualization of the imaged object. The sRGB values were computed from colorimetric computation on the spectral data (see Sec. 10.3.7).

10.3.3 Acquisition setup

We used the HySpex VNIR-1800 hyperspectral camera [51], manufactured by Norsk Elektro Optikk AS. It is a line-scan camera, consisting of a focusing mirror that allows only a narrow line of light in the scene to pass. This line is projected on a collimating mirror and then passed through a transmission grating. This grating separates the incoming light into different wavelengths and each wavelength is focused onto a detector array (sensor). The camera is coupled with a close-up lens with a working distance of 30 cm. The output from that array consists of a slice of hyperspectral image, having spatial information in one direction while spectral in-

formation in the other. A combination of these line scans provide a full resolution hyperspectral image of the object or scene. The process of line scan is illustrated in Fig. 10.1.

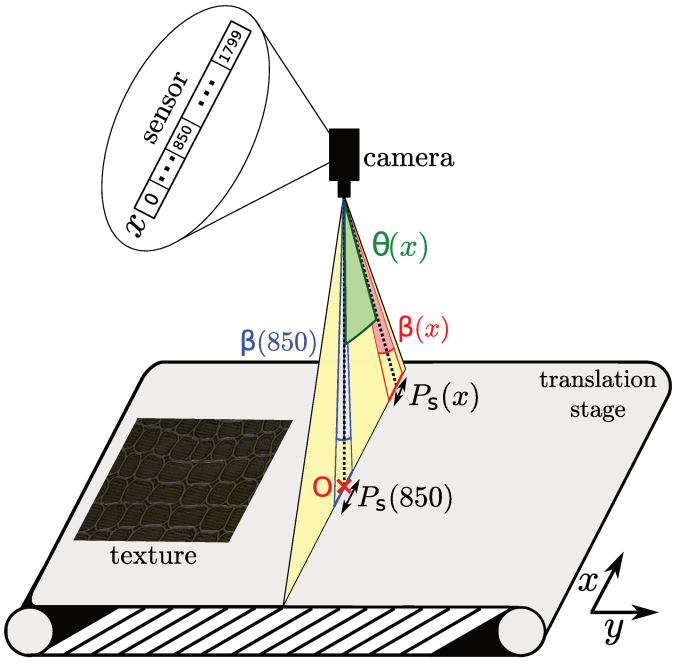


Figure 10.1: HySpex VNIR-1800 line scan acquisition process and cross-track distortion illustration on the x-axis. The camera acquires hyperspectral lines of pixels (x-axis). The hyperspectral image is obtained by translation of the object on the y-axis.

The output of HySpex VNIR-1800 camera consists of a radiance image $\mathbf{D}$ made of $K = 186$ channels associated to the spectral bands centered at wavelengths $\lambda_k \in \mathcal{B}$. The width of the acquired image is $M = 1800$ which corresponds to the number of photodetectors along the spatial dimension of the detector array of the camera (x-axis) while the height N is the number of pixels on the scanning direction (y-axis) and is physically bounded by the size of the scene and the translation stage. During the image acquisition, we observed spatial distortions in the output radiance images, which are addressed in the next Section.

10.3.4 Corrections of spatial distortions

We observed a *cross-track* distortion on the x-axis and a *shear* distortion on the y-axis on output images. These distortions were observed during the acquisition of test targets (grid images containing perfect squares, Macbeth ColorChecker, rulers). The correction of these distortions can be modeled by two functions h :

$[M] \times [N] \rightarrow [M]$ and $\mathrm{v} : [M] \times [N] \rightarrow [N]$, so that we set the value of pixel given by $(\mathrm{h}(x, y), \mathrm{v}(x, y))$ of the distorted image to pixel (x, y) of corrected image, formally:

$$\forall (x, y) \in [M] \times [N],\ \mathbf{I}(x, y) = \mathbf{D}(\mathrm{h}(x, y), \mathrm{v}(x, y)). \tag{10.2}$$

The function h considers cross-track distortion and is called *sensor correction* since the distortion only depends on the physical properties of the sensor of the camera and pointed out by the manufacturer as sensor model (see Sec. 10.3.4). The function v considers the shear distortion. This distortion is due to a non orthogonality between the sensor and the scan direction given by the translation stage. The function v is called *affine correction* since the shear distortion can be modeled by an affine transformation (see Sec. 10.3.4).

Cross-track distortion

The acquired images present a distortion across the scan direction called cross-track distortion. This distortion is due to the fact that each pixel is associated to a surface element whose size varies and is the same for each line of a channel, and for each channel. Figure 10.1 shows the whole field of view of the camera for a given line of pixels (yellow triangle). Point **O** (red cross) is denoted as the intersection of the optical axis of the camera and the translation stage. The camera associates 850 pixels to the left size of **O**, and 949 pixels to its right side as can be seen on the sensor of the camera. However, the surface area observed on each side of **O** is the same. Consequently pixels in the left side will be associated with higher surface elements (compression) than pixels in the right side (expansion). To overcome this distortion we use the sensor model function θ provided by the camera manufacturer that represents the field of view of each pixel in an acquired line of pixels. The whole field of view of the camera can be deduced as $\theta(1800) - \theta(0) = 0.29$ rad. We use $\theta(x)$ to compute the width of surface element $\mathrm{P_s}(x)$ (in mm) observed by each pixel as:

$$\forall x \in [M],\ \mathrm{P_s}(x) = H \cdot \left(\tan\left(\theta(x) + \beta(x)/2\right) - \tan\left(\theta(x) - \beta(x)/2\right)\right), \tag{10.3}$$

where $H = 340$ mm denotes the distance between the camera lens and **O**, and $\beta(x)$ represent the angular pixel size. The pixel size is undistorted at **O** ($x = 850$) and $\beta(850)$ equals to $1.6 \cdot 10^{-4}$ rad (data provided by the camera manufacturer). The angular pixels sizes $\beta(x)$ can be computed using the following recurrence:

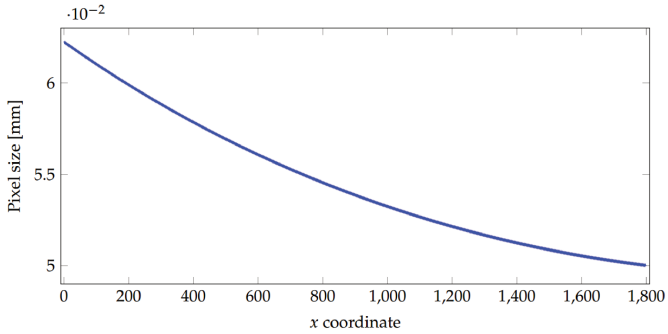


Figure 10.2: Size of surface element $\mathrm{P_s}(x)$ (in mm) observed by each pixel of coordinate x. Values are computed using Eq. (10.3) where values of $\theta(x)$ have been provided by the camera manufacturer.

$$\forall x \in [M],\ \beta(x) = \begin{cases} 1.6 \cdot 10^{-4},\ \text{if } x = 850, \\ 2\theta(x+1) - 2\theta(x) - \beta(x+1),\ \text{if } x < 850, \\ 2\theta(x) - 2\theta(x-1) - \beta(x-1),\ \text{if } x > 850. \end{cases} \tag{10.4}$$

Figure 10.2 shows the pixel size P_s w.r.t. x coordinate. Note that $\sum_{x=0}^{1799} \mathrm{P_s}(x) = 98.6$ mm is the width observed by our camera. To correct this distortion on acquired images, we measure the cumulative ratio $\mathrm{C_r}(x)$ with respect to $\mathrm{P_s}(850)$ as:

$$\forall x \in [M],\ \mathrm{C_r}(x) = \sum_{i=0}^{x} \left(1 - \frac{\mathrm{P_s}(i)}{\mathrm{P_s}(850)}\right). \tag{10.5}$$

Finally the sensor correction h is given by:

$$\forall (x, y) \in [M] \times [N],\ \mathrm{h}(x, y) = x + \lfloor \mathrm{C_r}(x) \rceil. \tag{10.6}$$

The pixel coordinate is rounded to the nearest integer to avoid artifacts due to interpolation of pixel values.

Shear distortion

After the sensor correction, the image still exhibits a distortion on the y-axis due to geometry set-up. Indeed, the direction to which the material moves and the line sensor are not exactly perpendicular. This implies that a rectangle in the scene becomes a parallelogram in the acquired image (shear distortion). This can be

assumed as an affine transform, leading to the correction:

$$\forall (x, y) \in [M] \times [N],\ \mathrm{v}(x, y) = \lfloor ax + by \rceil, \tag{10.7}$$

and the transform only depends on two parameters a and b. The pixel coordinate is again rounded here to the nearest integer to avoid interpolation artifacts. These parameters are experimentally estimated using the following method: we acquire 10 images of a grid with different rotations (the rotation angle ranges from 0 to $\pi/4$). Note that the grid is originally made of perfect squares: four equal sides and four equal angles ($\pi/2$ rad). Then we apply the transform given by Eq. (10.7) on the grid images after sensor correction for different values for a and b. Finally, we measure the mean of the angles γ between "horizontal" and "vertical" lines obtained after Otsu thresholding [52] and Hough transform [53] for all sensor-corrected grid images and three spectral channels, and retain the values of a and b that minimize $|\gamma - \pi/2|$ (i.e. values of a and b that create right angles on grid images). We use here Otsu and Hough techniques since they require few parameters and are adequate to estimate lines in these simple grid images which are just made of dark straight lines on a light background. The angular resolution for Hough transform is set to $\arctan(2/1800) = 1.2 \cdot 10^{-3}$ rad and represents the smallest angle between two lines of pixels in an image acquired by our camera. Optimal values of a and b give:

$$\forall (x, y) \in [M] \times [N],\ \mathrm{v}(x, y) = \lfloor -0.021x + 1.032y \rceil, \tag{10.8}$$

Table 10.2 shows mean and standard deviation of $|\gamma - \pi/2|$ before and after applying the correction of Eq. (10.8). As can be seen, after correction, the mean γ of the angles of the squares of the grid images is close to $\pi/2$ with a difference of $6 \cdot 10^{-4}$ rad which is lower than the angular resolution ($1.2 \cdot 10^{-3}$ rad).

Table 10.2: Mean and standard deviation of $|\gamma - \pi/2|$ (in radians) on 10 grid images with different rotations before and after applying affine correction of Eq. (10.8), where γ is the angle between a "horizontal" and a "vertical" line. As can be seen, the difference between γ and $\pi/2$ is strongly reduced (i.e. γ close to $\pi/2$) after affine correction and is lower than the angular resolution ($1.2 \cdot 10^{-3}$ rad).

	Value of $\lvert\gamma - \pi/2\rvert$
after sensor correction	0.0167 (0.0098)
after sensor and affine correction	0.0006 (0.0003)

Figure 10.3 represents the vector field of the correction on an image of size 1800×2000 pixels. The origin of each vector is $(\mathrm{h}(x, y), \mathrm{v}(x, y))$, and its end is (x, y). Both corrections are translations along x-axis (Eq. (10.6)) or y-axis (Eq. (10.8)), so

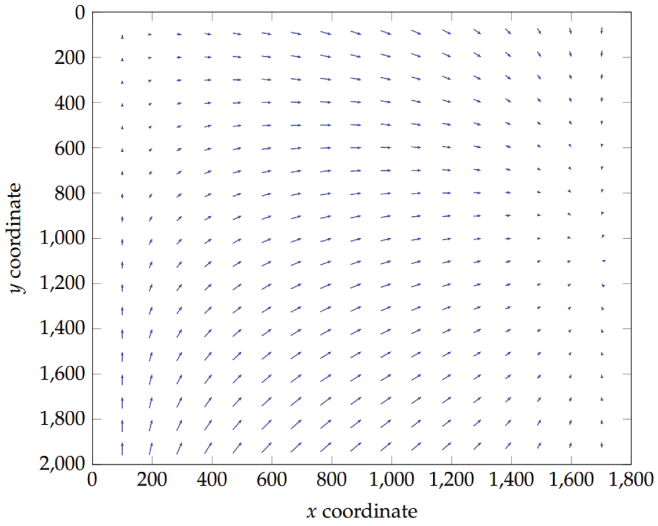


Figure 10.3: Vector field of both x-axis and y-axis corrections. The origin of each vector is $(\mathrm{h}(x,y),\mathrm{v}(x,y))$, and its end is (x,y). For clarity sake, only vector whose final coordinates (x,y) are multiple of 100 are shown here.

they can be done jointly or separately. At the end, we obtain, for each acquisition, a corrected radiance image $\mathbf{I}$ made of $K = 186$ channels.

10.3.5 Impact of the corrections on pixel resolution

To quantify the impact of the sensor and affine correction on pixel resolution, we acquire an image of a ruler with equally spaced markings of 1 mm along x-axis, then we measure the pixel resolutions in pixels.mm^{-1} thanks to the spaces between the markings of the single-channel image resulting from an average on 19 spectral channels without corrections, after sensor correction and after sensor and affine correction. Table 10.3 shows mean and standard deviation of these resolutions.

As can be seen, the pixel resolutions are similar on average before and after applying sensor correction (without corrections, pixel widths are higher to the left of the intersection $\mathbf{O}$ between the translation stage and the optical axis than to the right, see Sec. 10.3.4) but the standard deviation is reduced by 2 along x-axis which means that the sensor correction correctly decreases the cross-track distortion. Moreover, we notice that the pixel resolution on average is close to the theoretical one given by $1/\mathrm{P}_s(850) = 18.38$ pixels.mm^{-1} (see Eq. (10.3)) and that the affine correction does not impact pixel resolution along x-axis.

To illustrate the impact of the corrections, images of the 49^{th} channel (centered at

Table 10.3: Mean and standard deviation of pixel resolutions (in pixels.mm^{-1}) measured between each marking or a ruler for a single-channel image resulting from an average on 19 spectral channels without corrections, after sensor correction and after sensor and affine corrections. As can be seen, the pixel resolutions are similar on average before and after applying sensor correction but the standard deviation is reduced by 2 along x-axis which means that the sensor correction correctly decreases the cross-track distortion. Note the affine correction does not impact pixel resolution along x-axis.

	Ruler along x-axis
without corrections	18.30 (1.21)
after sensor correction	18.28 (0.60)
after sensor and affine correction	18.31 (0.61)

561.76 nm) of the acquisition of the grid made of perfect squares without corrections, after sensor correction and after sensor and affine correction are shown in Fig. 10.4. As can be seen, without corrections, the pixel resolution is increasing w.r.t. x coordinate (sides of square are higher to the right of the image than to the left). After sensor correction or both corrections, the pixel resolution is constant w.r.t. x coordinate (side of squares remains constant) which means that the cross-track distortion has been reduced. Moreover, we observe on the image after both correction that the shear distortion has been reduced (squares have right angles).

10.3.6 Reflectance computation

In this Section, we describe how we obtain the reflectance image $\mathbf{R}$ of a texture from the corrected radiance image $\mathbf{I}$. To properly estimate the reflectance on each surface element of a texture, we first estimate the spectra $\mathbf{S} = \{\mathrm{S}(\lambda^k)\}_{\lambda^k \in \mathcal{B}}$ of the illumination system provided with the HySpex VNIR-1800. This illumination system (3900e DC Regulated ER Light source) is manufactured by Illumination Technologies, Inc. [54].

For the estimation of the spectral power distribution of this illumination, we use the SG-3051 SphereOptics Diffuse Reflectance Tile [55] which provides a reflexivity of 99% in the range $[400, 1000]$ nm. We then acquire each texture simultaneously with the SG-3051 tile to produce uncorrected radiance images $\mathbf{D}$, then we correct it using Eq. (10.2) to obtain radiance images $\mathbf{I}$. We noticed that the SG-3051 tile has a rough texture that produces shaded areas in the acquired images. To produce robust estimation of the illumination, we adopt the following strategy for each acquired texture: we select a sub-image $\mathbf{P} = \{P^k\}_{k \in [K]}$ of 550×550 pixels centered in the SG-3051 tile in the radiance image $\mathbf{I}$, then we compute the summation image P of all channels: $P = \sum_{k \in [K]} P^k$ and retain the set $\mathcal{Q}$ of pixels coordinates (x, y) whose values in P are greater than the 95th percentile. Finally, the estimation of

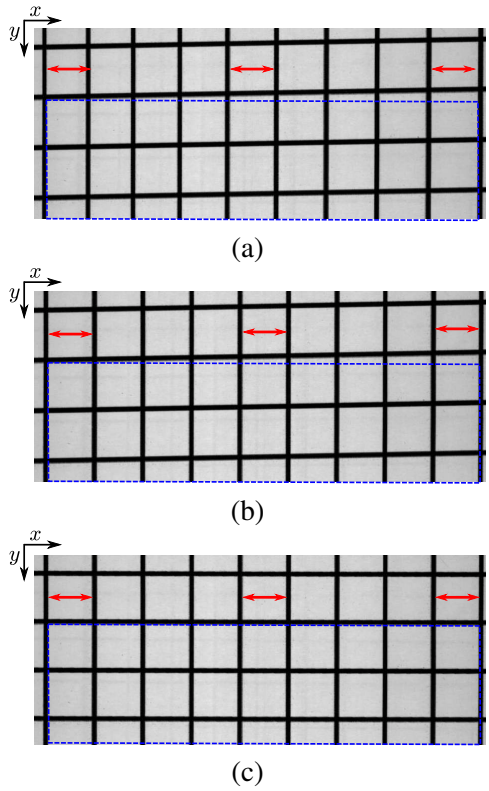


Figure 10.4: Images of the 49^{th} channel of the acquisition of the grid made of perfect squares without corrections (a), after sensor correction (b) and after sensor and affine correction (c). Without corrections, the pixel resolution is increasing w.r.t. x coordinate (sides of square are higher to the right of the image than to the left). After corrections, the pixel resolution is constant w.r.t. x coordinate (side of square remains constant). Moreover, the shear distortion is reduced after sensor and affine correction since square have right angles. Red arrows and blue rectangles have been added as overlays to obtain a better view on distances and angles. Note that they are identical (same lengths) from one image to another.

the illumination is given by taking, for each channel P^k, its median value among pixels in $\mathcal{Q}$ (to avoid impact of saturated values), formally:

$$\forall k \in [K],\ \mathrm{S}(\lambda^k) = \mathrm{Median}\left(\left\{P^k(x,y)\right\}_{(x,y)\in\mathcal{Q}}\right). \tag{10.9}$$

Now we can compute the reflectance image $\mathbf{R}$ by dividing each channel of radiance image $\mathbf{I}$ by its corresponding value in $\mathbf{S}$, inverting Eq. (10.1):

$$\forall k \in [K],\ \forall (x,y) \in [M] \times [N],\ R^k(x,y) = \frac{I^k(x,y)}{\mathrm{S}(\lambda^k)}. \tag{10.10}$$

Reflectance images are cropped to 1024×1024 pixels that corresponds to the minimum of the most relevant regions of interest (ROIs) of the 112 selected textures (ROIs that really contain texture elements). Figure 10.5 shows a sRGB rendering of two radiance images. For each radiance image $\mathbf{I}$, the subimage $\mathbf{P}$ selected for illumination estimation is displayed as red dashed square and the texture region that we retain as reflectance image $\mathbf{R}$ support is displayed as green solid square. Note that pixel values that undergone specular reflection of the light in the radiance image can have values that are higher than the SG-3051 tile and produce values over one in the reflectance image. We decide to keep them unchanged in the final dataset so that the original corrected radiance image can be retrieved by a multiplication with the original illumination spectra (which is provided in the metadata of the reflectance image file, see Sec. 10.3.7).

As a reference, we also compute the reflectance image of a Macbeth ColorChecker. Figure 10.6 shows the reflectances of the 24 color patches on average on 200×200 pixels for each patch. As for the reflectance measurements in [56], reflectance is mostly flat in the NIR.

10.3.7 Dataset description

The dataset will be available on the websites of the team/laboratory of the authors[2]. Each one of the 112 texture images is provided as both (i) hyperspectral reflectance image in *ENVI* format and (ii) sRGB reconstructed image in *PPM* format. (i) The *ENVI* format consists of two files: a binary file (extension *.raw*) of single-precision floating-point variables (32-bit) that stores the raw hyperspectral reflectance image in the *BSQ* (Band SeQuential) data organization (channel by channel) and a text file (extension *.hdr*) which contains the metadata of the image. The *.hdr* text file holds:

[2] http://color.univ-lille.fr/datasets and https://www.ntnu.edu/web/colourlab/software.

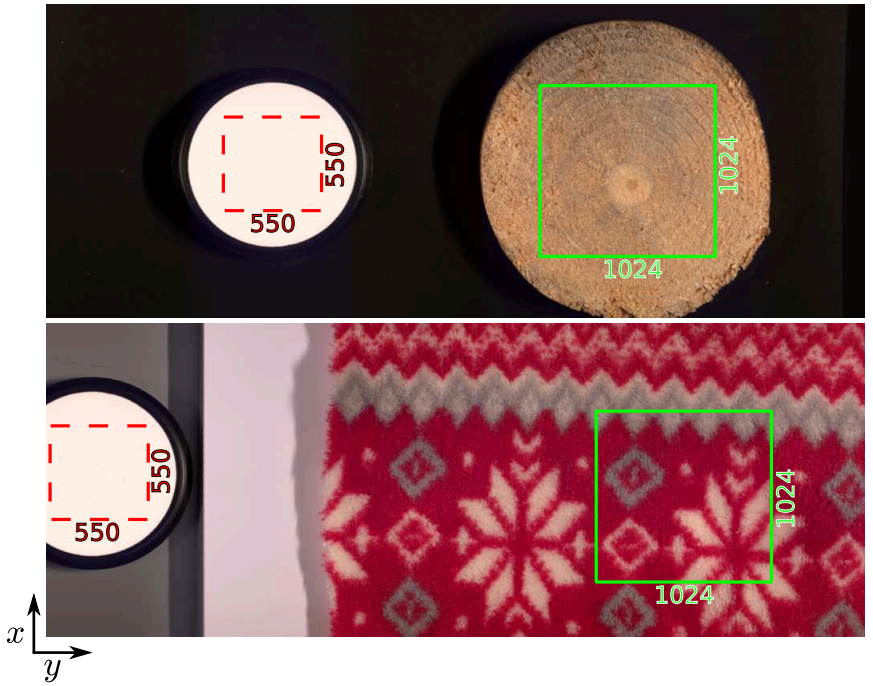


Figure 10.5: sRGB rendering of two corrected radiance images. The subimage (red dashed square) inside the SG-3051 reflectance tile is selected for illumination estimation and the texture region (green solid square) is retained as final reflectance image support.

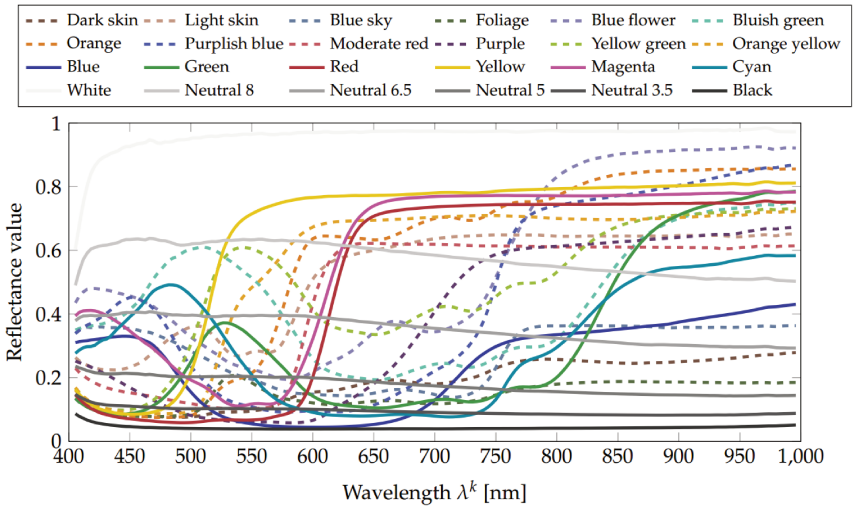


Figure 10.6: Spectral reflectances of the 24 patches of a Macbeth ColorChecker computed from the data acquired by the HySpex VNIR-1800 (on average on 200×200 pixels for each patch). The color of a curve is the sRGB color of the corresponding patch in the ColorChecker.

- camera parameters description (i.e. camera ID, integration time, aperture size, etc.). These metadata have been written by the HySpex acquisition software and we decided to keep them unchanged.

- required informations to open the associated binary raw file (i.e. image size, number of bands, data type, etc.),

- default bands (65, 45, and 18) whose band centers match with primary Red, Green and Blue of sRGB standards. Generally used to generate false color RGB image using 3 channels.

- wavelength array that contain values of $\lambda^k \in \mathcal{B}$ in nanometers such as center wavelength of channel R^k is λ^k,

- illumination array that contain the values of illumination $\{\mathrm{S}(\lambda^k)\}_{\lambda^k \in \mathcal{B}}$ provided by Eq. (10.9). The illumination can be used to compute the corrected radiance channels I^k using Eq. (10.10). The average value of this illumination over the 112 images is provided in Fig. 10.7. As can be seen, illumination around 400 is weak and consequently, respective channels in reflectance images are likely to undergo noise [57]. We then measure noise power by analyzing the standard deviation on 200×200 pixels on each gray patch

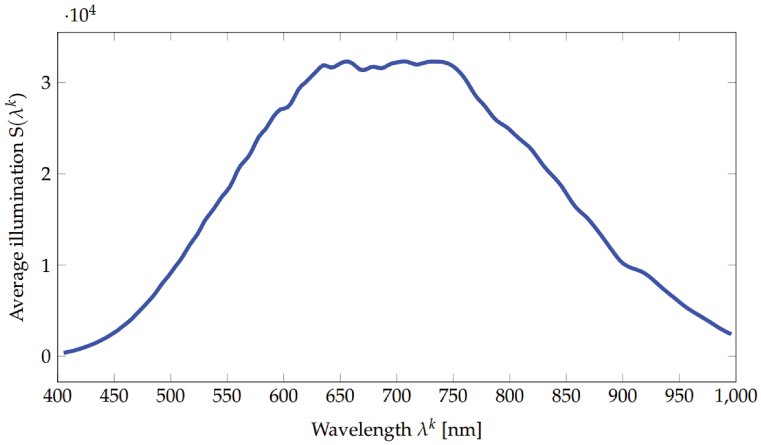


Figure 10.7: Average illumination computed over the 112 radiance images.

of the reflectance image of the Macbeth ColorChecker on Fig. 10.8. In addition to illumination effect, the problem of having a good signal in low wavelengths is classically due to optics and low sensor sensitivity in this area where we are at the limit of the optical model we use. Nevertheless we chose to provide these data and let the users decide if they want to use them in their simulations.

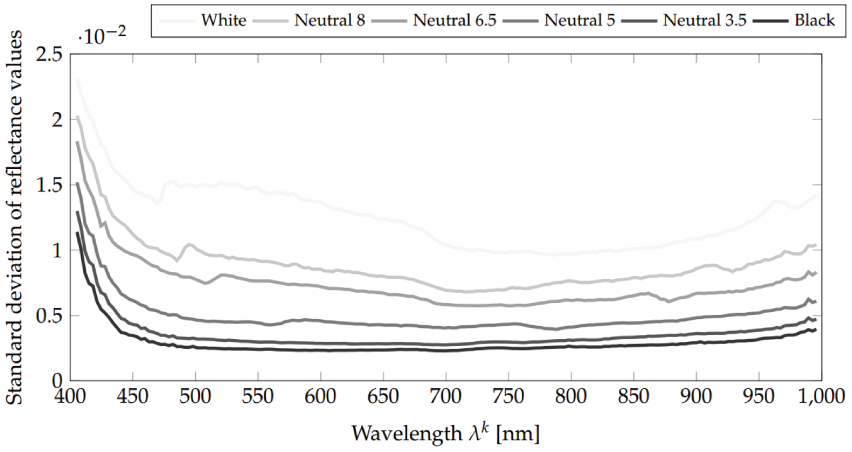


Figure 10.8: Standard deviations of spectral reflectances for 200×200 pixels of each gray patch of a Macbeth ColorChecker.

(ii) the reflectance images are converted from the 186-channel domain to the sRGB color space. This transformation is performed using the CIE XYZ 2° standard

observer functions. Then the reference white D65 is used to convert values from XYZ to the sRGB color space. The resulting sRGB images are provided in *PPM* format. Figures 10.9 and 10.10 show the 112 sRGB samples of textures with filenames. Moreover, we provide ImageJ plugins and Python code for opening and processing the data. These codes are available as supplementary material.

10.4 Spectral dimension analysis

In hyperspectral images, the adjacent spectral bands are highly correlated [58]. Even though there are many good arguments for processing hyperspectral images at full spectral resolution [59], often dimensionality reduction may be beneficial. It is performed by approximating the original hyperspectral data to a varying degree of precision with a smaller number of dimensions. The dimensionality or complexity of acquired reflectance spectra define the dimension of the vector space that is required to describe a spectrum.

There are various methods for computation of effective dimension of spectral data, which can mainly be divided into supervised and unsupervised techniques. Supervised methods aim at preserving the a priori information while the unsupervised methods do not assume any prior information about the input data. Among unsupervised, principal component analysis (PCA) [60, 61] and independent component analysis (ICA) [62] are the most popular. PCA aims to maximize the variance of input data by projecting in into the orthogonal subspace while ICA looks for the source of changes in the input data and these sources are assumed to be statistically independent from each other. Among the supervised methods, Fisher's linear discriminant analysis (LDA) [63] and discriminative locality alignment (DLA) [64] are the prominent techniques and many variants of these techniques are proposed in literature. A review of linear and non-linear dimensionality reduction methods can be found in [65].

Despite several studies, there is still no agreement on the number of effective dimensions to represent the spectral data. There are different conclusions by various authors for the number of effective dimensions that can represent the spectral information. As an example, for the Munsell reflectance dataset, different authors have provided different dimensions that effectively represent it, ranging from 3 to 18 dimensional representations [66]. The required dimensionality of data mainly depends on the statistical properties of the spectra being considered and the information that one wants to preserve or emphasize. A review of the work done for finding the dimension of spectral data is provided in [66].

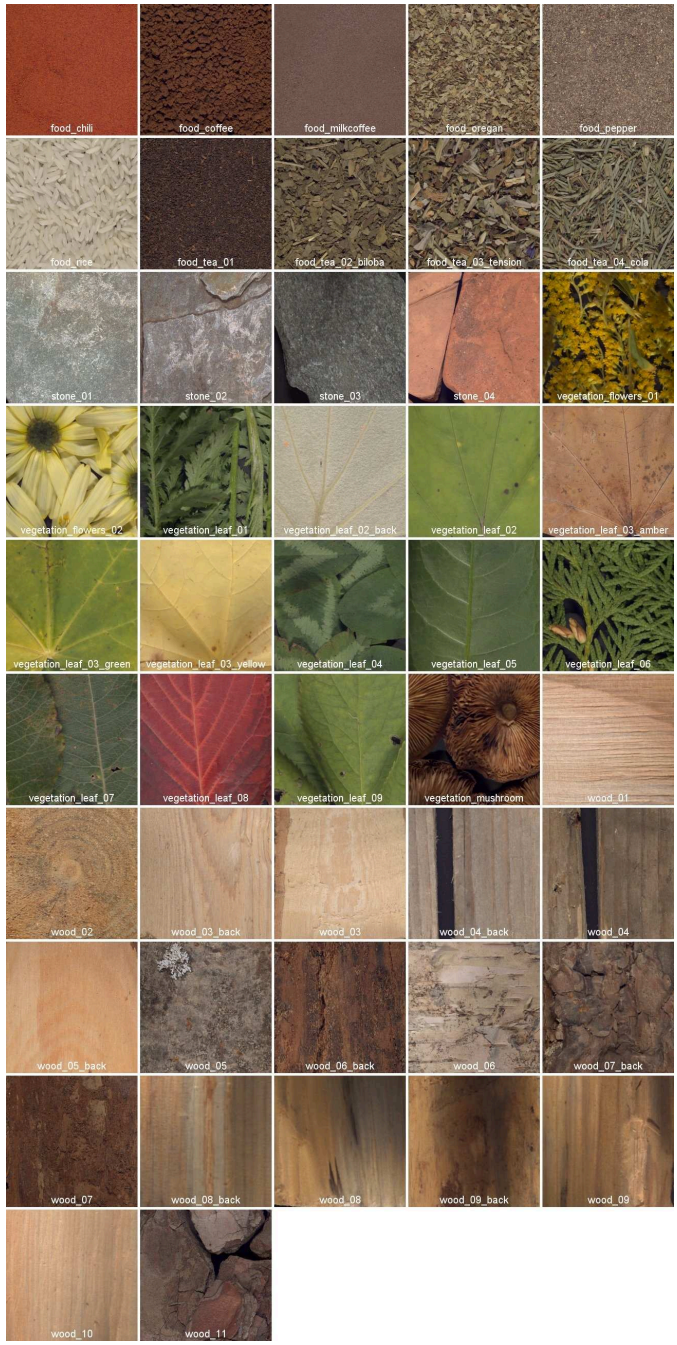


Figure 10.9: Reconstructed sRGB images from food, stone, vegetation and wood categories of the dataset.

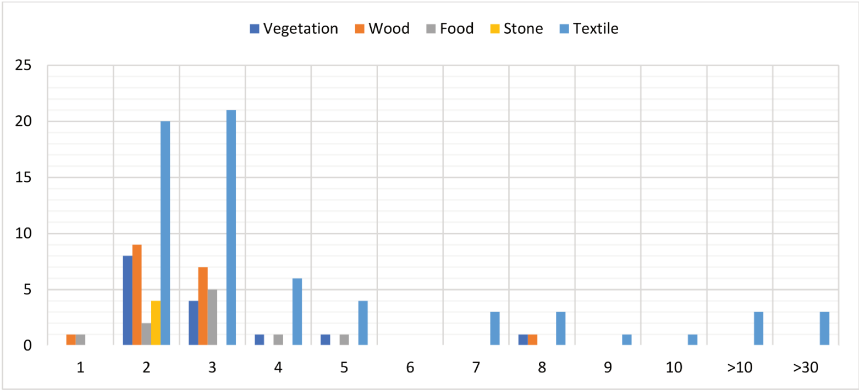
Figure 10.10: Reconstructed sRGB images from textile category of the dataset.

10.4.1 Spectral analysis of the proposed dataset

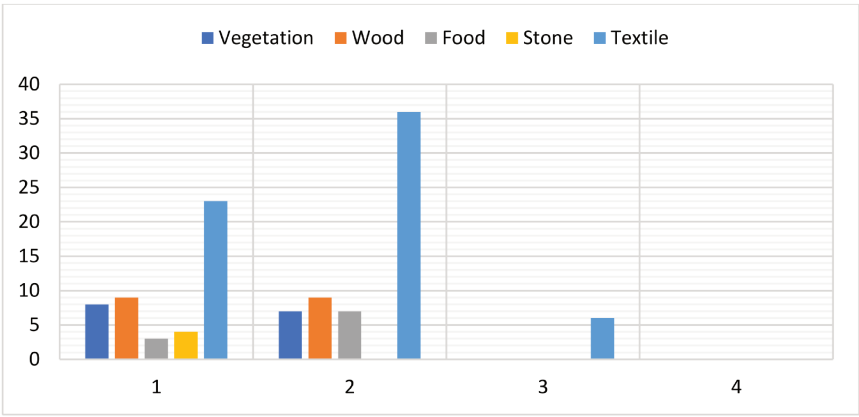
In this section we investigate on the effective spectral dimension of the acquired hyperspectral images. We do not assume any a priori knowledge about the data and provide a benchmark analysis in spectral domain. Having these goals in mind, we use the principal component analysis (PCA) technique, which is mainly used to reduce the dimensionality of data consisting of a large number of inter-related variables, while retaining the variance in data [61]. PCA has the ability of de-correlating the spectral data and providing the variance information through simple computational methods. This ability makes it one of the most popular techniques for computation of effective spectral dimension in hyperspectral imaging [67].

For the computation of effective spectral dimensions, each image is processed individually. First, a covariance matrix is computed and then eigen decomposition is applied on this matrix for obtaining the eigenvalues and eigenvectors. The eigenvectors are arranged in the order relative to their eigenvalues in descending order. These eigenvectors consist of the characteristic reflectances of the input hyperspectral image. For each eigenvalue, there exists corresponding eigenvectors and the strong concentration of variance is present in the first few eigenvectors. If reflectance spectra are linearly independent, then the eigenvalues will all be significantly above zero. It represents that each wavelength contain independent information and are not redundant. However, in case of hyperspectral imaging, typically only a few eigenvalues are significantly higher while the remaining values are close to zero, which indicates that their corresponding eigenvectors contain less variance and are mostly redundant. For an effective representation of the spectral data, we can select only the first few eigenvalues and the projection of spectra on the corresponding eigenvectors are able to provide adequate representation of the acquired data. It means that if we select the principal components containing 95% variance, the rest of principal components can be discarded with a loss of 5% of information. The decision of keeping a certain percentage of variance depends on the application for which a reduced set of representative data is required. For this analysis, we compute the effective dimension for each hyperspectral image which retain 99% and 95% of the total variance [66].

Figure 10.11(a) shows the number of components required to retain 99% of variance and Fig. 10.12 gives information about the effective spectral dimensions of the whole dataset and also on each of the five categories of materials. It can be observed that most of the hyperspectral images show an effective dimension of two and three. All of the “Stone” category images have effective dimension of two, while most of the “Food” category show effective dimension of three. Most of the “Textile” category images have effective dimension between two and five, while some of the images have the effective spectral dimension going beyond seven.



(a) Required No. of components that retain 99% variance



(b) Required No. of components that retain 95% variance

Figure 10.11: Result of PCA analysis on each image from the 5 categories of the hyperspectral dataset. The x-axis represent the effective dimension while the y-axis shows the number of images having that effective dimension. Most of the spectral data in our dataset can be represented by 1, 2 or 3 dimensions while retaining 95% of the variance.

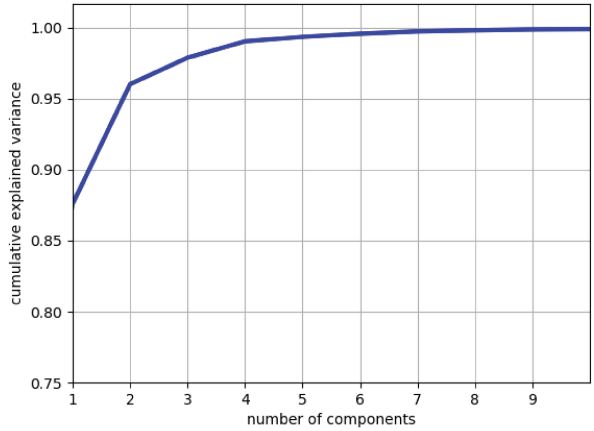
Among them, three images show effective dimension of 33, 47 and 58. We further examined the image that showed the highest number of dimension (*textile_05* image) and found that 82.55% of variance is retained by the 1st PC while 12.02% variance is retained by the 2nd component. The rest of components add very little to the information, for example the 3rd PC contain 0.7% of variance in this image. If the desired value of retained variance is reduced to 95%, then most of the images in our dataset can be defined with one or two principal components while six images from the "textile" require third component. Fig. 10.11(b) represent the number of images that require a certain number of dimensions for retaining 95% of variance. From these figures, it can be observed that the number of effective dimension is dependent on the nature of materials being imaged and may be used as a feature for estimating the category of imaged object. The reason that some of the images require higher dimension for the description of their spectral properties is due to the presence of anomalies, for example a tiny part with specular reflection, which causes the effective dimension to rise. Fig. 10.11 shows that by reducing the threshold for retained variance, the number of dimensions for representation of the spectral data is reduced.

10.4.2 Interpretation of the effective dimension

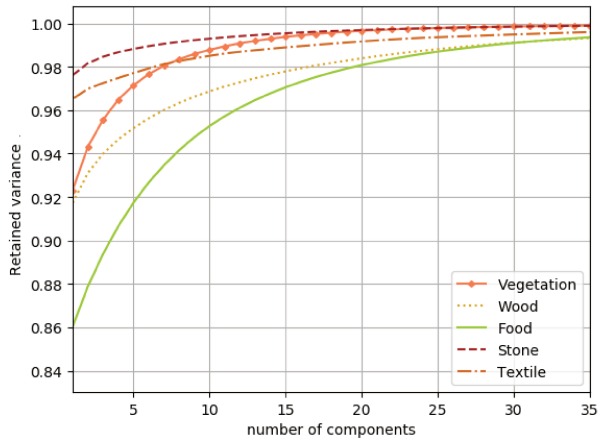
The effective dimension analysis of our dataset on the basis of PCA, as in in Sec. 10.4.1, provides a number indicating the number of dimensions required to retain a certain percentage of variance. However, it does not provide a physical meaning to the retained spectra since the principal components contain both positive and negative values. In order to obtain the information about the retained spectra, we use the non-negative matrix factorization (NMF) [68] based decomposition of the hyperspectral images in our dataset. NMF is used in remote sensing hyperspectral images for spectral unmixing of the acquired data [69, 70]. The advantage of NMF over PCA that it extracts sparse and meaningful features from the given data and the non-negativity constraint ensures the output to be interpretable in terms of its physical meaning.

NMF is able to compute the spectral signature of end-members in a given hyperspectral image and the presence of each end-member in each pixel, in the image in form of abundance maps. The assumption behind this computation is the *linear mixing model* which states that the spectral signature of each pixel is a linear combination of the spectral signature of end members at a particular spatial location.

In the NMF based decomposition, the number of components to be retained is provided by the user, which is mostly based on the prior-knowledge of the end-members. For our dataset, we use PCA to compute the required number of components. For the illustration of NMF based decomposition results, we provide

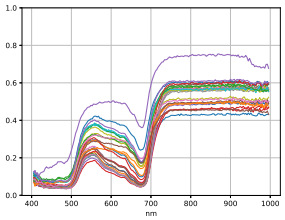


(a) Spectral analysis of whole dataset



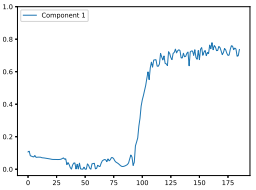
(b) Spectral analysis of each category of items in the dataset.

Figure 10.12: No. of components (x-axis) required to retain variance percentage (y-axis). (a) show the result for whole dataset while (b) shows the result for each of the five categories in the dataset. "Food" category images require more dimensions for their representation due to the variety of objects in that category, while "Stone" category require the least number of dimensions since the surfaces exhibit similar spectral properties.



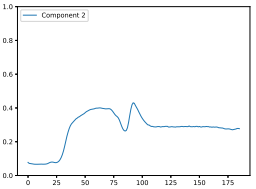
(a) Spectral samples from *vegetation_leaf_03_green* image

(b) Color rendering of *vegetation_leaf_03_green* image

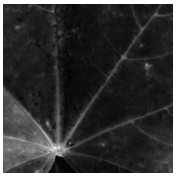


(c) First component from NMF decomposition

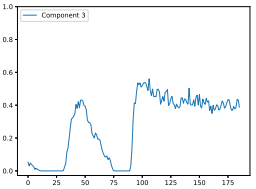
(d) Abundance map of first component, indicating the presence of chlorophyll in leaf.



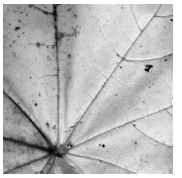
(e) Second component from NMF decomposition



(f) Abundance map of second component consisting of veins and the lower left part of leaf.

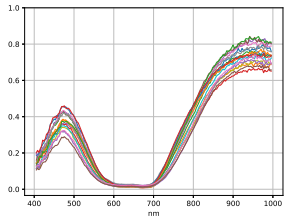


(g) Third component from NMF decomposition

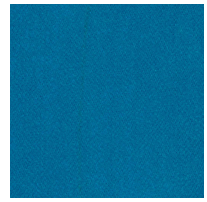


(h) Abundance map of third component indicating the green color of leaf

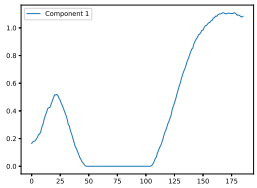
Figure 10.13: NMF based decomposition of *vegetation_03_green* image in 3 components. Each component exhibits the spectral composition of leaf, showing the spectra of chlorophyll, veins and the color of leaf.



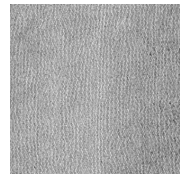
(a) Spectral samples from *textile_22_cyan* image



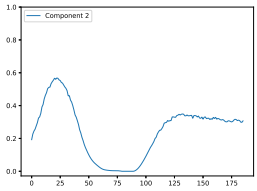
(b) Color rendering of *textile_22_cyan* image



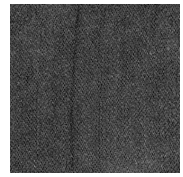
(c) First component from NMF decomposition



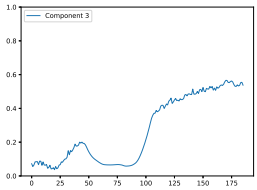
(d) Abundance map of first component showing the regions in direct contact with illumination



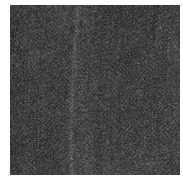
(e) Second component from NMF decomposition



(f) Abundance map of second component consisting of areas in shadow.



(g) Third component from NMF decomposition



(h) Abundance map of third component indicating the presence of a different material in the textile sample.

Figure 10.14: NMF based decomposition of *textile_22_cyan* image in 3 components. The extracted components consist of areas under direct illumination, shadow areas and the presence of different material in the textile sample.

two examples in Fig. 10.13 and 10.14, containing samples of leaf and textile, respectively. Fig. 10.13(b) is the color rendering of *vegetation_leaf_03_green* image from our dataset and Fig. 10.13(a) shows spectra from randomly selected pixels in this hyperspectral image. We decompose the image in three components and the spectral signatures of computed components are shown in Fig. 10.13(c), 10.13(e) and 10.13(g). The corresponding abundance maps in Fig. 10.13(d), 10.13(f) and 10.13(h) show the presence of extracted components in the image. The first component in Fig. 10.13(c) is significantly reflective in the near-infrared region while showing lower response in the visible region. This spectra suggests that this component is providing the information of chlorophyll in the leaf and the abundance map in Fig. 10.13(d) show the areas where the presence of chlorophyll is prominent. The second component is formed by the veins and lower left part of the leaf as can be seen in Fig. 10.13(f). The green color of leaf corresponds the the third component as seen in Fig. 10.13(g) and the abundance map in Fig. 10.13(h) provides the information about distribution of this component in the leaf.

Fig. 10.14 shows the decomposition results of *textile_22_cyan* image. The first and second components contain information of the reflectance from textile surface which causes the particular color appearance of it. It is interesting to note that although the spectral signatures of first two components are almost the same in shape in the visible region but have difference in intensity in the near-infrared region. Upon a closer inspection of the abundance maps in Fig.10.14(d) and 10.14(f), it is observed that the first component consist of directly illuminated regions while the second component is composed of regions in shadows. There is a third component in this sample which corresponds to a line in the abundance map (Fig.10.14(h)) and is probably caused due to some component in the textile which is not easily visible in a color image.

Although we have provided only two examples from NMF based decomposition of the spectral data, the results indicate that such an analysis can provide interesting information about the spectral composition of the materials present in an image. For example, the amount of chlorophyll in vegetation, information of pigments and dyes in textile, and the quality of food materials on the basis of their spectral analysis. We have used PCA to get information about the spectral dimension before decomposing it through NMF. From the comparison of spectral behavior and the effective dimension, it can be seen that effective spectral dimension of most of the objects in our dataset is within two to eight. However, this dimension should not be directly related to the required number of spectral channels needed for acquisition. Although there are some studies which aim at finding the number of filters that can best serve for acquiring the maximum spectral information [71], there is no agreement on the optimum number of channels since it is dependent on the nature

of surface spectral properties. Nevertheless, the knowledge of effective spectral dimension is useful for dimensionality reduction and obtaining the components of spectra by decomposing it. In the next Section, we study the impact of dimensionality reduction on classification performances on the hyperspectral dataset using Local Binary Pattern (LBP)-based descriptors.

10.5 Texture classification

A texture image is a characterization of the spatial and spectral properties of the physical structure of a material or an object. Texture analysis can be seen as a set of measures that quantify and provide information about the spatial arrangement of the spectral responses of an object. In texture classification the goal is to assign an unknown sample texture image to a known texture class. For this purpose discriminant texture features are extracted from test images and compared to those extracted from training images whose classes are known. To extract a feature from an image, a descriptor is needed. Local Binary Pattern (LBP) is one of the most robust feature descriptor which characterizes the local level variation in a neighborhood of each pixel. Due to its discrimination power and computational efficiency, LBP has become a popular approach in various applications [72]. LBP-based texture classification has first been performed on gray-level images since the original operator only uses the spatial information of texture [73]. Later, [74] has shown that classification based on a color analysis outperforms the single use of intensity spatial information. Thus, texture feature extraction is extended to the color domain by taking both spatial and color textural information into account. Here we propose to extend the color LBP-based descriptors to any K-channel radiance image $\mathbf{I}$, then we propose to study the discriminative power of these descriptors on our database with respect to increasing number of channels.

10.5.1 Texture features based on local binary patterns

The LBP operator [73] characterizes the level variations in the neighborhood $\mathcal{N}^{P,d}$ made of P pixels at spatial distance d from each pixel p of a gray-level image. The outputs of the LBP operator over the multi-channel image are summarized as a concatenated histogram to form a texture feature. Here we extend four color LBP operators to hyperspectral LBP operators, so that, the LBP operator is applied on a K-channel radiance image $\mathbf{I}$. This radiance image $\mathbf{I}$ is rendered using both a reflectance image $\mathbf{R}$ whom values above 1 (that correspond to specular surfaces) are clipped to 1 and quantized on 8-bit, and the standard illuminant E that light the whole spectrum uniformly. Note that we only hold the reflectance information by using illuminant E to be independent of illumination properties in our study. Moreover, data are quantized on 8 bits since channels of digital radiance and color images are often quantized on integer values.

We now detail each LBP operator and how it compares the radiance value I_p^k of a channel k with those in the neighborhood $\mathcal{N}^{P,d}$ with respect to the value of p.

- ***Marginal LBP (LBP):*** The basic LBP operator is applied marginally to each channel I^k at each pixel p as [75]:

$$\mathrm{LBP}^k[\mathbf{I}](p) = \sum_{q \in \mathcal{N}^{P,d}} s(I_q^k, I_p^k) \cdot 2^{\epsilon(q)}, \tag{10.11}$$

  where $\epsilon(q) \in [P]$ is the index of each neighboring pixel q of p in $\mathcal{N}^{P,d}$, and the sign function s is defined as: $s(\alpha, \beta) = 1$ if $\alpha \geq \beta$, 0 otherwise. The final texture feature results from the concatenation of the K 2^P-bin unnormalized histograms of $\{LBP^k[\mathbf{I}]\}_{k \in [K]}$ and its size is $K \cdot 2^P$.

- ***The Opponent Band LBP (OBLBP):*** [76] improves the LBP operator by taking the inter-channel correlation into account. For this purpose, they consider the opponent band (OBLBP) operator of each pair of channels (I^k, I^l), $(k, l) \in [K]^2$:

$$\mathrm{OBLBP}^{(k,l)}[\mathbf{I}](p) = \sum_{q \in \mathcal{N}^{P,d}} s(I_q^l, I_p^k) \cdot 2^{\epsilon(q)}. \tag{10.12}$$

  The final texture feature results from the concatenation of the K^2 histograms of $\{\mathrm{OBLBP}^{(k,l)}[\mathbf{I}]\}_{k,l \in [K]}$ and its size is $K^2 \cdot 2^P$.

- ***Luminance-Local Color Contrast LBP (L-LCCLBP):*** This approach considers an image as both the spatial information of luminance and the inter-channel information of different bands. The spatial information of the luminance results from the LBP operator applied to the pseudo panchromatic image (PPI) I^{PPI}, which is computed as the average value over all channels at each pixel [77]:

$$I^{PPI} = \frac{1}{K} \sum_{k \in [K]} I^k. \tag{10.13}$$

  Regarding the inter-channel content, [78] define the local color contrast (LCC) operator that depends on the angle Θ between the value of a pixel p and the average value $\bar{\mathbf{I}}_p = \frac{1}{P} \sum_{q \in \mathcal{N}^{P,d}} \mathbf{I}_q$ of its neighbors in the spectral domain:

$$\Theta[\mathbf{I}](p) = \arccos\left(\frac{\langle \mathbf{I}_p, \bar{\mathbf{I}}_p \rangle}{||\mathbf{I}_p|| \cdot ||\bar{\mathbf{I}}_p||} \right). \tag{10.14}$$

LCC operator is then given by:

$$\mathrm{LCC}[\mathbf{I}](p) = \begin{cases} \left\lceil \frac{2 \cdot 255}{\pi} \Theta[\mathbf{I}](p) \right\rceil & \text{if } 0 \leq \Theta[\mathbf{I}](p) \leq \pi/2, \\ 255 & \text{otherwise.} \end{cases} \tag{10.15}$$

The final texture feature is the concatenation of the histogram of $\mathrm{LBP}[I^{PPI}]$ and the histogram of $\mathrm{LCC}[\mathbf{I}]$ and its size is $2 \cdot 2^P$.

- ***Luminance-Opponent Band Angles LBP (L-OBALBP):*** As for L-LCCLBP, this approach first applies the LBP operator to the PPI I^{PPI}, then [79] considers the angle between each pair of bands $(k, l) \in [K]^2, k \neq l$ as:

$$\mathrm{OBALBP}[\mathbf{I}]^{(k,l)}(p) = \sum_{q \in \mathcal{N}^{P,d}} s\left(\arctan\left(\frac{I_q^k}{I_q^l + \eta} \right), \arctan\left(\frac{I_p^k}{I_p^l + \eta} \right) \right) \cdot 2^{\epsilon(q)}, \tag{10.16}$$

where η is a constant of small value. The final texture feature is the concatenation of the histogram of $\mathrm{LBP}[I^{PPI}]$ and the $K(K-1)$ histograms of $\{\mathrm{OBALBP}^{(k,l)}[\mathbf{I}]\}_{k \neq l \in [K]}$. Its size is $(1 + K(K-1)) \cdot 2^P$.

10.5.2 Assessment on our proposed dataset

Covariance analysis

Since 186 channels are too high to perform the proposed LBP descriptors, we propose to spectrally down-sample the 112 radiance images. We select, for each radiance image $\mathbf{I}$, L channels whose indexes uniformly range in $[K]$. The choice of a uniform distribution of channels is made to comply with datasets and hyperspectral devices presented in Tab. 10.1 that provide spectral channels whose associated bands are uniformly distributed in their respective spectral range. The output is a spectrally sampled radiance image $\mathbf{I}_L = \{I^k\}_{k \in \mathcal{C}_L}$ where $\mathcal{C}_L = \left\{ (l+1) \left\lceil \frac{K-1}{L+1} \right\rceil \right\}_{l \in [L]}$. Note that respective band centers $\{\lambda^k\}_{k \in \mathcal{C}_L}$ then range uniformly in $]405.37, 995.82[$ nm.

In order to determine the minimum number of channels L required to provide the best discrimination between channels, we study the normalized covariance matrix between each couple of channels over the whole database computed in Sec. 10.4. Figure 10.15 shows that close channels in term of band center wavelength are more correlated than distant ones, and that infrared channels are highly correlated. The area inside the two red lines represents the covariance between channels distant

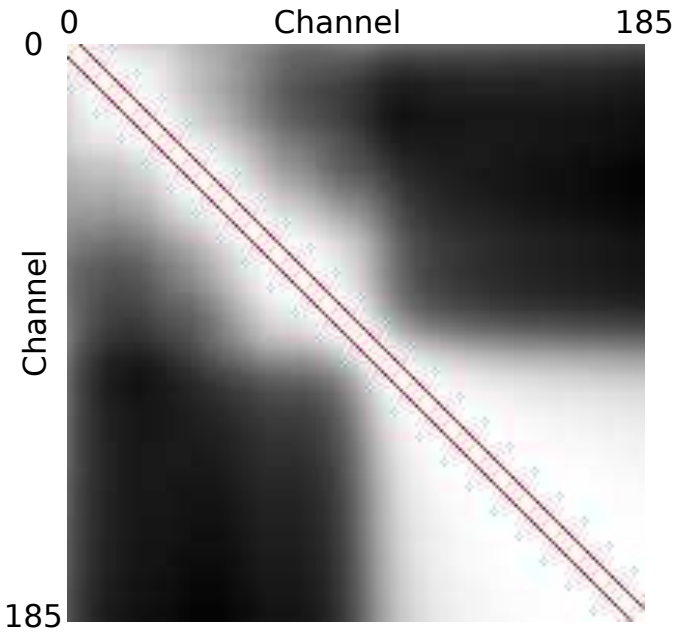


Figure 10.15: Normalized covariance between each couple of channels over the whole database. Values range between 31% (black) and 100% (white). The two red lines are separated by 7 channels, and inside the red lines the covariance is above 95%.

of less than 7 channels. In this area, the inter-channel covariance is above 95%. Stating this, we consider that a spacing of 7 channels (22.33 nm) is enough, so that in our study we consider a number of channels L that ranges from 1 to 27 channels. In colorimetry and in the colour imaging communities, it has been accepted that sampling spectra at 10 nm would provide enough discrimination [80], it is interesting to note that in our case, such large steps would provide a good description.

Classification scheme

Texture feature assessment can be performed on our database by considering the 112 texture images as 112 different classes. Each 8-bit hyperspectral radiance texture image $\mathbf{I}_L$ of size $1024 \times 1024 \times L$ is divided into 25 subimages of size $204 \times 204 \times L$ (without overlap), among which 12 are randomly considered as training images and the 13 others as test images. Supports of subimages according to this random sampling and corresponding classes are available as supplementary material as two text files (*train.txt* and *test.txt*). In a learning phase, LBP histogram features are extracted from each training image. Then, to assign a test image to one of the classes, the same features are extracted from each test image and compared

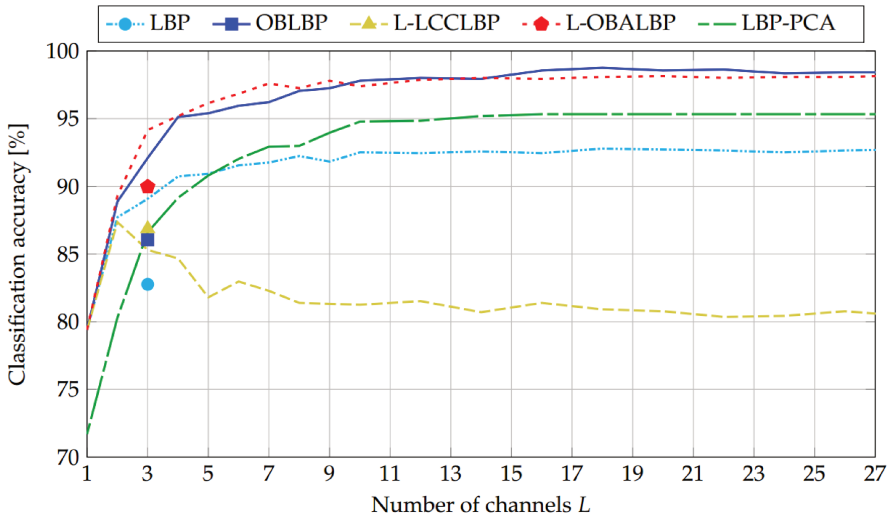


Figure 10.16: Classification accuracy reached by each LBP descriptor with respect to the number of channels L. LBP [73] (Eq. (10.11)), OBLBP [76] (Eq. (10.12)), L-LCCLBP [78] (Eq. (10.15)), L-OBALBP [79] (Eq. (10.16)). LBP-PCA refers to the marginal LBP operator applied to PCA channels. Classification performances are stabilized after $L = 10$ channels: about 92% accuracy for LBP, 98% accuracy for OBLBP and L-OBALBP, 81% accuracy for L-LCCLBP, 95% accuracy for LBP-PCA. The best performances are given by OBLBP with a 98.76% accuracy for $L = 18$ channels. The markers refer to classification accuracies of descriptors applied on the $L = 3$ channels of simulated RGB images from the three spectral sensitivities of the tri-CCD Basler L301kc color camera [82].

to those of each training image. This comparison is performed using a similarity/dissimilarity measure between test and train features. Finally, each test image is assigned to the class of the training image with the best match by using a decision algorithm. The performances of a classification algorithm are determined by the rate of well classified test images, and depend on three main parts of classification, namely the choice of discriminating textural features, the feature similarity measure, and the decision algorithm. In order to determine the most discriminant texture features with respect to the number of channels, we propose to retain the 1-Nearest Neighbor decision algorithm coupled with the similarity measure based on intersection between histograms [81] since this classification scheme requires no additional parameter. Moreover, for simplicity, we propose to use the LBP neighborhood $\mathcal{N}^{P,d}$ composed of $P = 8$ pixels at spatial distance $d = 1$.

Classification accuracy

Figure 10.16 shows the classification performances reached by each texture feature presented in Sec. 10.5.1 with respect to the number of selected channels L (see Sec. 10.5.2). Results show that, except for L-LCCLBP, increasing the number of channels improves the performances, and that after 10 channels the performances are stabilized. Marginal LBP operator that takes only spatial correlation into account reaches about 92% accuracy after 10 channels. By taking spectral correlation into account using opponent bands information, OBLBP and L-OBALBP operators reaches about 98% accuracy after 10 channels. Regarding L-LCCLBP operator, the performances are reduced since instead of using spatial and spectral correlation over each channel, this operator uses only projected information on the panchromatic intensity. Overall, the results are improved from 79.40% (LBP, OBLBP, L-LCCLBP and L-OBALBP with $L = 1$) to 98.76% (OBLBP with $L = 18$). In order to highlight the interest of hyperspectral classification against color classification, we have simulated color images using the three spectral sensitivities of the tri-CCD Basler L301kc color camera [82] and E illumination. Results show that, for three operators, select three narrow spectral bands (two in the visible domain and one in the infrared domain) improves the performances compared to select three wide bands in the visible domain. We can conclude that by increasing the number of channels, LBP descriptors provides more discriminative information about spatial and spectral correlation at the expense of feature size and computation time and, except for L-LCCLBP, hyperspectral imaging improves the discriminatory power of descriptors compared to color imaging.

In order to study the influence of channel distribution, we have selected 10 sets of 10 random channels. All have provided performances that are less than or close to performances obtained with a uniform distribution of 10 channels for marginal LBP operator. We also considered the 10 channels with the highest variance among our dataset. They provide weak classification performances compared to uniform or random distributions since all the related band centers are in the near infrared domain and consequently do not take into account the textural information observed in the visible range.

In order to study the discrimination of PCA for hyperspectral texture classification, we propose to apply the marginal LBP operator over the projections of the 186-channel hyperspectral images on principal axis of the covariance matrix (denoted as PCA channels in the following). Note that OBLBP, L-LCCLBP and L-OBALBP operators are not relevant for this study since they use spectral correlation which by definition does not exist between PCA channels. PCA channels are computed for each image of the dataset by first projecting the 186-channel reflectance image onto each principal axis such as the first PCA channel is associated

with the eigenvector with the highest eigenvalue and so on. For a fair comparison with the previous operators, PCA channels are then quantized on 8 bits using the same quantization function for all PCA channels of all 112 images. LBP-PCA accuracy is shown Fig. 10.16 where L is the number of sorted (according to decreasing eigenvalues) PCA channels that are considered for the LBP operator. By taking only the PCA channel with the highest variance (1^{st} channel), classification performances are lower than by taking one real channel (93^{th} centered at 699 nm). This can be explained by a specificity of our database. Since PCA is a projection of all channels into one channel, two similar textures with different spectral properties are hardy discriminated using the PCA. For illustration purpose of this result Fig. 10.17 shows the sRGB renderings, the 93^{th} channels and the 1^{st} PCA channels of two textures that are spatially similar but spectrally different (*textile_14_orange* and *textile_14_blue*). As can be seen, images of the 93^{th} channel of the two textures have different spatial properties and intensities, so they can be easily discriminated whereas the 1^{st} PCA channel of both textures have similar spatial properties and intensities, so they can be hardly discriminated. By increasing the number of PCA bands, useful information over all channels is used so that it becomes more discriminative than taking each channel separately. From Fig. 10.16 we can see that after 5 PCA channels, performances exceed marginal LBP and after 10 PCA channels the performances are stabilized at about 95% accuracy. To conclude, using the PCA allows to improve the performances by using only spatial correlation. However, in opposition to OBLBP and L-OBALBP that provides the best performances, it ignores the spectral correlation that allows to compare the different channels, and thus, to discriminate textures using their spectral properties.

10.6 Conclusions

In this paper, we introduced and described a hyperspectral reflectance dataset of close range textured objects surfaces. The spectral range of this dataset is from 405 to 995 nm, thus providing the spectral information in both visible and NIR regions. There are five classes of objects and a total of 112 images are provided. We discussed the hyperspectral image acquisition protocol and the corrections that are applied to the data.

We performed a spectral analysis to quantify the spectral complexity of the samples and show that a limited number of components permits a good description of their reflectance. Furthermore, we extended local binary pattern operators to hyperspectral texture analysis. We observed that increasing the number of bands permits better texture classification. We also show that the opponent band local binary pattern performed the best amongst the tested texture descriptors. Indeed, such descriptor

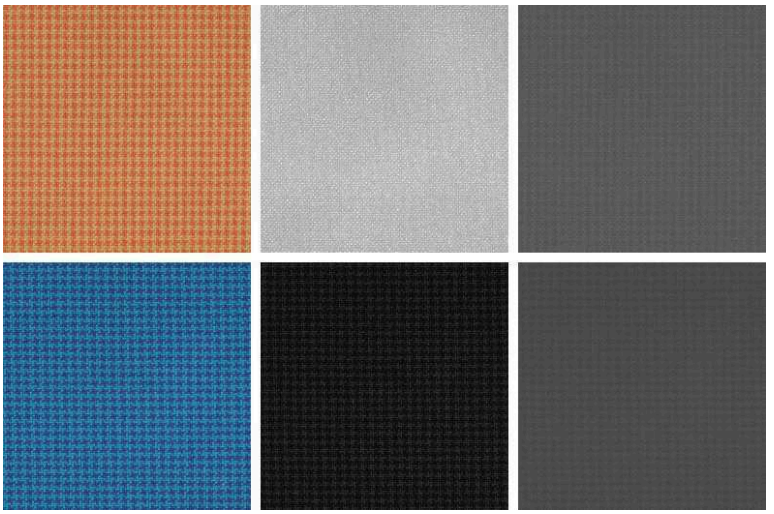


Figure 10.17: Texture image sample: *textile_14_orange* (top) and *textile_14_blue* (bottom). From the left to the right : sRGB rendering, 93^{th} channel (centered at 699 nm), 1^{st} PCA channel. Images of the 93^{th} channel of the two textures have different spatial properties and intensities, so they can be easily discriminated whereas the 1^{st} PCA channel of both textures have similar spatial properties and intensities, so they can be hardly discriminated.

uses inter-channel correlation, to discriminate textures using their spectral properties.

This dataset is available for scientific uses and simulations and it provides a benchmark to test various computer vision algorithms that are related to object classification and material identification. This dataset will also help in designing optimal spectral sensors, computational imaging systems and spectral reconstruction algorithms.

Bibliography

[1] W. L. Wolfe, *Introduction to imaging spectrometers*, vol. 25. SPIE Press, 1997.

[2] P. J. Miller, "Use of tunable liquid crystal filters to link radiometric and photometric standards," *Metrologia*, vol. 28, no. 3, p. 145, 1991.

[3] N. Neumann, M. Ebermann, K. Hiller, M. Seifert, M. Meinig, and S. Kurth, "MEMS Tunable Fabry-Pérot Filters for Infrared Microspectrometer Applications," in *Imaging and Applied Optics 2016*, p. AIT4B.2, Optical Society of America, 2016.

[4] F. Sigernes, M. Syrjäsuo, R. Storvold, J. Fortuna, M. E. Grøtte, and T. A. Johansen, "Do it yourself hyperspectral imager for handheld to airborne operations," *Opt. Express*, vol. 26, pp. 6021–6035, Mar 2018.

[5] X. Kang, X. Xiang, S. Li, and J. A. Benediktsson, "PCA-Based Edge-Preserving Features for Hyperspectral Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, pp. 7140–7151, Dec 2017.

[6] T. V. Bandos, L. Bruzzone, and G. Camps-Valls, "Classification of hyperspectral images with regularized linear discriminant analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 862–873, March 2009.

[7] J. M. Amigo, C. Ravn, N. B. Gallagher, and R. Bro, "A comparison of a common approach to partial least squares-discriminant analysis and classical least squares in hyperspectral imaging," *International Journal of Pharmaceutics*, vol. 373, no. 1, pp. 179 – 182, 2009.

[8] W. Song, S. Li, X. Kang, and K. Huang, "Hyperspectral image classification based on KNN sparse representation," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 2411–2414, July 2016.

[9] G. Mercier and M. Lennon, "Support vector machines for hyperspectral image classification with spectral-based kernels," in *IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477)*, vol. 1, pp. 288–290 vol.1, July 2003.

[10] F. Liu, X. Ye, Y. He, and L. Wang, "Application of visible/near infrared spectroscopy and chemometric calibrations for variety discrimination of instant milk teas," *Journal of Food Engineering*, vol. 93, no. 2, pp. 127 – 133, 2009.

[11] E. Merényi, W. H. Farrand, J. V. Taranik, and T. B. Minor, "Classification of hyperspectral imagery with neural networks: comparison to conventional tools," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, p. 71, May 2014.

[12] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural networks for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, pp. 3639–3655, July 2017.

[13] M. Paoletti, J. Haut, J. Plaza, and A. Plaza, "A new deep convolutional neural network for fast hyperspectral image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, 2017.

[14] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, pp. 6–36, June 2013.

[15] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *Journal of Biomedical Optics*, vol. 19, p. 010901, 2014.

[16] C. Fischer and I. Kakoulli, "Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications," *Studies in Conservation*, vol. 51, no. sup1, pp. 3–16, 2006.

[17] J. Y. Hardeberg, S. George, F. Deger, I. Baarstad, and J. E. H. Palacios, "Spectral scream: Hyperspectral image acquisition and analysis of a masterpiece," in *Public Paintings by Edvard Munch and His Contemporaries: Change and Conservation Challenges* (T. Frøysaker, N. Streeton, H. Kutzke, F. Hanssen-Bauer, and B. Topalova-Casadiego, eds.), London: Archetype Publications, 2015.

[18] Z. Pan, G. Healey, M. Prasad, and B. Tromberg, "Face recognition in hyperspectral images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1552–1560, Dec 2003.

[19] A. Gowen, C. O'Donnell, P. Cullen, G. Downey, and J. Frias, "Hyperspectral imaging – an emerging process analytical tool for food quality and safety control," *Trends in Food Science & Technology*, vol. 18, no. 12, pp. 590 – 598, 2007.

[20] T. Eckhard, M. Klammer, E. M. Valero, and J. Hernández-Andrés, "Improved spectral density measurement from estimated reflectance data with kernel ridge regression," in *Image and Signal Processing*, (Cham), pp. 79–86, Springer International Publishing, 2014.

[21] L. G. Coppel, S. Le Moan, P. Z. Elıas, R. Slavuj, and J. Y. Harderberg, "Next generation printing–towards spectral proofing," in *Advances in Printing and Media Technology*, vol. 41, pp. 19–24, 2014.

[22] A. Majda, R. Wietecha-Posłuszny, A. Mendys, A. Wójtowicz, and B. Łydżba-Kopczyńska, "Hyperspectral imaging and multivariate analysis in the dried blood spots investigations," *Applied Physics A*, vol. 124, p. 312, Mar 2018.

[23] J.-H. Cheng, H. Jin, Z. Xu, and F. Zheng, "NIR hyperspectral imaging with multivariate analysis for measurement of oil and protein contents in peanut varieties," *Anal. Methods*, vol. 9, pp. 6148–6154, 2017.

[24] J.-H. Cheng, B. Nicolai, and D.-W. Sun, "Hyperspectral imaging with multivariate analysis for technological parameters prediction and classification of muscle foods: A review," *Meat Science*, vol. 123, pp. 182 – 191, 2017.

[25] B. Zhang, J. Li, S. Fan, W. Huang, C. Zhao, C. Liu, and D. Huang, "Hyperspectral imaging combined with multivariate analysis and band math for detection of common defects on peaches (prunus persica)," *Computers and Electronics in Agriculture*, vol. 114, pp. 14 – 24, 2015.

[26] G. J. Brelstaff, A. Parraga, T. Troscianko, and D. Carr, "Hyperspectral camera system: acquisition and analysis," in *Proc.SPIE. Geog. Inf. Sys. Photogram. and Geolog./Geophys. Remote Sensing*, vol. 2587, pp. 2587 – 2587 – 10, 1995. http://www.cvc.uab.es/color_calibration/Bristol_Hyper/.

[27] D. H. Foster, K. Amano, and S. M. C. Nascimento, "Time-lapse ratios of cone excitations in natural scenes," *Vision Research*, vol. 120, pp. 45–60, 2016. http://personalpages.manchester.ac.uk/staff/d.h.foster/Time-Lapse_HSIs/Time-Lapse_HSIs_2015.html.

[28] B. Arad and O. Ben-Shahar, "Sparse recovery of hyperspectral signal from natural RGB images," in *Proceedings of the 14th European Conference on Computer Vision (ECCV'16)*, vol. 9911 of *Lecture Notes in Computer Science*, (Amsterdam, The Netherlands), pp. 19–34, Springer-Verlag, Oct. 2016. http://icvl.cs.bgu.ac.il/hyperspectral.

[29] S. M. C. Nascimento, F. P. Ferreira, and D. H. Foster, "Statistics of spatial cone-excitation ratios in natural scenes," *Journal of the Optical Society of America A*, vol. 19, pp. 1484–1490, Aug. 2002. http://personalpages.manchester.ac.uk/staff/d.h.foster/Hyperspectral_images_of_natural_scenes_02.html.

[30] D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster, "Frequency of metamerism in natural scenes," *Journal of the Optical Society of America A*, vol. 23, pp. 2359–2372, Oct. 2006. http://personalpages.manchester.ac.uk/staff/d.h.foster/Hyperspectral_images_of_natural_scenes_04.html.

[31] S. M. C. Nascimento, K. Amano, and D. H. Foster, "Spatial distributions of local illumination color in natural scenes," *Vision Research*, vol. 120,

pp. 39–44, Mar. 2016. https://personalpages.manchester.ac.uk//staff/d.h.foster/Local_Illumination_HSIs/Local_Illumination_HSIs_2015.html.

[32] J. Eckhard, T. Eckhard, E. M. Valero, J. L. Nieves, and E. G. Contreras, "Outdoor scene reflectance measurements using a Bragg-grating-based hyperspectral imager," *Applied Optics*, vol. 54, pp. D15–D24, May 2015. http://colorimaginglab.ugr.es/pages/Data#__doku_ugr_hyperspectral_image_database.

[33] A. Chakrabarti and T. Zickler, "Statistics of real-world hyperspectral images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'11)*, (Colorado Springs, USA), pp. 193–200, June 2011. http://vision.seas.harvard.edu/hyperspec.

[34] R. M. H. Nguyen, D. K. Prasad, and M. S. Brown, "Training-based spectral reconstruction from a single RGB image," in *Proceedings of the 13th European Conference on Computer Vision (ECCV'14)*, (Zürich, Switzerland), pp. 186–201, Springer-Verlag, Sept. 2014. https://sites.google.com/site/hyperspectralcolorimaging/dataset.

[35] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum," *IEEE Transactions on Image Processing*, vol. 19, pp. 2241–2253, Sept. 2010. http://www.cs.columbia.edu/CAVE/databases/multispectral/.

[36] S. Hordley, G. Finalyson, and P. Morovic, "A multi-spectral image database and its application to image rendering across illumination," in *Proceeding of the 3rd International Conference on Image and Graphics (ICIG'04)*, (Hong Kong, China), pp. 394–397, Dec. 2004. http://www2.cmp.uea.ac.uk/Research/compvis/MultiSpectralDB.htm.

[37] D. H. Brainard, "Hyperspectral image data." http://color.psych.upenn.edu/hyperspectral.

[38] "Hyperspectral images of illustrated manuscripts." http://personalpages.manchester.ac.uk/staff/d.h.foster/Hyperspectral_Images_of_Illustrated_Manuscripts.html. [Online; accessed 26-Apr-2018].

[39] A. Mirhashemi, "Introducing spectral moment features in analyzing the SpecTex hyperspectral texture database," *Machine Vision and Applications*,

vol. 29, pp. 415–432, Apr. 2018. http://www.uef.fi/web/spectral/spectex.

[40] S. Le Moan, S. T. George, M. Pedersen, J. Blahová, and J. Y. Hardeberg, "A database for spectral image quality," in *Proceedings of the SPIE-IS&T Electronic Imaging: Image Quality and System Performance XII*, vol. 9396, (San Francisco, California, USA), p. 93960P, Feb. 2015. https://www.ntnu.edu/web/colourlab/software.

[41] A. Noviyanto and W. H. Abdullah, "Honey dataset standard using hyperspectral imaging for machine learning problems," in *25th European Signal Processing Conference (EUSIPCO)*, pp. 473–477, Aug 2017.

[42] A. Zacharopoulos, K. Hatzigiannakis, P. Karamaoynas, V. M. Papadakis, M. Andrianakis, K. Melessanaki, and X. Zabulis, "A method for the registration of spectral images of paintings and its evaluation," *Journal of Cultural Heritage*, vol. 29, pp. 10–18, Jan. 2018. http://www.ics.forth.gr/cvrl/msi.

[43] M. Nouri, N. Gorretta, P. Vaysse, M. Giraud, C. Germain, B. Keresztes, and J.-M. Roger, "Near infrared hyperspectral dataset of healthy and infected apple tree leaves images for the early detection of apple scab disease," *Data in Brief*, vol. 16, pp. 967–971, 2018.

[44] T. Hirvonen, J. Orava, N. Penttinen, K. Luostarinen, M. Hauta-Kasari, M. Sorjonen, and K.-E. Peiponen, "Spectral image database for observing the quality of nordic sawn timbers," *Wood Science and Technology*, vol. 48, no. 5, pp. 995–1003, 2014. http://www.uef.fi/web/spectral/spectral-image-database-of-nordic-sawn-timbers.

[45] T. Skauli and J. Farrell, "A collection of hyperspectral images for imaging systems research," in *Proceedings of the SPIE Electronic Imaging Annual Symposium (SPIE'13): Digital Photography IX*, vol. 8660, (Burlingame, California, USA), pp. 86600C–86600C–7, Feb. 2013. http://www.imageval.com/scene-database/.

[46] H. Shin, N. H. Reyes, A. L. Barczak, and C. S. Chan, "Colour object classification using the fusion of visible and near-infrared spectra," in *PRICAI: Trends in Artificial Intelligence* (B.-T. Zhang and M. A. Orgun, eds.), pp. 498–509, Springer Berlin Heidelberg, 2010.

[47] H. Steiner, O. Schwaneberg, and N. Jung, "Advances in active near-infrared sensor systems for material classification," in *Imaging and Applied Optics Technical Papers*, p. ITu2C.2, Optical Society of America, 2012.

[48] W. Guifang, M. Hai, and P. Xin, "Identification of varieties of natural textile fiber based on Vis/NIR spectroscopy technology," in *IEEE Advanced Information Technology, Electronic and Automation Control Conference*, pp. 585–589, Dec 2015.

[49] B. H. Horgan, E. A. Cloutis, P. Mann, and J. F. Bell, "Near-infrared spectra of ferrous mineral mixtures and methods for their identification in planetary surface spectra," *Icarus*, vol. 234, pp. 132 – 154, 2014.

[50] J. Lehtonen, J. Parkkinen, and T. Jaaskelainen, "Optimal sampling of color spectra," *J. Opt. Soc. Am. A*, vol. 23, pp. 2983–2988, Dec 2006.

[51] "HySpex VNIR-1800." https://www.hyspex.no/products/vnir_1800.php. [Online; accessed 26-Apr-2018].

[52] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[53] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.

[54] "Illumination Technologies Inc. 3900e DC Regulated ER Lightsouce." http://bit.ly/IT3900e. [Online; accessed 21-May-2018].

[55] "SG-3051 SphereOptics Diffuse Reflectance Tile." http://sphereoptics.de/wp-content/uploads/2014/03/Zenith_Product-Brochure.pdf. [Online; accessed 21-May-2018].

[56] P.-J. Lapray, J.-B. Thomas, and P. Gouton, "A Database of Spectral Filter Array Images that Combine Visible and NIR," in *Computational Color Imaging Workshop*, pp. 187–196, 2017.

[57] F. Deger, A. Mansouri, M. Pedersen, J. Y. Hardeberg, and Y. Voisin, "A sensor-data-based denoising framework for hyperspectral images," *Opt. Express*, vol. 23, pp. 1938–1950, Feb 2015.

[58] D. Gillis, J. H. Bowles, and M. E. Winter, "Dimensionality reduction in hyperspectral imagery," in *Proc. SPIE, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery IX*, vol. 5093, pp. 45–56, 2003.

[59] H. Deborah, N. Richard, and J. Y. Hardeberg, "A comprehensive evaluation of spectral distance functions and metrics for hyperspectral image processing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, pp. 3224–3234, June 2015.

[60] H. Hotelling, "Analysis of a complex of statistical variables into principal components.," *Journal of educational psychology*, vol. 24, no. 6, p. 417, 1933.

[61] I. Jolliffe, *Principal Component Analysis*. John Wiley & Sons, Ltd, 2014.

[62] J. Wang and C.-I. Chang, "Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, pp. 1586–1600, June 2006.

[63] L. Zhang, Y. Zhong, B. Huang, J. Gong, and P. Li, "Dimensionality reduction based on clonal selection for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 4172–4186, Dec 2007.

[64] T. Zhang, D. Tao, and J. Yang, "Discriminative locality alignment," in *proceedings of the 10th European Conference on Computer Vision*, (Marseille, France), pp. 725–738, 2008.

[65] J. Khodr and R. Younes, "Dimensionality reduction on hyperspectral images: A comparative review based on artificial datas," in *4th International Congress on Image and Signal Processing*, vol. 4, pp. 1875–1883, Oct 2011.

[66] J. Y. Hardeberg, "On the spectral dimensionality of object colours," in *Conference on Colour in Graphics, Imaging, and Vision*, pp. 480–485, 2002.

[67] R. Dusselaar and M. Paul, "Hyperspectral image compression approaches: opportunities, challenges, and future directions: discussion," *J. Opt. Soc. Am. A*, vol. 34, pp. 2170–2180, Dec 2017.

[68] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Computational Statistics & Data Analysis*, vol. 52, no. 1, pp. 155 – 173, 2007.

[69] J. Li, J. M. Bioucas-Dias, A. Plaza, and L. Liu, "Robust collaborative nonnegative matrix factorization for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, pp. 6076–6090, Oct 2016.

[70] W. Bao, Q. Li, L. Xin, and K. Qu, "Hyperspectral unmixing algorithm based on nonnegative matrix factorization," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 6982–6985, July 2016.

[71] A. Alsam, D. Connah, and J. Hardeberg, "Multispectral imaging: How many sensors do we need?," *Journal of Imaging Science and Technology*, vol. 50, no. 1, pp. 45–52, 2006.

[72] A. Verma, D. Tyagi, and S. Sharma, "Recent advancement of lbp techniques: A survey," in *2016 International Conference on Computing, Communication and Automation (ICCCA)*, pp. 1059–1064, April 2016.

[73] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971–987, Jul 2002.

[74] C. Palm, "Color texture classification by integrative Co-occurrence matrices," *Pattern Recognition*, vol. 37, no. 5, pp. 965 – 976, 2004.

[75] T. Ojala, M. Pietikainen, and D. Harwood, "Performance evaluation of texture measures with classification based on kullback discrimination of distributions," in *Proceedings of 12th International Conference on Pattern Recognition*, vol. 1, pp. 582–585 vol.1, Oct 1994.

[76] T. Mäenpää, M. Pietikainen, and J. Viertola, "Separating color and pattern information for color texture discrimination," in *Object recognition supported by user interaction for service robots*, vol. 1, pp. 668–671, 2002.

[77] S. Mihoubi, O. Losson, B. Mathon, and L. Macaire, "Multispectral demosaicing using pseudo-panchromatic image," *IEEE Transactions on Computational Imaging*, vol. 3, pp. 982–995, Dec 2017.

[78] C. Cusano, P. Napoletano, and R. Schettini, "Combining local binary patterns and local color contrast for texture classification under varying illumination," *J. Opt. Soc. Am. A*, vol. 31, pp. 1453–1461, Jul 2014.

[79] S. H. Lee, J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Local color vector binary patterns from multichannel face images for face recognition," *IEEE Transactions on Image Processing*, vol. 21, pp. 2347–2353, April 2012.

[80] H. J. Trusseli and M. S. Kulkarni, "Sampling and processing of color signals," *IEEE Transactions on Image Processing*, vol. 5, pp. 677–681, Apr 1996.

[81] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, pp. 11–32, Nov 1991.

[82] "L301kc - Basler L300." https://www.baslerweb.com/en/products/cameras/line-scan-cameras/l300/l301kc/. [Online; accessed 14-May-2018].

Chapter 11

Discussion and Conclusion

In this chapter, we will discuss the overall research carried out during the work presented in this dissertation, the choices made during this research and what other options could be exploited. Finally we will conclude this dissertation along with perspectives on future.

11.1 Discussion on research work

During this research work, we made several choices based on various factors. In this section, we will discuss about those methods, some of the reasons of making the choice for particular techniques and what alternatives could be done.

11.1.1 Natural scene statistics based illuminant estimation methods

In this work, we presented the concept of multispectral constancy, which is based on spectral adaptation transform and the knowledge of scene illumination. Scene illumination is required in the sensor dimension and it can either be manually extracted from a white diffuser placed in the scene, or through illuminant estimation. We discussed about illuminant estimation in detail in **Article B** (Chapter 5). We proposed the extension of four illuminant estimation algorithms from color to multispectral images. These methods are based on natural scene statistics and are able to perform well when the assumptions of having a natural scene with variety of surfaces, are fulfilled.

We used the natural scene statistics based illuminant estimation algorithms due to their performance for color images. They are able to perform well in most of the cases. We tested our proposed algorithms on natural scenes containing diffuse surfaces. However, the assumption of natural scene statistics, despite being popular,

can be challenged in the context of scenes that we encounter in this modern world. Most of the natural objects and natural light sources have smooth spectra, while the modern lighting sources that we use in offices and houses are not smooth in the wavelength spectrum. They can consist of sharp peaks and may have no response in some of the wavelength regions. If the spectra is not smooth then such illuminant estimation methods tend to fail as the basic assumption is not fulfilled. Another reason which alleviates the efficiency of natural statistics based methods is the nature of scenes itself that we expect to see for a system with multispectral camera fitted on it. Its not necessary that an outdoor scene means a scene with natural materials. A scene having plastic materials or metals with strong specular reflection and street view with concrete buildings all around, do not fulfill the basic assumption of natural scene statistics. Also, for a scene after a disaster like an earthquake, the statistical properties may not be close to the natural scene statistics. Another example can be of a robot with camera fitted on it to get the information of a road. In such a scenario, the scene statistics will be different and a statistics based illuminant estimation method will not perform well. It is possible to redefine scene statistics for a particular type of images but they may not be generic for all type of conditions.

To summarize, we made a choice of using natural scene statistics based illuminant estimation methods for multispectral imaging, but we are aware that they can fail in certain conditions. There is another class of illuminant estimation algorithms which is based on the physical properties of materials and do not assume presence of a variety of surfaces in a scene. In the following section, we discuss such methods.

11.1.2 Physics based illuminant estimation methods

The physics based illuminant estimation methods use the physical properties of materials. Light falling on a surface is reflected back in form of either diffuse or specular reflection. Diffuse reflection occurs when the incident light is absorbed by the material and only a part of it is reflected back while in case of specular reflection, all of the incident light is reflected back. The specular reflection consists of information about the incident light, otherwise known as the scene illumination. Physics based illuminant estimation methods use this property of materials and are also known as highlights based illuminant estimation methods. Such methods do not require a variety of surfaces and can perform well even in the presence of a single surface in the scene.

The first thing to do in highlights based illuminant estimation method is to detect the highlight regions in an image. It is not a trivial task and there is a lot of research on the topic of highlights detection itself. We have covered some of this research

in **Article C** (Chapter 6). Assuming that highlights are detected in an image, the next step is to estimate the illuminant from that region containing the specular reflection. In **Article D** (Chapter 7), we have provided a review of such methods and the feasibility of extension of some of the physics based illuminant estimation methods from color to multispectral images.

There are some issues related with such illuminant estimation methods. The first issue is assumption for these methods that a linear imaging system is used and there is no non-linearity included in the imaging pipeline. The second issue, which is not easily avoidable in most cases, is the saturated pixels. Specular reflection reflects back almost all of the incident light and as a result, pixels in that region gets saturated in the image. If the exposure time is minimized to avoid saturation, almost all of the other regions may become dark. We faced this problem while processing the multispectral images that are taken in the work described in Chapter 9. We used objects having specular reflection property so that we can use information from their surface for estimation of scene illumination. However, those areas got saturated in most of the cases while we were trying to set the exposure time for each channel individually. Due to this reason, we did not use the physics based illuminant estimation methods in our research, although we have provided literature review on them.

However, we found an interesting observation that the spectral gray-edge method (described in Chapter 5) is able to perform very well for the multispectral images that we took for experiments. This algorithms also performed well in the experiments as described in **Article B**. The reason that the edge based illuminant estimation method works well is that the incident light on a material causes sheen reflection around its edges. This reflection is very useful for determination of scene illumination since it is not saturated like specular reflection, but contains the information of scene illumination. Although we found this sheen reflectance very helpful, our findings can not be generalized since they are based on only a few images. It is still to be investigated that how such a reflection is helpful in various conditions.

11.1.3 Use of linear methods

In the proposed multispectral constancy pipeline, we have used linear methods for data transformation. For spectral reconstruction, we have used Wiener estimation method, which is fundamentally based on pseudo inverse of matrix. Such an inverse is useful for minimization of root mean square error and this is what we get when the input data have minimum influence of illumination. The choice for using linear methods in our research is to investigate the overall effectiveness of the concept of illuminant invariant representation of multispectral images. We could

have used non-linear methods for those tasks but then our concentration would be diverted towards the tool for conversion instead of the overall concept itself. As an example, we did not use machine learning based methods for illuminant estimation and data conversion, although they are used for illuminant estimation in color images and have provided promising results. In case of color images, there are standard color spaces and the image data is more or less in a standardized form. A change in imaging camera does not significantly influence the final output as there are device independent representations for color images which can standardize the data taken from various color imaging devices. Machine learning algorithms work very well in the presence of large data for training and when the test input have the same characteristics as the training data.

Having a standardized representation for multispectral data is not available in our case. The reason for it is that there is no standard multispectral camera. These cameras come with a variety of number of channels and sensor configurations. In case we want to develop a non-linear machine learning based method to perform any or all of the tasks that are part of the multispectral constancy pipeline, a huge set of training data acquired from a particular imaging device will be required. A system trained on such data may work well for that particular imaging device, but may fail if the imaging device is changed. Lack of training data is the bottleneck for machine learning methods and since we do not have sufficient data at this stage, we use linear methods in our proposed multispectral constancy pipeline.

11.1.4 Design of multispectral imaging system

We used the simulation of multispectral sensors in **Article B** (Chapter 5) to find the optimal number of sensors and sensor configuration. We found that eight sensors with overlapping configuration are able to provide best results in terms of illuminant estimation. However, we also found during spectral reconstruction that overlapping sensors cause errors due to the lower spectral resolution and the linear methods developed for this purpose assume that there are no inter-channel dependencies in the image. We proposed spectral adaptation transform for addressing this error and found good results in terms of spectral reconstruction.

Despite this study for optimal number of sensors and sensor configuration, we did not attempt to propose a set of multispectral sensors that can serve the best for illuminant estimation and spectral reconstruction. The reason for it is because of the lack of sufficient data that could be generalized for all the materials and scenarios. We could have attempted to propose multispectral sensors on the basis of the data that we have, but then those sensors would be valid for our data only and would lack the generalization.

The aim of this work is to enable the use of multispectral imaging in uncontrolled imaging conditions and we used off the shelf multispectral camera for the demonstration of our proposed framework in Chapter 9. Our proposed concept can be utilized during the designing of multispectral imaging sensors, but due to the limitation of data and lack of generalization, we did not attempt it.

11.1.5 Colorimetry in multispectral imaging

In **Article E** (Chapter 8), we used a simulated multispectral camera to acquire the colorimetric information from a scene. Here we examined two methods; one is to directly acquire colorimetric data through calibration matrix and other is to get the spectra reconstructed and then use the standard CIE color matching functions for acquisition of colorimetric data. We got good results when the calibration matrix was used on the acquired multispectral data, while the spectral reconstruction method did not provide better results. The results are very interesting and show that a multispectral camera can be used to acquire colorimetric information of a scene.

We have used this technique in Chapter 9 for visualization of the acquired multispectral data. Although the visualization results are good in our case, we can not conclude that this method should be used for visualization of multispectral data in all the cases. We did not include the visual perception quality metrics in our experiments and provided results on the basis of per pixel matching. In order to make this method of visualization as a standard for multispectral imaging, extensive psychometric tests need to be performed where observes can select the image of their preference on the basis of factors like color and contrast. We did not perform such tests in our work since our aim is to get an illuminant invariant representation of multispectral data for compute vision applications, which may not be directly related to visual perspective of humans.

11.1.6 Analysis on reflectance dataset

We attempted to perform surface identification on the basis of its spectral properties during the work described in Chapter 9. However, the results were not as promising as expected. The spectral signature of some materials can be close enough with each other, making it challenging to distinguish among them. In the experiments, we used linear imaging model and there are several limitations being implied due to it. The framework of multispectral constancy proposed in this work can be improved through the use of more complex and non-linear models (e.g. Bi-directional reflectance distribution function). The experimental results also implies that apart from the spectral information, the spatial characteristics of a surface are also a distinct feature. The spatial features include texture of a material's surface.

We felt the need for high resolution spectral and spatial data of various material surfaces. To fulfill this need, we created collaboration with a team from University of Lille in France, who are specialized in texture analysis. After discussions, we defined the image acquisition protocol and formed a hyperspectral reflectance image dataset of materials consisting of five different categories. We performed spectral analysis of the acquired data while the other team performed spatial analysis of texture of acquired images. Both teams found very interesting results in terms of both spectral and spatial analysis. It was found that texture classification results are improved through the use of spectral information and inter-channel correlation. However, those results can not be generalized since our analysis is limited to those materials and that reflectance data only.

The advantage of the hyperspectral reflectance dataset provided in this work, compared with the already available databases is the availability of high resolution spectral and spatial data. The spectral range in our dataset covers both the visible and near infrared region. Each image in the dataset contains a unique material and it can be used as ground-truth spectral data for surface identification applications. The areas where this data could be used are in the fields of image processing, computational imaging and computer vision, such as surface identification, spatio-spectral analysis of textured surfaces, image sensor simulation, color reproduction, image relighting and so on.

11.2 Conclusion

In this dissertation, we proposed multispectral constancy for illuminant invariant representation of multispectral images. Our aim is to enable the use of multispectral imaging in uncontrolled imaging environments. Multispectral constancy is used to transform the multispectral images taken under unknown illumination conditions, into a representation as if they were taken with known illuminant. Such an illuminant is called as canonical illuminant and the representation of multispectral data taken with this known illuminant is termed as canonical representation.

Multispectral constancy is based on spectral adaptation transform and information of scene illumination. Spectral adaptation transform is used to incorporate the intrinsic properties of imaging sensor and minimize the error during transformation of multispectral data. The inter-channel dependencies and overlapping of sensitivities of imaging filters are addressed in the spectral adaptation transform. This transform is unique for each imaging system and needs to be calculated only once. The computed spectral adaptation transform for a given imaging sensor is used during image transformation afterwards. The second part of multispectral constancy is the information of scene illumination in the sensor domain. This information can either be obtained by placing a white diffuser in the scene, or by estimating

the scene illumination.

For estimation of scene illumination, we proposed the extension of four image statistics based illuminant estimation methods from color to spectral images. We had extensive analysis and evaluation of these algorithms by using simulated multispectral sensors. We found that edge based illuminant estimation provides the best result. We also provide literature review on highlights detection in images, and use of those highlights for illuminant estimation.

We demonstrate the effectiveness of the concept of multispectral constancy by taking images of scenes with varying illuminations. Spectra of surfaces is reconstructed from the multispectral data after the transformation of this data into the canonical form. We found promising results when the multispectral constancy framework is applied on the acquired multispectral images.

We found in our study that although spectral information is important for distinguishing between two material surfaces, it is not enough for material identification. We realized that not only spectral, but spatial information in the form of texture is also required for enabling a system to identify the material. For this purpose, we collaborated with a team which is specialized in texture analysis and created a hyperspectral reflectance dataset consisting for five different categories of materials. This dataset is publicly available for research purposes and will be helpful for many imaging and computer vision tasks.

11.3 Perspectives

This study will make a step closer towards the use of multispectral imaging for computer vision applications. The concept of multispectral constancy proposed in this work is based on linear imaging model and the processing of data is also performed linearly. The framework of multispectral constancy can be improved at many levels, starting from a more complex imaging model and the use of nonlinear methods for illuminant estimation, incorporation of the intrinsic properties of sensors and spectral reconstruction. We used statistics based illuminant estimation methods in this work. An overview of specular reflection based illuminant estimation methods is provided in this dissertation but is not used in the experiments. These methods are yet to be explored and more complex models, such as the bidirectional reflectance distribution function can be used for improvement of illuminant estimation and spectral reconstruction. Our proposed framework for multispectral constancy is yet to be analyzed for material identification and clarification tasks through the use of both spatial and spectral features.

A challenge faced in the designing of spectral imaging system is the lack of data. We have provided a high resolution spectral and spatial data of various material

surfaces, which can be used for simulation of optimal multispectral sensors. Although the data provided by us is limited and a result obtained from it cannot be generalized, it can still provide a good approximation. The quality of this hyperspectral reflectance dataset has to be analyzed through various experimental setups.

Using the concepts of color imaging on the higher dimensional data is challenging, but the results from our analysis are promising. By using the concept of multispectral constancy, the limitation of having known imaging conditions for multispectral imaging is addressed. With the use of multispectral imaging in outdoor environment, many new applications of computer vision will emerge. The higher spectral information obtained from multispectral cameras and their ability to work in ultraviolet and infrared regions of wavelength spectrum means that they may even replace color cameras for some conventional computer vision applications. However, it is too early to anticipate the precise impact of using multispectral imaging at this stage, since the sensor technology and is still under development. it is only recently that snapshot multispectral cameras have become available in the market. It will take some time for such systems to get accepted at a larger scale and after that, with our proposed concept of multispectral constancy, new horizons for research in multispectral imaging domain will open up...