UNIVERSITÄT OSNABRÜCK

Physics-Aware Shadow Compensation for Hyperspectral Imagery Based on Spectral Unmixing and Data Fusion

Dissertation

Zur Erlangung
des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
des Fachbereichs Mathematik/Informatik
der Universität Osnabrück

Vorgelegt von

Guichen Zhang, M. Sc.

Prüfer der Dissertation:

Prof. Dr. Peter Reinartz, Universität Osnabrück
Prof. Dr. Paul Scheunders, Universität Antwerpen, Belgien

Tag der mündlichen Prüfung: 18. Juli 2024

Osnabrück, 2024

ABSTRACT

Imaging spectrometers, also known as hyperspectral sensors, acquire reflectance spectra from targets on ground in up to hundreds of narrow spectral bands over a wide range of the electromagnetic spectrum. Having an increased spectral resolution with respect to other kinds of optical imagery, such as panchromatic and multispectral, hyperspectral data can discriminate materials more accurately and thus allow new applications in remote sensing. Nevertheless, the Earth's surface topography causes the occlusion of incoming illumination for ground targets, leading to shadow effects in the acquired images. Shadows can considerably decrease the performance of image analysis algorithms, and have thus drawn growing attention in the literature. Although shadow issues have been discussed for some kinds of optical sensors, few works have addressed these effects and their unique challenges in hyperspectral imagery, demanding novel shadow-aware methods. This dissertation proposes three robust algorithms, built on one another, for tackling shadow effects in hyperspectral images.

The dissertation first presents a shadow detection and removal framework based on physical assumptions and spectral unmixing. The main idea is to model the shadow formation using a few physically interpretable shadow-related parameters, and apply them in order to detect and remove shadows. Specifically, a novel physics- and shadow-aware spectral mixing model is proposed, which considers how material spectra and shadows contribute to the individual pixel spectrum measured by the sensor.

The described spectral mixing model can tackle only simple scenarios because it assumes simplified optical interactions and illumination conditions on ground. Thus, the following work improves the model to handle more complicated and generalized cases. The improved model regards the entire radiative propagation process, from illumination sources to the backscattered signals recorded by the sensor, using a discrete-time stochastic process and physical assumptions.

As the model's complexity increases, resolving unknown parameters via spectral unmixing becomes an ill-posed problem. Hence, a novel spectral unmixing approach for a robust estimation based on the Alternating Direction Method of Multipliers (ADMM) and data fusion is proposed. The ADMM decomposes a complex optimization problem into sub-problems, each of which is easier to solve. Digital Surface Models (DSMs) are also employed in this approach, as they are insensitive to shadow effects. In addition, spatial relationships between neighboring pixels are considered.

The proposed methods have been extensively evaluated using several simulated and real datasets in small and large regions. Results demonstrate that the proposed research works are effective and superior to state-of-the-art methods, both qualitatively and quantitatively.

ZUSAMMENFASSUNG

Abbildende Spektrometer, auch bekannt als hyperspektrale Sensoren, erfassen Reflexionsspektren von der Erdoberfläche in bis zu Hunderten von schmalen Spektralbändern über einen breiten Bereich des elektromagnetischen Spektrums. Aufgrund der höheren spektralen Auflösung im Vergleich zu anderen Arten optischer Bilder, wie panchromatischen und multispektralen, können Hyperspektraldaten Materialien genauer unterscheiden und ermöglichen so neue Anwendungen in der Fernerkundung. Allerdings führt die Topographie der Erdoberfläche dazu, dass die einfallende Beleuchtung für Bodenziele teilweise verdeckt wird, was zu Schatteneffekten in den aufgenommenen Bildern führt. Schatten können die Leistung von Bildanalysealgorithmen erheblich beeinträchtigen und haben daher in der Literatur zunehmende Beachtung gefunden. Obwohl Schattenprobleme für einige Arten von optischen Sensoren bearbeitet wurden, haben sich nur wenige Arbeiten mit diesen Effekten und ihren besonderen Herausforderungen bei hyperspektralen Bildern befasst, die neue Methoden zur Schattenberücksichtigung erfordern. In dieser Dissertation werden drei robuste, aufeinander aufbauende Algorithmen zur Bewältigung von Schatteneffekten in hyperspektralen Bildern vorgeschlagen.

In der Dissertation wird zunächst ein Verfahren zur Schattenerkennung und -entfernung vorgestellt, das auf physikalischen Annahmen und spektraler Entmischung basiert. Die Hauptidee besteht darin, die Schattenbildung anhand einiger physikalisch interpretierbarer schattenbezogener Parameter zu modellieren und diese zur Erkennung und Entfernung von Schatten anzuwenden. Insbesondere wird ein neuartiges physikalisches und schattenberücksichtigendes spektrales Mischungsmodell vorgeschlagen, das berechnet, wie Materialspektren und Schatten zu dem vom Sensor gemessenen individuellen Pixelspektrum beitragen.

Das beschriebene spektrale Mischungsmodell kann nur einfache Szenarien bewältigen, da es von vereinfachten optischen Interaktionen und Beleuchtungsbedingungen am Boden ausgeht. Daher wird in der folgenden Arbeit das Modell verbessert, um kompliziertere und allgemeinere Fälle zu behandeln. Das verbesserte Modell betrachtet den gesamten Strahlungsausbreitungsprozess, von den Beleuchtungsquellen bis zu den vom Sensor aufgezeichneten rückgestreuten Signalen, unter Verwendung eines zeitdiskreten stochastischen Prozesses und physikalischer Annahmen.

Da die Komplexität des Modells zunimmt, wird die Lösung unbekannter Parameter durch spektrale Entmischung zu einem schlecht gestellten Problem. Daher wird ein neuartiger spektraler Entmischungsansatz für eine robuste Schätzung auf der Grundlage der Alternating Direction Method of Multipliers (ADMM) und von Datenfusion vorgeschlagen. Die ADMM zerlegt ein komplexes Optimierungsproblem in Teilprobleme, von denen jedes einfacher zu lösen ist. Bei diesem Ansatz werden auch digitale Oberflächenmodelle verwendet, da sie unempfindlich gegenüber Schatteneffekten sind. Darüber hinaus werden räumliche Beziehungen zwischen benachbarten Pixeln berücksichtigt.

Die vorgeschlagenen Methoden wurden anhand mehrerer simulierter und realer Datensätze in kleinen und großen Regionen eingehend bewertet. Die Ergebnisse zeigen, dass die vorgeschlagenen Forschungsarbeiten sowohl qualitativ als auch quantitativ wirksam und den aktuellen state-of-the-art Methoden überlegen sind.

ACKNOWLEDGEMENTS

My doctoral journey has been one of the most memorable experiences of my life. This would not have been possible without the support and encouragement of many people.

First of all, I would like to thank Prof. Dr. Peter Reinartz for the opportunity to work in the Department of Photogrammetry and Image Analysis at the German Aerospace Center (DLR) and to carry out this dissertation at the University of Osnabrück. I am grateful for his guidance and support throughout my research. His constructive feedback, the team's collaborative and supportive working atmosphere, and the possibility to participate in international conferences, have significantly contributed to my success.

My heartfelt thanks go to Rupert Müller, who supported my research in hyperspectral imaging techniques. Whether it was data acquisition, preprocessing, or technical discussions, Rupert has always been there to guide me. His expertise and dedication have advanced my work and deepened my understanding of the field.

I would like to express my sincerest appreciation to Dr. Daniele Cerra for his invaluable guidance in many aspects throughout my research. His expertise in hyperspectral data analysis and our insightful discussions have been vital in shaping my work. Additionally, he dedicated countless hours to proofreading my journal papers and this dissertation, considerably enhancing the quality of my work.

I am particularly grateful to Prof. Dr. Paul Scheunders from the University of Antwerp. Getting to know and collaborating with his Vision Lab was an invaluable experience. Prof. Scheunders invested significant time in our technical discussions and shared his extensive knowledge of spectral unmixing. His ideas gave me a unique perspective on research problems and encouraged me to explore new approaches. His detail-oriented mindset has greatly enhanced the quality of my work. I also appreciate the time he invested as a reviewer of my dissertation and as a member of my dissertation committee.

I would like to thank the German Academic Exchange Service (DAAD) for their financial support of my PhD study and DLR for providing the platform and resources.

I would like to sincerely thank all my colleagues at DLR. In particular, I would like to thank Dr. Rudolf Richter for his technical guidance on atmospheric correction of hyperspectral data, Dr. Pablo d'Angelo and Dr. Franz Kurz for their assistance on the Digital Surface Models, the OpAIRS team at the Remote Sensing Technology Institute (IMF) at DLR for their help with the HySpex dataset, Dr. Uta Heiden for her insights on hyperspectral image analysis, Dr. Eleonóra Víg for her suggestions on Deep Learning, Dr. Danfeng Hong for his helpful tips, and Peter Schwind and Maximilian Langheinrich for their technical support. Additionally, I am grateful to my peers in the scientific community, Dr. Rob Heylen, Dr. Vera Andrejchenko, Christian Trainotti, and many others, for sharing their time and profound knowledge with me and greatly enriching my research.

Finally, my deepest gratitude goes to my family: my parents, who brought me to this wonderful world and support me at all costs, and my husband, who loves me with his heart and shares his life with me. I feel great of having you all along my journey.

Contents

1

Introduction

"The more I live, the more I learn. The more I learn, the more I realize, the less I know."

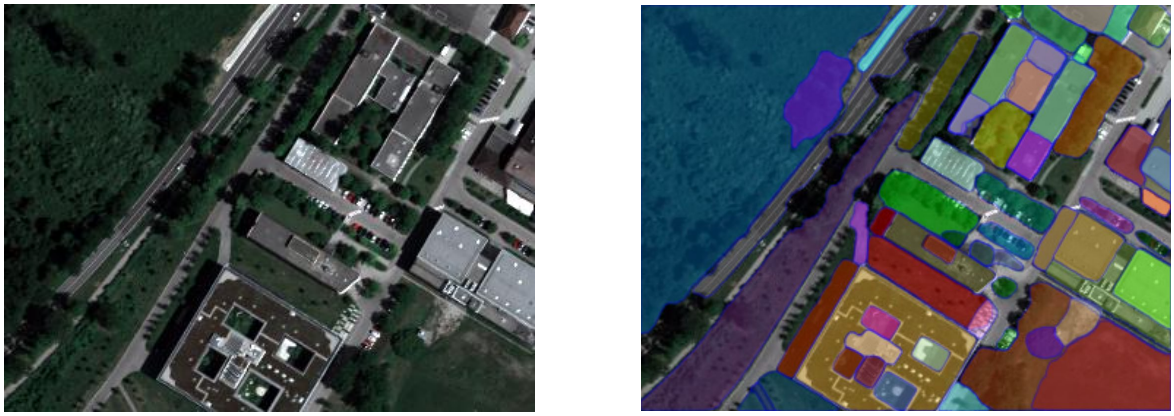
- Michel Legrand

Contents

1.1 Motivation and Challenges

Over the past century, remote sensing has become an important technique to observe Earth's surface from a distance, as it allows monitoring large regions of the ground surface time-efficiently from airborne and spaceborne sensors. Hyperspectral remote sensing, which has evolved from hyperspectral imaging or imaging spectroscopy, has attracted increasing attention for Earth observation [1]. Compared to previous optical systems, such as panchromatic and multispectral, hyperspectral imaging collects spectral signals using typically hundreds of spectral bands with narrow spectral bandwidth, e.g., 10 nm, thus acquiring more precise spectral information of a target on ground. This brings opportunities for the use of hyperspectral imaging in various applications in Earth observation, such as precision agriculture, food security, mineralogy, water constituents analysis, and precise urban and natural resources monitoring [2–4].

Fig. 1.1 (a) presents an airborne hyperspectral image as a true color composite in a suburban area and gives an example of shadow effects. Shadows, formed when an object (partly) blocks light from an illumination source, is one of the main inevitable phenomena in optical imagery and can have an undesired impact in image analysis, such as inaccurate classification and segmentation results [5]. For instance, the state-of-the-art image segmentation framework Segment Anything Model (SAM) was applied to Fig. 1.1 (a), partitioning the image into homogeneous regions as shown in Fig. 1.1 (b). This segmentation framework was trained on 11 million images and demonstrates good overall transferability to remote sensing imagery in segmentation tasks [6]. Nevertheless, several unsuccessfully segmented regions can be spotted, in particular in shadowed regions, where the same object, if partially shadowed, is segmented in two regions following the shadow boundaries.



(a) (b)

Figure 1.1: An example of image segmentation in the area of Oberphaffenhofen, Bavaria, Germany: (a) an airborne hyperspectral image shown as a true color composite acquired by the HySpex sensor (VNIR-1600); (b) image segmentation result of (a) generated by SAM [7].

One solution to this challenge is deriving shadow-removed images by simulating how affected pixels would appear if illuminated by direct sunlight, and using these data as input for image segmentation. Shadow detection and removal for optical images have been addressed since the 1990s. From then on, numerous methods have been proposed to detect, mitigate,

and remove shadow effects [8–10]. Previous works have concentrated on shadow issues for panchromatic and Red, Green, and Blue (RGB) imagery, because their spatial resolutions are typically very high and shadows can impact image analysis significantly. Thanks to the rapid development of sensor technology in the past decades, airborne hyperspectral imaging can also provide high-resolution images, making shadow effects non-negligible. Nevertheless, shadow effects in hyperspectral imagery have attracted limited attention, and several critical questions need to be sufficiently addressed in resolving shadow issues.

Question 1: How can we resolve shadow effects by considering the characteristics of hyperspectral imagery?

Featuring up to hundreds of narrow spectral bands, hyperspectral imagery offers a fine description of spectral features. This brings valuable information when tackling shadow issues but also raises several challenges.

- With the increase of spectral bands, the *curse of dimensionality* can happen, where analyzing data in high-dimensional spaces becomes difficult. Moreover, the inter-class similarity and intra-class variability make it challenging to extract meaningful information [11, 12].
- A large amount of bands offers a fine description of spectral features. However, this indicates that more spectral features need to be recovered in shadowed regions.
- Hyperspectral imagery is a type of raster data, where each pixel records signals of an area whose size is associated with the ground sampling distance. Thus, the signal at one pixel can be a complex mixture consisting of several ground objects under both sunlit and shadowed illumination conditions.
- Hyperspectral remote sensing imagery is typically acquired above the Earth's surface from a distance of hundreds to thousands of kilometers. In passive optical remote sensing images, in addition to shadow effects, other factors can affect the path from the sensor to the Earth's surface during the image acquisition. In particular, the atmosphere plays an essential role, as it affects the propagation of electromagnetic radiation. This brings a coupled and complicated influence on the imaging system.
- Training datasets are necessary when dealing with shadow effects using machine learning, especially for deep learning frameworks. For hyperspectral imagery, the increased dimensionality along the spectral direction requires larger amounts of training data. However, it is challenging and expensive to build training datasets for shadow detection and removal tasks for hyperspectral remote sensing imagery. Thus, approaches that require a limited amount of training samples are desired.

Question 2: How can we take advantage of three-dimensional (3D) data when dealing with shadow effects in hyperspectral imagery?

In addition to two-dimensional (2D) imagery, 3D representation of the Earth's surface is an additional data source in spaceborne and airborne image analysis [13]. A representative type of 3D data is Digital Surface Models (DSMs), which provides a detailed description of surface topography by capturing the elevation of the Earth's surface including natural

and artificial features, with respect to a fixed elevation reference, such as a reference geoid [13, 14].

Previous works have demonstrated the advantages of applying DSMs together with spectral information in answering remote sensing questions [15, 16]. In particular, the occlusion of light causing shadow effects can be explained by DSMs and auxiliary data, such as the acquisition time and the geo-location of the study area [17].

Question 3: How can we jointly consider spectral and spatial information in resolving shadow issues in hyperspectral imagery?

Shadow detection and removal are fundamental and challenging tasks in image analysis, and resolving shadow issues requires an understanding of the image semantics [18]. Spatial information captures relationships among pixels or regions. Previous works have demonstrated that jointly considering spatial and spectral information can significantly boost performance in various remote sensing tasks, such as image classification [19], segmentation [20], and feature learning [21, 22]. Specifically, different approaches have been proposed for different tasks. In order to jointly consider spectral and spatial information in resolving shadow problems, it becomes essential to understand the underlying relations between shadow effects, spectral, and spatial information.

1.2 Research Scope

In order to answer the above-mentioned research questions, the goal of this dissertation is to develop advanced and robust algorithms for resolving shadow issues in hyperspectral images, whose results potentially support widespread image analysis methods in presence of shadow effects. In detail, the research objectives are:

Objective 1: According to *Question 1*, shadow detection and removal for hyperspectral remote sensing imagery face additional challenges, compared to traditional panchromatic and RGB imagery. Specifically, coupled factors should be considered, namely the curse of dimensionality, the retrieval of information at the sub-pixel level, atmospheric effects, and limited training data. Recently, physics-based approaches have attracted attention for tackling shadow problems through physical assumptions. If one can model the shadow formation using a few physical-interpretable and shadow-related parameters, detecting and removing shadow effects via those parameters is promising. Nevertheless, existing methods lack studies on modeling shadow effects, especially for hyperspectral images. This leads to the first objective of this dissertation: *developing a physics-aware shadow detection and removal framework exclusively for hyperspectral images.*

Objective 2: Understanding shadow formation and modeling shadow effects accurately are essential in resolving shadow issues. In practice, modeling shadow effects is difficult. Existing works deal with shadow effects using simplified physical assumptions and can not represent all typical situations of optical interactions. Hence, the second object of this dissertation focuses on *improving the physics-aware modeling of shadows by considering more comprehensive scenarios in the presence of shadows and nonlinear optical interactions.*

Objective 3: According to **Objective 2**, an improved shadow-aware model accounts for typical ground scenarios. Nevertheless, resolving unknown parameters in the model is an inverse and ill-posed problem [23]. Incorporating auxiliary information can improve the performance of the inverse process and improve the robustness of the parameter estimation. In order to answer *questions 2 and 3*, the third objective of this dissertation is *developing a spatial-spectral shadow-aware spectral unmixing method that jointly considers spectral, spatial, and topographical information in dealing with shadow issues.*

1.3 Outline of this Dissertation

This is a *cumulative* dissertation for hyperspectral image analysis focusing on shadow issues.

Chapter 2 introduces fundamental knowledge associated with the objectives of this dissertation. It starts with different categories of remote sensing data, with a focus on the characteristics of passive optical images. Later, it introduces important physical quantities in radiometric propagation, which becomes the foundation to present two important radiometric effects in passive remote sensing images, namely shadows and atmosphere. Afterward, it continues with the essential technique in hyperspectral image analysis, i.e., spectral unmixing, and related several classic solvers.

Chapter 3 gives a comprehensive summary and discussion of the state-of-the-art methods. It analyzes shadow-aware approaches for different types of optical remote-sensing imagery. The methods are divided into three groups, according to image categories: panchromatic, RGB, and multi- and hyperspectral. Each group discusses shadow detection and removal methods separately. Finally, it summarizes the contributions of this dissertation.

Chapter 4 introduces three datasets used in this dissertation, including two datasets acquired by airborne sensors and one dataset by ground measurement.

Chapter 5 presents a summary of main contributions associated with three peer-reviewed journal publications, corresponding to three **Objectives** in Section 1.2. The full papers can be found in the attached Appendix. Later, an overall discussion is presented to address the remaining open points.

Chapter 6 concludes this dissertation and unveils potential research directions in the future.

2

Theoretical Background

"A picture is worth a thousand words. Is this true, and if so, why?"

- James B. Campbell and Randolph H. Wynne

Contents

2.1 Remote Sensing Imagery

Remote sensing refers to the acquisition of physical characteristics of an object, area, or phenomenon from a distance. In the scope of this dissertation, remote sensing refers to acquiring information about Earth, including its surface and atmosphere, using sensors mounted on platforms [24]. Typical remote sensing platforms include ground-based (mobile and tripod), airborne (manned aircraft and unmanned aerial vehicles), and satellite platforms [25]. In earlier times, aerial photographs were acquired and utilized to produce landscape measurements with extensive human labor. In 1972, the launch of Landsat-1 marked a milestone for systematic and repetitive observation of Earth's surface and introduced an era of digital image analysis for remote sensing [24]. According to image acquisition techniques [26], remote sensing can be categorized into *active* remote sensing and *passive* remote sensing (Fig. 2.1). In active remote sensing, sensors emit and receive electromagnetic radiation to and from observed targets, respectively, such as radar, Synthetic Aperture Radar (SAR), and Light Detection and Ranging (LiDAR). In passive remote sensing, sensors capture the radiation that is scattered, transmitted, and emitted from targets, typically originating from solar illumination.

In images of the Earth's surface collected by passive sensors, each pixel captures the solar radiation scattered and emitted from an area associated with the Instantaneous Field of View (IFOV) of the sensor at one or more regions in the electromagnetic spectrum. Each region in the spectrum is referred to as a spectral band. Spatial and spectral resolutions are two fundamental measures of images. Spatial resolution can be defined as the smallest spatial area on the ground surface measured by a sensor, and its value depends on the imaging system and the distance between the sensor and the targets. Spatial resolution is often associated with Ground Sampling Distance (GSD) that measures the distance between the centers of two neighboring pixels on the ground surface. Spectral resolution defines the ability to resolve features in the electromagnetic spectrum and can be identified as the width of a spectral band. The weighted average of wavelengths against spectral response functions is referred to as the central wavelength [27].

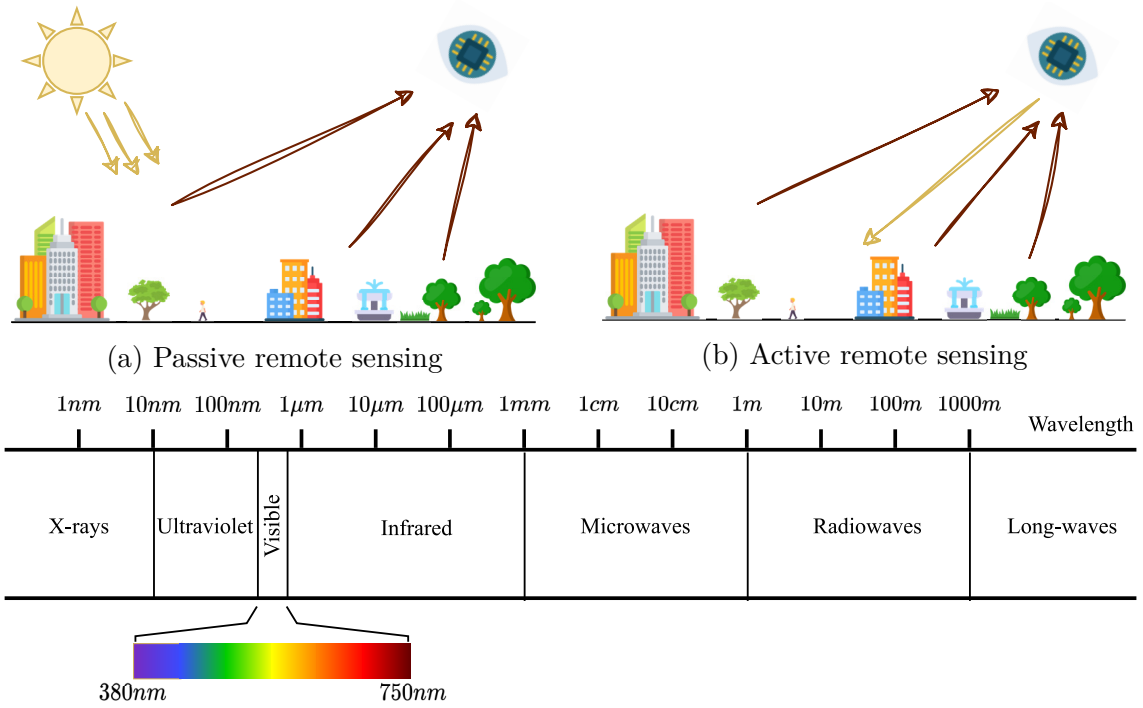
Typical remote sensing sensors are panchromatic, multispectral, and hyperspectral. Fig. 2.2 illustrates a comparison of their spectral characteristics.

2.1.1 Panchromatic

Panchromatic imagery captures back-scattered radiation using one single spectral band, which covers several regions of the electromagnetic spectrum, namely green, red, and near-infrared, ranging typically from 450 nm to 900 nm. Panchromatic sensors compensate their limited spectral information with a higher spatial resolution [24].

2.1.2 Multispectral

Multispectral imagery captures images at more than one spectral band, typically from 3 to 15, and may acquire radiation in the electromagnetic spectrum ranges of ultra-violet, visible,



(a) Passive remote sensing

(b) Active remote sensing

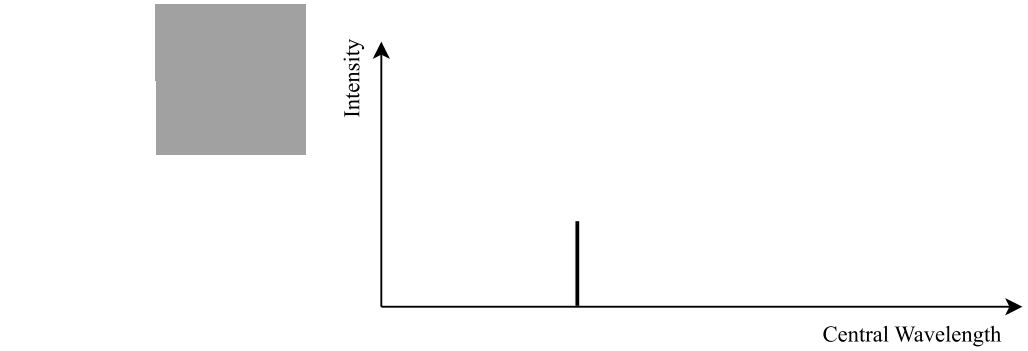
(c) Electromagnetic spectrum

Figure 2.1: Remote sensing image acquisition with the passive form in (a) and the active form in (b). Typical types of imaging with electromagnetic spectrum from small to large wavelengths in (c).

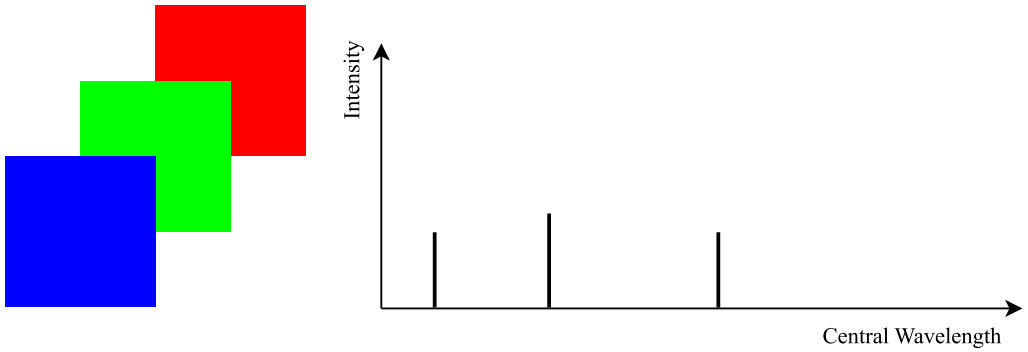
and infrared. Each spectral band typically spans from tens to thousands of nanometers. Compared to panchromatic imagery, the higher number of spectral bands in multispectral imagery results in a decrease in spatial resolution [28].

The increased number of bands provides enhanced spectral information in multispectral imagery [29]. For instance, Near-infrared (NIR) bands have been used to compute Normalized Difference Vegetation Index (NDVI), whose results indicate the health condition of vegetation [30]. Combined usage of Short-wave Infrared (SWIR), NIR, and red band shows the vegetation density [31]. SWIR, NIR, and the blue band can be used to monitor crop health [32]. Moreover, Long-wave Infrared (LWIR), also named as Thermal Infrared (TIR), has been applied to observe object temperatures. In this case the measured radiation is related to heat emitted from the targets instead of backscattered solar radiation. Typical applications are smoke monitoring, forest fire mapping, and sea surface temperature estimation [33].

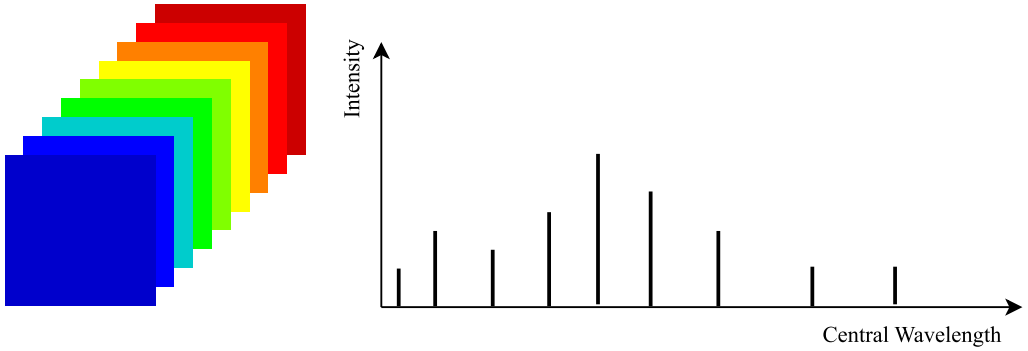
A specific category of multispectral imagery is named RGB [28]. The RGB data consist of three spectral bands that represent the ranges to which the human eye is sensitive, i.e., red, green, and blue, whose central wavelengths are typically around 640 nm, 510 nm, and 470 nm. RGB can thus be easily interpreted by humans and has been widely used in computer vision and remote sensing.



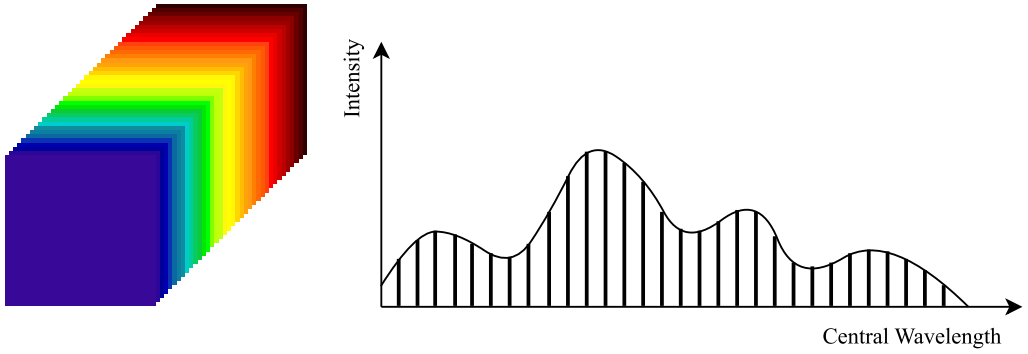
(a) Panchromatic imagery



(b) RGB imagery



(c) Multispectral imagery



(d) Hyperspectral imagery

Figure 2.2: An illustration of spectral information at one pixel in (a) panchromatic imagery, (b) RGB imagery, (c) multispectral imagery, and (d) hyperspectral imagery.

2.1.3 Hyperspectral

Hyperspectral imagery, also referred to as imaging spectrometer data, captures the radiation from the Earth's surface across the electromagnetic spectrum using typically hundreds of spectral bands [34–36]. The spectral resolution of hyperspectral data can range from 2 nm to 2000 nm, whose typical values are below 20 nm. Hence, hyperspectral imagery captures radiation at higher spectral resolution and can collect more spectral features of ground objects compared to multispectral images. The higher spectral resolution allows for distinguishing materials using narrow bands, which may not be possible using multispectral data [37].

Hyperspectral sensors have been mounted on different platforms. Representative airborne hyperspectral sensors are Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), Compact Airborne Spectrographic Imager (CASI) and Hyperspectral Mapper (HyMap), while representative spaceborne hyperspectral sensors are Hyperion, Compact High Resolution Imaging Spectrometer (CHRIS), Moderate Resolution Imaging Spectroradiometer (MODIS), Deutsches Zentrum für Luft- und Raumfahrt Earth Sensing Imaging Spectrometer (DESIS), and Environmental Mapping and Analysis Program (EnMAP).

Previous works have demonstrated the advantages of hyperspectral imagery over multispectral data in different applications [2–4]. In order to provide a broad overview, we list not only applications in Earth observation but also in other fields.

- Agriculture: Many works have used hyperspectral images for precision agriculture to estimate Leaf Area Index (LAI) [38], leaf chlorophyll [39], biomass [40], nitrogen content [40], water content [41], and to identify types and health status of plants [41–43].

- Urban and Natural Resources Monitoring: A number of works have demonstrated the superiority of hyperspectral images in the precise classification of urban regions [44], change detection [45], water resource management [35], assessment and warning of flood [46], and mining [47].

- Food Security: Food products require high quality and safety, demanding an objective, fast, and accurate determination system for quality identification in the food industry. Hyperspectral images present advantages in measuring, inspecting, sorting, and grading food products efficiently and effectively [48]. With a hyperspectral imaging system, one can evaluate meat quality and ripening of fruits [49], identify wheat kernel types [50], and detect defects [36].

- Medical Diagnosis: Hyperspectral imaging has also proven its advantages in medical image analysis, such as the diagnosis of tumors and classification of different body tissues, whose spectral characteristics present strong patterns at specific ranges of wavelengths [51, 52].

- Waste Sorting and Recycling: As hyperspectral imagery is able to distinguish chemical constituents of materials, it has also been found helpful in sorting wastes [53, 54], such as cardboards, glass, metal, and plastics.

2.2 Physical Quantities in Radiometry

Radiometry is formally defined as the science of characterizing or measuring how much electric magnetic energy is present at, or associated with, some location or direction in space [55]. Physical terms are used to describe different radiometric concepts. This section introduces several physical quantities related to remote sensing imaging.

2.2.1 Radiant Energy

Energy is carried by one or more photons in a ray or beam. Given the frequency index i, frequency ν_i, number of photons n_i, and reduced Planck's constant $\hbar$, the total energy $\mathcal{Q}$ can be written as:

$$\mathcal{Q} = 2\pi \sum_i n_i \hbar \nu_i \qquad [\mathrm{J}] \tag{2.1}$$

2.2.2 Radiant Flux

Typically, it is more convenient to consider a ray in terms of the rate of flow of energy, resulting in radiant flux Φ:

$$\Phi = \frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}t} \qquad [\mathrm{W}] \tag{2.2}$$

2.2.3 Irradiance

Irradiance is defined as the radiance flux *delivered to* a surface per unit area:

$$\mathcal{E} = \frac{\mathrm{d}\Phi_{\mathrm{in}}}{\mathrm{d}A} \qquad [\mathrm{W\,m^{-2}}] \tag{2.3}$$

where $\mathrm{d}A$ is an area element on the surface of interest.

2.2.4 Radiant Exitance

Radiant exitance measures the radiant flux per unit area *exiting from* a surface:

$$M = \frac{\mathrm{d}\Phi_{\mathrm{out}}}{\mathrm{d}A} \qquad [\mathrm{W\,m^{-2}}] \tag{2.4}$$

2.2.5 Radiant Intensity

The radiant intensity describes the radiant flux for each element of solid angle from a point source into a specific direction. The element of solid angle $\mathrm{d}\Omega$ is defined as the conic angle encompassing the area element $\mathrm{d}A$ on the surface of a sphere (Fig. 2.3).

$$I = \frac{\mathrm{d}\Phi}{\mathrm{d}\Omega} \qquad [\mathrm{W\,sr^{-1}}] \tag{2.5}$$

where the element of solid angle $\mathrm{d}\Omega = \sin\theta \mathrm{d}\theta \mathrm{d}\phi$, and θ and ϕ are orientation angles in the spherical coordinate system.

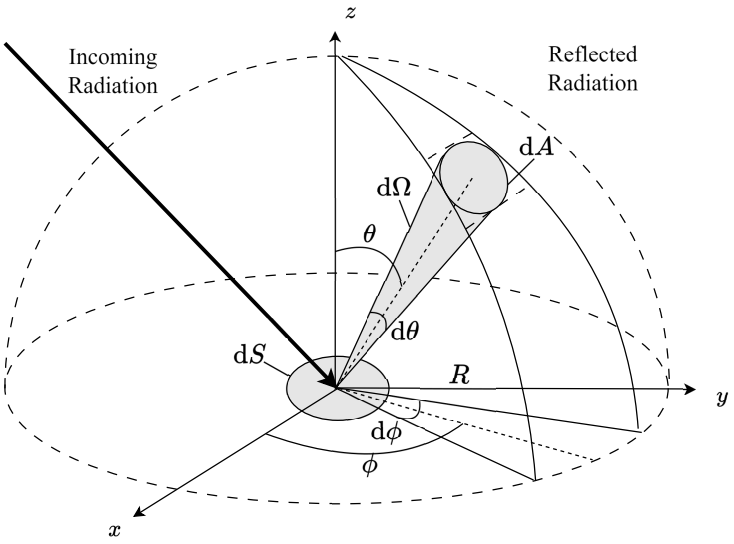


Figure 2.3: Geometric representation of the element of solid angle Ω in the spherical coordinate system. Unlike elsewhere in this dissertation, please note that (x, y, z) in this figure defines the spherical coordinate system whose polar angle is noted as θ, the azimuthal angle is noted as ϕ, and the radius is noted as R. $\mathrm{d}S$ and $\mathrm{d}A$ are area elements on the surfaces of interest.

2.2.6 Radiance

Radiance $\mathcal{L}$ describes the flux for each element of solid angle $\mathrm{d}\Omega$ in a certain direction θ concerning the plane of interest at the area element $\mathrm{d}A$.

$$\mathcal{L} = \frac{\mathrm{d}^2\Phi}{\mathrm{d}(A\cos\theta)\mathrm{d}\Omega} \qquad [\mathrm{W\,m^{-2}\,sr^{-1}}] \tag{2.6}$$

2.2.7 Optical Properties of Materials

Three processes can occur when the energy is incident on the surface of a material: reflection, absorption, and transmission. The fractions of energy reflected, absorbed, and transmitted by the material are defined respectively as reflectance, absorptance, and transmittance, indicating the property of the material itself.

The reflectance r is the unitless ratio of the reflected radiant exitance M_r to the incoming irradiance $\mathcal{E}$.

$$r = \frac{M_r}{\mathcal{E}} \tag{2.7}$$

The absorptance describes the ability of a material to convert irradiance to another form of energy, such as thermal energy. The absorptance α is defined as the ratio of the flux converted to another form of energy per unit area M_α to the incoming irradiance $\mathcal{E}$.

$$\alpha = \frac{M_\alpha}{\mathcal{E}} \tag{2.8}$$

The transmittance is defined as the ratio of the radiant exitance from the back of a material M_τ, to the incoming irradiance $\mathcal{E}$.

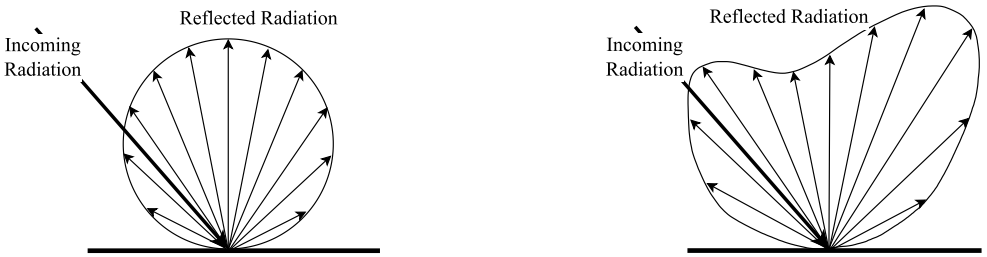
$$\tau = \frac{M_\tau}{\mathcal{E}} \tag{2.9}$$

Following the conservation of energy, all incident flux needs to be reflected, absorbed, or transmitted, thus:

$$r + \alpha + \tau = 1 \tag{2.10}$$

2.2.8 Lambertian and Non-Lambertian Surfaces

When a surface reflects energy, its radiant intensity is angularly distributed into the hemisphere above it. In an ideal case, one can assume that the radiant intensity is proportional to the cosine of the angle between the reflected radiation and the surface normal (the Lambert's cosine law [56]). This ideal surface is referred to as the Lambertian surface (Fig. 2.4). Following the relationship between radiant intensity and radiance (Eqs. (2.5) and (2.6)), the reflected radiance from a Lambertian surface will be the same in all directions. The isotropic feature of the Lambertian surface can significantly simplify the computational process and has become a reasonable approximation in many practical applications [55]. Nevertheless, the Lambertian assumption does not hold for some surfaces whose reflected radiation strongly depends on the reflection angle. These surfaces are referred to as non-Lambertian surfaces, and can be described by the Bidirectional Reflectance Distribution Function (BRDF) [57].



(a) Lambertian surface (b) Non-Lambertian surface

Figure 2.4: Two surface types: (a) Lambertian surface and (b) non-Lambertian surface.

2.3 The Shadow Effect in Optical Imaging

Shadows appear due to the occlusion in the light path from the illumination source(s) to the target object(s). The shadow formation consists of indoor and outdoor scenarios. Figure 2.5 illustrates a classic indoor case generated by image simulation via ray tracing. The scene comprises a point light source and a cube in a Cornell box [58]. Self-cast shadow occurs on the part of the object, where the object itself occludes the illumination source. Cast shadow is projected instead on nearby objects and consists of *umbra* and *penumbra* [59]: umbra, alternatively referred to as fully shadowed region, occurs where the illumination is completely blocked by the object; penumbra, also known as partly shadowed region, occurs where the illumination is partly blocked by the object. It is worth noting that, the reflected light of the illumination source, namely the ambient light, can reach cast shadow regions. As ambient light contains much less energy than direct illumination, shadowed regions scatter considerably less radiation to the sensor and appear much darker compared to sunlit areas [60].

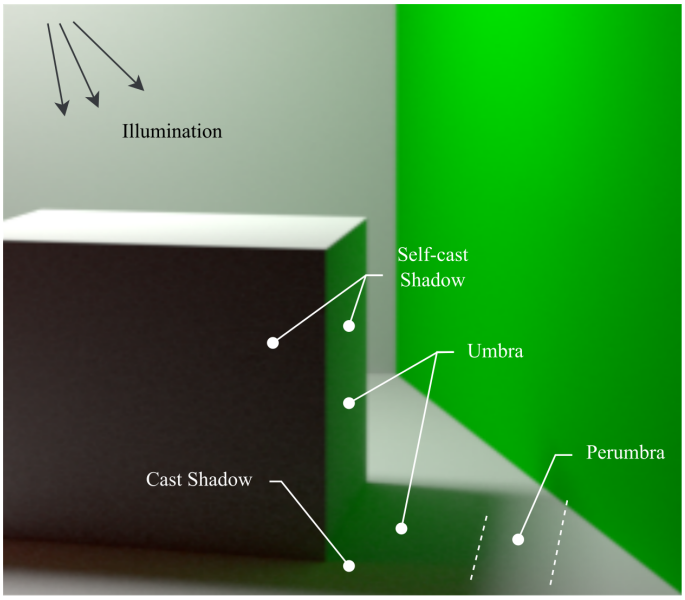


Figure 2.5: Indoor shadow formation illustrated in the ray traced image with point light source, adapted from [61].

.

The outdoor scenario consists of the solar illumination and occluding objects. Similar to indoor shadows, outdoor shadows can be categorized into self-cast and cast shadows, together with umbra and penumbra regions. However, outdoor scenes are typically more complicated. The shadow formation needs to be explained by the imaging chain containing all optical interactions associated with the sun, atmosphere, ground objects, and the sensor. According to the properties of occluding objects, the outdoor shadows can be further divided into object shadows and cloud shadows (see Fig. 2.6).

Following the radiation propagation theory [55], the solar energy paths describe possible radiation propagation from the illumination source, namely the sun, to the sensor. The

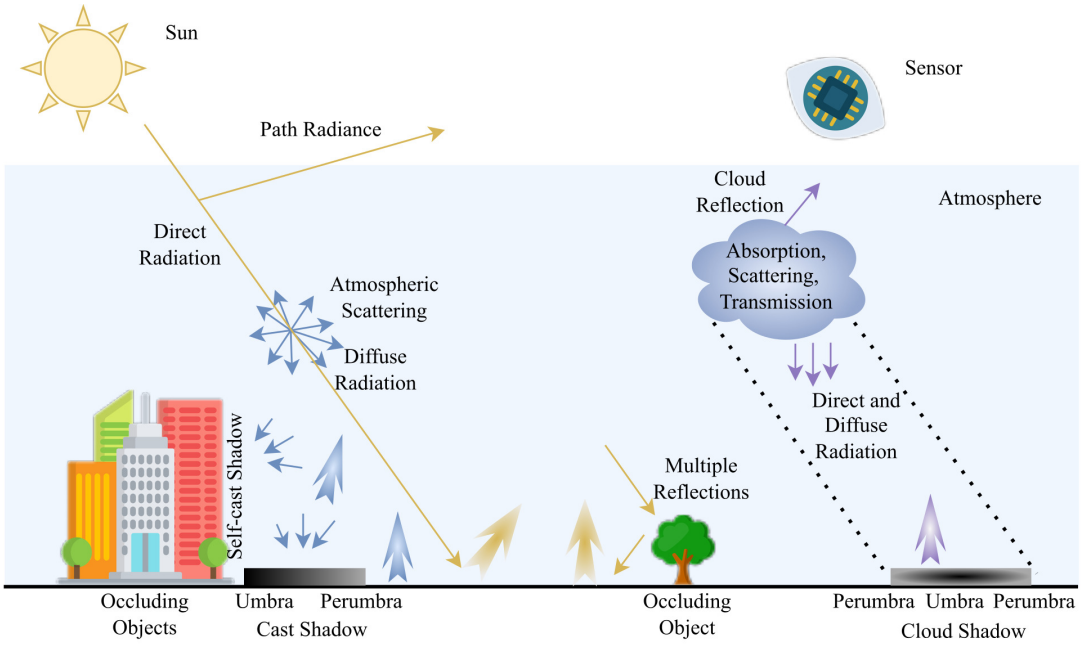


Figure 2.6: An illustration of outdoor shadow formation in optical remote sensing imagery.

dominant solar energy path is that the photons originate from the sun, pass through the atmosphere, interact with the ground surface, and are back-scattered to the sensor via the atmosphere. These photons reaching the ground surface form an important illumination source named direct solar radiation (direct sunlight). Specifically, direct solar radiation can not reach those regions completely blocked by occluding objects in the energy path, resulting in shadows (umbra). The other essential energy path is that the photons originating from the sun are scattered in all directions by the atmosphere and interact with the ground surface before being scattered back to the sensor. These scattered photons form the other important illumination source named diffuse solar radiation (skylight), which is considerably smaller than the direct solar radiation. It is worth noticing that the diffuse solar radiation scattered in all directions in the atmosphere can reach the ground surface, with and without the object's occlusion, thus becoming the dominant illumination source in shadowed areas.

In another possible energy path, the photons interact with more than one ground surface via multiple reflections before being scattered back to the sensor. This effect is considerably smaller than direct radiation but can also become significant in some situations, such as in the presence of shadows and vegetation [62]. Moreover, photons initiated from the sun can also interact with the atmosphere and be scattered back to the sensor without reaching the ground surface, leading to the path radiance [63].

Furthermore, cloud shadow is a typical scenario in outdoor optical imaging, especially in airborne and spaceborne remote sensing images. Compared to object shadows, clouds may not completely occlude but can only attenuate direct solar radiation. Thus, a part of attenuated direct radiation can transmit through clouds and reach the ground surface [64]. In particular, the impact of clouds on solar radiation depends on cloud types, leading to

complicated effects by cloud shadows. Moreover, it is worth noticing that diffuse radiation takes part in cloud shadows.

2.4 Atmospheric Correction

Remote sensing images are typically acquired above the ground surface using aircraft and satellites. The atmosphere introduces one of the largest impacts on radiation propagation [63, 65]. The atmosphere alters the spectral values of radiation reaching the sensor, leading to difficulties in interpreting target characteristics on the ground surface. The atmospheric correction aims to remove atmospheric effects on the reflectance values of images recorded by satellite and airborne sensors [63].

Typically, the sensor records the scaled radiance, namely digital number DN. The at-sensor radiance $\mathcal{L}$ can be obtained with the radiometric offset c_0 and c_1:

$$\mathcal{L} = c_0 + c_1 DN \tag{2.11}$$

In the range of wavelength spectrum between 0.35 µm and 2.5 µm [i], the at-sensor radiance consists of three components (Fig. 2.7):

$$\mathcal{L} = \mathcal{L}_{\mathrm{p}} + \mathcal{L}_{\mathrm{reflect}} + \mathcal{L}_{\mathrm{adj}} \tag{2.12}$$

where $\mathcal{L}_{\mathrm{p}}$ represents path radiance, in which photons from the sun interact with the atmosphere and are scattered into the sensor's IFOV without interacting with the ground surface. $\mathcal{L}_{\mathrm{reflect}}$ represents the reflected radiance of a single pixel that receives direct and diffuse solar radiation. $\mathcal{L}_{\mathrm{adj}}$ represents the reflected radiation from the neighborhood of the target pixel, scattered by the atmosphere into the IFOV of the target pixel.

Note that the ground reflectance value as r, and the ground-to-sensor atmospheric transmittance as τ, and the global irradiance on the ground as $\mathcal{E}_{\mathrm{g}}$, the reflected radiance $\mathcal{L}_{\mathrm{reflect}}$ can be computed as:

$$\mathcal{L}_{\mathrm{reflect}} = \frac{r\mathcal{E}_{\mathrm{g}}}{\pi} \tag{2.13}$$

where $\mathcal{E}_{\mathrm{g}} = \tau_{\mathrm{dir}}\mathcal{E}_{\mathrm{dir}} + \tau_{\mathrm{diff}}\mathcal{E}_{\mathrm{diff}}$, $\mathcal{E}_{\mathrm{dir}}$ represents the direct solar irradiance, and $\mathcal{E}_{\mathrm{diff}}$ represents the diffuse solar irradiance. The transmittances of the direct and diffuse solar radiation are τ_{dir} and τ_{diff}, respectively.

Thus, the atmospherically corrected ground reflectance $\hat{r}$ can be written as:

$$\hat{r} = \frac{\pi\left[(c_0 + c_1 DN) - \mathcal{L}_{\mathrm{p}}\right]}{\tau_{\mathrm{dir}}\mathcal{E}_{\mathrm{dir}} + \tau_{\mathrm{diff}}\mathcal{E}_{\mathrm{diff}}} \tag{2.14}$$

[i] The atmospheric correction for thermal spectral ranges follows different processes and is out of the scope of this thesis.

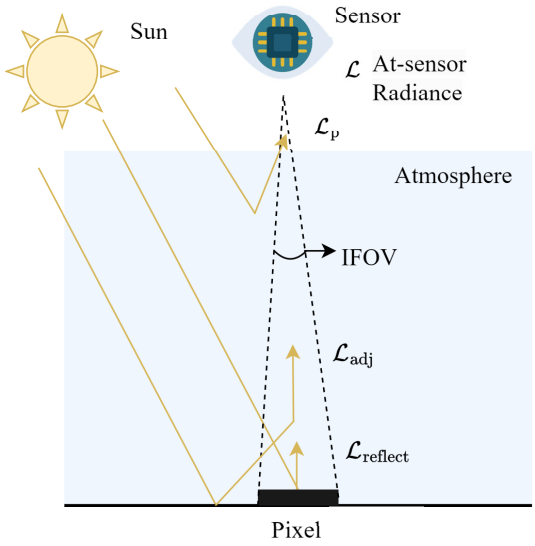


Figure 2.7: An illustration of solar radiation components in flat terrain. Please note that the reflection angles do not strictly correspond to reality, but rather a simplification for the sake of depiction.

2.5 Spectral Unmixing

In hyperspectral imaging, each single sensor detector captures the signal backscattered from one ground pixel associated with the sensor's IFOV. After atmospheric correction (see Section 2.4), each pixel corresponds to a surface reflectance spectrum. Because of limits imposed by spatial resolution, a pixel can contain one or more object(s) or material(s). Thus, spectral unmixing has become an essential technique to analyze material composition within a pixel [23, 62]. Specifically, spectral unmixing is the process to decompose a pixel into a collection of the spectral signatures of pure materials (i.e., endmembers) and their corresponding contributions (i.e., abundances) [23, 62, 66]. The unmixing process can be unsupervised and supervised. Unsupervised spectral unmixing aims at solving both the endmembers identification and abundances estimation. Typical steps include estimating the number of endmembers, endmember extraction, and abundance estimation. Representative methods for computing the number of endmembers are Hyperspectral Signal Identification by Minimum Error (HySime) [67], Harsanyi–Farrand–Chang (HFC) [68], and termed noise-whitened HFC [69]. Popular methods for endmemebr extraction are Pure Pixel Index (PPI) [70], Vertex Component Analysis (VCA) [71], and Simplex Growing Algorithm (SGA) [72]. In supervised unmixing, endmembers, usually collected in a so-called endmember library, are given in advance, reducing the unmixing process to abundance estimation.

Spectral unmixing is an inverse problem of spectral mixing, which investigates how materials contribute to the individual pixel spectrum measured by the sensor [73]. Fig. 2.8 illustrates the spectral mixing and unmixing processes, where a mixed pixel comprises three different materials with their corresponding fractions. Developing a good mixing model is one of the uttermost prior conditions for a successful spectral unmixing process.

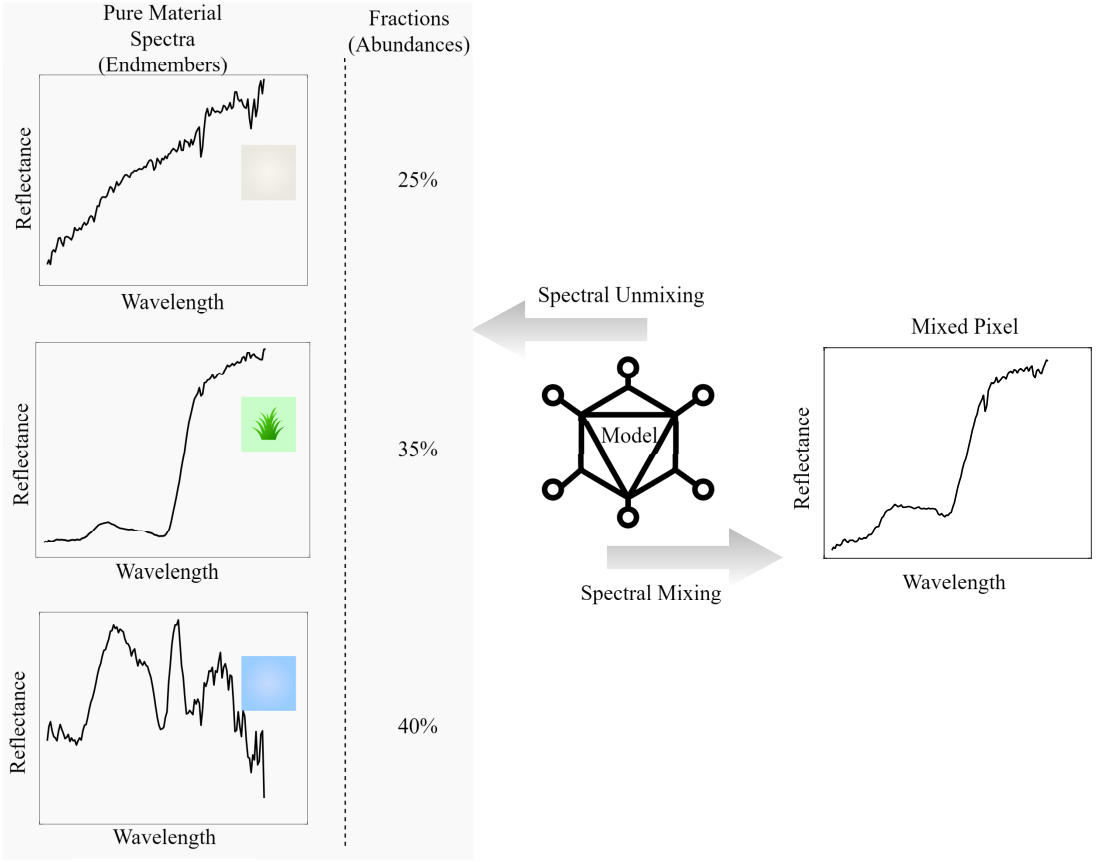


Figure 2.8: An example of the spectral mixing and unmixing processes, where a mixed pixel consists of three pure materials.

Methods based on radiative transfer model the optical interactions from the illumination sources to the sensor and can build an accurate mixing model [17, 74]. However, this is a non-trivial problem. Firstly, this model requires a detailed knowledge of ground geometry and radiometry data that are typically unavailable or expensive. Besides, this inverse problem is difficult to solve because of the complexity of the radiative transfer model [66, 75]. Thus, in the past decades some works have derived simplified physics-based mixing models following various assumptions. Fig. 2.9 illustrates four typical spectral mixing scenarios: pure pixel, linear mixing, bilinear mixing, and intimate mixing. A pure pixel consists of one material (i.e., endmember) with 100% abundance. In practice, most scenarios consist of more than one endmembers at one pixel, resulting in various mixing types: linear, bilinear, and intimate mixing models. These models are discussed in details in Sections 2.5.1 and 2.5.2.

For the reader's convenience, we highlight here essential mathematical notations utilized in the following sections. Please refer to the List of Symbols for a comprehensive description of notations used in this dissertation. Denote a hyperspectral image with B spectral bands and N pixels as $\boldsymbol{X} \in \mathbb{R}^{B\times N}$, with $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N]$, where pixel $\boldsymbol{x}_j = (x_{j,1}, x_{j,2}, \cdots x_{j,B})^T \in \mathbb{R}^{B\times 1}$. The reconstructed image of $\boldsymbol{X}$ after spectral unmixing is noted as $\hat{\boldsymbol{X}}$, with $\hat{\boldsymbol{X}} = [\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \cdots, \hat{\boldsymbol{x}}_N]$, where pixel $\hat{\boldsymbol{x}}_j = (\hat{x}_{j,1}, \hat{x}_{j,2}, \cdots \hat{x}_{j,B})^T \in \mathbb{R}^{B\times 1}$. An endmember library

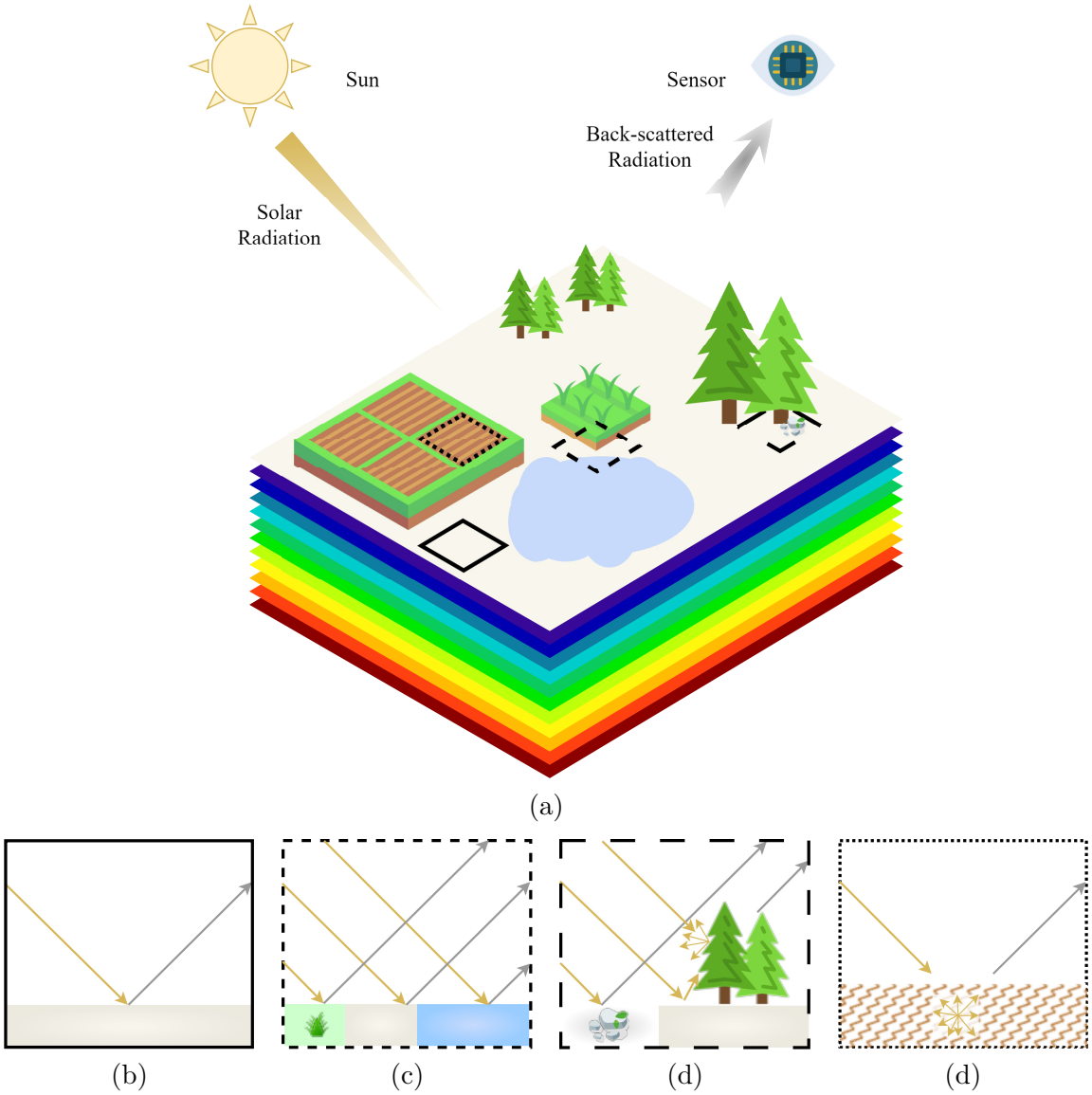


Figure 2.9: Four typical spectral mixing scenarios in outdoor condition: (a) pure pixel, (b) linear mixing, (c) bilinear mixing, and (d) intimate mixing.

$\boldsymbol{E} \in \mathbb{R}^{B\times p}$ consists of p endmembers, where the i-th endmember is denoted as $\boldsymbol{e}_i \in \mathbb{R}^{B\times 1}$. Denote an abundance matrix associated with $\boldsymbol{E}$ as $\boldsymbol{A} \in \mathbb{R}^{p\times N}$, with $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_N]$, where $\boldsymbol{a}_j \in \mathbb{R}^{p\times 1}$. $\boldsymbol{O} = [\boldsymbol{o}_1, \boldsymbol{o}_2, \cdots, \boldsymbol{o}_N] \in \mathbb{R}^{B\times N}$ represents the residual matrix of a spectral mixing model.

2.5.1 Linear Spectral Unmixing

The Linear Mixing Model (LMM) follows the straightforward assumption that the incoming solar illumination interacts with a pixel only once before being scattered back to the sensor [76]. The spectral mixture is then given by the sum of the material spectra, weighted by their spatial proportion within the pixel. This simple model assumes an ideal scene structure,

with a flat ground surface and spatially separable ground materials [62, 66].

Following the above assumptions, a mixed pixel $\boldsymbol{x}_j$ based on LMM is obtained as:

$$\boldsymbol{x}_j = \sum_{i=1}^{p} \boldsymbol{e}_i a_{j,i} + \boldsymbol{o}_j \tag{2.15}$$

In practice, we can write Eq. (2.15) to model all pixels $\boldsymbol{X} \in \mathbb{R}^{B \times N}$ of an image in the matrix form in Eq. (2.16).

$$\boldsymbol{X} = \boldsymbol{E}\boldsymbol{A} + \boldsymbol{O} \tag{2.16}$$

Following physical interpretations, different constraints have been applied to spectral unmixing.

- **Non-negative Constraint**
  Endmembers are reflectance values of materials that are positive by definition. Besides, the abundances represent the proportions of endmembers, and no endmembers can have negative proportions. Thus, the non-negative constraint on endmembers and abundances, i.e., Abundance Nonnegative Constraint (ANC), is commonly applied.

  $$\boldsymbol{E} > \boldsymbol{0}, \boldsymbol{A} > \boldsymbol{0} \tag{2.17}$$

- **Abundance Sum-to-one Constraint (ASC)**
  If an endmember library covers all materials in a scene, a pixel j can be completely decomposed in endmember contributions. In this case, the summation of abundances at pixel j follows:

  $$\boldsymbol{1}^T \boldsymbol{a}_j = 1 \tag{2.18}$$

- **Sparsity Constraints**
  Typically, one pixel contains one or several materials (i.e., endmembers), considerably less than the total number of materials in the endmember library. This phenomenon indicates a naturally sparse property of abundances at each pixel. Hence, sparsity constraints have been applied to spectral unmixing to promote the robustness of spectral unmixing [77–79]. A common strategy is embedding different types of regularization terms, such as the L2-norm $\|\boldsymbol{A}\|_{2,1}$, L1-norm $\|\boldsymbol{A}\|_{1,1}$, and L0-norm $\|\boldsymbol{A}\|_0$ [79]. Specifically, the L2-norm computes the Euclidean norm, such as in the Ridge regression; L1-norm computes the sum of the absolute values, such as in Least Absolute Shrinkage and Selection Operator (LASSO) [80]; L0-norm considers the number of non-zero elements, but it is non-convex and can be challenging to solve [81].

- **Total Variation (TV) Constraint**
  The TV constraint accounts for the spatial homogeneity of pixels in a local neighborhood. In particular, it is very likely that two neighboring pixels have similar material composition [82]. TV measures the norm of the finite differences of pixels in a local

neighborhood $\mathcal{N}(j)$ of pixel j:

$$\sum_{j=1}^{N} \sum_{m \in \mathcal{N}(j)} \|\boldsymbol{a}_j - \boldsymbol{a}_m\|_{1,1} \tag{2.19}$$

where $\mathcal{N}(j)$ denotes the set of pixels belonging to the local neighborhood of pixel j.

In practice, one or several constraints can be applied alone or combined to the LMM, resulting in several representative spectral unmixing methods. For example, the ANC and ASC are often applied along with the LMM, resulting in Fully Constrained Least Squares (FCLS) unmixing [23]. Nevertheless, it may be a strong assumption that an endmember library covers all materials in a scene. Thus, some later work relaxes ASC, allowing some materials in a pixel not to be present in the endmember library, resulting in Non-negative Least Squares (NNLS) unmixing [23]. Sparsity constraint has been jointly used together with the ANC in Sparse Unmixing via Variable Splitting Augmented Lagrangian (SUnSAL) [83]. Later, TV constraint has been added to SUnSAL, leading to SUnSAL-TV [82]. In addition, the TV constraint has been considered in the FCLS unmixing framework [84].

2.5.2 Nonlinear Spectral Unmixing

The linear mixing model depends on strict physical assumptions that are often not fulfilled in practice. In many situations, nonlinear optical interactions are non-negligible [62, 66, 85]. Nonlinear mixtures can occur at both microscopic and macroscopic levels and, depending on the size of the particles under investigation, different categories of models exist [62].

In the macroscopic scenario, nonlinear optical interactions can occur because of height differences between ground objects [62, 66]. In order to allow incoming light to interact more than once before being scattered back to the sensor, some nonlinear models use higher-order terms through the term-wise product of spectra. Fig. 2.9 (d) illustrates a typical scenario with second-order optical interactions. Specifically, after an incoming light interacts with a ground material at a pixel, it can either be scattered back to the sensor or interact with other materials at the pixel. These optical interactions can continue unlimited times, but most works regard only up to two interactions, resulting in a group of bilinear models in Eq. (2.20). Note that we omit the pixel index j in Eq. (2.20) for simplification while highlighting the fact that $\boldsymbol{x}$, a, m, and $\boldsymbol{o}$ are pixel-dependent as in LMM.

$$\boldsymbol{x} = \sum_{i=1}^{p} \boldsymbol{e}_i a_i + \sum_{i_1=1}^{p} \sum_{i_2=1}^{p} m_{i_1,i_2} \boldsymbol{e}_{i,1} \odot \boldsymbol{e}_{i,2} + \boldsymbol{o} \tag{2.20}$$

In addition to the linear interaction and the residual terms, the bilinear model contains additional terms representing the second order of optical interactions.The double scattering can be modeled through the element-wise multiplication of endmembers $\boldsymbol{e}_{i,1}$ and $\boldsymbol{e}_{i,2}$, where i_1 and i_2 represent endmember indices. Moreover, the bilinear models include parameters m, which can be associated with abundances and may differ according to physical assumptions, leading to various types of bilinear models, such as Nascimento, Fan, Polynomial Post-nonlinear Model (PPNM), and Generalized Bilinear Model (GBM) in Table 2.1.

Table 2.1: Constraints on parameters in bilinear models

Model	m_{i_1,i_2} in Eq. (2.20)	Constraints
Nascimento [86]	$\forall i_1 \geq i_2 : m_{i_1,i_2} = 0$ $\forall i_1 < i_2 : m_{i_1,i_2} \geq 0$	$\sum_{i=1}^{p} a_i + \sum_{i_1=1}^{p-1} \sum_{i_2=i_1+1}^{p} m_{i_1,i_2} = 1$ $\forall i : a_i \geq 0$
Fan [87]	$\forall i_1 \geq i_2 : m_{i_1,i_2} = 0$ $\forall i_1 < i_2 : m_{i_1,i_2} = a_{i_1} a_{i_2}$	$\sum_{i=1}^{p} a_i = 1$ $\forall i : a_i \geq 0$
PPNM [88]	$\forall i_1, i_2 : m_{i_1,i_2} = \gamma a_{i_1} a_{i_2}$	$\sum_{i=1}^{p} a_i = 1$ $\forall i : a_i \geq 0$ $\gamma \in \mathbb{R}$
GBM [89]	$\forall i_1 \geq i_2 : m_{i_1,i_2} = 0$ $\forall i_1 < i_2 : m_{i_1,i_2} = \gamma_{i_1,i_2} a_{i_1} a_{i_2}$	$\sum_{i=1}^{p} a_i = 1$ $\forall i : a_i \geq 0$ $\gamma_{i_1,i_2} \in [0, 1]$

- **Nascimento**
  The Nascimento model [86] regards the bilinear terms as new endmembers whose abundances are denoted as m_{i_1,i_2}, indicating that m_{i_1,i_2} is decorrelated with a_i. Assuming that a pixel can be fully decomposed into endmembers, the Nascimento model follows the sum-to-one constraint. In addition, it assumes that the incoming light can not interact with the same endmember twice, leading to zero values when $i_1 \geq i_2$. Thus, we write the sum-to-one constraint in Eq. (2.21). Besides, the abundances are assumed to be positive values, resulting in $a_i \geq 0$ and $m_{i_1,i_2} \geq 0$.

$$\sum_{i=1}^{p} a_i + \sum_{i_1=1}^{p-1} \sum_{i_2=i_1+1}^{p} m_{i_1,i_2} = 1 \tag{2.21}$$

- **Fan**
  The Fan model [87] considers the relationship between abundances of linear terms a_i and nonlinear terms m_{i_1,i_2}. Specifically, the model assumes that abundances of bilinear terms m_{i_1,i_2} should be proportional to the abundances of endmembers a_i. In addition, it excludes self-interactions, resulting in:

$$\boldsymbol{x} = \sum_{i=1}^{p} \boldsymbol{e}_i a_i + \sum_{i_1=1}^{p-1} \sum_{i_2=i_1+1}^{p} a_{i_1} a_{i_2} \boldsymbol{e}_{i_1} \odot \boldsymbol{e}_{i_2} + \boldsymbol{o} \tag{2.22}$$

  subject to

$$\sum_{i=1}^{p} a_i = 1, \forall i : a_i \geq 0 \tag{2.23}$$

- **Polynomial Post-nonlinear Model (PPNM)**
  Similar to the Fan model, PPNM retains the relationship of abundances between linear and bilinear terms. In addition, it allows self-interactions of endmembers with a hyperparameter γ. One can also interpret the PPNM as the nonlinear transformation

of LMM at the second order.

$$\boldsymbol{x} = \sum_{i=1}^{p} \boldsymbol{e}_i a_i + \gamma \sum_{i_1=1}^{p} \sum_{i_2=1}^{p} a_{i_1} a_{i_2} \boldsymbol{e}_{i_1} \odot \boldsymbol{e}_{i_2} + \boldsymbol{o} \tag{2.24}$$

subject to

$$\sum_{i=1}^{p} a_i = 1, \quad \forall i : a_i \geq 0, \quad \gamma \in \mathbb{R} \tag{2.25}$$

- **Generalized Bilinear Model (GBM)**
  GBM applies additional free parameters to bilinear terms. Compared to the Fan model, the GBM contains endmember-dependent parameters γ_{i_1,i_2} to describe the strength of nonlinear interactions between endmembers $\boldsymbol{e}_{i_1}$ and $\boldsymbol{e}_{i_2}$.

$$\boldsymbol{x} = \sum_{i=1}^{p} \boldsymbol{e}_i a_i + \sum_{i_1=1}^{p-1} \sum_{i_2=i_1+1}^{p} \gamma_{i_1,i_2} a_{i_1} a_{i_2} \boldsymbol{e}_{i_1} \odot \boldsymbol{e}_{i_2} + \boldsymbol{o} \tag{2.26}$$

  subject to

$$\sum_{i=1}^{p} a_i = 1, \forall i : a_i \geq 0, \gamma_{i_1,i_2} \in [0,1] \tag{2.27}$$

At the microscopic level, intimate mixtures occur, in which the optical interactions with grains or particles are smaller than the path length followed by the photons [62]. Typical examples of intimate mixtures are sand and mineral mixtures [90]. Several methods aim to model the intimate mixtures, such as the Hapke model [91], the Kulbelka-Munk model [92], and the Shkuratov model [93]. The main idea of these models is to describe the measured reflectances as a function of parameters intrinsic to the mixtures, such as the mass fractions, the density and size of the individual particles, and the single-scattering albedo [62]. The Hapke, a representative model for intimate spectral unmixing, can be written in Eq. (2.28).

$$x(\mu_e, \mu_i, \omega, g) = \frac{w\mu_i}{4\pi(\mu_i + \mu_e)} \left[(1 + B(g))p(g) + H(\mu_i)H(\mu_e) - 1\right] \tag{2.28}$$

where x represents the wavelength-dependent bidirectional reflectance (see Fig. 2.10). $\mu_i = cos(\theta_i)$ and $\mu_e = cos(\theta_e)$ represent the cosine of the angles of the incoming radiation and back-scattered radiation, respectively. g is the angle between the incoming and back-scattered radiations in the scattering plane, $p(g)$ is the normalized phase function of the material describing the angular scattering dependence, $B(g)$ is a correction factor for the opposition effect, w is Single Scattering Albedo (SSA) of the mixture, and H is the Chandrasekhar's isotropic scattering function [66].

2.5.3 Evaluation Metrics

Several metrics are available to quantitatively evaluate the performance of abundance estimation.

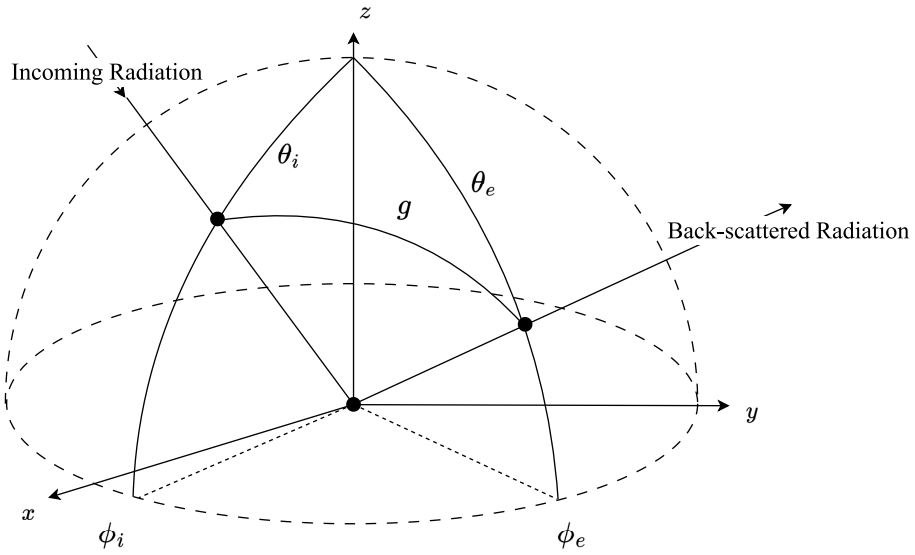


Figure 2.10: An illustration of geometries used in the bidirectional reflectance in Eq. (2.28), adapted from [66]. Unlike elsewhere in this dissertation, please note that (x, y, z) in this figure defines the spherical coordinate system whose polar angle is noted as θ and the azimuthal angle is noted as ϕ.

- **Mean Reconstruction Error (RE)**
  The mean reconstruction error RE is computed as the mean value of the Euclidean distance between input pixels $\boldsymbol{x}$ and its reconstructed pixels $\hat{\boldsymbol{x}}$.

$$RE = \frac{1}{N}\sum_{j=1}^{N}\sqrt{\sum_{b=1}^{B}(x_{j,b} - \hat{x}_{j,b})^2} \tag{2.29}$$

- **Mean Abundance Error (AE)**
  The mean abundance error (AE) is computed as the mean value of the Manhattan distance between the reference abundances and estimated abundances.

$$AE = \frac{1}{Np}\sum_{j=1}^{N}\sum_{i=1}^{p}|a_{j,i} - \hat{a}_{j,i}| \tag{2.30}$$

- **Spectral Reconstruction Error (SRE)**
  In order to evaluate the spectral characteristics of the reconstruction errors, the spectral reconstruction error SRE is computed for each spectral band b, averaged over N pixels:

$$SRE(b) = \frac{1}{N}\sum_{i=1}^{N}|x_{i,b} - \hat{x}_{i,b}| \tag{2.31}$$

2.5.4 Optimization Methods for Spectral Unmixing

Supervised spectral unmixing is an ill-posed optimization problem. Typically, abundances are solved by minimizing the pixel reconstruction error, while taking into account constraints.

Different strategies have been proposed to solve spectral unmixing robustly and accurately.

2.5.4.1 Least Squares

The least squares is a classic approach to derive unknown variables in regression problems [94, 95]. This approach consists of two categories: linear least squares and nonlinear least squares, depending on whether the mixing model contains a linear combination of unknown variables. Specifically, the linear least squares have a closed form, while the nonlinear least squares are typically solved iteratively.

Given the LMM in Eq. (2.16), its loss function can be written as:

$$\begin{aligned} \text{Loss} &= \frac{1}{2}\|\boldsymbol{E}\boldsymbol{A} - \boldsymbol{X}\|_F \\ &= \frac{1}{2}(\boldsymbol{E}\boldsymbol{A} - \boldsymbol{X})^T(\boldsymbol{E}\boldsymbol{A} - \boldsymbol{X}) \end{aligned} \tag{2.32}$$

where $\| \bullet \|_F$ denotes the Frobenius norm.

The closed-form solution of $\boldsymbol{A}$ can be derived by minimizing the loss function. If and only if $\boldsymbol{A}$ is full rank, the partial derivative of the loss function with respect to $\boldsymbol{A}$ can be written as:

$$\frac{\partial \text{Loss}}{\partial \boldsymbol{A}} = \boldsymbol{A}^T\boldsymbol{E}^T\boldsymbol{E} - \boldsymbol{E}^T\boldsymbol{X} \tag{2.33}$$

The closed-form solution can be computed by setting $\frac{\partial \text{Loss}}{\partial \boldsymbol{A}} = 0$, resulting in:

$$\hat{\boldsymbol{A}} = (\boldsymbol{E}^T\boldsymbol{E})^{-1}\boldsymbol{E}^T\boldsymbol{X} \tag{2.34}$$

Different strategies have been applied in order to account for one or more constraints in this optimization problem. A typically strategy is to iteratively enforce these constraints given the results from the least squares estimation, until the convergence condition is met.

2.5.4.2 Gradient Descent

The gradient descent solves the optimization problem in more general scenarios [96]. Given a general spectral mixing model, which can be either be linear or nonlinear, its loss function can be written as:

$$\text{Loss} = \frac{1}{2}\|f(\boldsymbol{E}, \boldsymbol{\xi}) - \boldsymbol{X}\|_F \tag{2.35}$$

where $\boldsymbol{\xi}$ are the unknown variables, i.e., abundances and hyperparameters of a spectral mixing model.

The concept of Gradient Descent Algorithm (GDA) is to search for the optimal unknown variables by minimizing the loss function in the gradient direction. GDA firstly initializes

the unknown variables and then solved them interactively using the update formula (2.36) until a convergence condition is met.

$$\boldsymbol{\xi} = \boldsymbol{\xi} - \beta \frac{\partial Loss}{\partial \boldsymbol{\xi}} \tag{2.36}$$

The parameter β represents the learning rate. The ANC and ASC can be enforced by setting the bounding range during the optimization process.

2.5.4.3 Sequential Quadratic Programming (SQP)

Sequential Quadratic Programming (SQP) is a group of methods for general constrained (nonlinear) optimization problems. The key idea of SQP is solving a sequence of quadratic programming problems. Each quadratic problem consists of a quadratic objective functions subject to constraints.

The general nonlinear optimization problem can be written as:

$$\min_{z} f(z) \qquad \text{subject to} \begin{cases} c(z) = 0 \\ d(z) \geq 0 \end{cases} \tag{2.37}$$

where $c(z)$ and $d(z)$ represent the equality and non-equality constraints on unknown variables z for the objective function f.

The above-mentioned problem can be written in the quadratic form:

$$\min_{\rho} f_k + \nabla f_k^T \rho + \frac{1}{2} \rho^T \nabla_{zz}^2 \mathcal{L}_k \rho \tag{2.38}$$

$$\text{subject to} \begin{cases} \nabla c(z_k)^T \rho + c(z_k) = 0 \\ \nabla d(z_k)^T \rho + d(z_k) \geq 0 \end{cases} \tag{2.39}$$

where $\mathcal{L}_k = f(z) - \lambda_k c(z) - \lambda_k^{'} d(z)$, λ and $\lambda^{'}$ are Lagrange multipliers, k represents for the index of the interactive process, ρ denotes the search direction at the iteration of k.

Given initial values of x_0, λ_k, and $\lambda_k^{'}$, the optimization problem computes the search direction d_k at the $k-th$ iteration and updates the variables using $[x_{k+1}, \lambda_{k+1}, \lambda_{k+1}^{'}]^T = [x_k, \lambda_k, \lambda_k^{'}]^T + d_k$, for $k = 0, 1, 2, \cdots$, until the convergence condition is met.

2.5.4.4 Alternating Direction Method of Multipliers (ADMM)

The alternating direction method of multipliers (ADMM) aims to solve optimization tasks by separating a complex problem into smaller sub-problems. It is suitable for optimization problems whose objective function is separable and subject to multiple constraints and regularization terms. ADMM can be typically converged to solutions with modest accuracy within a few tens of iterations [97]. This approach has demonstrated its success in various regression tasks, such as ridge regression, lasso regression, and regression problems with TV and low-rank constraints [83, 98, 99].

The basic form solved by the ADMM can be written as:

$$\min_{\boldsymbol{z}\in\mathbb{R}} f_1(\boldsymbol{z}) + f_2(\boldsymbol{G}\boldsymbol{z}) \tag{2.40}$$

where $\boldsymbol{G}, \boldsymbol{z} \in \mathbb{R}$.

A key step in ADMM is the splitting of variables. For the problem in Eq. (2.40), a new variable $\boldsymbol{y}$ can be created subject to $\boldsymbol{y} = \boldsymbol{G}\boldsymbol{z}$:

$$\min_{\boldsymbol{z}\in\mathbb{R},\boldsymbol{y}\in\mathbb{R}} f_1(\boldsymbol{z}) + f_2(\boldsymbol{y}), \quad \text{subject to} \quad \boldsymbol{y} = \boldsymbol{G}\boldsymbol{z} \tag{2.41}$$

Thus, the ADMM form of the original problem in Eq. (2.40) can be written into:

$$\min_{\boldsymbol{z}\in\mathbb{R},\boldsymbol{y}\in\mathbb{R}} f_1(\boldsymbol{z}) + f_2(\boldsymbol{y}) + \frac{\mu}{2}\|\boldsymbol{G}\boldsymbol{z} - \boldsymbol{y}\|_F \tag{2.42}$$

where $\mu > 0$ represents the penalty parameter for each updates and can be either a constant value or updated at each iteration [97].

At the first iteration ($t = 0$), the primal variables $\boldsymbol{x}$ and $\boldsymbol{y}$ are initialized with $\boldsymbol{x}^{(t)}$ and $\boldsymbol{y}^{(t)}$. In addition, the dual variable $\boldsymbol{u}$ is initialized with $\boldsymbol{u}^{(t)}$. Afterwards, each variable is updated alternatively by fixing the others. Specifically, the updating formulas at the t-th iteration can be written as:

$$\begin{cases} \boldsymbol{z}^{(t+1)} = \arg\min_{\boldsymbol{z}} f_1(\boldsymbol{z}) + f_2(\boldsymbol{y}^{(t)}) + \frac{\mu}{2}\|\boldsymbol{G}\boldsymbol{z} - \boldsymbol{y}^{(t)}\|_F \\ \boldsymbol{y}^{(t+1)} = \arg\min_{\boldsymbol{y}} f_1(\boldsymbol{z}^{(t)}) + f_2(\boldsymbol{y}) + \frac{\mu}{2}\|\boldsymbol{G}\boldsymbol{z}^{(t)} - \boldsymbol{y}\|_F \\ \boldsymbol{u}^{(t+1)} = \boldsymbol{u}^{(t)} + \mu(\boldsymbol{G}\boldsymbol{z}^{(t+1)} - \boldsymbol{y}^{(t+1)}) \end{cases} \tag{2.43}$$

The updating process continues until some stopping criterion is satisfied. Typical criterion can be on the primal residual $\|\boldsymbol{G}\boldsymbol{z} - \boldsymbol{y}\|_F < \varepsilon_{\text{primal}}$, on the dual residual $\mu\|\boldsymbol{y}^{(t+1)} - \boldsymbol{y}^{(t)}\|_F < \varepsilon_{\text{dual}}$, and on the number of iterations t.

3

State-of-the-Art

"If I have seen further, it is by standing on the shoulders of giants."

- Isaac Newton

Contents

3.1 Shadow Detection and Removal in Optical Image Analysis

Tackling shadow effects in optical image analysis has attracted attention since the 1990s. Previous shadow-aware methods have been proposed to detect, mitigate, and remove shadow effects. These methods typically correct specific types of shadows, such as cast shadow, self-cast shadow, and cloud shadow [100]. For a detailed explanation of shadow categories, please refer to Section 2.3. In this dissertation we discuss exclusively cast shadows.

Shadow-aware methods have been proposed for specific categories of optical imagery. In particular, the number and wavelength range of spectral bands play an essential factor in developing shadow-aware solutions. This chapter discusses state-of-the-art shadow-aware methods for three categories of optical imagery: panchromatic with one spectral band, RGB with three spectral bands, and multi- and hyperspectral with more than three spectral bands [59, 101, 102]. Please refer to Section 2.1 for a detailed description of optical imagery. For shadow detection tasks, one can apply methods proposed for imagery with fewer spectral bands to those with more spectral bands. In this case, dimensionality reduction is necessary as a preprocessing step for the imagery with more spectral bands.

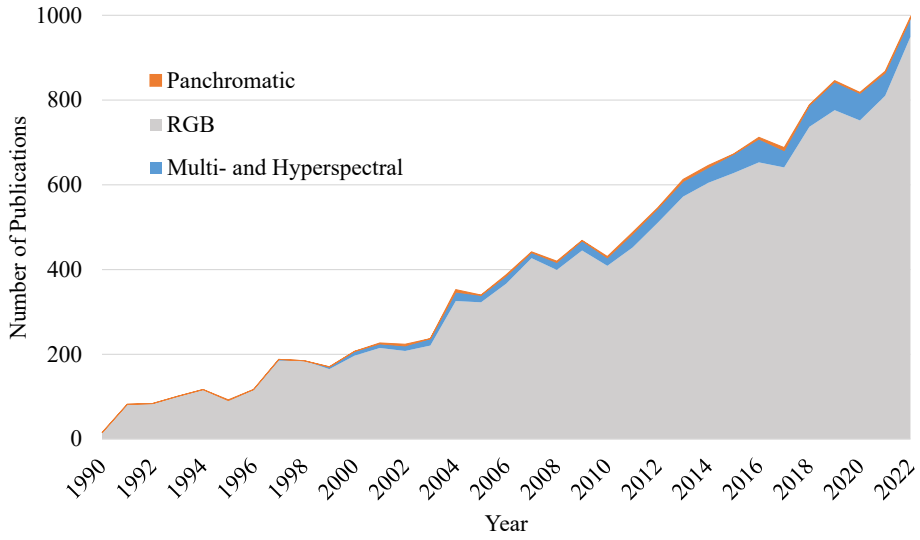


Figure 3.1: The total number of publications regarding shadow-aware methods in the Web of Science Core Collection database from 1990 to 2022, grouped into three categories of optical images: panchromatic, RGB, and multi- and hyperspectral.

Fig. 3.1 illustrates the total number of publications relevant to the shadow topic in the Web of Science Core Collection database from 1990 to 2022, grouped into three categories of optical images: panchromatic, RGB, multi- and hyperspectral. Given specified conditions, queries were executed within the searching fields of “title”, “abstract”, and “keywords”. In the panchromatic group, the searching conditions include “shadow”, “image”, and “panchromatic”, and exclude “RGB”, “multispectral”and “hyperspectral”. In the RGB group, the searching conditions include “shadow”, “image”, “RGB”, but exclude “multispectral” and “hyperspectral”. In the multi- and hyperspectral group, the searching conditions include “shadow”, together with at least one of the words among “multispectral” and “hyperspectral”.

Compared to other optical imagery, many works deal with shadow issues for RGB imagery. In contrast, limited studies on shadows exist for multi- and hyperspectral imagery.

Previous approaches typically conduct shadow detection and removal in two successive steps. Shadow detection is frequently used as a preliminary step before shadow removal, thus shadow removal can highly depend on shadow detection results. Recently, one-step solutions are becoming popular to solve both tasks simultaneously.

In the next sections, we discuss shadow detection and removal methods for panchromatic (Section 3.2), RGB (Section 3.3), and multi- and hyperspectral images (Section 3.4). In each category of optical imagery, we separately explore shadow detection and removal approaches. One-step solutions for shadow detection and removal have been categorized into the subsections of shadow removal.

3.2 Panchromatic

3.2.1 Shadow Detection

3.2.1.1 Thresholding

Since shadowed pixels typically obtain considerably lower values than sunlit areas, one category of simple but popular approaches detects shadows using thresholding. One can empirically set a threshold by visual interpretation. However, it requires human supervision and can be subjective in some cases. Therefore, many works have been proposed to identify threshold values automatically. The bimodal histogram splitting is a typical method to compute threshold values automatically [101, 103]. For a panchromatic image with shadowed regions, a valley region can be observed in the histogram, whose left side represents dark shadow pixels and the right side indicates sunlit pixels. The threshold value is computed by searching for a point in the valley region whose gradient is close to zero. Later, authors in [104] proposed a locally adaptive thresholding method, where it determines the threshold relative to the average intensity in a local window. Results demonstrate that this adaptive approach improves the robustness and accuracy of shadow detection. Nevertheless, the bimodal splitting method generates binary shadow masks while ignoring the penumbra areas. In order to tackle this problem, authors in [105] took the bimodal splitting method as the first step, followed by morphological operation and image matting technique to generate soft shadow masks [106, 107]. This operation can remove noisy pixels but also small shadowed regions. The thresholding method retains many advantages, including computational speed, explainability, and simplicity. However, selecting suitable thresholds can be challenging, especially in complicated scenes [108].

3.2.1.2 Machine Learning

Supervised and unsupervised machine learning methods have been proposed for shadow detection. Authors in [109] conducted unsupervised clustering and considered shadowed pixels as one cluster with the lowest intensity values. Later, an object-based shadow

detection algorithm has been proposed in [110] based on image segmentation and thresholding. Specifically, the image segmentation extracts homogeneous regions, followed by the progressive merging of neighboring segments according to some criteria. After that, thresholding is applied to image segments to generate shadow masks. Later, a shadow detection framework was proposed by combining image segmentation, thresholding, and classification [111]. Firstly, the framework sets strict conditions to extract underdetected shadowed regions based on image segmentation and multiple thresholding steps. Afterwards, it uses these underdetected shadowed regions as training data and produces the final shadow detection result.

3.2.1.3 Topography

This category of methods detects shadows based on auxiliary information, such as ground surface topography and the sun's position [112, 113]. These methods do not require optical imagery as input and thus have been used for shadow detection regardless of categories of optical images. DSMs is a typical data source that describes the ground surface topography. Given the sun's position and DSMs, shadow masks can be derived using geometric calculations [112, 113]. Nevertheless, these algorithms depend on the availability of DSMs that are typically expensive to acquire with high precision [114]. In addition, precise geo-coregistration between DSMs and optical imagery is challenging, affecting the quality of shadow masks. Previous work in [114] alleviates this problem. Firstly, it generates shadow masks using ground surface topography as an initial result. After that, it uses preliminary shadow masks and input images as training datasets and trains a supervised classifier via Support Vector Machine (SVM) to improve detection results.

3.2.2 Shadow Removal

To the author's knowledge, a limited amount of works has been proposed to remove shadows in panchromatic imagery. Existing works assume that pixel intensities are similar in a local neighborhood. Thus, they model the relationship of intensities between neighboring shadowed and sunlit regions, using histogram matching and linear regression. Afterwards, shadowed pixels are corrected by the derived models [110, 111].

3.3 RGB

3.3.1 Shadow Detection

3.3.1.1 Thresholding

Similar to panchromatic images, thresholding has been applied for detection shadows in RGB images. One common technique is computing the grayscale imagery by averaging RGB imagery along the spectral dimension [115, 116]. Afterward, the bimodal histogram splitting method is applied to the grayscale image to compute the optimal threshold. Besides, additional thresholding steps can be individually applied to red, green, and blue bands to improve the detection results [116]. Nevertheless, dark pixels in sunlit regions and bright

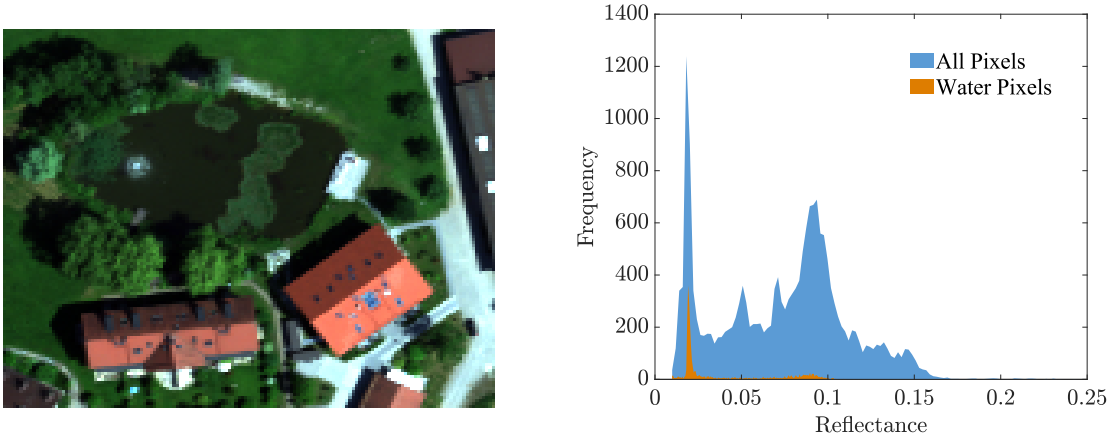


Figure 3.2: An example of histograms of averaged reflectances in a study region with a water body. (a) The study region shown in the true color composite, (b) histograms of the averaged reflectances in the full region in blue color and in the water region in orange color. Reflectance values are computed by averaging four spectral bands: red, green, blue, and NIR.

materials in shadowed areas can be incorrectly detected [117]. For example, water bodies can be often identified as shadows, and bright objects in the shaded areas may be recognized as sunlit pixels. Fig. 3.2 (a) presents a scene with a large water body area in the true color composite. Fig. 3.2 (b) compares two histograms from the averaged reflectance image in Fig. 3.2 (a). The histogram in blue is computed from all pixels, and its left peak represents shadows and water regions. The histogram in orange is computed from water pixels, and its peak overlaps with that of shadowed pixels, indicating the challenges of distinguishing water from shadows. In order to tackle this issue, authors in [116] and [118] applied water masks by computing spectral indices before shadow detection.

3.3.1.2 Color Transformations

Shadowed pixels comprise strong color characteristics, such as higher saturation at short blue-violet bands and increased hue values due to atmospheric scattering [119]. Previous methods have taken advantage of the color characteristics and detected shadows in converted feature space after color transformations. Representative color spaces can be Hue-Intensity-Saturation (HIS), Hue-Saturation-Value (HSV), and Hue-Chroma-Value (HCV) [119, 120]. Authors in [119] used the ratio of hue and intensity for shadow detection. Similarly, authors in [120, 121] detected shadows in new feature space derived by spectral computation. Specifically, a new index for shadow detection was proposed in [121], namely Normalized Saturation Value Difference Index (HSVDI) in HSV color space. Pixels with higher HSVDI values than the predefined threshold are identified as shadows. Later, a joint usage of different indices, i.e., spectral index in HIS color space and NDVI, yields more accurate results than using one index only [122]. In particular, NDVI can eliminate the confusion between dark vegetation and shadowed pixels.

3.3.1.3 Machine Learning

In unsupervised methods, authors in [102] applied K-means clustering, considering the shadow as one cluster with the lowest intensity values. The Gaussian mixture model is utilized in [123] to detect shadowed regions. In supervised methods, classification has been applied to separate shadowed from sunlit pixels assisted by training samples [124]. The performance of machine-learning-based methods depends on the selection of training samples and models.

3.3.1.4 Deep Learning

Recently, deep-learning-based approaches have attracted attention to detect shadows. To the best of the author's knowledge, the first work for shadow detection based on deep learning has been proposed in [125]. Their work automatically learns the most relevant features of the objects and along object boundaries in a supervised manner using multiple Convolutional Neural Networks (ConvNets). The learned features are then fed to a conditional random field model to generate smooth shadow contours [126] . Deep-learning-based approaches often require large amounts of training data. In practice, acquiring datasets with shadow annotation at pixel-level precision is challenging and expensive, and inaccurate training data can lead to imprecise shadow detection. To alleviate this challenge, the work in [127] allows the usage of quickly but imperfectly labeled images, followed by the automatic correction of a part of inaccurate annotations for shadow detection. This process significantly eases the data annotation and can create a larger training dataset with less human labor. In addition, this work is built upon ConvNets for shadow detection while incorporating image-level semantic information. To further ease the requirement of the training dataset, authors in [128] proposed a framework to generate synthetic shadows by integrating the shadow illumination model, 3D models, and shadow-free images. The training dataset contains an arbitrary combination of a shadow-free image, a shadow mask, and shadow attenuation parameters.

Later, the Generative Adversarial Network (GAN)-based frameworks have become popular in solving shadow problems in RGB images [129, 130]. To the best of the author's knowledge, the first GAN-based framework, namely Shadow Conditional Generative Adversarial Network (SCGAN), was proposed in [129] for shadow detection. Given an input image, the generator is trained to output shadow masks, and the discriminator has to decide if a shadow mask comes from the training data or the generator. Moreover, shadow detection is a binary classification problem with highly unbalanced classes, where shadowed pixels are much less than sunlit pixels. Hence, the framework embeds an additional sensitivity parameter to balance the shadow and sunlit samples. Recently, more deep-learning-based frameworks have been proposed to solve shadow detection and removal in one step. Please refer to Section 3.3.2.

3.3.2 Shadow Removal

3.3.2.1 Property-Based Methods

As the color and texture are typically assumed to be consistent in a local neighborhood, existing approaches aim to recover shadowed pixels using their neighboring sunlit areas at the pixel or object level. One group of methods operates in the gradient domain and derives illumination invariant images, i.e., intrinsic images, using the log-chromaticity projection [8, 131]. Later, it detects shadow edges using input and intrinsic images. Shadows are removed by nullifying the gradients on shadow edges. Authors in [132] recovered the shadowed pixels using their sunlit neighborhoods in a two-dimensional manner. Nevertheless, it is challenging to detect shadow boundaries accurately. For example, boundary regions between two objects can often be inaccurately detected [131]. Besides, the loss of textural information happens in shadowed areas.

Moreover, color transformation can help shadow removal. Assuming shadows decrease the intensity values while retaining color information, it is straightforward to separate the intensity values from chromaticity information and recover the intensity values exclusively. This partition can be done by converting RGB images to various color spaces, such as HIS, HSV, and HCV [119].

Furthermore, local or non-local matching methods have been applied for shadow removal. In [120], shadowed pixels are corrected based on the relationship of paired sunlit and shadowed regions using three models (gamma model, linear model, and histogram matching). In later works, shadow and sunlit areas are matched based on texture similarity before applying correction models [133]. Recently, approaches include other types of external data. For instance, depth data aid the non-local matching, assuming that pixels with similar chromaticity, normals, and spatial locations have similar intensities [134]. Nevertheless, it proves challenging to correlate regions in large and complex scenes automatically. In addition, it is difficult to apply a single correction model to an entire image indiscriminately, because the radiometry of the image can vary in the spatial domain [101].

3.3.2.2 Machine Learning

As in shadow detection, one can also translate shadow removal into a classification task. Earlier works select training samples separately from sunlit and shadowed areas, followed by two classification procedures [112, 135]. Afterward, they correct shadowed pixels, corresponding to their sunlit ones in the same class. In addition, morphological filters and interpolations can be applied to generate natural transitions in the penumbra areas [135].

The shadow removal methods based on classification can conveniently match non-local but similar ground objects in an image and thus may cope with larger and more complicated scenes. Nevertheless, acquiring training samples can be expensive and time-consuming, and training data may not be transferable from one scene to the other. Besides, generating the training set for shadow removal requires interpreting ground objects in shadowed regions, which may be challenging in deep shadows.

3.3.2.3 Deep Learning

Deep learning methods that learn information without hand-crafted features have become popular. Compared to applications where deep learning has demonstrated its advantage over classic machine learning approaches, such as classification and semantic segmentation, limited works based on deep learning have been proposed for shadow removal.

Given a large amount of training data, authors in [9] proposed an automatic and end-to-end deep neural network (DeshadowNet) based on multi-context ConvNets for simultaneous shadow detection and removal. By jointly training three networks, this framework can integrate high-level semantic information, mid-level appearance information, and local image details in the final prediction. Later GAN-based networks learn an image-to-image mapping function using an encoder-decoder architecture [130, 136]. To the best of the author's knowledge, the first end-to-end approach that simultaneously tackles shadow detection and removal was proposed in [130]. The proposed framework Stacked Conditional Generative Adversarial Network (STCGAN), consists of two stacked Conditional Generative Adversarial Networks (CGANs). The first generator produces a shadow mask, used as input for the second generator to produce shadow-free images. The stacked adversarial components can preserve the global scene characteristics hierarchically, resulting in a fine-grained and natural recovery of shadow-free images. Later, authors in [136] proposed an Attentive Recurrent Generative Adversarial Network (ARGAN) consisting of a generator and discriminator. The generator consists of the shadow-aware encoder and decoder. By incorporating an attention mechanism, the network chooses what it wants to observe and locate shadows of the input image. In addition, the encoder and decoder generate shadow attention maps and shadow-reduced maps, respectively. The discriminator classifies the output shadow-free image as real or fake.

Nevertheless, deep learning-based shadow removal requires a large training dataset to learn the relationship between shadow and sunlit regions [137–140]. Firstly, the training set must contain pairs of images with and without shadows that meet strict requirements, such as identical locations, camera conditions, and near-simultaneous acquisition times. Due to the changing environment light and camera exposure, a training pair may have inconsistent colors and luminosity [141]. Thus, one challenge is learning a physically plausible transformation, regardless of the semantic or illumination inconsistency [142]. Besides, ground-truth binary shadow masks need to be manually labeled. In order to meet the strict requirements, existing works investigate scenes taken from ground-based platforms. Pairs of shadow and shadow-free images have been acquired by setting and removing artificial objects. Consequently, applying this group of methods to airborne and spaceborne remote sensing imagery is challenging.

In order to relieve the requirement of training data, recently, authors in [142] proposed a self-supervised shadow removal method based on Cycle GAN [143], which learns image-to-image translations between two image domains from both paired and unpaired images. Unpaired images indicate a set of shadow and shadow-free images that are not associated with each other. Specifically, Cycle GAN contains two pairs of GANs with cycle consistency constraints: one aims to translate the shadow image to the shadow-free image, and the other takes the shadow-free image and translates it back to the shadow image.

3.4 Multispectral and Hyperspectral

3.4.1 Shadow Detection

3.4.1.1 Thresholding

More spectral information in multi- and hyperspectral imagery improves the performance of the thresholding methods in shadow detection. Specifically, NIR bands have been chosen over RGB, as they contain lower reflectance values at shadowed pixels but higher values at dark objects, thus improving separability between dark objects and shadowed regions [59, 117]. Moreover, some works jointly consider multiple spectral bands to distinguish shadows from water body. For example, previous works use NIR and green bands, because water has lower reflectance at NIR but higher reflectance in the green band [59, 144]. In addition, reflectances at the wavelengths of 0.85 µm and 1.6 µm have been proven helpful to separate water from shadows. Besides, five different indices are derived through band computation for shadow detection [116]. Nevertheless, it is challenging to apply thresholding methods to remote sensing images that typically contain complicated ground scenarios. In particular, the histogram of input images can be multi-modal; thus, the bimodal splitting can not be applied.

3.4.1.2 Machine Learning

In the unsupervised scenario, one can convert the hyperspectral data from a Cartesian space to a hypespherical coordinate system, where one pixel consists of a magnitude value and spectral angle vector [102]. Then, the K-means clustering is applied to the magnitude values [145]. The cluster with the lowest average values is determined as shadows. Furthermore, previous works conduct shadow detection using supervised classification. Authors in [135] proposed a binary classification framework followed by post-processing. Multispectral imagery and their extracted features through wavelet transform are used as input of the SVM classifier [146]. After that, morphological operators are applied for noise reduction and the generation of penumbra regions, creating a natural transition on shadow boundaries. Inspired by [135], an object-based supervised classification framework was proposed in [147]. The multispectral imagery is segmented into appropriate sizes before supervised classification, taking advantage of spectral and contextual information.

3.4.1.3 Physics-Based Methods

This group of methods considers the formation of shadows based on physical assumptions. Typically, one assumes that shadowed regions do not receive solar illuminations due to the occlusion of ground objects. Thus, the observed reflectance of shadowed pixels is assumed to be zero or a very small value. Earlier works apply a matched filter to detect pixels with (nearly) zero reflectances [148]. Later, a group of methods detects shadows using spectral unmixing [148–150]. In brief, spectral unmixing assumes that each pixel is composed of a mixture of pure material spectra, i.e., endmembers, associated with their corresponding proportions, i.e., abundances. Given a pixel and endmembers, spectral unmixing estimates

the abundances. Please refer to Section 2.5 for a detailed explanation of spectral unmixing. In order to handle shadow effects, one strategy is adding a virtual endmember whose reflectance values are constantly zeros to represent the shadow spectrum [149]. The abundance of the virtual endmember corresponds to the shadow fraction in that pixel. As an equivalent process, authors in [151] introduced a shadow-related parameter, associated with pixelwise fractional values of cast shadows, into spectral mixing models.

Authors in [150] detected shadows by estimating the order of multiple optical reflections. Specifically, they assume that incoming light undergoes an infinite number of optical interactions with objects in shadowed pixels and thus will not reflect light to the sensor. A multilinear spectral mixing model was proposed, where a parameter P indicates the order of multiple reflections for each pixel. A larger P value implies that a pixel is more likely to be shaded.

Following the radiative transfer theory, direct and diffuse solar radiations are the main illumination sources for outdoor scenes, where direct radiation is significantly larger than diffuse radiation [55]. Fully sunlit pixels receive both of them, while fully shadowed pixels, i.e., umbra of shadowed areas, receive solely diffuse radiation due to the occlusion by ground objects. Partly sunlit or shadowed pixels, i.e., penumbra of shadowed areas, receive a part of direct and diffuse solar radiations. Please refer to Section 2.3 for more details about shadow formation. Previous works detect shadows by computing the proportion of diffuse radiation with respect to total radiation in a pixel [152]. In other words, diffuse radiation accounts for 100 % of the total radiation for a fully shadowed pixel, while it consists of a considerably lower percentage for a fully sunlit pixel.

3.4.2 Shadow Removal

3.4.2.1 Property-Based Methods

This group of methods aims to match shadowed pixels with corresponding sunlit pixels using spectral similarity measures, among which Spectral Angle Distance (SAD) has been commonly used due to its shadow insensitivity [153]. Authors in [154] extracted endmembers separately in sunlit and shadowed regions using VCA, followed by abundance estimation in the two regions through NNLS. Then, the pairwise similarity of endmembers extracted from the two regions is computed using SAD. Finally, each shadowed pixel is reconstructed using abundances and their paired endmembers in sunlit regions. The shadow restoration result naturally embeds material composites at the sub-pixel level and thus allows for multiple ground objects in one pixel. Similar work in [147] matches endmembers extracted in shadowed regions with their corresponding endmembers extracted in sunlit areas using a different similarity measure. For the same material, the construction of the spectral scatter plot in shadowed and sunlit regions shares the same shape in a two-dimensional spectral space but differs in size. Later, authors in [155] introduced a similar framework as in [147] using spectral unmixing techniques. Additionally, they investigated the impacts of various nonlinear unmixing approaches. The investigation indicates that the choice of spectral mixing models is an important factor for the performance of shadow removal. Moreover, shadow removal based on spectral matching can be combined with classification. The work in [156]

firstly classifies sunlit regions using SVM and computes the spectral center for each class [146]. Later, it removes shadows by matching shadowed pixels with one of those spectral centers. Nevertheless, these methods assume shadow effects to be wavelength-independent. In practice, shadow effects not only alter the spectra magnitude but also apply a wavelength-dependent impact on spectra [55, 103]. This spectral distortion caused by shadows can lead to unsuccessful spectral matching, especially in complicated scenes. In addition, spectral matching can also confuse different materials with similar spectra. Besides, this group of methods requires accurate shadow masks as prior knowledge.

3.4.2.2 Machine Learning & Deep Learning

Authors in [157] estimated shadow effects and then generated a series of multi-exposure images based on the intrinsic decomposition model. After that, they designed a two-stage image fusion process to generate shadow-removed images. In the first stage, the obtained over-exposure images are utilized to recover spectral information in shadowed regions. In the second stage, an interpolation step is employed to generate a natural transition at shadow boundaries.

Furthermore, existing works train models to learn the feature relationships between sunlit and shadowed areas. Authors in [158] generated shadow-invariant classification maps based on Recurrent Neural Network (RNN). This network is trained on a shadow-aware semantic annotation database. Specifically, 103 image patches with six land-cover classes are labeled in sunlit and shadowed regions. Nevertheless, acquiring a large amount of training data for remote-sensing images is challenging. Recently, a shadow removal framework based on Cycle GAN has been proposed to alleviate the scarcity of training data [143, 159]. Specifically, Cycle GAN can learn from both paired and unpaired images for shadow removal. The framework contains two generators and two discriminators. One set of generators and discriminators maps shadowed pixels to sunlit pixels, while the other set maps sunlit pixels to shadowed pixels. Original sunlit pixels are preserved in restored images without prior information via a consistent loss.

3.4.2.3 Physics-Based Spectral Unmixing Methods

Physics-based spectral unmixing methods have demonstrated their advantages in shadow removal. Compared to machine learning-based approaches, unmixing-based methods often require no or limited prior knowledge of shadows. In addition, spectral unmixing can analyze images at the sub-pixel level, making it particularly effective in handling bordering pixels where material mixtures and partial shadows often occur. This section starts with general spectral unmixing methods, followed by those specifically proposed for shadow removal.

Classic linear and nonlinear spectral unmixing methods, including LMM [76], Nascimento [86], Fan [87], PPNM [88], GBM [89], and Hapke [91], have been discussed in the previous chapter. Interested readers are referred to Section 2.5. Apart from linear and bilinear models, some works have attempted to consider all orders of interactions using harmonic functions [160] and polynomial functions [161]. Recently, a Multilinear Mixing (MLM) model [150]

based on the stochastic process and physical assumptions has been proposed. This model traces the path of a single light ray and extends the optical interactions to the infinite order.

Several spectral mixing models account for shadow effects. The main idea is that if the shadow formation can be modeled with certain parameters, shadow effects can be corrected by corresponding parameters during pixel reconstruction. Some works conduct spectral unmixing by including an additional "shade" endmember, whose reflectance values at all wavelengths are zeros [149, 162, 163]. Their unmixing results at each pixel contain the abundances of regular endmembers and the abundance of the shade endmember. Then, shadows are removed through pixel reconstruction, by re-scaling the regular abundances with the abundance of the shade endmember. This method of shadow removal has been also applied in [163] as a preprocessing step for the classification of tree species. In practice, the shadowed regions receive scattered solar illumination, resulting in low but not zero reflectance. Hence, authors in [162] conducted the endmember extraction and selected the endmember with the lowest average spectra as the shade endmember [162]. Furthermore, the shade endmember strategy is equivalent to modeling a shadowed pixel by scaling its measured spectrum exposed to sunlight using a shadow-related parameter, such as in the Shadow Multilinear Mixing (SMLM) model [151]. The shadow-removed pixel reconstruction is conducted using the spectral mixing model by setting the shadow-related parameter to zero [151].

Recent works argue that shadows can introduce wavelength-dependent distortions. The spectral distortion requires a more precise spectral mixing model to accurately compute the ground pixels in the presence of shadow effects. In an early work, shadows have been treated as a nonlinear effect in a specific scenario, where trees block the direct solar illumination on a region covered by grass [86]. This work models the shadowed spectrum using the element-wise product of the tree and grass spectra. The shadowed spectrum is used as an additional endmember in the unmixing process. Later, a group of methods includes wavelength-dependent spectral distortion in shadowed regions, allowing illumination conditions to vary over a scene. Authors in [164] developed a spectral unmixing method using images with radiance values and DSMs. This work models the optical imaging chain considering direct and diffuse solar radiations, shadow effects, atmospheric impact, and topography [55]. After solving unknown variables in the model using optimization techniques, they remove shadows by correcting shadow-related parameters during pixel reconstruction. Shadow-compensated Bilinear Mixing Model (SCBMM) was proposed in [165] that solves the abundances and diffuse solar radiation simultaneously based on Global Particle Swarm Optimization (GPSO) [166, 167]. Nevertheless, as the complexity of spectral mixing models increases to tackle shadow effects, spectral unmixing as an inverse process can become an ill-posed problem, which can not be easily solved accurately and robustly. For example, results can become noisy, and confusion between similar materials may appear in shadow-removed images.

3.4.2.4 Incorporating Auxiliary Information in Spectral Unmixing

Incorporating auxiliary information has shown notable potential to improve the performance of spectral unmixing for shadow removal [168, 169]. Specifically, ground topography is closely related to shadow formation and provides essential information when modeling radiations from illumination sources to the ground surface. Among different categories of topographic data, DSMs captures the height of natural and artificial objects on the ground surface and thus has been widely used in spectral unmixing. DSMs can be acquired using different techniques. One approach is using multi-view stereo images acquired by spaceborne sensors, such as WorldView-2, and airborne sensors, such as in [170]. The other most common approach is using airborne LiDAR sensors [171]. Generating DSMs from stereo images is less expensive and can cover larger areas. However, it can suffer from errors in dense matching and noises due to complicated geometric structures, occlusions, and shadows [172, 173]. On the other hand, LiDAR data are insensitive to weather and illumination conditions but are more expensive and challenging to acquire in large regions [172, 173].

Authors in [16] proposed a physics-based spectral mixing method to model at-sensor radiance. Specifically, the model uses the directions of incident sunlight and ground topography to compute global solar irradiance and shadow situations in the study regions. In addition, it considers distances, elevations, and relative angles of different ground objects to model local environment illumination. Recently, some works embed topographic information in a simplified manner for shadow removal. Authors in [174] derived a shadow-insensitive representation using hyperspectral imagery and DSMs. In particular, DSMs aid in modeling ground surface illumination. For example, this work computes direct solar radiation using the angle between the direct solar beam and the surface normal. In addition, it computes the diffuse solar radiation using the sky view factor that indicates the percent of visible sky hemisphere at a specific location. Despite not learning shadow-free radiance or reflectance, this shadow-insensitive representation considerably improves segmentation results.

Furthermore, some works incorporate spatial information with spectral unmixing, improving the accuracy and robustness of unmixing performance. Spatial-spectral unmixing approaches regard the spatial dependence in local and non-local neighborhoods: in local neighborhoods, pixels are assumed to be strongly correlated to their neighbors [169, 175], while in non-local neighborhoods, similar patches in a larger region are assumed to share similar texture [176, 177].

Earlier works [178] conduct spectral unmixing by iteratively selecting and removing endmembers. The optimization criteria are the minimization of the Root Mean Square (RMS) and spatial structure of the RMS image, improving the local homogeneity of abundance maps. Later, a hierarchical Bayesian model incorporates spatial dependencies in local neighborhoods using Markov random fields for spectral unmixing [179]. The input image is assumed to be partitioned into classes where the statistical properties of the abundances are homogeneous. The unmixing problem is then solved by Markov Chain Monte Carlo (MCMC) [180]. Recently, convolutional operations have been implemented to account for spatial-contextual information in spectral unmixing [78, 181]. This approach sequentially decomposes the hyperspectral image from local attention to global aggregation and preserves the approximately continuous spectral and spatial proprieties.

Moreover, a group of methods applies various spatial regularizations to spectral mixing models. A linear spectral unmixing approach has been proposed in [182] by embedding spatial autocorrelation of abundances in local neighborhoods. A spatial-spectral coherence is implemented in spatial regularization terms [183]. Later, the spatial regularizers in local neighborhoods has been demonstrated via the L2-norm [184] and TV [82]. TV regularization has attracted attention, as it promotes piecewise smooth abundance maps and better preserves edges [82, 84]. In particular, the SUnSAL-TV incorporates the TV and sparse regularizations into the LMM method. This optimization problem is challenging, as it contains nonsmooth terms and high dimensional data. Hence, SUnSAL-TV introduces new variables per regularizer and convertes the initial problem into a sequence of much simpler problems using the ADMM method [82, 97]. This work conducts a detailed experiment on simulated and real datasets at different noise levels, indicating the superiority of TV in estimating noise-robust abundances.

Despite that TV-based regularization better preserves edges in abundance maps, it may also introduce shrinkage effects, particularly on the boundary pixels across similar yet different materials. Some studies consider edge properties as weighting factors for spatial regularization in spectral unmixing [84, 185, 186]. The weights describe feature similarities between a target pixel and its neighboring pixel. Given a target pixel, its neighboring pixels with higher feature similarities will acquire larger weights, thus impacting the target pixel more. Common similarity measures include spectral distance, principal components, and abundance distance [185, 186]. In addition, combined weights derived by two or more similarity measures may further improve the unmixing performance. Authors in [185] compute two weights associated with spectral and spatial similarities, respectively. These similarity measures assist in segmenting the homogeneous and transition areas. The segmented image promotes endmember extraction, assuming that pure pixels are more likely to appear in the spatially homogeneous areas and less likely to be found in transition areas [187]. In addition to endmember extraction, segmented results promote pixels in a homogeneous region to have similar abundances based on a weighted nonnegative matrix factorization algorithm. Furthermore, similarity measures of elevations have been derived from DSM data in [84]. Compared to spectral information, elevation data can significantly improve abundance estimation thanks to its illumination-insensitive property [84]. Generally, the more accurately the weights describe the ground features, the better the abundance estimation performance. On the other hand, inaccurate weights may also decrease performance. For instance, height and spectral similarity measures can be inaccurate in the presence of shadows.

3.5 Contribution of this Dissertation

The objective of this dissertation, as mentioned in Chapter 1, is to develop advanced and robust algorithms for resolving shadow issues in hyperspectral imagery.

This chapter has provided a comprehensive literature review for shadow detection and removal in panchromatic, RGB, and multi- and hyperspectral imagery. Shadow-aware solutions have been mostly proposed for RGB data, and these methods, especially for shadow removal, can not be applied to hyperspectral images. Recently, shadow-aware

solutions specifically aiming at hyperspectral images have been proposed. However, notable research gaps remain to achieve better results. Firstly, most methods detect and remove shadows in two successive steps, and thus, the results of shadow detection directly impact the performance of shadow removal. This challenge demands a one-step solution to solve detection and removal simultaneously. Secondly, supervised learning-based methods typically require large training datasets. However, this dataset is expensive and difficult to acquire for shadow detection and removal tasks, especially for airborne and spaceborne images. Hence, a method requiring no or limited training samples is essential.

Recently, spectral unmixing-based approaches have demonstrated their advantages in partly solving the above-mentioned problems. Unfortunately, several research gaps still exist and demand further investigation. Firstly, some previous methods have yet to consider the physical characteristics of shadows, which can lead to inaccurate results. Some other methods suggest comprehensive physics-based models to regard shadow effects, making their models difficult to solve. Consequently, a physics-based model leveraging simplicity and accuracy is desired. Secondly, some methods have incorporated auxiliary data into spectral unmixing. However, none of them focused on shadow problems. Hence, further studies are required to connect shadow effects with data fusion in spectral unmixing.

This dissertation aims to solve the mentioned research gaps. Three solutions have been proposed in three peer-reviewed articles by the author. These works are presented in detail in Chapter 5 and are summarized as:

- **Contribution 1:** To address the **Objective 1** in Chapter 1, a one-step approach for shadow detection and restoration is presented in Section 5.1. This framework simultaneously models the material mixtures and shadow effects at the sub-pixel level based on physical assumptions and spectral unmixing. Results contain shadow-removed images and soft shadow detection masks that follow the natural transition on the shadow boundaries. In addition, this framework generates material abundance maps as a valuable by-product, indicating shadow-insensitive material compositions.
- **Contribution 2:** Modeling optical interactions and shadow effects is typically complicated in outdoor scenes. Linking with the **Objective 2** in Chapter 1, Section 5.2, built upon the first contribution, proposes an improved physics-based spectral mixing model that accounts for more general ground scenarios in the presence of shadow, nonlinear optical interactions, and ground mixtures. The improved model yields more physically interpretable parameters with higher accuracy, and better shadow-removed images.
- **Contribution 3:** As a typically inverse and ill-posed problem, spectral unmixing methods often suffer from inaccurate parameter estimation. To address the **Objective 3** in Chapter 1, Section 5.3 proposes a novel shadow-aware spectral unmixing method by incorporating DSMs and spatial information.

Finally, a general discussion is given to further investigate shadow effects using a field study and to demonstrate state-of-the-art approaches in larger regions.

4

Datasets

"You can have data without information, but you cannot have information without data."

- Daniel Keys Moran

Contents

4.1 HySpex/3K Dataset

This dataset was acquired by two imaging spectrometers, i.e., the HySpex cameras, and an 3K camera system, which were mounted simultaneously on the Deutsches Zentrum für Luft- und Raumfahrt (DLR) research aircraft D-CFFU, a Dornier 228-212 modified for Earth observation research [188]. HySpex and 3K cameras collect hyperspectral and RGB imagery, respectively. The data were acquired at the same time over Oberpfaffenhofen, Bavaria, Germany, between 8:40 and 9:50 a.m. on June 4th, 2018, according to Central European Summer Time (CEST) (see Fig. 4.1 (a)). A total of 12 flight lines were recorded at two different altitudes, namely 1 km and 1.9 km above ground level [188].

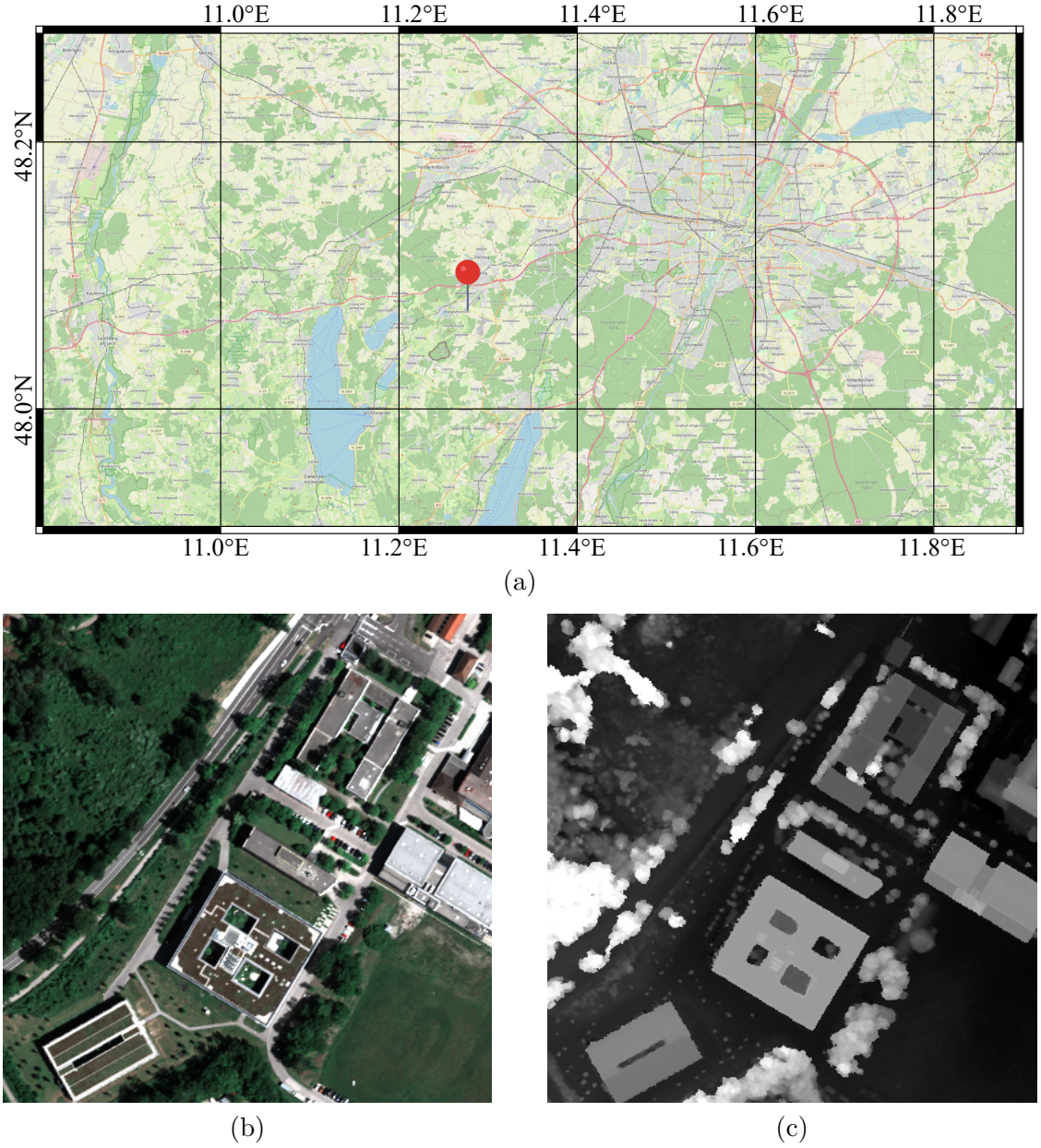


Figure 4.1: An example of the HySpex/3K dataset: (a) study location near the city of Munich; (b) hyperspectral image in the true color composite; (c) DSM geometrically co-registered with the image in (b).

4.1.1 Hyperspectral Imagery

The HySpex system consists of two pushbroom imaging spectrometers manufactured by the Norwegian company Norsk Elektro Optikk [189]. Table 4.1 presents selected technical specifications. A visible and NIR camera, i.e., VNIR-1600, covers the wavelength range of 416 nm and 992 nm with a total of 160 spectral bands. A SWIR camera, i.e., SWIR-320m-e, covers the wavelength range of 968 nm and 2498 nm with a total of 256 spectral bands. The spatial resolution depends on flight altitude and is thus configurable between 0.3 m and 3 m. For example, at the flight altitude of 1 km above ground level, the VNIR-1600 produces a spatial resolution between 0.5 m and 1 m along track and between 0.3 m and 0.5 m across-track. The SWIR-320m-e has a spatial resolution between 1.1 m and 1.7 m at 1 km above ground level.

Table 4.1: Selected technical specifications of two HySpex sensors

	VNIR-600	SWIR-320m-e
Imaging	Pushbroom	Pushbroom
Detector	Si CCD Array	MCT Array
Number of spectral bands	160	256
Spectral range [nm]	416 - 992	968 - 2498
Spectral sampling distance [nm]	3.6	6
IFOV across- /along- track [mrad]	0.18 / 0.36*	0.75 / 0.75*
Mass [kg]	4.6	7.5
Dimensions [mm]	335 × 84 × 138	390 × 140 × 152
Digitization [bit]	12	14

* can be doubled with field expander.
Further characteristics of the sensors can be found on the producer's webpage (Norsk Elektro Optikk).

The hyperspectral images have been preprocessed using the software Catena, developed at DLR [190]. Processing steps include systematic correction [191], co-registration of Visible and Near-infrared (VNIR) and SWIR data via Binary Robust Invariant Scalable Keypoints (BRISK) matching [192], orthorectification [193] , and atmospheric correction using the software Atmospheric and Topographic Correction (ATCOR) [63]. Interested readers are referred to [188, 190, 194] for a detailed description of the flight campaign, DLR HySpex system, and preprocessing.

The hyperspectral imagery used in this dissertation was acquired by the visible and NIR camera, i.e., VNIR-1600 at the flight altitude of 1.6 km above ground and with a spatial resolution of 0.7 m, and have been processed to surface reflectance. After removing the water vapor bands, a total of 101 spectral bands were kept for further processing. Fig. 4.1 (b) shows a subset of hyperspectral data in the true color composite.

4.1.2 RGB Imagery and DSMs

RGB images were acquired by the DLR 3K system, which consists of three 35 mm Canon EOS cameras [188]. At flight altitude of 1 km above ground level, the ground sampling

distance of the 3K camera is around 13 cm. The camera was positioned to look sideways, i.e., left and right, and nadir. Interested readers are referred to [195] for a detailed description of the 3K camera system.

Multi-view stereo images acquired with the 3K camera system were employed to generate DSMs [170], whose values represent surface elevation above the ellipsoid. In addition, elevation values were normalized between 0 and 1. The DSMs have been geometrically co-registered and re-sampled so that DSMs and hyperspectral images share the same geo-coordinates and ground sampling distance (i.e., 0.7 m).

4.2 Extended DLR HyperSpectral Unmixing (HySU) Dataset

4.2.1 DLR HySU

The DLR HySU, as a benchmark dataset for spectral unmixing, contains a hyperspectral image, endmember library, and ground truth of abundances [188]. The hyperspectral image is a subset of the hyperspectral data acquired by the VNIR-1600 spectrometer in Section 4.1. Please refer to Section 4.1 for the detailed description of sensor characteristics, acquisition conditions, and preprocessing steps. Specifically, the hyperspectral image in the DLR HySU dataset features a ground sampling distance of 0.7 m. Besides, it comprises 135 spectral bands ranging from 417 nm to 903 nm, while spectral bands beyond 903 nm were discarded due to low Signal-to-noise Ratio (SNR). Fig. 4.2 (a) presents an overview of the study region in the true color composite.

The study area on the ground surface contains artificial and square targets with different sizes in the range of 0.25 m and 3 m. Each size contains five targets: bitumen, red-painted metal sheets, blue fabric, red fabric, green fabric, and a background material (grass) [188]. This dissertation considers the square targets with the side lengths of 3 m, and the target area can be translated in number of pixels in the image, and is going to be used as the ground truth for the abundances (Table 4.2). The reflectance of six materials has been

Table 4.2: Ground truth of abundances in the HySU dataset

Endmember	Target area [pixel]
bitumen	18.429
red-painted metal sheets	18.061
blue fabric	18.245
red fabric	18.798
green fabric	18.521
Total	92.054

extracted from the image, constructing the endmember library for this dataset (Fig. 4.2 (b)). Reference abundance maps were generated using the endmember library and FCLS [196]. In addition, the ground target sizes in Table 4.2 was used to evaluate abundance estimation errors [188]. Results demonstrated that reference abundance maps have an average error of 2.3 % [188], which is low enough to justify their use as ground truth.

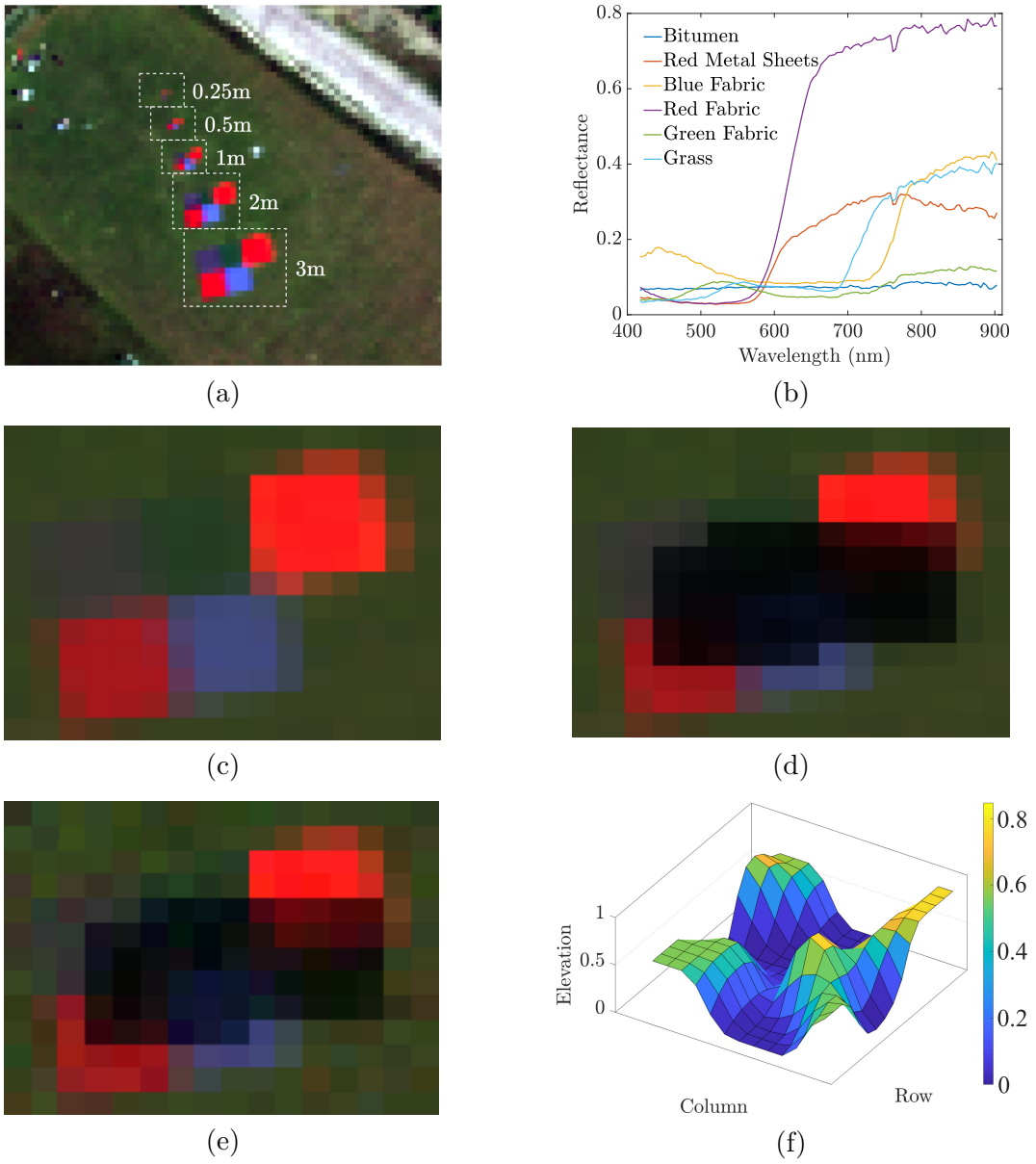


Figure 4.2: Extended DLR HySU dataset. (a) the study region with different sizes of targets, adapted from [188], (b) reflectance of six materials: bitumen, red metal sheets, blue fabric, red fabric, green fabric, and grass, (c) hyperspectral image of 3 m targets in true color composite, (d) hyperspectral image with simulated shadows, (e) hyperspectral image with simulated shadows and additional random noise ($SNR = 30$), (f) an example of simulated DSM, normalized between 0 and 1.

4.2.2 Simulated Shadows and DSMs

In order to validate shadow-aware spectral unmixing methods on the DLR HySU dataset, we simulate shadows partially covering the shadow-free image in Fig. 4.2 (c). Firstly, a shadowed region is manually drawn in the center of the shadow-free image to shade a part of all targets, resulting in a binary shadow mask. Then, a 3-by-3 Gaussian filter is applied to the shadow mask and generates a soft shadow mask Q. Given a pixel $\boldsymbol{y}_j$ in the shadow-free image, we simulate the shadow-included image in Fig. 4.2 (d) using Eq. (4.1).

$$\boldsymbol{x}_j = (1 - Q_j)\boldsymbol{y}_j + Q_j T(s_{0_{\text{diff}}})\boldsymbol{y}_j \tag{4.1}$$

where

$$T(s_{0_{\text{diff}}}) = \frac{F_j \cdot (k_1 \boldsymbol{\lambda}^{-k_2} + k_3)}{1 + F_j \cdot (k_1 \boldsymbol{\lambda}^{-k_2} + k_3)} \tag{4.2}$$

with $F_j = 1$, and k_1, k_2, k_3 are set as in the individual contribution (Sections 5.2 and 5.3) in Chapter 5.

It is worth noticing that Eq. (4.1) implicitly embeds nonlinear effects contained in pixel $\boldsymbol{y}_j$, so we do not add additional nonlinear effects in this image. However, since the study area is located in the middle of a field with flat terrain, and we quantitatively validate solely the abundances of the ground targets made of synthetic materials, the nonlinear effect plays a minor role.

In order to validate the noise robustness of the proposed methods, additional random noise ($SNR = 30$) is simulated and added to the shadow-included image (Fig. 4.2 (e)). Moreover, evaluating the third contribution of this dissertation (Section 5.3) requires DSMs together with hyperspectral images. Since the HySU dataset does not contain a DSM, a synthetic DSM is simulated through a piecewise homogeneous distribution using the Potts model [179]. Besides, an additional Gaussian filtering with a size of 3 has been applied to the simulated DSMs for gradual elevation changes at the borders between objects (see Fig. 4.2 (f)).

4.3 Cubert Measurements

This dataset was acquired with the snapshot sensor system Cubert UHD-185 Firefly manufactured by the company Cubert. This imaging system consists of two sensors, i.e., one spectrometer and one panchromatic sensor, and simultaneously records a full-frame hyperspectral image and a panchromatic image. The spectrometer features 138 spectral bands in the range of 450 nm and 998 nm with a spectral resolution of 4 nm. The acquired hyperspectral image consists of 50-by-50 pixels. The panchromatic sensor provides images with higher spatial resolution. Each panchromatic image consists of 1000-by-1000 pixels. In addition to hyperspectral and panchromatic images, the processing software provided by the manufacturer produces pan-sharpened hyperspectral images, whose spatial resolution is equal to that of the panchromatic images. In this dissertation, the Cubert sensor was

mounted on the ground-based platform to measure material spectra, while it is possible to mount the camera system on Unmanned Aerial Vehicles (UAVs) [197].

The experiment was conducted between 11 a.m. and 12:30 a.m. (CEST) on August 16th, 2022. The weather was characterized by clear skies with very few intermittent cumulus clouds. A dark current image was taken at the beginning of the experiment. A white Lambertian reflectance panel was measured approximately every 20 minutes for calibration. The integration times were set according to the illumination conditions of ground targets to avoid over- and under-exposure. Fig. 4.3 presents an example of the measurements.

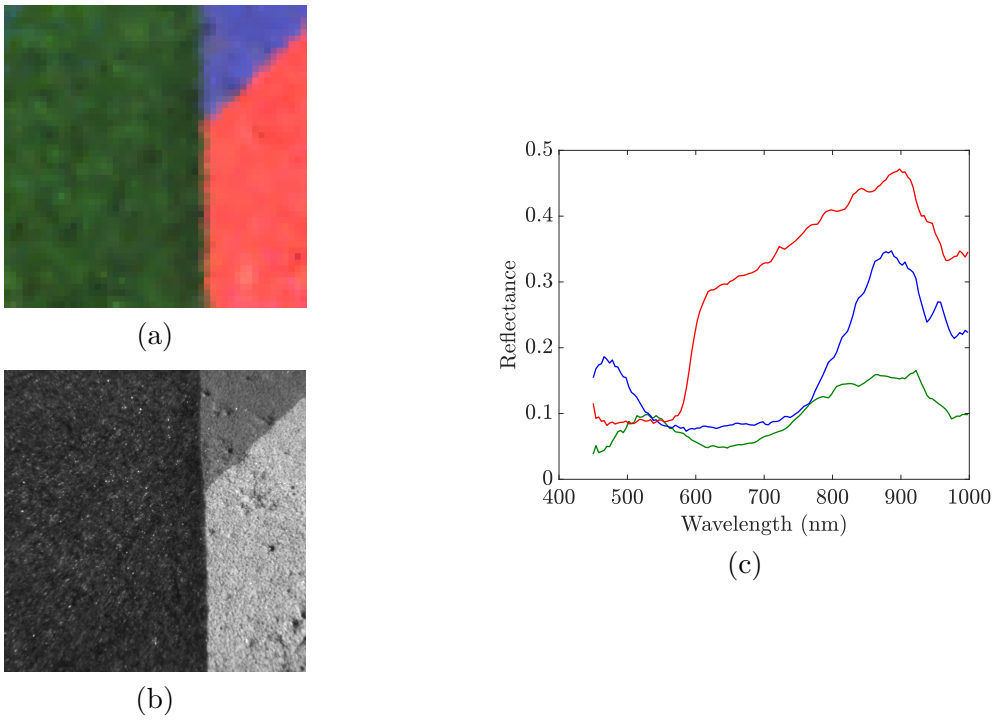


Figure 4.3: An example of the Cubert measurement, consisting green, red, and blue targets. (a) hyperspectral image in true color composite, (b) panchromatic image of (a), (c) reflectance of three targets in (a), shown in their corresponding colors.

5

Summary of the Contributions

"To raise new questions, new possibilities, to regard old problems from a new angle, requires creative imagination and marks real advance in science."

- Albert Einstein

Contents

Linking to the research *Questions* and **Objectives** in Chapter 1, this chapter introduces three novel methods associated with aforementioned **three Contributions** in Section 3.5, to resolve shadow issues in hyperspectral image analysis. An overview of the datasets utilized in this chapter is provided in Chapter 4, such as sensors, preprocessing steps, and acquisition conditions. For a more in-depth understanding of the datasets employed in each **Contribution**, please refer to the corresponding sections, namely Sections 5.1, 5.2, 5.3. In summary:

- **Contribution 1:**
  This work addresses **Objective 1**, thus answering *Questions 1.*
  [198] Guichen Zhang, Daniele Cerra, and Rupert Müller. Shadow detection and restoration for hyperspectral images based on nonlinear spectral unmixing. *Remote Sensing*, 12(23):3985, 2020

- **Contribution 2:**
  This work addresses **Objective 2**, thus answering *Questions 1.*
  [199] Guichen Zhang, Paul Scheunders, Daniele Cerra, and Rupert Müller. Shadow-aware nonlinear spectral unmixing for hyperspectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15:5514–5533, 2022

- **Contribution 3:**
  This work addresses **Objective 3**, thus answering *Questions 2 and 3.*
  [200] Guichen Zhang, Paul Scheunders, and Daniele Cerra. Shadow-aware nonlinear spectral unmixing with spatial regularization. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–16, 2023

- **General Discussion:** This section gives additional discussion, which has not been published elsewhere, offering valuable additional insights to supplement the contributions above.

5.1 Shadow Detection and Restoration for Hyperspectral Images

Shadows are frequently observed in high-resolution images, raising challenges in image interpretation, such as classification and object detection. This work presents a novel framework for shadow detection and restoration of atmospherically corrected hyperspectral images based on nonlinear spectral unmixing. The mixing model is applied pixelwise as a nonlinear combination of endmembers related to both pure sunlit and shadowed spectra, where the former are manually selected from scenes and the latter are derived from sunlit spectra following physical assumptions. Shadowed pixels are restored by simulating their exposure to sunlight using sunlit endmembers weighted by abundance values. The proposed framework is demonstrated on real airborne hyperspectral images. A comprehensive assessment of is carried out both visually and quantitatively. With respect to binary shadow masks, our framework can produce soft shadow detection results, keeping the natural transition of illumination conditions on shadow boundaries. Results show that the framework can effectively detect shadows and restore information in shadowed regions.

5.1.1 Introduction

In images with high spatial resolution, shadows are frequently visible [108]. When an object occludes the direct solar illumination outdoors, self-cast shadow occurs on the part of the object with no direct solar illumination. Cast shadow, which this work considers, is projected instead on nearby objects, and consist of umbra and penumbra [59]. Interested readers are referred to Section 2.3 for more details about shadow formation.

Previous works studying shadow detection or shadow removal from optical images use optical earth observation data, including RGB, multispectral, and hyperspectral images [59, 101, 102]. In particular, imaging spectrometer data, also referred to as hyperspectral[i], provide spectral measurements with near-continuous acquisition wavelengths. These data convey rich spectral information related to the physical properties of ground materials and their chemical composition, compared with RGB and multispectral images, and are extremely valuable for different remote sensing applications [201]. Nevertheless, most shadow detection and removal methods are proposed for RGB data. These methods typically cannot be adapted to hyperspectral images.

This work aims to mitigate the following open problems regarding shadow detection and removal in hyperspectral images.

- Shadow detection is commonly required before shadow removal, and strongly influences shadow removal results. Therefore, it is important to acquire a reliable shadow mask before the shadow removal process [108, 156].

[i]We are aware that the terms “imaging spectroscopy” and “imaging spectrometer data” are more exact than “hyperspectral imaging” and “hyperspectral data” respectively, and therefore should be preferred. Nevertheless, in this work we also use the term “hyperspectral” for sake of briefness [34].

- Most shadow detection methods generate binary shadow detection results, where one pixel is either sunlit or shadowed [9, 118]. In fact, pixels located on shadow boundaries are neither completely sunlit nor shadowed [202, 203]. A natural transition on shadow boundaries should be considered.
- Despite numerous methods proposed for shadow detection or removal, there have been only few attempts at recovering full spectra from hyperspectral data [147, 149, 153, 154].
- Some methods require a large number of training samples as input, which are difficult and time-consuming to acquire [9, 127, 131].
- Shadow restoration may introduce spectral distortion in sunlit pixels.
- Precise 3D data may be effective at assisting shadow detection and removal, but their availability is usually limited and their cost high [112, 117].
- Most of the experiments have been carried out on simple scenes, usually consisting of a single shadowed area with limited materials [9, 120]. In reality, airborne or satellite images present more complicated scenarios.

The proposed framework can contribute to some extent to the reported open problems. As an extension of our previous work [204, 205], we propose a shadow detection and restoration method for high-resolution hyperspectral reflectance images based on nonlinear unmixing, considering both umbra and penumbra. Our proposed framework restores reflectance data in shadowed regions without the requirement of shadow detection results as an additional input. In addition to the restored images, the framework computes soft shadow detection maps ranging from 0 to 1 which, unlike binary masks, yield a natural restoration on the shadow boundaries. As an optional step, our method iteratively refines the initial spectral library by automatically including undetected materials. We tested the proposed framework on airborne data acquired by an imaging spectrometer in the VNIR spectral ranges.

The remainder of this section is organized as follows. Section 5.1.2 proposes a shadow detection and restoration method based on radiative transfer and a nonlinear unmixing model. Section 5.1.3 introduces the test data acquired by an imaging spectrometer and Section 5.1.4 analyzes experimental results, followed by detailed discussions in Section 5.1.5. Finally, Section 5.1.6 summarizes this work and gives directions for possible future extensions.

5.1.2 Methodology

The proposed framework for simultaneous shadow detection and removal is reported in Fig. 5.1. The input contains a hyperspectral image and a spectral library consisting of pure spectra from sunlit regions, i.e. sunlit endmembers. In particular, the spectral library should only include endmembers selected from fully sunlit areas, while it should not include any endmember from shadows or penumbra regions. In order to fully satisfy these requirements, we manually select endmembers from sunlit regions in this work. The output of the framework consists of a sunlit factor map and a restored shadow-free hyperspectral image.

Direct and diffuse solar radiations are the main illumination sources for outdoor scenes [55]. Sunlit regions receive both of them, while the umbra in shadowed regions receives mainly the diffuse solar irradiance due to occlusion. Despite different illumination conditions between sunlit and shadowed regions, reflectance as a physical property should remain theoretically unchanged for a material. In reality, though, reflectances derived by atmospheric correction in shadowed regions are much lower than those in sunlit regions for the same material. The reason is, atmospheric correction incorrectly assumes that both sunlit and shadowed areas receive direct and diffuse solar radiations.

In the proposed framework, we model the reflectance of a shadowed material given the reflectance of the same material under sunlight, following the assumptions in atmospheric correction (Section 5.1.2.1). Subsequently, we regard both sunlit and shadowed spectra as endmembers and present a nonlinear spectral unmixing approach (Section 5.1.2.2). Finally, sunlit spectra and abundances are used to restore the shadow-free image (Section 5.1.2.2). The proposed framework generates as an additional output a soft shadow mask, i.e. sunlit factor map, by residual analysis of the mixing models (Section 5.1.2.3). The sunlit factor map can locate sunlit pixels, where values of the restored image are then replaced by their input pixels.

5.1.2.1 Shadowed Spectra Model

Assuming the ground surface to be Lambertian, the reflectance of a sunlit pixel can be computed as:

$$\hat{r}(\lambda) = \frac{\pi \mathcal{L}_{\mathrm{l}}(\lambda)}{\mathcal{E}_{\mathrm{l}}(\lambda) + \mathcal{E}_{\mathrm{s}}(\lambda)} \tag{5.1}$$

where $\mathcal{L}_{\mathrm{l}}(\lambda) = \frac{(\mathcal{E}_{\mathrm{l}}(\lambda)+\mathcal{E}_{\mathrm{s}}(\lambda))r(\lambda)}{\pi}$ is the radiance of the sunlit pixel at wavelength λ, $\mathcal{E}_{\mathrm{l}}$ is the direct solar irradiance at the sunlit pixel at wavelength λ, and $\mathcal{E}_{\mathrm{s}}$ is the diffuse solar irradiance at the sunlit pixel at wavelength λ.

Following the atmospheric correction, the observed reflectance for a shadow pixel $\hat{r}_{\mathrm{s}}$ can be written in Eq. (5.2). We use the term "observed" as Eq. (5.2) follows the computation of the atmospheric correction step. Nevertheless, such observed reflectance is incorrect in terms of physics.

$$\hat{r}_{\mathrm{s}}(\lambda) = \frac{\pi \mathcal{L}_{\mathrm{s}}(\lambda)}{\mathcal{E}_{\mathrm{l}}(\lambda) + \mathcal{E}_{\mathrm{s}}(\lambda)} + \frac{\pi \mathcal{L}'_{\mathrm{s}}(\lambda)}{\mathcal{E}_{\mathrm{l}}(\lambda) + \mathcal{E}_{\mathrm{s}}(\lambda)} \tag{5.2}$$

where $\mathcal{L}_{\mathrm{s}}(\lambda) = \frac{\mathcal{E}_{\mathrm{s}}(\lambda)r(\lambda)}{\pi}$ is the radiance of the shadowed pixel contributed at wavelength λ by the linear part, i.e. diffuse solar irradiance, while $\mathcal{L}'_{\mathrm{s}}(\lambda)$ is the radiance of the shadowed pixel at wavelength λ contributed by the nonlinear part, i.e. multiple reflections of direct solar irradiance caused by surrounding objects.

Modeling nonlinear effects for spectral unmixing has been explored for decades. In this section, we compute $\mathcal{L}'_{\mathrm{s}}(\lambda)$ using the Fan model [87], which forms nonlinear interactions

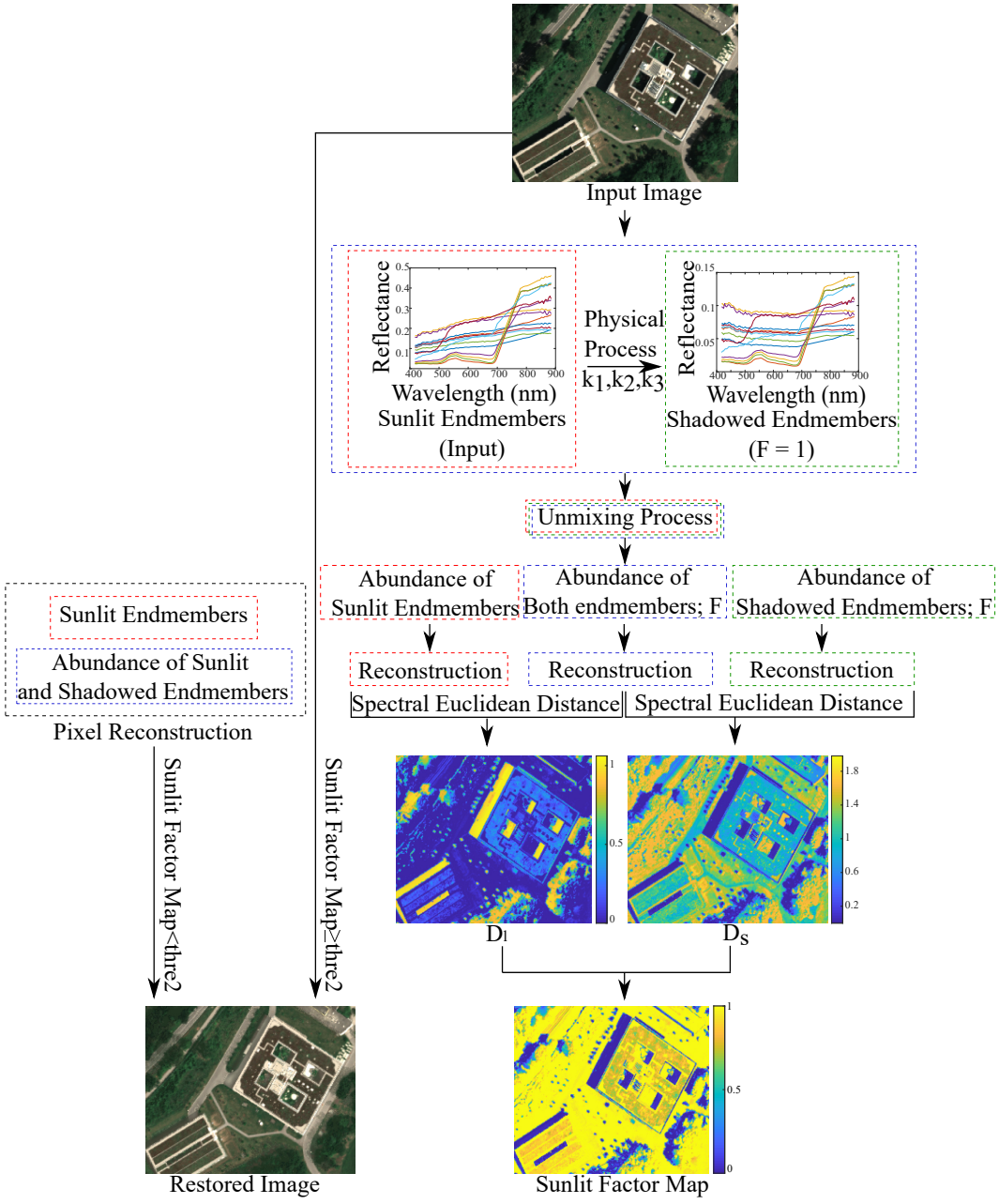


Figure 5.1: The proposed framework. The inputs are a hyperspectral image, the physical parameters k_1, k_2, and k_3, and a spectral library containing manually selected endmembers in sunlit regions, i.e. sunlit endmembers. After the unmixing process, the restored image is reconstructed by a nonlinear combination of the sunlit endmembers, using the abundances of the same materials in the shadow. The framework outputs the sunlit factor map, computed by spectral Euclidean distances of the reconstruction results. Finally, in order to avoid introducing spectral distortions, sunlit pixels in the restored image are replaced by their original values.

through the multiplication of reflectances using abundances as coefficients:

$$\mathcal{L}'_{\mathrm{s}}(\lambda) = (\mathcal{E}_{\mathrm{s}}(\lambda) + (\mathcal{E}_{\mathrm{l}}(\lambda)) \frac{\sum_{i_1=1}^{p-1} \sum_{i_2=i_1+1}^{p} a_{i_1} a_{i_2} \hat{r}_{i_1}(\lambda) \hat{r}_{i_2}(\lambda)}{\pi} \tag{5.3}$$

where p is the number of materials (endmembers) in one pixel, $\hat{r}_i(\lambda)$ is the reflectance of the i-th sunlit material (endmember) at wavelength λ, and a_i is the i-th abundance corresponding to $\hat{r}_i$.

After combining Eqs. (5.1), (5.2), and (5.3), $\hat{r}_{\mathrm{s}}$ can be written as:

$$\hat{r}_{\mathrm{s}}(\lambda) = \frac{\mathcal{E}_{\mathrm{s}}(\lambda)}{\mathcal{E}_{\mathrm{l}}(\lambda) + \mathcal{E}_{\mathrm{s}}(\lambda)} \hat{r}(\lambda) + \sum_{i_1=1}^{p-1} \sum_{i_2=i_1+1}^{p} a_{i_1} a_{i_2} \hat{r}_{i_1}(\lambda) \hat{r}_{i_2}(\lambda) \tag{5.4}$$

The ratio $\frac{\mathcal{E}_{\mathrm{s}}(\lambda)}{\mathcal{E}_{\mathrm{l}}(\lambda)}$ indicates the proportion of the diffuse solar irradiance to the direct solar irradiance on the ground surface. For the same time and location, this ratio becomes smaller at longer wavelengths. Besides, this ratio depends on atmospheric conditions such as aerosol, humidity, and dust content [206]. Consequently, we model the ratio $\frac{\mathcal{E}_{\mathrm{s}}(\lambda)}{\mathcal{E}_{\mathrm{l}}(\lambda)}$ as a power function $k_1 \lambda^{-k_2} + k_3$. By assuming atmospheric conditions to be constant across a single airborne image, all parameters k_1, k_2, and k_3 are constants. Another free parameter F, representing how much diffuse irradiance a pixel receives out of a certain direct solar irradiance, is estimated pixelwise. The described ratio is then computed as:

$$\frac{\mathcal{E}_{\mathrm{s}}(\lambda)}{\mathcal{E}_{\mathrm{l}}(\lambda)} = F(k_1 \lambda^{-k_2} + k_3) \tag{5.5}$$

where λ is a wavelength, k_1, k_2, k_3 are positive quantities, and F ranges from 0 to 1.

By combining Eqs. (5.4) and (5.5), we have:

$$\hat{r}_{\mathrm{s}}(\lambda) = \frac{F(k_1 \lambda^{-k_2} + k_3)}{F(k_1 \lambda^{-k_2} + k_3) + 1} \hat{r}(\lambda) + \sum_{i_1=1}^{p-1} \sum_{i_2=i_1+1}^{p} a_{i_1} a_{i_2} \hat{r}_{i_1}(\lambda) \hat{r}_{i_2}(\lambda) \tag{5.6}$$

The parameters k_1, k_2, and k_3 in Eq. (5.6) can be solved by using manually selected pairs of sunlit spectra $\hat{r}(\lambda)$ and shadowed spectra $\hat{r}_{\mathrm{s}}(\lambda)$ for selected materials in the scene. In high-resolution images, shadow boundaries appear between sunlit and shadowed regions and may span more than one pixel. As an example in Fig. 5.2, the selected pixels in each pair should therefore be located close to but not directly on the shadow boundary.

5.1.2.2 Nonlinear Mixing Model

We write Eq. (5.6) in vector form, in order to solve for all wavelengths simultaneously and construct a nonlinear mixing model to allow more materials to be present in one pixel. Note that $\boldsymbol{e}_i$ is the i-th sunlit endmember, $\boldsymbol{e}_{\mathrm{s}_i}$ is the i-th shadowed endmember, where $i = 1, 2, ...p$, with p the total number of endmembers, a_i is the i-th abundance corresponding to $\boldsymbol{e}_i$, and a_{s_i} is the i-th abundance corresponding to $\boldsymbol{e}_{\mathrm{s}_i}$. Given $\boldsymbol{e}_i$, a corresponding shadowed endmember

Figure 5.2: An example of selecting pure sunlit (with blue marker) and shadowed pixels (with red marker) for the same material.

$\boldsymbol{e}_{s_i}$ can be written as:

$$\boldsymbol{e}_{s_i} = \frac{F(k_1\boldsymbol{\lambda}^{-k_2}+k_3)}{F(k_1\boldsymbol{\lambda}^{-k_2}+k_3)+1}\boldsymbol{e}_i + \sum_{i_1=1}^{p-1}\sum_{i_2=i_1+1}^{p} a_{i_1}a_{i_2}\boldsymbol{e}_{i_1}\boldsymbol{e}_{i_2} \tag{5.7}$$

Considering both $\boldsymbol{e}_i$ and $\boldsymbol{e}_{s_i}$, a pixel $\boldsymbol{x}$ can be written in Eq. (5.8):

$$\boldsymbol{x} = \sum_{i=1}^{p} a_i\boldsymbol{e}_i + \sum_{i=1}^{p}\frac{F(k_1\boldsymbol{\lambda}^{-k_2}+k_3)}{F(k_1\boldsymbol{\lambda}^{-k_2}+k_3)+1}a_{s_i}\boldsymbol{e}_i + \sum_{i_1=1}^{p-1}\sum_{i_2=i_1+1}^{p} a_{i_1}a_{i_2}\boldsymbol{e}_{i_1}\boldsymbol{e}_{i_2} \tag{5.8}$$

where $\sum_{i=1}^{p}(a_i + a_{s_i}) = 1$, $a_i \geq 0$, and $a_{s_i} \geq 0$. In order to account for physical considerations, abundances are positive values. In addition, we apply the sum-to-one constraint by assuming that all endmembers are recognized for each pixel. Since spectral values of shadowed pixels are much lower than those of sunlit pixels, the sum-to-one constraint assures that shadowed pixels yield large abundances of shadowed endmembers, instead of small abundances of sunlit endmembers. Moreover, when solving this equation, we additionally apply a Total Generalized Variation (TGV) algorithm to the parameter F for spatial smoothness in an iterative manner [207]. In the first iteration, we solve all unknown parameters in Eq. (5.8). After that, F is spatially filtered through the TGV algorithm, and then used as a known parameter in the second iteration.

With a_{s_i} and a_i representing respectively the abundance of shadowed and sunlit endmember for the same material, the shadow-restored pixel $\boldsymbol{x}_{\text{restore}}$ is computed as:

$$\boldsymbol{x}_{\text{restore}} = \sum_{i=1}^{p}(a_i + a_{s_i})\boldsymbol{e}_i + \sum_{i_1=1}^{p-1}\sum_{i_2=i_1+1}^{p} a_{i_1}a_{i_2}\boldsymbol{e}_{i_1}\boldsymbol{e}_{i_2} \tag{5.9}$$

5.1.2.3 Sunlit Factor Map

From Section 5.1.2.1 and 5.1.2.2, endmembers can be either sunlit $\boldsymbol{e}_i$ or shadowed $\boldsymbol{e}_{\mathrm{s}_i}$. We decompose Eq. (5.8) into two equations by separating the $\boldsymbol{e}_i$ and $\boldsymbol{e}_{\mathrm{s}_i}$ terms, resulting in Eqs. (5.10) and (5.11).

$$\boldsymbol{x} = \sum_{i=1}^{p} a_{\mathrm{s}_i} \boldsymbol{e}_{\mathrm{s}_i} \tag{5.10}$$

$$\boldsymbol{x} = \sum_{i=1}^{p} a_i \boldsymbol{e}_i + \sum_{i_1=1}^{p-1} \sum_{i_2=i_1+1}^{p} a_{i_1} a_{i_2} \boldsymbol{e}_{i_1} \boldsymbol{e}_{i_2} \tag{5.11}$$

After spectral unmixing, the reconstructed pixels using Eqs. (5.8), (5.10), and (5.11) are noted as $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{x}}_{\mathrm{s}}$, and $\hat{\boldsymbol{x}}_{\mathrm{l}}$, respectively. Both sunlit and shadowed pixels can be reconstructed with small reconstruction errors using Eq. (5.8). Shadowed pixels can be reconstructed using Eq. (5.10), while sunlit pixels can be reconstructed using Eq. (5.11) with small reconstruction errors. Therefore, in a B-dimensional space spanned by B spectral bands, the Euclidean distance D_{s} between $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{x}}_{\mathrm{s}}$ is small in shadowed pixels and large in sunlit pixels. On the other hand, the Euclidean distance D_{l} between $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{x}}_{\mathrm{l}}$ is large in shadowed pixels and small in sunlit pixels. We therefore compute a sunlit factor map pixel by pixel according to equation $\frac{D_{\mathrm{s}}}{D_{\mathrm{l}}+D_{\mathrm{s}}}$. The sunlit factor map ranges from 0 to 1. In this experiment, we use two fixed thresholds set as $thre1 = 0.1$ and $thre2 = 0.9$, respectively. When sunlit factor values are smaller than $thre1$, pixels are assumed to be fully shadowed pixels (umbra). When sunlit factor values are larger than $thre2$, pixels are assumed to be fully sunlit pixels.

5.1.3 Dataset

The proposed method was validated on six subsets (see Figs. 5.3 and 5.4) from hyperspectral images in the HySpex/3K dataset (Section 4.1, Chapter 4). Six subsets consist of common ground objects, such as buildings, grass, and trees. The workflow for all six subsets is kept unaltered, including the fourth containing a large pond of water, for which no additional water mask was used. Such targets are usually challenging for this kind of application, as water can be confused with shadows due to its low albedo. A spectral library is given as an input by manually selecting pure pixels of relevant materials in sunlit regions for each subset (second row in Fig. 5.4). In addition, ten pairs of pixels have been selected in the experiment to compute parameters k_1, k_2, and k_3 in Eq. (5.6). We solve the parameters $k_1 = 1.296$, $k_2 = 6.068$, and $k_3 = 0.442$ according to Eq. (5.6) as described in Section 5.1.2.1, and these parameters are assumed to be constant for all the processed subsets.

5.1.4 Results

5.1.4.1 Reconstruction Error

We compare the proposed mixing model in Eq. (5.8) with two well-known models, i.e. LMM[76] and the Fan model [87]. The mean reconstruction errors REs are computed for

Figure 5.3: Six subsets selected from hyperspectral images in the HySpex/3K dataset (see Section 4.1, Chapter 4) in the study area of Oberpfaffenhofen, Bavaria, Germany.

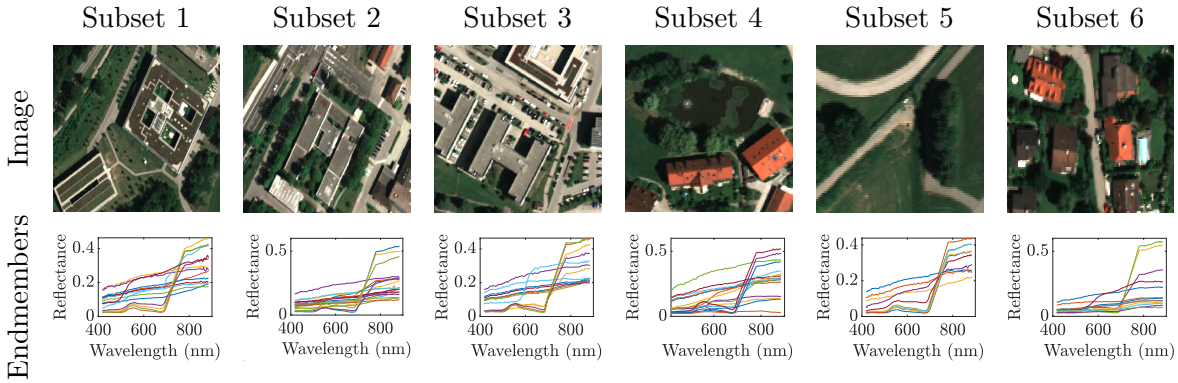


Figure 5.4: Six subsets with manually selected sunlit endmembers.

Table 5.1: Mean reconstruction errors *REs* for six subsets

Subset	Region	LMM	Fan	Proposed
1	sunlit regions	0.113	0.083	0.077
	shadowed regions	0.414	0.421	0.026
	both	0.191	0.171	0.064
2	sunlit regions	0.088	0.077	0.071
	shadowed regions	0.209	0.210	0.021
	both	0.114	0.106	0.060
3	sunlit regions	0.092	0.090	0.081
	shadowed regions	0.708	0.738	0.023
	both	0.290	0.298	0.062
4	sunlit regions	0.059	0.044	0.039
	shadowed regions	0.099	0.100	0.018
	both	0.064	0.052	0.037
5	sunlit regions	0.088	0.079	0.063
	shadowed regions	0.0685	0.732	0.030
	both	0.199	0.200	0.057
6	sunlit regions	0.126	0.108	0.117
	shadowed regions	0.156	0.158	0.025
	both	0.132	0.118	0.084

each subset. In addition, we individually compute *RE*s for sunlit and shadowed regions. Please refer to Section 2.5.3 about evaluation metrics for spectral unmixing. Table 5.1 shows *RE*s in subsets 1 to 6. In sunlit regions, we observe a small change of errors among the three models, where the difference of errors remains within 0.04. Compared with the Fan model and the proposed model, the LMM model presents slightly higher errors in sunlit regions. This indicates that the proposed model shows similar reconstruction results with other models in sunlit regions. However, our model exhibits significant improvements in shadowed regions, yielding considerably lower errors with respect to the other two models. This improvement demonstrates that our method can effectively model shadowed pixels.

5.1.4.2 Spectral Distance

An important criteria of shadow restoration is the spectral distance between sunlit and shadowed pixels belonging to the same material. Ideally, the reflectance is an intrinsic property of materials, and should not change between sunlit and shadowed areas. Thus, the spectral distance between sunlit and shadowed pixels for one material in restored images should be significantly smaller than in the input images. In this experiment, we compute the spectral distance using $\|\boldsymbol{x}_\mathrm{l} - \boldsymbol{x}_\mathrm{s}\|_2$ for the input images and $\|\hat{\boldsymbol{x}}_\mathrm{l} - \hat{\boldsymbol{x}}_\mathrm{s}\|_2$ for the restored images, respectively.

We select pairs of sun-shade pixels in each subset, as shown in the first row of Fig. 5.5. For each pair of pixels, the yellow and cyan markers represent sunlit and shadowed pixels, respectively. Both markers in each pair are close to each other and to the shadow boundary, so we assume that the selected sunlit and shadowed pixels belong to the same material. The second row in Fig. 5.5 shows the spectral distance between sunlit and shadowed pixels for each pair of pixels. The bars in blue and orange represent the spectral distances computed from the original and restored images, respectively. After shadow restoration, the spectral distances decrease significantly. One exception is represented by the sixth sample of subset 4, where the spectral distance increases by 0.1 after shadow restoration. This sample belongs to water, for which reflectances are small in both sunlit (lower than 0.035) and shadowed regions (lower than 0.025), as it shows in Fig. 5.6. In addition, the shadowed water pixels are affected by nonlinear effects known to be relevant in water, and are shadowed also by trees. This causes the restored pixels to contain a small abundance value of the material "trees", in the spectral range known as the red edge (Fig. 5.6).

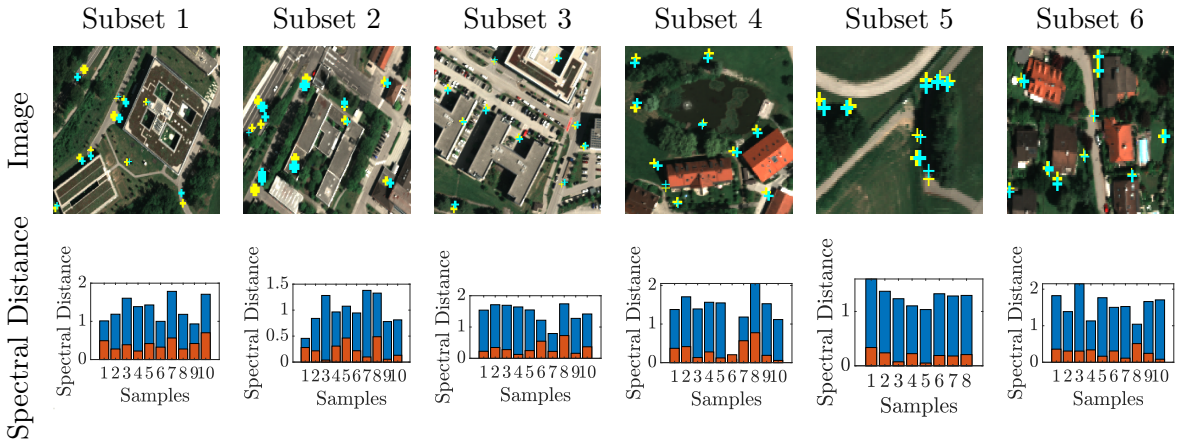


Figure 5.5: Comparison of the spectral Euclidean distance between input and restored images. First row: the six subsets. Second row: spectral distance of up to 10 pairs of samples in each subset (input and restored images in blue and orange, respectively).

5.1.4.3 Restoration and Classification Results

Fig. 5.7 compares input and restored images, along with their classifications. A total of 6565 training samples are manually selected from sunlit regions, while a total of 5927 test samples are selected in comparable quantities from both sunlit and shadowed regions. There are seven classes in six subsets, including tree, grass, impervious, bare soil, tiled roof, objects painted in red, i.e. red material, and water. As an example, Fig. 5.8 reports a detailed comparison of pixelwise classifications highlighting improvements after shadow restoration. The hyperspectral images used in this study were acquired with nearly the same acquisition time and solar zenith angles. Therefore, we can not validate our restored images with additional acquisitions with shadows occupying smaller areas. As an alternative, we compare the results with Google Earth images at same locations with the acquisition date of July 10th, 2016 in Fig. 5.9, with the assumption that most ground objects did not change within a two years time span.

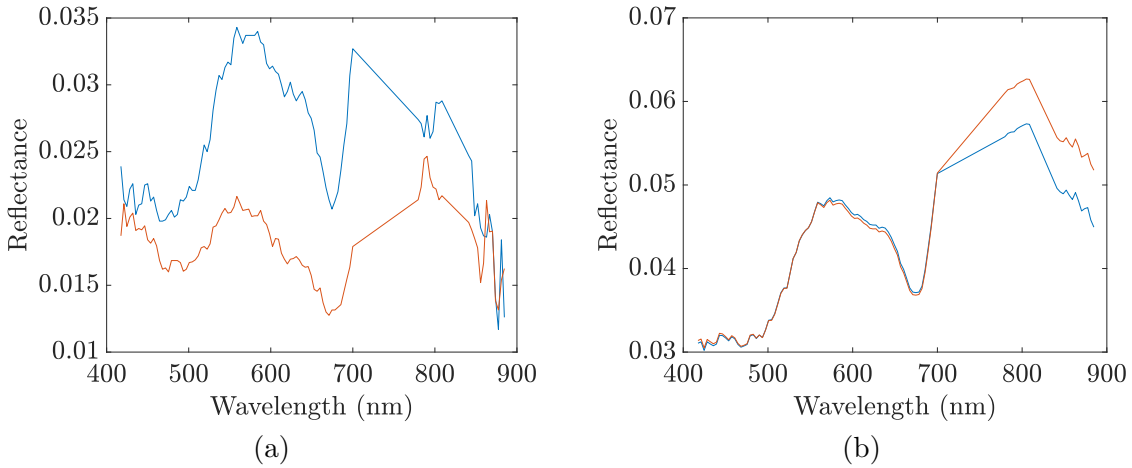


Figure 5.6: Mean reflectance of water regions in subset 4 of Fig. 5.5. The blue and red color represents mean reflectance of sunlit and shadowed pixels, respectively. Pixels are selected from (a) the input image and (b) the restored image in subset 4.

The classifications of the input images are inaccurate for most of the shadowed regions. When the water class is not present in a subset, shadowed impervious surfaces are mostly classified as vegetation (subset 1, 3, and 5) or tiled roof (subset 2). When water pixels are included in the training samples (subset 4), most shadowed regions are classified as water. The tree and grass pixels, both in input and restored images, are mostly classified as vegetation because the discriminative "red edge" feature typical of vegetation is visible in shadowed areas. Besides large and homogeneous areas, smaller objects are also recovered in shadowed regions. For example, subset 3 contains trees in the shadow, with tree crowns becoming visible in the restored image. A white car on the left side of the "H"-shape building in subset 3 is an example for other isolated objects being restored. Compared to white and red cars, dark objects, e.g. black cars, are considered as shadowed pixels in our proposed framework, as their reflectance values are small and comparable with shadowed pixels. In subset 3, these are restored as impervious surfaces.

Impervious surfaces shadowed by trees are sometimes classified as vegetation (e.g. on the top left side in subset 1). When pixels are shadowed by trees, especially in deep shadows, their spectra contain the "red edge" feature, due to incoming light interactions with the nearby trees. Thus, the abundance values of vegetation at these impervious surfaces are larger than zero, resulting in a mixture of impervious and vegetation materials in the reconstruction.

Table 5.2 presents the overall accuracies (OA) and Kappa (K) values of classification results. Both figures of merit increase by more than 10 % in subsets 2, 4, 5, and 6, and increase by more than 20 % in subsets 1 and 3. The increase in performance is due to the improved classification results in shadowed regions.

5.1.4.4 Sunlit Factor Map

Sunlit factor maps in Fig. 5.10 present an additional output of the proposed framework. The values range from 0 to 1. Instead of a binary mask, Fig. 5.10 shows a smooth transition

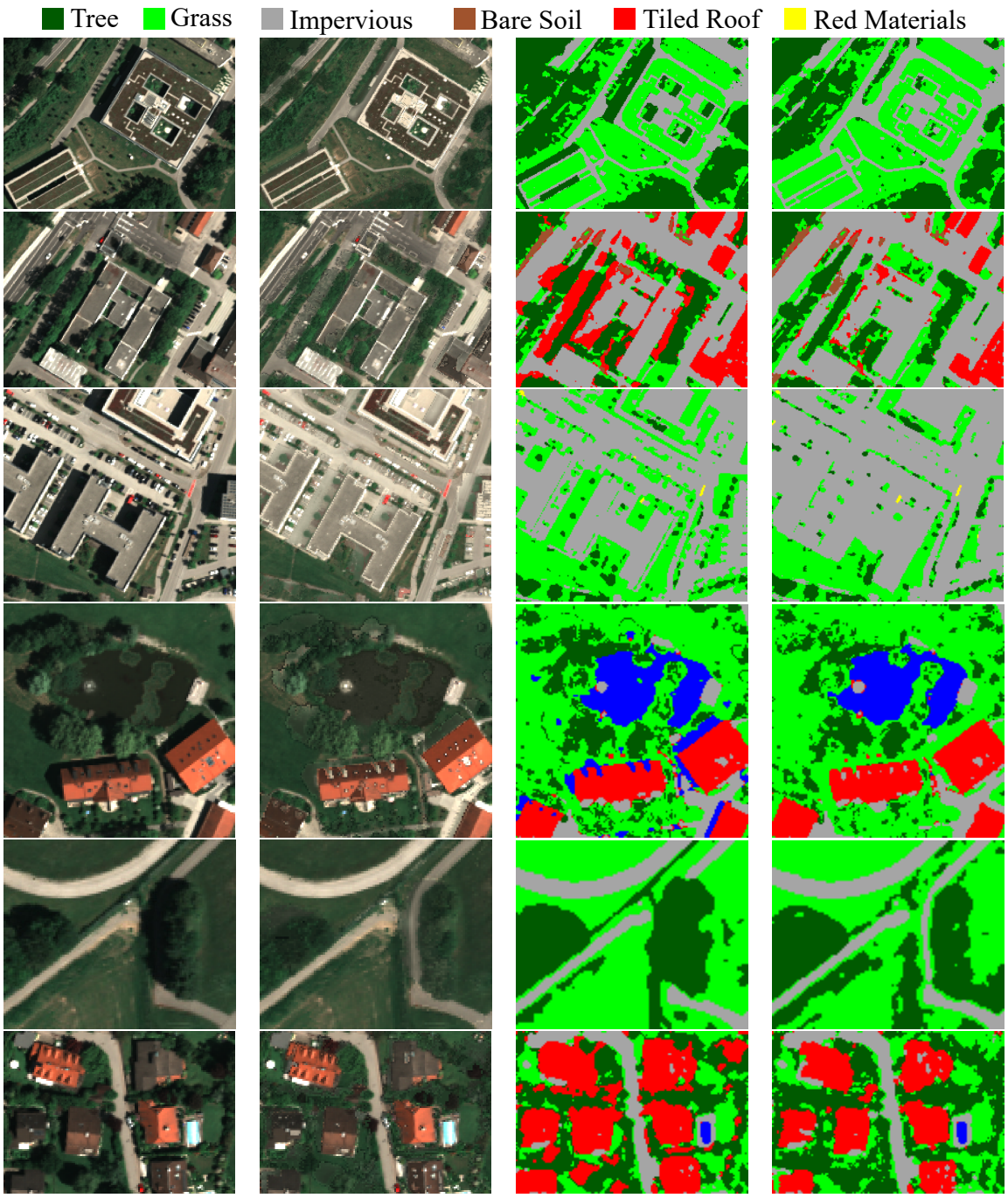
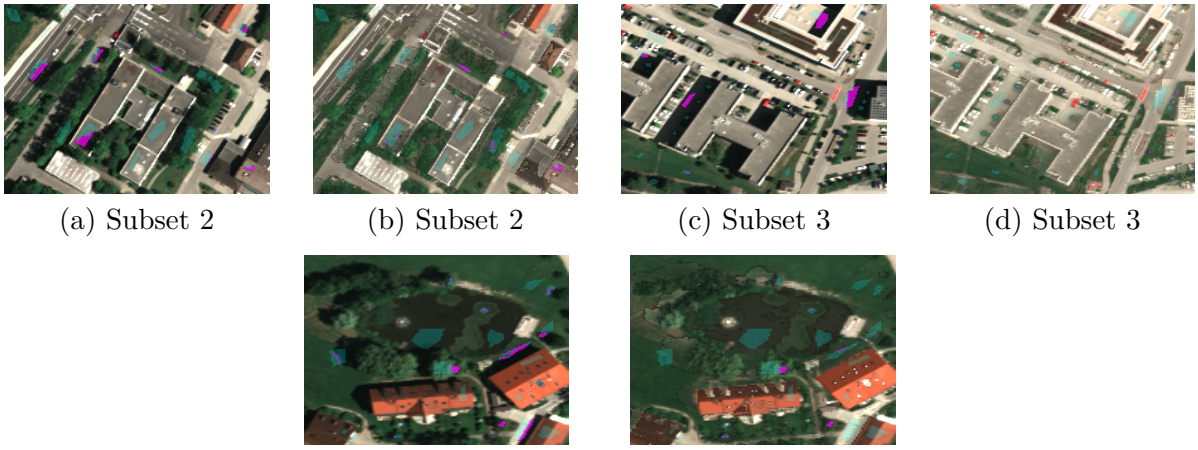


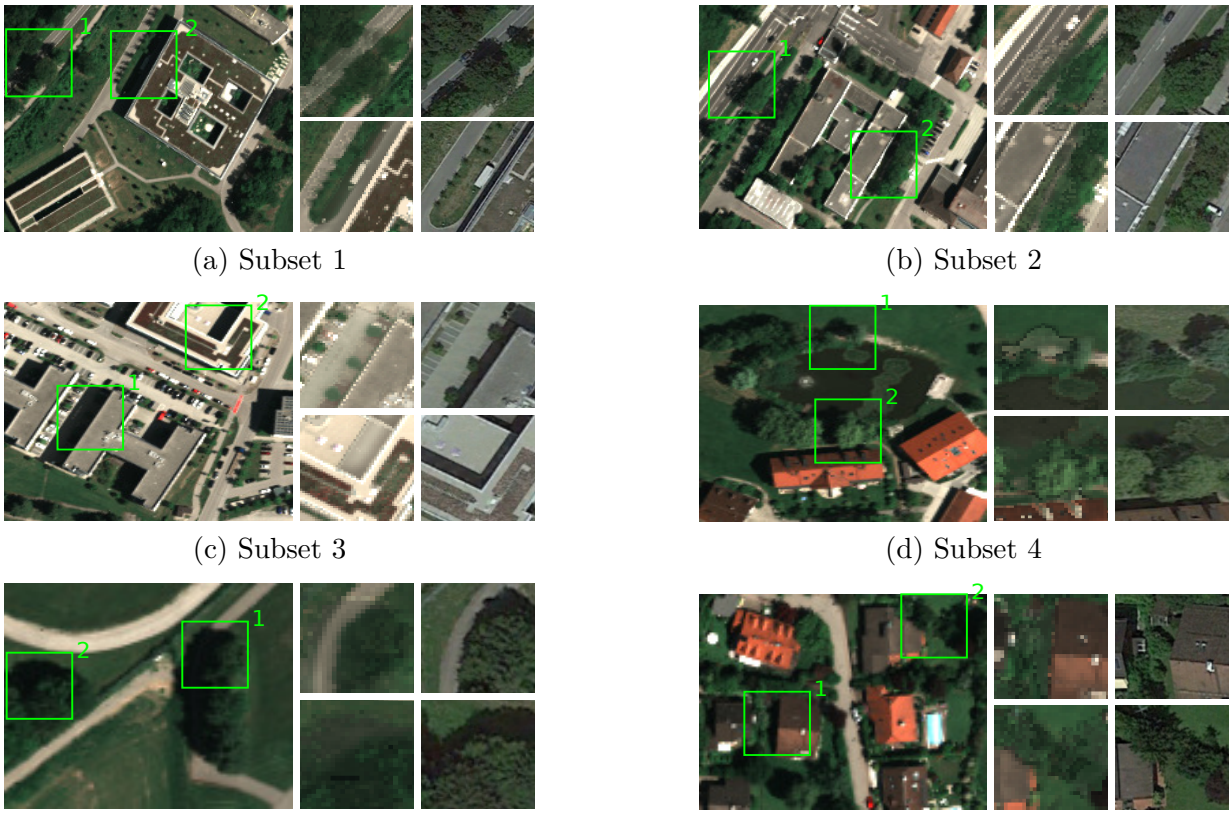
Figure 5.7: Rows: Six subsets. First column: input images; second column: restored images; third column: classification maps of the input images; last column: classification maps of the restored images.



(a) Subset 2 (b) Subset 2 (c) Subset 3 (d) Subset 3

(e) Subset 4 (f) Subset 4

Figure 5.8: Comparison of classification results in Table 5.2 for input images (a, c, e) and restored images (b, d, f) in subsets 2, 3, and 4. Correctly and incorrectly classified areas are marked in cyan and magenta, respectively.



(a) Subset 1 (b) Subset 2

(c) Subset 3 (d) Subset 4

(e) Subset 5 (f) Subset 6

Figure 5.9: Comparison between restored subsets and Google Earth images. For each subset, on the left: input image with two selected regions of interest; rows on the right: regions of interest from restored image and screenshots from Google Earth data in which shadowed areas are partially sunlit.

between sunlit and shadowed areas, yielding a more realistic representation of shadows. This work sets two thresholds $thre1 = 0.1$ and $thre2 = 0.9$ to identify pure shadowed pixels (value

Table 5.2: Comparison of classification accuracies using input and restored images

Data	Input	Restored
Subset 1	OA=73.472% K=0.552	OA=95.366% K=0.927
Subset 2	OA=82.203% K=0.715	OA=93.553% K=0.883
Subset 3	OA=55.0% K=0.366	OA=93.939% K=0.880
Subset 4	OA=84.495% K=0.799	OA=95.138% K=0.937
Subset 5	OA=80.340% K=0.703	OA=90.170% K=0.852
Subset 6	OA=85.373% K=0.80	OA=93.284% K=0.908

$< thre1$) and pure sunlit pixels (value $> thre2$). The values between $thre1$ and $thre2$ are regarded as transition areas between sunlit and shadowed pixels, i.e. shadow boundaries. When an area is shadowed by man-made objects, i.e. buildings, the transition areas are smaller. When an area is shadowed by vegetation, i.e. trees, the shadow boundaries span larger regions.

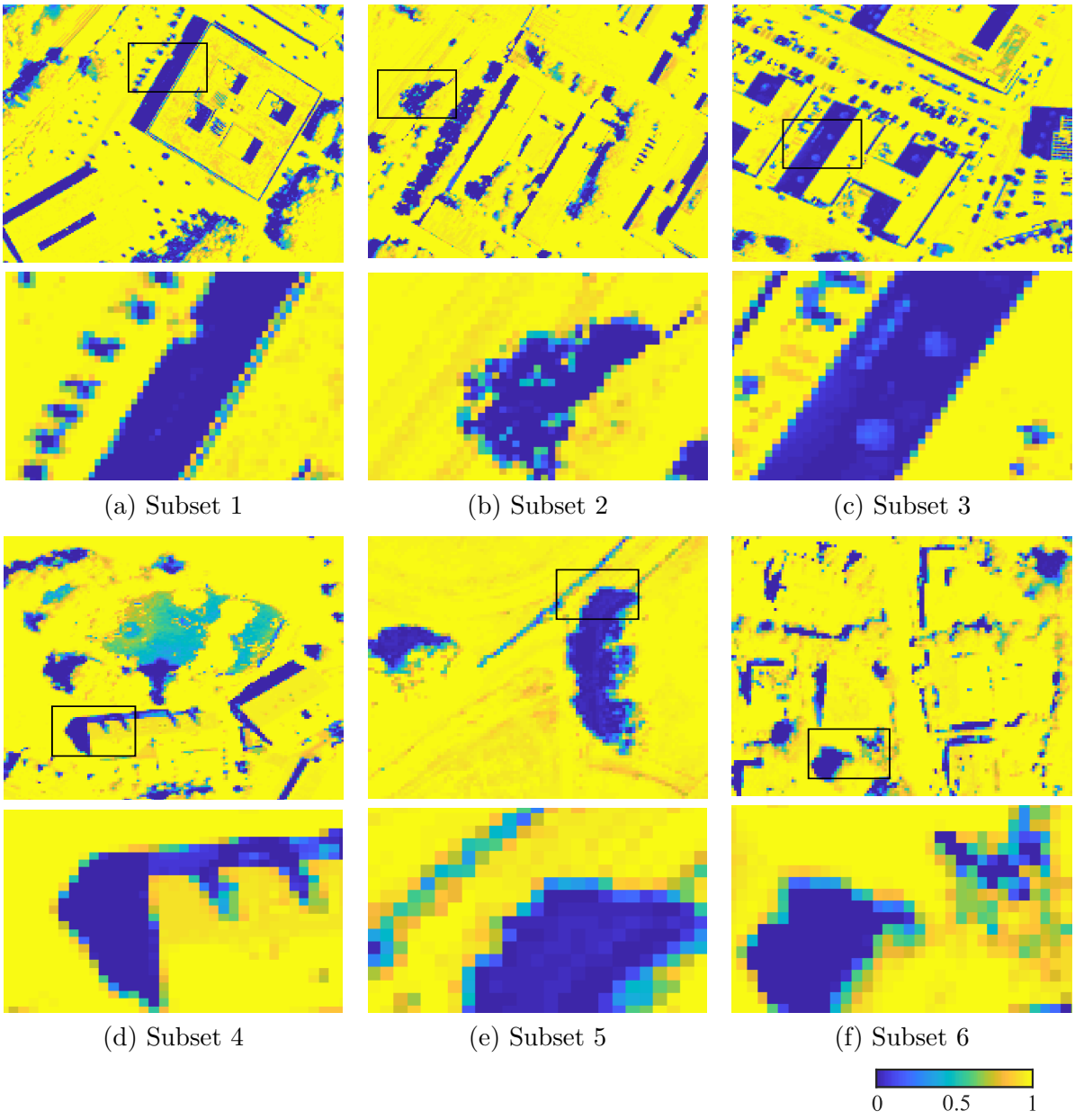
5.1.4.5 The F Parameter

For a pixel on the ground surface, the diffuse solar irradiance come isotropically from the sky [55]. For a given location and acquisition time, the proportion of diffuse to direct solar irradiance is constant. However, at a shadowed pixel where the sky is partially occluded, the diffuse solar irradiance decreases because the pixel can not see the sky from all directions. The F parameter (Fig. 5.11) represents the scale of the proportion of diffuse to direct solar irradiance in Eq. (5.5). We set F values at sunlit pixels to zeros, as F is relevant for the shadowed terms in Eq. (5.8). The F values remain approximately homogeneous within one shadowed region and slightly increase on the shadow boundaries. Among different shadowed regions, pixels shadowed by vegetation show moderately larger values with respect to pixels shadowed by man-made objects.

5.1.5 Discussion

5.1.5.1 Level of Automatism

The framework runs automatically giving as input a hyperspectral image, the selected endmembers, and the relevant parameters. This implies that our method so far depends on manually selected endmembers, as the input spectral library is composed by pure pixels



(a) Subset 1 (b) Subset 2 (c) Subset 3

(d) Subset 4 (e) Subset 5 (f) Subset 6

Figure 5.10: Sunlit factor maps ranging from 0 to 1. Values smaller than $thre1$ are considered as pure shadowed pixels. Values larger than $thre2$ are regarded as pure sunlit pixels. For each subset, top: sunlit factor map marked with the region of interest; bottom: zoomed-in image of the region of interest.

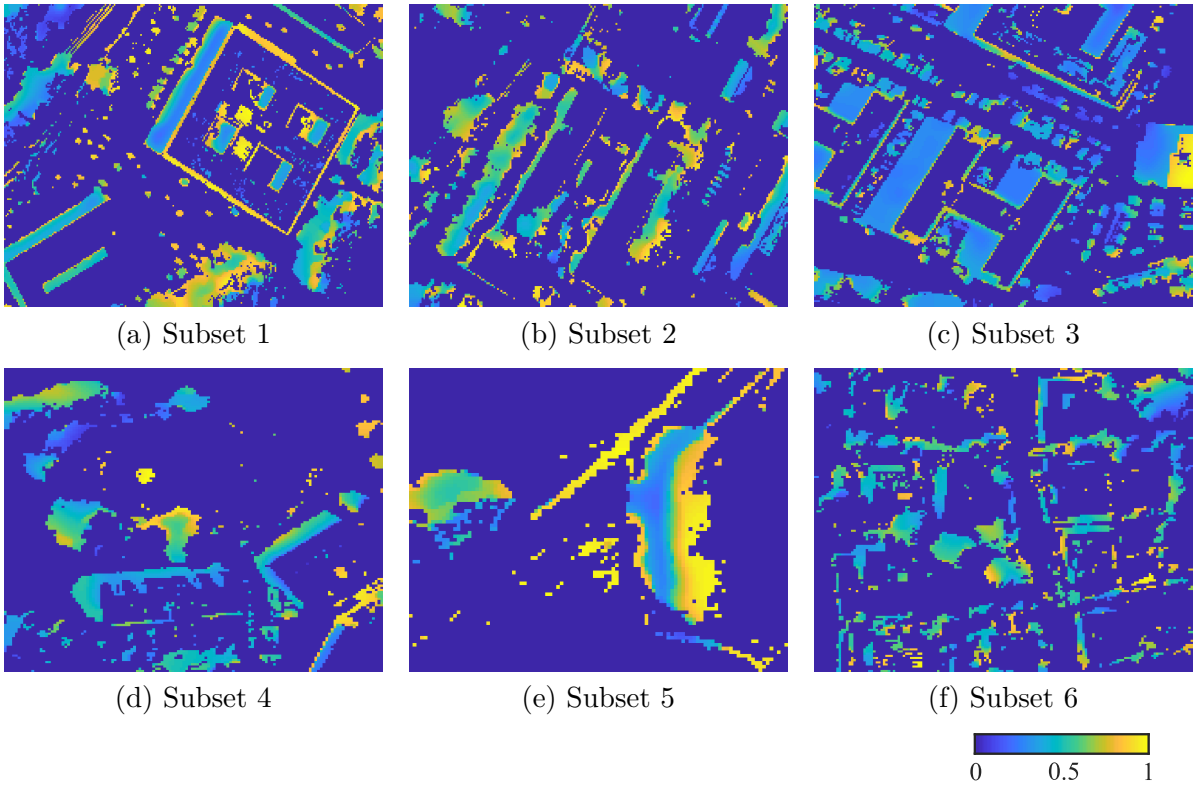


Figure 5.11: The F parameter with the range of values from 0 to 1.

exclusively selected in sunlit regions. However, to the best of our knowledge, existing endmember extraction methods either do ignore shadowed regions, or regard shadowed regions as an additional dark endmember. Thus, their extracted endmembers usually contain pixels in shadowed regions or on the shadow boundaries, which cannot be used in our framework. In addition, the input spectral library should consider the fact that the observed values of the same material in hyperspectral images may vary, due to the spectral variability effect [208], which has been taken into account by manually selecting endmembers.

An endmember extraction method that excludes shadowed regions and shadow boundaries would not only help our specific framework, but also yield a more consistent physical representation of a scene, as the reflectance of a specific material should not change according to illumination conditions. Therefore, we introduce a simple but effective way of extracting endmembers automatically by taking into account shadows.

A straightforward way of selecting sunlit pixels is thresholding. In our experiment in subset 1, all the pixels having mean reflectance larger than an empirical threshold (set to 0.08 in this experiment) are selected as candidate sunlit pixels. However, this may include some pixels located at shadow boundaries. Thus, a Canny edge detector [209] has been applied to detect and remove all boundary pixels from sunlit pixels candidates. In addition, considering the endmember variability effect, we apply the method in [210] to extract endmember bundles based on Vertex Component Analysis (VCA) [71]. By merging similar endmembers, we show the automatically extracted endmembers in Fig. 5.12 (a). By using endmembers reported in Fig. 5.12 (a) and our proposed framework, our results are shown in Fig. 5.12 (b, d, e).

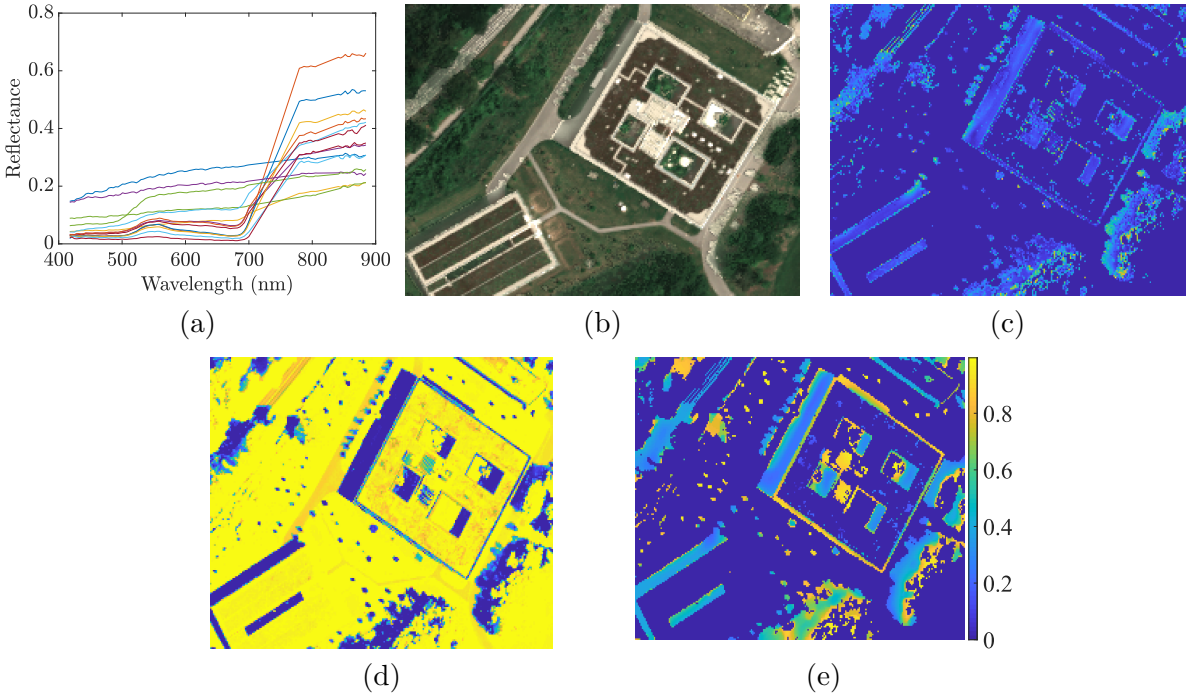


Figure 5.12: Shadow detection and restoration using automatically extracted endmembers in subset 1. (a) extracted endmembers, (b) restored image, (c) Euclidean distance between restored images using manually and automatically extracted endmembers, (d) sunlit factor map, and (e) F parameter.

Both restoration and computed parameters are visually similar to the results obtained by employing the manually selected endmembers. Fig. 5.12 (c) depicts the Euclidean distance of the images of subset 1 restored by manual and automatic endmember extraction, having a maximum value of 0.13. This slight difference is due to the slightly different sets of endmembers selected.

5.1.5.2 Computational Cost

All algorithms were developed in MATLAB and run on an Intel Core i7 – 8650U CPU, 1.90GHz machine with 4 Cores and 8 Logical Processors. We use the MATLAB function FMINCON to perform nonlinear optimization. The processing time depends on the number of input pixels and endmembers. If a shadow map is unknown, the algorithm requires 2445 seconds to restore the subset 1, having a size of 181×245 pixels. Otherwise, the algorithm needs 1031 additional seconds to produce a sunlit factor map. On the other hand, if a shadow detection map is given, then the algorithm processes only shadowed pixels, requiring 948 seconds for shadow restoration.

5.1.5.3 Benefits and Challenges

The proposed framework shows promising results on detecting shadows and restoring spectral information in shadowed regions for hyperspectral imagery. Methods proposed for shadow restoration in RGB and multispectral images are difficult to adapt to hyperspectral images,

as their characteristics pose specific challenges [135, 211]. For example, shadow removal methods may use ground images as training and test data, which would not work in the case of airborne images [202, 211]. Besides, simple scenes are often used as test data, where a single shadowed region exists in one test image [9]. This assumption often does not hold for airborne images containing more complicated scenarios. The proposed framework contributes to the open problems in the following aspects.

As a first aspect, some previous works assume diffuse solar irradiance to have zero [149] or constant values [153, 212] across all wavelengths. These assumptions simplify real scenarios and may introduce errors in modeling shadowed spectra. The proposed framework considers diffuse irradiance and multiple reflections of direct solar irradiance as the illumination sources in shadowed regions, following physical assumptions. Second, several previous studies develop shadow detection and restoration methods in two separate frameworks, indicating that accurate shadow detection results are required to achieve satisfying restoration [154, 213]. Our proposed framework computes shadow detection maps based on the residual analysis of pixel reconstruction through spectral unmixing, thus it does not require a shadow map as additional input. Third, a soft shadow detection yields a more realistic representation of shadows with respect to a binary shadow mas as, from a physical point of view, shadow boundaries are usually neither pure sunlit nor pure shadowed pixels. In addition, soft shadow masks allow some flexibility as they can be thresholded by an user to generate conservative or complete binary masks. Fourth, our framework does not require a large amount of training data, usually scarcely available and expensive to derive.

The proposed framework still contains several open problems. First, despite the correct classification results, we observe spectral distortions of shadowed pixels for some impervious surfaces, if slightly different materials are present in the scenes. An area in subset 2 (Fig. 5.13) shows an impervious surface shadowed by a building. The relative spectra appear distorted with respect to the neighbouring sunlit pixels belonging to the same material, as it is assumed that pixels on opposite edges of a shadow boundary usually exhibit similar reflectance spectra. Thus, we investigate the abundance maps of endmembers dominating the sunlit regions (in Fig. 5.13 (c)) and shadowed regions (in Fig. 5.13 (d)). In Fig. 5.13 (e), we show the reflectances of the endmembers corresponding to Fig. 5.13 (c) as a solid line and Fig. 5.13 (d) as a dashed line, respectively. The spectral angle between the two reflectances in Fig. 5.13 (e) is equal to 0.035, indicating that the related two materials are highly similar. Besides, the spectral angle assumes a value of 0.032 between two reflectances in sunlit and shadowed pixels marked with a "+" in the restored image (Fig. 5.13 (f)). This implies that, when the spectral angle between two spectra is small, the restored results may not distinguish the related materials.

Second, endmembers used in the framework do not include black objects, such as cars in subset 3, because spectra of black objects are similar to shadowed pixels. Thus, the proposed framework regards black objects as shadows, as their sunlit factor values are low (Fig. 5.10).

Third, although the sunlit factor values of water regions are higher with respect to shadowed pixels (Fig. 5.10 (d)), water can still be confused with shadows due do its low albedo. Thus, D_{l} and D_{s} for sunlit water pixels are comparable and considerably smaller than 0.1.

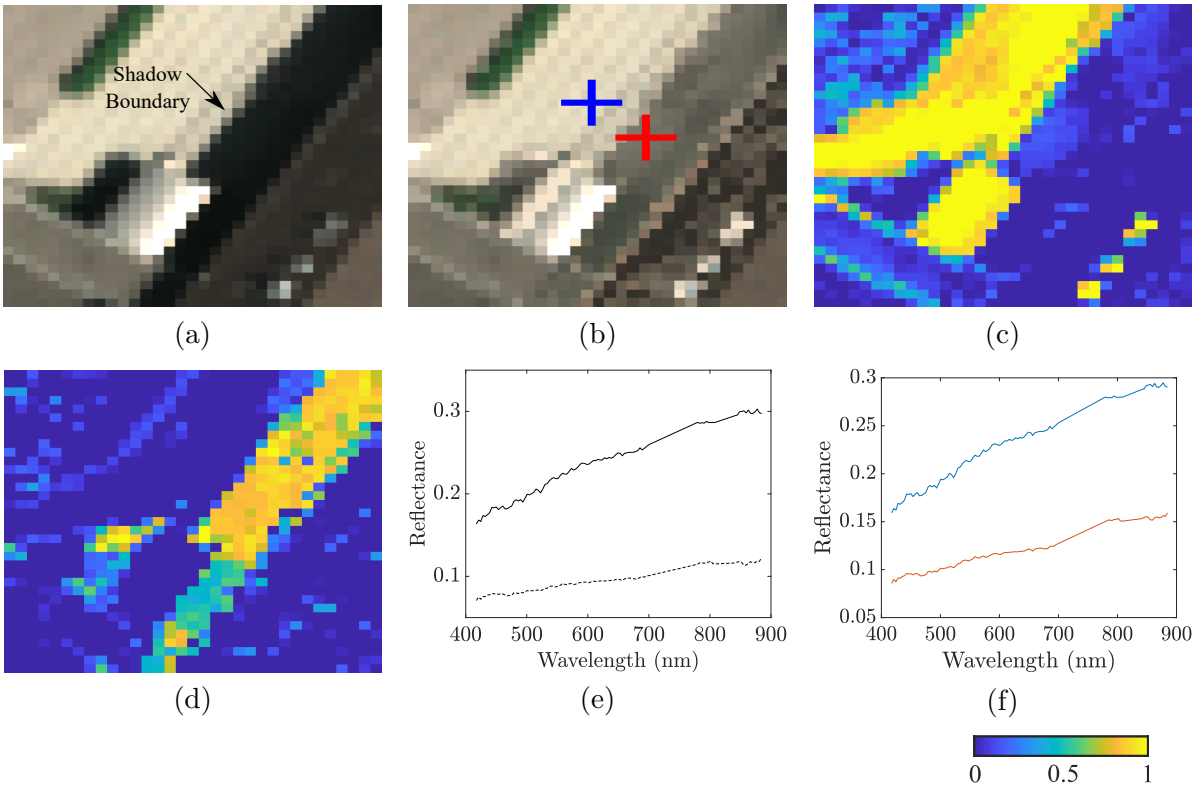


Figure 5.13: An example of spectral inconsistency in the neighborhood of a shadow boundary in a subarea of subset 2: (a) input image; (b) shadow-restored image; (c) abundance map for a material dominating the sunlit region; (d) abundance map for a material dominating the shadowed region; (e) endmembers corresponding to the abundance maps of (c) as a solid line and (d) as a dashed line; (f) reflectance of sunlit (blue) and shadowed (red) pixels in (b).

Fourth, local texture in restored shadowed regions can be lost (Fig. 5.9 (f)), or present a high level of noise (Fig. 5.9 (b)). These examples of information loss are partly due to the signal to noise ratio for shadowed pixels being lower with respect to sunlit pixels.

5.1.6 Summary

In this work, motivated by the fact that reflectance values for a given material should be independent from illumination conditions, we have proposed a novel framework for shadow detection and restoration of hyperspectral images based on nonlinear unmixing. The framework regards pure sunlit and shadowed spectra as sunlit and shadowed endmembers, respectively. Pure sunlit spectra are manually selected from the input images, while pure shadowed spectra are computed from sunlit spectra based on physical assumptions. Subsequently, the algorithm solves abundances related to sunlit and shadowed endmembers through a nonlinear mixing model. Then, we reconstruct restored images pixelwise using abundance maps and only the sunlit endmembers. As a byproduct, the proposed framework can generate sunlit factor maps that can locate sunlit pixels. Finally, sunlit pixels in the restored images are replaced by their original values. The proposed framework is validated on real airborne hyperspectral images using visual analysis and quantitative assessments. Compared with two

well-known mixing models, i.e. LMM and Fan, the proposed mixing model can reconstruct shadowed pixels with significantly lower errors. After restoration, shadowed regions become visually alike to adjacent sunlit regions, and exhibit similar reflectance values. In addition, classification results are visually more convincing and accuracies increase by more than 10 % for the investigated subsets after shadow restoration. The derived sunlit factor maps can produce soft shadow masks, representing natural transitions around shadow boundaries. We also demonstrate the possibility of detecting and including new materials in the input endmember library.

The work carried out so far raises open issues which are of interest for further investigation. Embedding spatial information may decrease the spectral distortion in shadow-removed pixels. In addition, black objects and water regions require further investigation. Future works could consider spectral bands that can increase the distinction between shadow and dark objects. Finally, the nonlinear mixing model in this work allows the interactions of up to two endmembers. Higher-order nonlinear models could be included to model more accurately the physical interactions in the scene.

5.2 Shadow-Aware Spectral Unmixing Method

In hyperspectral imagery, differences in ground surface structures cause a large variation in the optical scattering in sunlit and (partly) shadowed pixels. The complexity of the scene demands a general spectral mixing model that can adapt to the different scenarios of the ground surface. This section proposes a physics-based spectral mixing model, i.e., the Extended Shadow Multilinear Mixing (ESMLM) model that accounts for typical ground scenarios in the presence of shadows and nonlinear optical effects, by considering multiple illumination sources. Specifically, the diffuse solar illumination alters as the wavelength changes, requiring a wavelength-dependent modeling of shadows. Moreover, we allow different types of nonlinear interactions for different illumination conditions. The proposed model is described in a graph-based representation, which sums up all possible radiation paths initiated by the illumination sources. Physical assumptions are made to simplify the proposed model, resulting in material abundances and four physically interpretable parameters. Additionally, shadow-removed images can be reconstructed. The proposed model is compared with other state-of-the-art models using one synthetic dataset and two real datasets. Experimental results show that the ESMLM model performs robustly in various illumination conditions. In addition, the physically interpretable parameters contain valuable information on the scene structures and assist in performing shadow removal that outperforms other state-of-the-art works.

5.2.1 Introduction

Spectral unmixing, as an inverse process of spectral mixing, aims to quantitatively analyze material composition at the sub-pixel level. An introduction of spectral unmixing is detailed in Section 2.5. In summary, most existing spectral mixing models, i.e., LMM, Fan, GBM, MLM, simply ignored the shadow effect and thus typically output inaccurate estimation at shadowed pixels. Later, shadows have been treated as a wavelength-independent scaling effect in the Shadow Linear Mixing Model (SLMM) and SMLM, by including another paramter Q, representing the pixelwise fractional value of cast shadow [151]. However, shadow not only scales a spectrum, but also causes wavelength dependent distortions [103]. Some works have paid attention to the spectral distortions caused by shadow, based on nonlinear modeling and allowing multiple illuminations in a scene (see Section 3.4.2.3).

Although it has been shown that embedding multiple illumination sources in a model improves unmixing performances in (partly) shadowed pixels, nonlinearity has been modeled similarly for all pixels, regardless of the illumination conditions [198]. However, the nonlinearity in shadowed areas can behave quite differently from sunlit areas, due to the light attenuation caused by occluding objects.

In this work, we present a novel nonlinear mixing model, that is an extension of the SMLM model [151], and that overcomes the mentioned problems in the following aspects:

- We consider two illumination sources, i.e., direct and diffuse solar illuminations. Following physical assumptions, we allow variable illumination conditions over the scene, where sunlit regions receive direct as well as diffuse solar illuminations, while shadowed

regions receive diffuse solar illumination and possibly reflected direct solar illumination. Specifically, one pixel can be composed of shadowed areas along with fully sunlit areas, and can therefore be treated as a partly shadowed pixel, resulting in a better representation of shadow boundaries.

- Our proposed model allows two different types of nonlinear interactions. Besides the nonlinear optical interactions caused by the direct incoming light from both illumination sources, a pixel can receive secondary reflections from its neighboring pixels. In this way, the proposed model can produce reconstructed pixels with spectral values larger than those obtained through LMM, without losing the physical meaning of the parameter P, while energy conservation still holds.
- We describe our model using a graphical representation with multiple illumination sources. The mixing result is computed as the sum of all light contributions, weighted by their probabilities. In addition to the abundances values, our model generates four pixelwise physically interpretable parameters: Q (spatial fraction of shadow in a pixel), F (sky view factor, which denotes the fraction of the sky hemisphere that is visible from the ground surface [214]), P (the probability of higher-order interactions of the incoming light ray), and K (a strength factor of neighbor interactions, denoting the fraction of the scattered light from the neighborhood that is received by the pixel).

The remainder of this section is organized as follows. In Section 5.2.2, we describe the impact of shadow on the observed reflectance, based on radiative transfer and atmospheric correction. Section 5.2.3 introduces several physics-based mixing models using graphical representations, while Section 5.2.4 describes the proposed mixing model. Section 5.2.5 introduces the experimental setup, including three datasets for the evaluation of the spectral mixing models, the unmixing procedure, and the experimental design. Section 5.2.6 demonstrates the experimental results quantitatively using a synthetic dataset, and section 5.2.7 and 5.2.8 present both quantitative and qualitative results using real datasets. Finally, we conclude our work and give future prospects in Section 5.2.9.

5.2.2 The Impact of Shadow on Reflectance

This section is built upon the theory of radiation propagation and atmospheric correction. Interested readers are referred to Sections 2.3 and 2.4 to better understand this section.

Let us assume that the ground targets are located on a flat terrain, and behave as a Lambertian surface. Then, at each wavelength λ, the at-sensor radiance $\mathcal{L}(\lambda)$ for a ground pixel with reflectance $r(\lambda)$ can be written as:

$$\mathcal{L}(\lambda) = \mathcal{L}_{\mathrm{p}}(\lambda) + \frac{\delta\tau_{\mathrm{dir}}(\lambda)\mathcal{E}_{\mathrm{l}}(\lambda)r(\lambda)}{\pi} + \frac{\tau_{\mathrm{diff}}(\lambda)\mathcal{E}_{\mathrm{s}}(\lambda)r(\lambda)}{\pi} \tag{5.12}$$

where $\mathcal{L}_{\mathrm{p}}(\lambda)$ is the path radiance, $\mathcal{E}_{\mathrm{l}}(\lambda)$ the direct solar irradiance and $\mathcal{E}_{\mathrm{s}}(\lambda)$ the diffuse solar irradiance on the ground target. The transmittances of the direct and diffuse solar radiation are $\tau_{\mathrm{dir}}(\lambda)$ and $\tau_{\mathrm{diff}}(\lambda)$, respectively, while δ is a binary value indicating if the ground surface receives any direct solar irradiance. In traditional atmospheric correction algorithms, δ is

set to 1 , as the ground surface is usually assumed to be horizontal and unobstructed, i.e., it "sees" the entire hemisphere above. Thus, given the at-sensor radiance $\mathcal{L}$, atmospheric correction is applied and $\hat{r}(\lambda)$ is derived by inverting Eq. (5.12):

$$\hat{r}(\lambda) = \frac{\pi(\mathcal{L}(\lambda) - \mathcal{L}_{\mathrm{p}}(\lambda))}{\tau_{\mathrm{dir}}(\lambda)\mathcal{E}_{\mathrm{l}}(\lambda) + \tau_{\mathrm{diff}}(\lambda)\mathcal{E}_{\mathrm{s}}(\lambda)} = r(\lambda) \tag{5.13}$$

However, a pixel may not or only partly receive direct sunlight, due to occlusion by ground objects. The at-sensor radiance $\mathcal{L}_{\mathrm{s}}(\lambda)$ of a fully shadowed pixel, containing one material with reflectance $r(\lambda)$, is given by Eq. (5.12) with $\delta = 0$:

$$\mathcal{L}_{\mathrm{s}}(\lambda) = \mathcal{L}_{\mathrm{p}}(\lambda) + \frac{\tau_{\mathrm{diff}}(\lambda)\mathcal{E}_{\mathrm{s}}(\lambda)r(\lambda)}{\pi} \tag{5.14}$$

When standard atmospheric correction (Eq. (5.13)) is applied on such a pixel, then the reflectance $\hat{r}_{\mathrm{s}}(\lambda)$ is computed as:

$$\begin{aligned} \hat{r}_{\mathrm{s}}(\lambda) &= \frac{\pi(\mathcal{L}_{\mathrm{s}}(\lambda) - \mathcal{L}_{\mathrm{p}}(\lambda))}{\tau_{\mathrm{dir}}(\lambda)\mathcal{E}_{\mathrm{l}}(\lambda) + \tau_{\mathrm{diff}}(\lambda)\mathcal{E}_{\mathrm{s}}(\lambda)} \\ &= \frac{\tau_{\mathrm{diff}}(\lambda)\mathcal{E}_{\mathrm{s}}(\lambda)r(\lambda)}{\tau_{\mathrm{dir}}(\lambda)\mathcal{E}_{\mathrm{l}}(\lambda) + \tau_{\mathrm{diff}}(\lambda)\mathcal{E}_{\mathrm{s}}(\lambda)} \end{aligned} \tag{5.15}$$

in which Eq. (5.14) is substituted.

As reflectance represents an intrinsic property of a material, and should not change according to illumination conditions, we expect that $\hat{r}_{\mathrm{s}}(\lambda) = \hat{r}(\lambda) = r(\lambda)$. However, during atmospheric correction, the direct solar irradiance is incorrectly assumed to be an illumination source in shadowed regions. As a consequence, the computed reflectance values in these areas are much smaller than their correct values, and a wavelength-dependent deviation exists between $\hat{r}_{\mathrm{s}}(\lambda)$ and $r(\lambda)$.

If no occlusion occurs on a ground pixel, the diffuse radiation comes from all directions of the sky dome. Otherwise, the diffuse irradiance decreases by the sky view factor $F \in [0, 1]$, representing the fraction of sky that a ground pixel can "see". Following previous works, we model the decrease of the diffuse-to-direct solar irradiance as the wavelength goes up through a power function, which describes stronger atmospheric scattering at short wavelengths [55, 198, 215, 216]:

$$\frac{\tau_{\mathrm{diff}}(\lambda)\mathcal{E}_{\mathrm{s}}(\lambda)}{\tau_{\mathrm{dir}}(\lambda)\mathcal{E}_{\mathrm{l}}(\lambda)} = F(k_1\lambda^{-k_2} + k_3) \tag{5.16}$$

with k_1, k_2, $k_3 > 0$.

Combining equations (5.15) and (5.16), we derived an expression for the reflectance of a ground material in fully shadowed regions with respect to the reflectance of the same material exposed to direct sunlight as:

$$\hat{r}_{\mathrm{s}}(\lambda) = \frac{F(k_1\lambda^{-k_2} + k_3)}{1 + F(k_1\lambda^{-k_2} + k_3)}\hat{r}(\lambda) \tag{5.17}$$

The derived relationship between the reflectances of the same material in different illumination conditions is only valid in the following simplified scenario. First, each ground pixel contains only one material. Second, illumination interacts only once with a ground pixel before being scattered back to the sensor. Third, a ground pixel can only be either fully sunlit or fully shadowed. In reality, a ground pixel can be composed of multiple materials, and illumination sources can interact multiple times with ground materials. In addition, pixels may only be partly shadowed. To include these situations, a shadow-aware nonlinear spectral mixing model is required.

5.2.3 Mixing Models and their Graph Representations

The mixing models describe the optical interactions to a certain degree of complexity in the imaging chain [62, 66]. Following the work in [150, 151], we introduce state-of-the-art spectral mixing models based on a ray-based approximation of light and a graph-based representation of the optical interactions. *Please note that this section is not a repetition of Section 2.5. Instead, it presents mixing models from a new perspective and serves as the basis for the proposed model.*

The entire process of the incoming light from the illumination sources undergoing optical interactions, and each sensitive element of the spectrometer recording the back-scattered light from the corresponding ground pixel can be described as a discrete-time stochastic process [151]. A light path is defined by the random variable $\{\Xi_n\}_{n\geq 0}$ with $\forall n\ \Xi_n \in S$, and the discrete set S contains all possible interactions that a light ray can undergo before reaching the observer. In passive optical imaging, the light path always starts from the illumination source $\Xi_0 = s_0$. States in which the light ray interacts with a ground material are indicated as $\{\Xi_i = s_i\}_{i=1}^{L}$. Since we consider only the scattered light eventually received by the observer, a light path ends with the observer state $\Xi_{L+1} = o$. $L \in [1, \infty]$ is the path length, indicating the number of optical interactions that a light ray underwent before being scattered back to the observer.

The probability of observing a certain *path* of length L is given by:

$$P(path) = P(\Xi_0 = s_0, \Xi_1 = s_1, \cdots, \Xi_L = s_L, \Xi_{L+1} = o) \tag{5.18}$$

It is assumed that this stochastic process follows the Markov property:

$$P(\Xi_{n+1}|\Xi_0, \cdots \Xi_n) = P(\Xi_{n+1}|\Xi_n) \tag{5.19}$$

At each state, the optical properties of the light ray will be altered. This alteration describes a relative change in the spectrum of the light ray, according to the reflectance of the object associated with that state. If $T(s_i)$ is the operator that acts on the light ray in state s_i, the total effect of $path = \{\ \Xi_0 = s_0, \Xi_1 = s_1, \cdots, \Xi_L = s_L, \Xi_{L+1} = o\}$ on a light ray is given by $\prod_{i=0}^{L} T(s_i)$. The operator $T(s_i)$ is associated with the state s_i.

- For states $\{\Xi_i = s_i\}_{i=1}^{L}$ representing ground materials, $T(s_i) = \boldsymbol{e}_i$.
- For the state $\{\Xi_{L+1} = o\}$ presenting the observer, $T(o) = \mathbf{1}$.

Table 5.3: Probabilities of light paths and their spectral contributions in different mixing models

Path	Probability	Contribution	Parameters	Model
$s_0 - \boldsymbol{e}_i - o$	a_i	$\boldsymbol{e}_i$	/	LMM [76]
$s_0 - \boldsymbol{e}_i - o$	a_i	$\boldsymbol{e}_i$	$\gamma_{i,j} = 0, if\ i \geq j$ $\gamma_{i,j} = 1, if\ i < j$	Fan [87]
$s_0 - \boldsymbol{e}_i - \boldsymbol{e}_j - o$	$\gamma_{i,j} a_i a_j$	$\boldsymbol{e}_i \odot \boldsymbol{e}_j$	$\gamma_{i,j} = b$	PPNM [88]
			$\gamma_{i,j} = 0, if\ i \geq j$ $\gamma_{i,j} \in [0,1], if\ i < j$	GBM [89]
$s_0 - \boldsymbol{e}_i - o$	$a_i(1-P)$	$\boldsymbol{e}_i$		
$s_0 - \boldsymbol{e}_i - \boldsymbol{e}_j - o$	$a_i a_j (1-P)P$	$\boldsymbol{e}_i \odot \boldsymbol{e}_j$		
$s_0 - \boldsymbol{e}_i - \boldsymbol{e}_j - \boldsymbol{e}_k - o$	$a_i a_j a_k (1-P)PP$	$\boldsymbol{e}_i \odot \boldsymbol{e}_j \odot \boldsymbol{e}_k$	$P \in [0,1]$	MLM [150]
$\cdots$	$\cdots$	$\cdots$		
$s_0 - \boldsymbol{e}_i - \boldsymbol{e}_j - \boldsymbol{e}_k \cdots \boldsymbol{e}_R - o$	$a_i a_j a_k \cdots a_R (1-P)P^{R-1}$	$\boldsymbol{e}_i \cdot \boldsymbol{e}_j \odot \boldsymbol{e}_k \cdots \boldsymbol{e}_R$		
$s_0 - \boldsymbol{e}_i - o$	$(1-Q)a_i$	$\boldsymbol{e}_i$	/	SLMM [151]
$s_0 - \boldsymbol{e}_i - o$	$(1-Q)a_i(1-P)$	$\boldsymbol{e}_i$		
$s_0 - \boldsymbol{e}_i - \boldsymbol{e}_j - o$	$a_i a_j (1-P)P$	$\boldsymbol{e}_i \odot \boldsymbol{e}_j$		
$s_0 - \boldsymbol{e}_i - \boldsymbol{e}_j - \boldsymbol{e}_k - o$	$a_i a_j a_k (1-P)PP$	$\boldsymbol{e}_i \odot \boldsymbol{e}_j \odot \boldsymbol{e}_k$	$P, Q \in [0,1]$	SMLM [151]
$\cdots$	$\cdots$	$\cdots$		
$s_0 - \boldsymbol{e}_i - \boldsymbol{e}_j - \boldsymbol{e}_k \cdots \boldsymbol{e}_R - o$	$a_i a_j a_k \cdots a_R (1-P)P^{R-1}$	$\boldsymbol{e}_i \odot \boldsymbol{e}_j \odot \boldsymbol{e}_k \cdots \boldsymbol{e}_R$		
$s_{0_g} - \boldsymbol{e}_i - o$	$(1-Q)a_i$	$\boldsymbol{e}_i$		
$s_{0_g} - \boldsymbol{e}_i - \boldsymbol{e}_j - o$	$a_i a_j$	$\boldsymbol{e}_i \odot \boldsymbol{e}_j$	/	Fansky [198]
$s_{0_{\text{diff}}} - \boldsymbol{e}_i - o$	Qa_i	$\boldsymbol{e}'_i$		

- For the state $\{\Xi_0 = s_0\}$ representing the illumination source(s), $T(s_0)$ is a constant vector and corresponds to the illumination source.

Thus, an observed pixel $\boldsymbol{x}$ is described as the weighted average over all possible paths:

$$\boldsymbol{x} = \sum_{L=1}^{\infty} \Big(\sum_{s_0 \in S} \sum_{s_1 \in S} \cdots \sum_{s_L \in S} \Big) P(path) \prod_{k=0}^{L} T(s_k) \tag{5.20}$$

Table 5.3 shows a summary of light paths, their corresponding probabilities and spectral contributions for different mixing models. In the next sections, we will describe in more detail these models and their graph representations.

5.2.3.1 Linear Mixing Model (LMM)

LMM assumes that the incoming light interacts only once with a set of endmembers before being scattered back to the sensor ($L = 1$). The probability of an incoming light ray from the illumination source s that interacts with the ground surface with endmember $\boldsymbol{e}_l, (l = 1, \cdots, p)$ and is scattered back to the observer o is proportional to the abundance $a_l, (l = 1, \cdots, p)$. Thus: $P(path) = P(\Xi_0 = s_0, \Xi_1 = s_1, \Xi_2 = o) = a_l$, and $\prod_{k=0}^{L} T(s_k) = T(s_0)T(s_1)$, with $T(s_0) = \mathbf{1}$ and $T(s_1) = \boldsymbol{e}_l$. According to Eq. (5.20), the LMM is written as:

$$\boldsymbol{x} = \sum_{s_0 \in S} \sum_{s_1 \in S} a_l T(s_0) T(s_1) = \sum_{i=1}^{p} a_i \boldsymbol{e}_i \tag{5.21}$$

where $\sum_{i=1}^{p} a_i = 1$ and $\forall i\colon\ a_i \geq 0$.

5.2.3.2 Bilinear Mixing Models

In bilinear models, a light ray from the illumination source can either interact with an endmember once before being scattered back the sensor, i.e., $L = 1$, or it can have multiple interactions with endmembers up to the second order, i.e., $L = 2$. Thus, we have two possible light paths, with probabilities: $P(path) = P(\Xi_0 = s_0, \Xi_1 = s_1, \Xi_2 = o) = a_l$ and $P(path) = P(\Xi_0 = s_0, \Xi_1 = s_1, \Xi_2 = s_2, \Xi_3 = o) = \gamma_{m,n} a_m a_n$. Different bilinear models can be derived by constraining the free parameter $\gamma_{m,n}$ (Table 5.3). Then:

$$\begin{aligned} \boldsymbol{x} &= \sum_{s_0 \in S} \sum_{s_1 \in S} a_l T(s_0) T(s_1) + \sum_{s_0 \in S} \sum_{s_1 \in S} \sum_{s_2 \in S} \gamma_{m,n} a_m a_n T(s_0) T(s_1) T(s_2) \\ &= \sum_{i=1}^{p} a_i \boldsymbol{e}_i + \sum_i \sum_j \gamma_{i,j} a_i a_j \boldsymbol{e}_i \odot \boldsymbol{e}_j \end{aligned} \tag{5.22}$$

5.2.3.3 Multilinear Mixing (MLM) Model

Recently, authors in [150] extended bilinear mixing models to the MLM model that regards all orders of optical interactions. Similar to the case of linear and bilinear models, this assumes that a light ray incoming from the illumination source will interact with at least one material. Besides, the MLM model introduces a new parameter P: after each interaction with a material, the light ray will have a probability P of undergoing further interactions and a probability $(1 - P)$ of escaping the scene and reaching the observer. Following these assumptions, a light ray from the illumination source can interact with ground objects up to an infinite amount of times before being scattered back to the sensor, i.e., $L \in [1, \infty]$. Given a light path $path = \{\Xi_0 = s_0, \Xi_1 = s_1, \Xi_2 = s_2, \cdots, \Xi_L = s_L, \Xi_{L+1} = o\}$, its probability is given by: $P(path) = (1 - P) P^{L-1} a_{i_1} a_{i_2} \cdots a_{i_L}$. The spectral contribution of this path is: $\prod_{k=0}^{L} \boldsymbol{e}_{i_k}$ with $\boldsymbol{e}_{i_k}$ representing the endmember of the material that the ray interacts with the k_{th} time, and a_{i_k} its abundance. Thus:

$$\begin{aligned} \boldsymbol{x} &= \sum_{L=1}^{\infty} \left(\sum_{i_1=1}^{p} \cdots \sum_{i_L=1}^{p} \right) (1 - P) P^{L-1} \prod_{k=1}^{L} (a_{i_k} \boldsymbol{e}_{i_k}) \\ &= (1 - P) \sum_{i=1}^{p} a_i \boldsymbol{e}_i + (1 - P) P \sum_{i=1}^{p} \sum_{j=1}^{p} a_i a_j \boldsymbol{e}_i \odot \boldsymbol{e}_j + \cdots \\ &= \frac{(1 - P) \sum_{i=1}^{p} a_i \boldsymbol{e}_i}{1 - P \sum_{i=1}^{p} a_i \boldsymbol{e}_i} \end{aligned} \tag{5.23}$$

5.2.3.4 Shadow Linear Mixing (SLMM) Model

The SLMM model extends the endmember library with a "zero-reflectance" spectrum. Numerically, this technique is equivalent to including a parameter $Q \in [0, 1]$, which represents the spatial fraction of shadow in a pixel [150]. Values of $Q = 0$ and $Q = 1$ indicate a fully sunlit and fully shadowed pixel, respectively, while $Q \in (0, 1)$ describes a partly shadowed pixel. This model can estimate abundances under the shadow by setting $Q = 0$ during pixel

reconstruction. Similar to LMM, the light path of the SLMM is $P(s_0, \boldsymbol{e_l}, o) = (1-Q)a_l$ with the spectral contribution of $\boldsymbol{e}_l$. Thus:

$$\boldsymbol{x} = \sum_{s_0 \in S} \sum_{s_1 \in S} (1-Q) a_l T(s_0) T(s_1) = \sum_{i=1}^{p} (1-Q) a_i \boldsymbol{e}_i \tag{5.24}$$

where $\sum_{i=1}^{p} a_i = 1$ and $\forall i\colon\ a_i \geq 0$.

5.2.3.5 Shadow Multilinear Mixing (SMLM) model

The SMLM model [151] extends the MLM model from [150] in order to deal with shadows. It is assumed that shadowed regions do not receive direct sunlight, but only multiple reflections of direct sunlight. The SMLM model uses the parameter Q to represent the shadow fraction within a pixel. Thus, the light paths and probabilities of the SMLM model are the same as those of the MLM model except for the first order, which is rescaled with $(1-Q)$, hereby subtracting the shadow fraction from the direct sunlight term in a spectrum. Thus:

$$\begin{aligned}
\boldsymbol{x} =& \sum_{L=1}^{1} \left(\sum_{i_1=1}^{p} \cdots \sum_{i_L=1}^{p} \right) (1-Q)(1-P) \prod_{k=1}^{L} (a_{i_k} \boldsymbol{e}_{i_k}) + \\
& \sum_{L=2}^{\infty} \left(\sum_{i_2=1}^{p} \cdots \sum_{i_L=1}^{p} \right) (1-P) P^{L-1} \prod_{k=2}^{L} (a_{i_k} \boldsymbol{e}_{i_k}) \\
=& (1-Q)(1-P) \sum_{i=1}^{p} a_i \boldsymbol{e}_i + (1-P) P \sum_{i=1}^{p} \sum_{j=1}^{p} a_i a_j \boldsymbol{e}_i \odot \boldsymbol{e}_j + \cdots \\
=& \frac{(1-P) \sum_{i=1}^{p} a_i \boldsymbol{e}_i}{1 - P \sum_{i=1}^{p} a_i \boldsymbol{e}_i} - Q(1-P) \sum_{i=1}^{p} a_i \boldsymbol{e}_i
\end{aligned} \tag{5.25}$$

5.2.3.6 Mixing Model with Multiple Light Sources

The work in [198] allows different illumination conditions in sunlit and shadowed regions and regards the shadow effect in a wavelength-dependent manner. In this work, sunlit areas receive the entire solar radiation, i.e, direct as well as diffuse solar radiation, while the shadowed regions only receive diffuse radiation. Hence, in the graph representation, two illumination sources, each with its own state, are considered: global radiation s_{0_g} and diffuse radiation s_{0_diff}. Specifically, s_{0_g} is equivalent to s_0, where only one illumination source is considered. The model accounts for the light paths from the global radiation up to the second order, leading to two possible light paths: $P(path_\mathrm{g}) = P(\Xi_0 = s_{0_\mathrm{g}}, \Xi_1 = s_1, \Xi_2 = o) = (1-Q)a_l$ for $L = 1$ and $P(path_\mathrm{g}) = P(\Xi_0 = s_{0_\mathrm{g}}, \Xi_1 = s_1, \Xi_2 = s_2, \Xi_3 = o) = a_m a_n$ for $L = 2$. The spectral contribution of two light paths is $T(s_{0_\mathrm{g}}) \prod_{k=1}^{L} \boldsymbol{e}_k$, with $L = 1$ and $L = 2$, respectively. Moreover, the model assumes that diffuse solar radiation interacts not more than once with ground materials, resulting in the light path $P(path_\mathrm{diff}) = P(\Xi_0 = s_{0_\mathrm{diff}}, \Xi_1 = s_1, \Xi_2 = o) = Qa_l$, where $L = 1$ and Q is the spatial fraction of shadow in a pixel. The spectral contribution is $\prod_{k=0}^{L} T(s_{0_\mathrm{diff}}) \prod_{k=1}^{L} \boldsymbol{e}_l$, where $L = 1$. Since the light paths at the second-order $(L = 2)$ is equivalent to the Fan model, we refer to this model as the Fansky model. From

Eq. (5.16) follows that $T(s_{0_{\text{diff}}}) = \frac{\boldsymbol{\tau}_{\text{diff}} \odot \boldsymbol{\mathcal{L}}_{\text{s}}}{\boldsymbol{\tau}_{\text{dir}} \odot \boldsymbol{\mathcal{L}}_{\text{l}} + \boldsymbol{\tau}_{\text{diff}} \odot \boldsymbol{\mathcal{L}}_{\text{s}}}$. Then, the mixing model can be written as the contribution of all possible light paths, initiated from two illumination sources:

$$\begin{aligned} \boldsymbol{x} = &\sum_{L=1}^{2} \Big(\sum_{s_{0_{\text{g}}} \in S} \cdots \sum_{s_L \in S} \Big) P(path_{\text{g}}) T(s_{0_{\text{g}}}) \prod_{k=1}^{L} T(s_k) + \\ &\sum_{L=1}^{1} \Big(\sum_{s_{0_{\text{diff}}} \in S} \cdots \sum_{s_L \in S} \Big) P(path_{\text{diff}}) T(s_{0_{\text{diff}}}) \prod_{k=1}^{L} T(s_k) \\ = &(1-Q) \sum_{i=1}^{p} a_i \boldsymbol{e}_i + \sum_{i=1}^{p} \sum_{j=i}^{p} a_i a_j \boldsymbol{e}_i \odot \boldsymbol{e}_j + Q \sum_{i=1}^{p} a_i \boldsymbol{e}'_i \end{aligned} \tag{5.26}$$

where $\boldsymbol{e}'_i = \frac{\boldsymbol{\tau}_{\text{diff}} \odot \boldsymbol{\mathcal{L}}_{\text{s}}}{\boldsymbol{\tau}_{\text{dir}} \odot \boldsymbol{\mathcal{L}}_{\text{l}} + \boldsymbol{\tau}_{\text{diff}} \odot \boldsymbol{\mathcal{L}}_{\text{s}}}$ with $\boldsymbol{\mathcal{L}}_{\text{l}}$ and $\boldsymbol{\mathcal{L}}_{\text{s}}$ are the vector forms of $\mathcal{L}_{\text{l}}(\lambda)$ and $\mathcal{L}_{\text{s}}(\lambda)$, respectively.

5.2.4 Proposed Method

We propose an extended SMLM (ESMLM) model by allowing multiple illumination sources, i.e., direct and diffuse solar radiation. Moreover, apart from the optical interactions occurring in a ground pixel determined by its IFOV, a pixel can also receive additional illumination from its neighboring pixels through secondary reflections. Fig. 5.14 depicts the occurring optical interactions for five different scenarios that are considered in this model. As the path radiance is assumed to be removed by atmospheric correction [63], the model describes three types of light paths, corresponding to three illumination sources: global solar illumination $s_{0_{\text{g}}}$, diffuse solar illumination $s_{0_{\text{diff}}}$ and neighboring illumination $s_{0_{\text{N}}}$. The light paths together with their probabilities and spectral contributions for the three illumination sources in the proposed model are presented in Table 5.4, followed by the physical assumptions and a detailed explanation for each illumination source in the remaining part of this section. In brief, the mixing model is computed as the sum of the contributions from all illumination sources in Eqs. (5.27) and (5.28), and contains four physically explainable parameters:

- P: the probability that a light ray undergoes additional interactions with endmembers,
- Q: the spatial fraction of shadow,
- F: the sky view factor,
- K: a strength factor of neighbor interactions, denoting the fraction of the scattered light from the neighborhood that is received by the pixel.

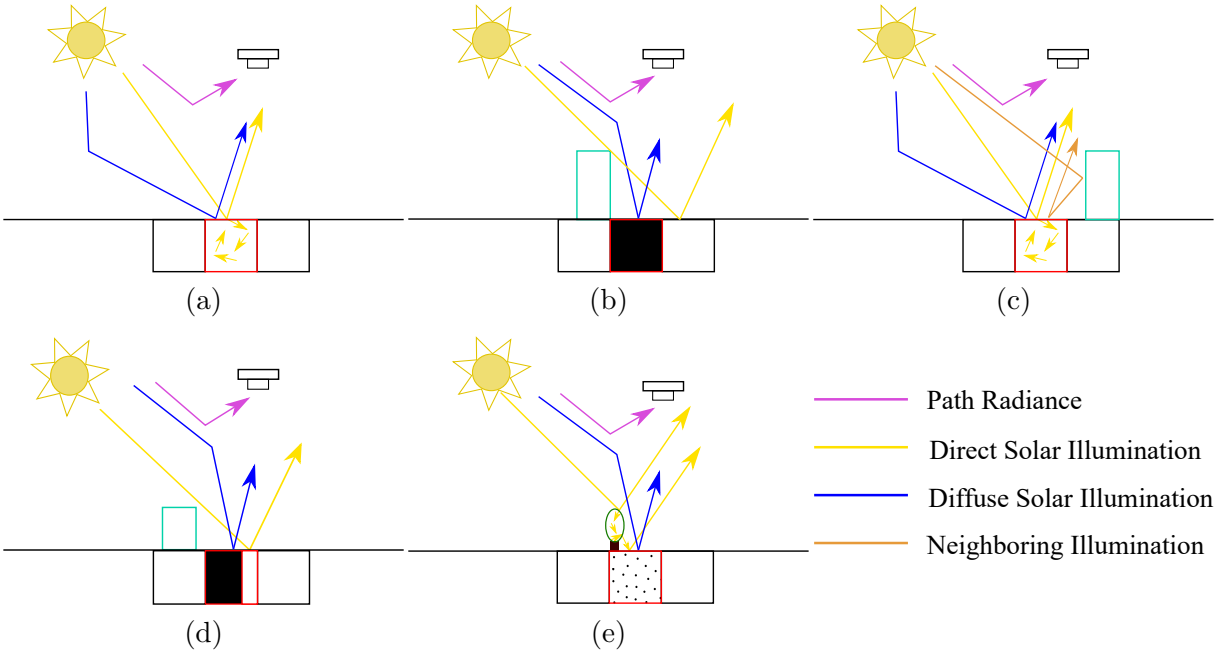


Figure 5.14: Solar radiation paths when a pixel is: (a) exposed to direct sunlight; (b) fully shadowed; (c) receiving secondary illumination from neighbors; (d) partly shadowed with sunlit and shadowed regions spatially separated in a pixel; (e) partly shadowed with sunlit and shadowed regions not spatially separable in a pixel.

Table 5.4: Probabilities of light paths and their spectral contributions in the proposed model

Path Category	Path	Probability	Contribution	parameters
$path_{\mathrm{g}}$	$s_{0_{\mathrm{g}}} - \boldsymbol{e}_i - o$	$(1-Q)(1-P)a_i$	$\boldsymbol{e}_i$	
	$s_{0_{\mathrm{g}}} - \boldsymbol{e}_i - \boldsymbol{e}_j - o$	Pa_ia_j	$\boldsymbol{e}_i \odot \boldsymbol{e}_j$	
$path_{\mathrm{diff}}$	$s_{0_{\mathrm{diff}}} - \boldsymbol{e}_i - o$	Qa_i	$T(s_{0_{\mathrm{diff}}}) \cdot \boldsymbol{e}_i$	$P, Q, K \in [0, 1]$
$path_{\mathrm{N}}$	$s_{0_{\mathrm{N}}} - \boldsymbol{e}_i - o$	$(1-Q)(1-P)a_i$	$K\boldsymbol{\chi} \odot \boldsymbol{e}_i$	

$$
\begin{aligned}
\boldsymbol{x} = &\sum_{L=1}^{2}\Big(\sum_{s_{0_{\mathrm{g}}}\in S}\cdots\sum_{s_L\in S}\Big)P(path_{\mathrm{g}})T(s_{0_{\mathrm{g}}})\prod_{k=1}^{L}T(s_k)+\\
&+\sum_{L=1}^{1}\Big(\sum_{s_{0_{\mathrm{N}}}\in S}\cdots\sum_{s_L\in S}\Big)P(path_{\mathrm{N}})T(s_{0_{\mathrm{N}}})\prod_{k=1}^{L}T(s_k)+\\
&\sum_{L=1}^{1}\Big(\sum_{s_{0_{\mathrm{diff}}}\in S}\cdots\sum_{s_L\in S}\Big)P(path_{\mathrm{diff}})T(s_{0_{\mathrm{diff}}})\prod_{k=1}^{L}T(s_k)\\
=&(1-Q)(1-P)\sum_{i=1}^{p}a_i\boldsymbol{e}_i + P\sum_{i=1}^{p}\sum_{j=1}^{p}a_ia_j\boldsymbol{e}_i\odot\boldsymbol{e}_j+\\
&(1-Q)(1-P)K\sum_{i=1}^{p}a_i\boldsymbol{e}_i\odot\boldsymbol{\chi} + QT(s_{0_{\mathrm{diff}}})\sum_{i=1}^{p}a_i\boldsymbol{e}_i
\end{aligned}
\tag{5.27}
$$

In sunlit regions, the global solar illumination $s_{0_{\mathrm{g}}}$ is the main illumination source, and the proposed model retains most of the assumptions for $s_{0_{\mathrm{g}}}$ made by the SMLM model. One difference is that we constrain the parameter P within $[0, 1]$, in order to preserve its physical

interpretation. Moreover, with the aim of keeping all types of nonlinear interactions up to the same order, the proposed model limits nonlinear interactions of an incoming light ray up to the second order. In the specific, the following assumptions for s_{0_g} are made:

- An incoming light ray from the global illumination source will interact with at least one material in a pixel. After each interaction with a material, the ray will have probabilities P and $(1-P)$ of undergoing further interactions within the current pixel or escaping the current pixel, respectively.
- As the shadowed part of a pixel does not have a direct line of sight to the sun, the probability that the reflected light escapes a partly shadowed pixel after the first interaction is re-scaled with $(1-Q)$, with $Q \in [0,1]$ the fractional value of the shadow in the pixel. Thus, after the first interaction with a material, the light ray will have a probability $(1-Q)(1-P)$ of escaping the current pixel. On the other hand, the shadowed part of a pixel can receive reflected light from s_{0_g}, thus the probability of a secondary reflection remains P without re-scaling with $(1-Q)$.

These assumptions for the global illumination source lead to two possible light paths: $P(path_\mathrm{g}) = P(\Xi_0 = s_{0_\mathrm{g}}, \Xi_1 = s_1, \Xi_2 = o) = (1-Q)(1-P)a_l, (l = 1, \cdots, p)$ for $L = 1$ and $P(path_\mathrm{g}) = P(\Xi_0 = s_{0_\mathrm{g}}, \Xi_1 = s_1, \Xi_2 = s_2, X_2 = o) = Pa_m a_n, (m, n = 1, \cdots, p))$ for $L = 2$. The spectral contribution of these two light paths is $T(s_{0,\mathrm{g}}) \prod_{k=1}^{L} \boldsymbol{e}_k$, with $L = 1$ and $L = 2$, respectively.

In addition to receiving global illumination, the target pixel receives secondary reflections from its neighborhood. The neighbor illumination source s_{0_N} follows the following assumptions:

- By keeping all types of nonlinear effects up to the second order, only neighboring regions having a direct view of the sun can contribute to the target pixel. Thus, the neighboring effect corresponds to the reflected light of a pixel after receiving the global illumination s_{0_g}.
- Following the Lambertian law, by escaping the pixel, the scattered light ray from s_{0_g} is reflected in all directions, including towards the sensor and neighboring pixels, with equal probability of $(1-Q)(1-P)$.
- By assuming a homogeneous local neighborhood, the probability that a pixel scatters light to its neighboring pixels is equal to the probability that the neighboring pixels scatter light to the pixel, and is given by $(1-Q)(1-P)$.
- We define an additional parameter $K \in [0,1]$, i.e., a strength factor of neighbor interactions, denoting the fraction of the scattered light from the neighborhood that is received by the pixel.

Thus, the probability of the light paths, corresponding to the neighbor illumination source are: $P(path_\mathrm{N}) = P(\Xi_0 = s_{0_\mathrm{N}}, \Xi_1 = s_1, \Xi_2 = o) = (1-Q)(1-P)a_l, (l = 1, \cdots, p)$ for $L = 1$. The spectral contribution is $T(s_{0_\mathrm{N}}) \prod_{k=1}^{L} \boldsymbol{e}_k$ where $T(s_{0_\mathrm{N}}) = \boldsymbol{\chi}$. The neighborhood is defined

by a radius R. The neighborhood spectrum $\boldsymbol{\chi}(i,j)$ of target pixel $\boldsymbol{x}(i,j)$ is computed as the average spectrum of its neighboring pixels, weighted by their inverse distance to the target:

$$\boldsymbol{\chi}(i,j) = \frac{\sum_{s=-R}^{R}\sum_{t=-R}^{R}\boldsymbol{\chi}(i+s,j+t)\omega(s,t)}{\sum_{s=-R}^{R}\sum_{t=-R}^{R}\omega(s,t)} \tag{5.28}$$

where $\omega(s,t) = \frac{\delta}{D((i,j),(i+s,j+t))}$ and D denotes the Euclidean distance between two pixels in the spatial domain. Since only neighboring regions having a direct view of the sun can contribute to the target pixel, we exclude (partly) shadowed pixels when computing $\boldsymbol{\chi}$, by using the δ symbol, where $\delta = 1$ in full sunlit pixels with $Q < 0.1$, and $\delta = 0$ otherwise.

Last but not least, the diffuse solar illumination $s_{0_{\text{diff}}}$ plays an important role in shadowed regions. The diffuse solar illumination is the scattered light by the atmosphere in all directions. Since we aim to keep all types of nonlinear interactions up to the same order, i.e., the second order, we regard only the linear interactions for $s_{0_{\text{diff}}}$. In the proposed model, the following assumptions hold for $s_{0_{\text{diff}}}$:

- A light ray from the diffuse solar illumination source will interact with at least one material. After the first interaction, the light ray will escaping the pixel with a fraction of Q and reaching the observer.

Hence, the light path corresponding to the diffuse solar illumination source is: $P(path_{\text{diff}}) = P(\Xi_0 = s_{0_{\text{diff}}}, \Xi_1 = s_1, \Xi_2 = o) = Qa_l, (l = 1, \cdots, p)$ with the spectral contribution for $L = 1$. The spectral contribution is $T(s_{0_{\text{diff}}}) \prod_{k=1}^{L} \boldsymbol{e_i}$ where $T(s_{0_{\text{diff}}}) = \frac{\boldsymbol{\tau}_{\text{diff}} \odot \boldsymbol{\mathcal{L}}_{\text{s}}}{\boldsymbol{\tau}_{\text{dir}} \odot \boldsymbol{\mathcal{L}}_{\text{l}} + \boldsymbol{\tau}_{\text{diff}} \odot \boldsymbol{\mathcal{L}}_{\text{s}}}$.

5.2.5 Experimental Setup

One difficulty for the quantitative evaluation of the shadow-aware unmixing methods is that the ground truth of abundances and shadow fractions is not available, and very difficult to acquire in the case of shadows. Thus, we first validate our method on a simulated dataset with known abundances and parameters for a quantitative evaluation of the performance of the mixing models. Furthermore, we compare the unmixing methods on a real image with simulated shadowed pixels. Finally, we show experimental results on real airborne hyperspectral imagery without ground truth data, both quantitatively and qualitatively.

5.2.5.1 Datasets

- **Synthetic Dataset**
  Considering that a validation dataset with (partly) shadows is not available and very difficult to acquire, we validate our method on a simulated dataset to evaluate the *RE* and *AE* quantitatively. We randomly select 10 endmembers from the United States Geological Survey (USGS) spectral library of minerals [ii], where each material comprises 224 spectral bands ranging from 383 nm to 2508 nm. Abundances are then randomly generated following the Dirichlet distribution that automatically enforces

[ii] https://speclab.cr.usgs.gov/spectral-lib.html

ANC and ASC. Then, hyperspectral data are generated following each of the considered mixing models, i.e., LMM, Fan, SLMM, SMLM, Fansky and the proposed ESMLM. Parameters P are randomly generated based on the half-normal distribution with $\sigma = 0.3$. Values larger than one are set to zero, following the work in [150]. Other parameters including $Q \in [0,1]$, $F \in [0,1]$, $K \in [0,1]$, $\gamma \in [0,1]$, $b \in [-1,1]$ are generated following the uniform distribution. k_1, k_2, k_3 are chosen the same as the ones used for the real hyperspectral imagery in Section 5.2.5.2. Furthermore, we add white noise with signal-to-noise ratios $SNR = [50, 100]$ to the simulated dataset to assess the noise impact on different unmixing methods.

- **Real Dataset: Extended DLR HySU**
  The proposed method was validated on the extended DLR HySU dataset described in Section 4.2, Chapter 4. In this experiment, we use the hyperspectral image with simulated shadows but without random noise.

- **Real Dataset: HySpex/3K**
  A subscene (see Fig. 5.15 (a)) is selected from the hyperspectral images in the HySpex/3K dataset. Please refer to Section 4.1 in Chapter 4 for a detailed description of this dataset. A spectral library of endmembers is generated by manually selecting pure pixels of relevant materials in fully sunlit pixels of the image (Fig. 5.15 (b)). We select endmembers manually, as the endmembers should be extracted from fully sunlit pixels. In other words, to be able to use an automatic endmember extraction method, one would have to find an shadow detection method that can distinguish fully sunlit pixels from other pixels with ideally 100% accuracy, which is not easy to achieve in reality. For validation and comparison of the proposed model, we selected two subsets from the subscene (Fig. 5.15 (c) and (d)), which are dominated by shadow effects and cover three different shadow types, described in Fig. 5.14 (b), (d), (e).

5.2.5.2 Unmixing Procedure

We compare the proposed model (ESMLM) with the following state of the art models: LMM [76], Fan [87], SLMM [151], SMLM [151], and Fansky [198]. All algorithms were developed in MATLAB and run on an Intel Core i7 −8650 U CPU, 1.90 GHz machine with 4 Cores and 8 Logical Processors. We use the MATLAB function FMINCON to perform the nonlinear optimization. The processing time depends on the number of input pixels and endmembers. Table 5.5 shows the running time of the compared models in the two subsets of the real airborne hyperspectral imagery. The function and constraint tolerance are set to 10^{-10} and 10^{-8}, respectively. The initial values of the abundances are set to $\frac{1}{p}$ and the initial values of the unknown parameters F, Q, P, K are set to $[1, 1, 0, 0]$.

For the methods considering skylight, i.e., the Fansky and ESMLM model, ten pairs of pixels have been selected in the scene to compute the parameters k_1, k_2, and k_3 by using Eq. (5.17). Specifically, we select fully sunlit and shadowed pixels of the same pure material on the high-resolution hyperspectral image, assuming that two pixels near a shadow boundary are composed of the same material. Besides, we avoid vegetation materials during the pixel selection to avoid nonlinear effects. Fig. 5.16 (a) shows an example of selecting one pair

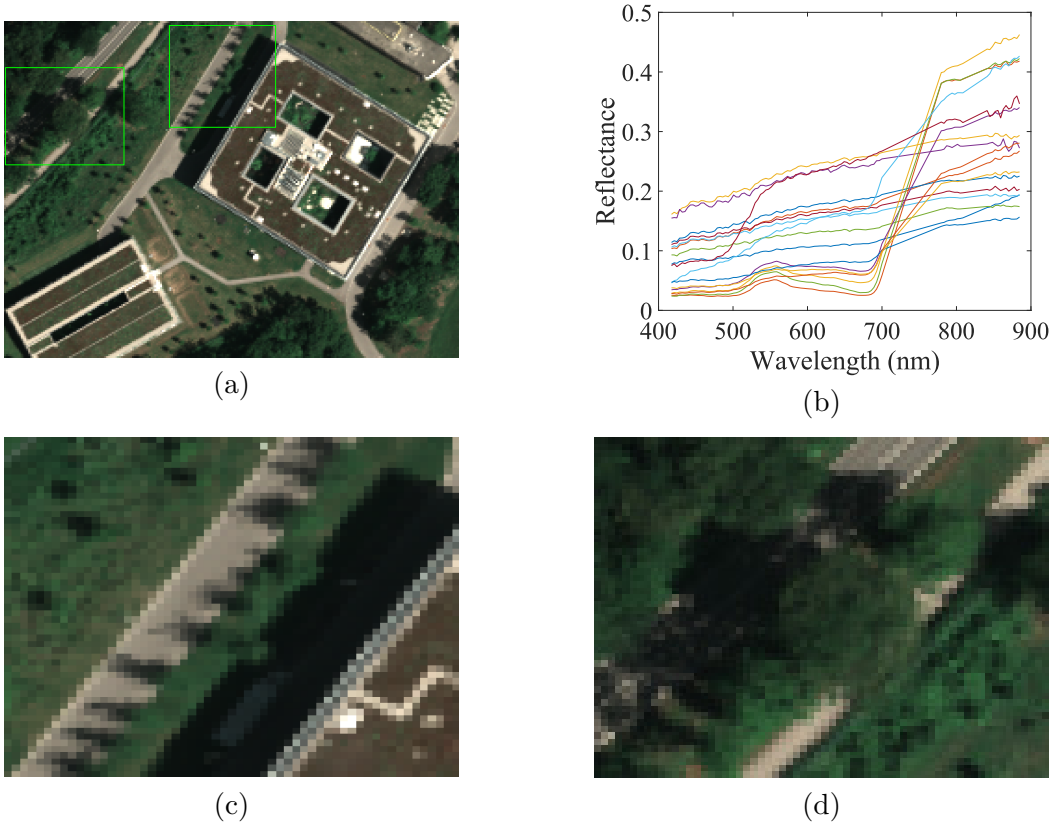
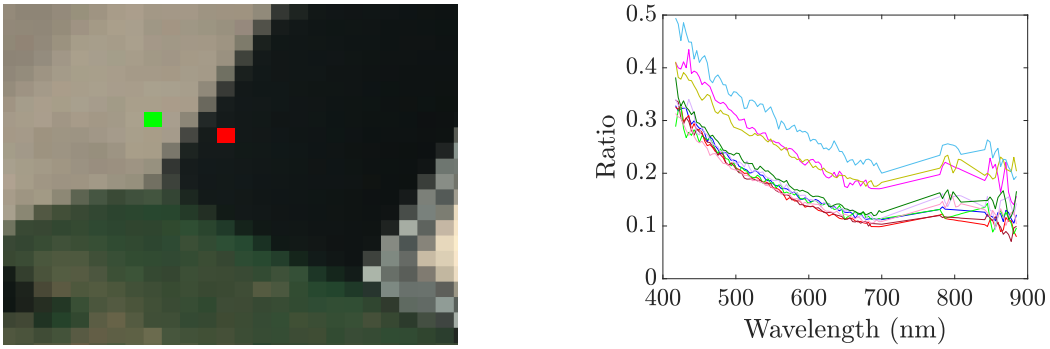


Figure 5.15: A subscene selected from the HySpex/3K dataset: (a) hyperspectral image as a true color composite acquired by the HySpex sensor in the study area of Oberpfaffenhofen, Bavaria, Germany; (b) endmember library, manually selected from (a); (c) and (d) true color composites of subsets selected from image (a).

Table 5.5: Running time of compared models in the two subsets

Model	Running time (s)	
	Subset1 (3135 pixels)	Subset 2 (3328 pixels)
LMM	26.04	20.71
Fan	55.89	41.99
SLMM	34.23	24.50
SMLM	57.47	45.31
Fansky	522.28	381.49
ESMLM	203.72	204.83

of pixels. Fig. 5.16 (b) presents the ratios computed by 10 pairs of fully sunlit and fully shadowed pixels. Assuming that the atmospheric conditions are constant in the entire region, these parameters are assumed to be constant, and were set as: $k_1 = 1.296$; $k_2 = 6.068$; $k_3 = 0.442$.



(a) (b)

Figure 5.16: Selection of 10 pairs of sunlit and shadowed pixels. (a) an example of the pixel selection for computing parameters k_1, k_2, and k_3 and (b) ratios computed from 10 pairs of pixels.

5.2.5.3 Experimental Design

Experimental results are shown in the following sections. In section 5.2.6 we perform a quantitative analysis of the mean reconstruction error (RE) and mean abundance error (AE) on the synthetic dataset simulated by the USGS spectral library. Quantitative evaluation metrics are introduced in Section 2.5.3. In Section 5.2.7, we evaluate the unmixing results on the extend DLR HySU dataset [188]. Specifically, the five ground targets are used to validate abundance errors. Section 5.2.8 evaluates the spectral mixing models on the real hyperspectral imagery without ground truth data, quantitatively and qualitatively. In Section 5.2.8.1, we perform a quantitative analysis of the spatial and spectral reconstruction errors. Moreover, we generate shadow-removed images in Section 5.2.8.2. This can be achieved by "lightening up" the shadow fraction in a mixing model, if applicable. Some of the unmixing methods output physically interpretable parameters, which provide valuable information about the observed surface. We discuss qualitative results of output parameters and abundances in 5.2.8.3 and 5.2.8.4. In section 5.2.8.5, we conduct an ablation study of the proposed model and analyze the impact of each parameter on the experimental results. Section 5.2.8.6 discusses the impact of endmember extraction methods on the unmixing results. Finally, we demonstrate our proposed model on the entire test image in section 5.2.8.7.

5.2.6 Synthetic Dataset

Table 5.6 and Table 5.7 present RE and AE of mixing models at different noise level following Eq. (2.29) and (2.30), respectively. The columns represent the spectral mixing models according to which mixtures are generated, and the rows correspond to the methods that were used to unmix the data. The last column conveys the mean performance of each unmixing method for all types of generated mixtures. For each type of mixture, the first and second best unmixing methods have been highlighted in red and green colors, respectively. For all mixtures generated by the different models, the proposed unmixing method obtained the best or second-best RE and AE among all comparing methods, and achieved the best

results on average. Results indicate that the ESMLM model can tackle different kinds of mixtures.

Table 5.6: Mean reconstruction error (RE) for the synthetic dataset

Noiselss	LMM	Fan	SLMM	SMLM	Fansky	ESMLM	Mean
LMM	0	0.9134	1.219	1.133	0.731	0.865	0.810
Fan	0.125	0	1.318	1.235	0.721	0.895	0.716
SLMM	0	0.913	0	0.122	0.262	0.248	0.258
SMLM	0	0.913	0.007	0.005	0.232	0.216	0.229
Fansky	0.054	0	0.050	0.052	0	0.060	0.036
ESMLM	0.002	0.050	0	0.013	0.018	0.002	0.014
$SNR100$	LMM	Fan	SLMM	SMLM	Fansky	ESMLM	Mean
LMM	0	0.915	0.038	0.242	0.723	0.606	0.421
Fan	0.124	0	0.159	0.353	0.043	0.232	0.152
SLMM	0	0.915	0	0.111	0.723	0.589	0.390
SMLM	0	0.915	0	0	0	0.565	0.367
Fansky	0.053	0	0.053	0.048	0	0.129	0.047
ESMLM	0.003	0.045	0.001	0.008	0.034	0.006	0.016
$SNR50$	LMM	Fan	SLMM	SMLM	Fansky	ESMLM	Mean
LMM	0.029	0.916	0.043	0.243	0.723	0.607	0.427
Fan	0.128	0.035	0.160	0.353	0.051	0.234	0.160
SLMM	0.029	0.916	0.017	0.114	0.723	0.590	0.398
SMLM	0.029	0.916	0.017	0.016	0.723	0.567	0.378
Fansky	0.061	0.035	0.056	0.052	0.022	0.134	0.060
ESMLM	0.029	0.064	0.017	0.020	0.045	0.029	0.034

5.2.7 Real Dataset: Extended DLR HySU

Fig. 5.17 presents the obtained RE in fully sunlit pixels, (partly) shadowed pixels, and the entire image, respectively, and Table 5.8 compares the AE of the ground targets made of synthetic materials.

In fully sunlit pixels, reconstruction errors of linear and nonlinear models are similar, indicating that the nonlinear effect plays a minor role in the study scene. On the other hand, RE largely varies in (partly) shadowed pixels. When the shadow effect is taken into account, RE significantly decreases in (partly) shadowed pixels and the lowest reconstruction errors are obtained when considering the skylight.

Using the ground truth data from Table 4.2, the total number of pixels for each material is representative for the total corresponding abundance in the image. Thus, we represent the abundance estimation error by the absolute difference in number of pixels between ground truth and estimated values over the five targets (Table 5.8). In addition, we present the total (absolute and in percentage) estimation error by summing up the errors of all endmembers. Besides, we compare abundance maps qualitatively for all materials in Fig. 5.18, where the first column shows the reference abundance maps for easier comparison. Specifically, the

Table 5.7: Mean abundance error AE for the synthetic dataset

Noiselss	LMM	Fan	SLMM	SMLM	Fansky	ESMLM	Mean
LMM	0	0.082	0.121	0.126	0.096	0.104	0.088
Fan	0.049	0	0.138	0.139	0.101	0.115	0.090
SLMM	0	0.082	0	0.031	0.044	0.046	0.034
SMLM	0	0.082	0.010	0.006	0.044	0.046	0.031
Fansky	0.032	0	0.067	0.036	0	0.031	0.028
ESMLM	0.001	0.010	0	0.007	0.013	0	0.005
$SNR100$	LMM	Fan	SLMM	SMLM	Fansky	ESMLM	Mean
LMM	0	0.082	0.121	0.126	0.096	0.104	0.088
Fan	0.049	0	0.137	0.139	0.101	0.115	0.090
SLMM	0	0.082	0	0.031	0.044	0.046	0.034
SMLM	0	0.082	0.010	0.006	0.044	0.046	0.031
Fansky	0.032	0	0.067	0.036	0	0.031	0.028
ESMLM	0	0.010	0	0.007	0.013	0	0.005
$SNR50$	LMM	Fan	SLMM	SMLM	Fansky	ESMLM	Mean
LMM	0.001	0.082	0.121	0.126	0.096	0.104	0.088
Fan	0.049	0.001	0.137	0.139	0.101	0.115	0.090
SLMM	0.001	0.082	0.004	0.031	0.044	0.046	0.035
SMLM	0.001	0.082	0.014	0.008	0.044	0.046	0.033
Fansky	0.032	0.001	0.067	0.036	0.001	0.031	0.028
ESMLM	0.002	0.010	0.005	0.008	0.014	0.003	0.007

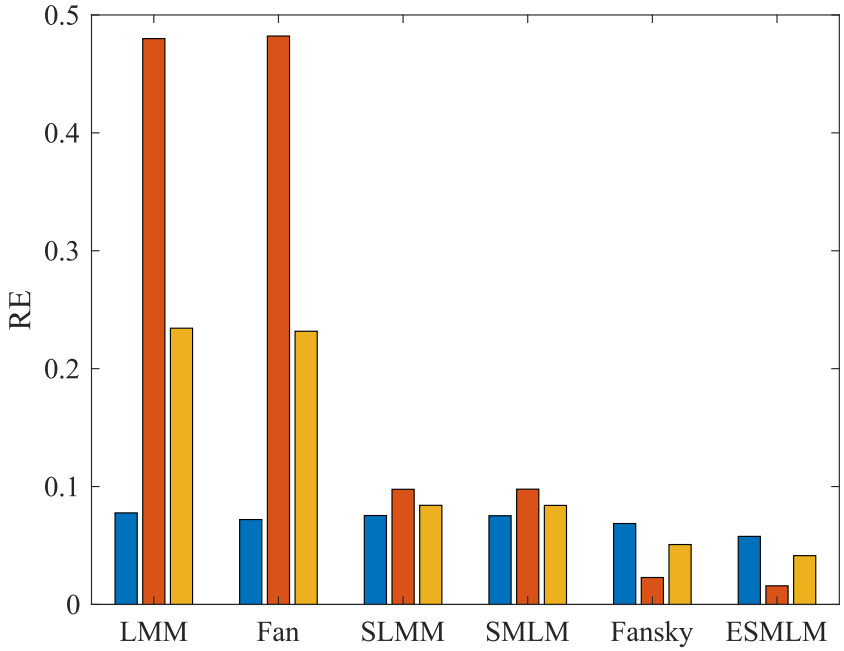


Figure 5.17: Mean reconstruction error RE of fully sunlit pixels (in blue), (partly) shadowed pixels (in orange), and the entire image (in yellow) in HySU dataset. Fully sunlit and (partly) shadowed pixels are identified as $Q \leq 0.1$ and $Q > 0.1$, respectively.

Table 5.8: Abundance error in number of pixels in the HySU dataset

Endmember	LMM	Fan	SLMM	SMLM	Fansky	ESMLM
Bitumen	50.422	53.184	43.019	43.430	9.504	1.366
Red-painted Metal Sheets	6.529	6.051	2.234	2.286	4.097	0.702
Blue Fabric	12.232	12.299	10.234	10.356	0.837	0.108
Red Fabric	9.111	9.932	7.570	7.628	3.719	0.443
Green Fabric	6.471	6.229	10.020	10.029	1.832	2.614
Total	84.765	87.695	73.076	73.728	19.989	5.233
Total(%)	92.08	95.26	79.38	80.09	21.71	5.68

reference abundance maps are estimated using the shadow-free image through non-negative least squares.

In sunlit regions, the LMM and the Fan model estimate correct abundances, but show high abundance estimation errors compared to other methods in (partly) shadowed pixels. Among all materials, the largest abundance error appears in bitumen, which has a relatively small reflectance and is therefore easily confused with shadows. Besides, more confusion between similar materials can be observed. An example appears in shadowed pixels of grass, where the LMM and Fan models confuse those regions with green fabric.

Compared to the LMM and the Fan model, the SLMM and SMLM model perform slightly better. In partly shadowed pixels, the SLMM and the SMLM model detect part of the correct materials. However, some shadowed pixels of red fabric, which have been estimated as bitumen by the LMM and Fan model, are confused with red metal sheets.

The Fansky and ESMLM models outperform SLMM and SMLM for the abundance estimation of all ground targets, indicating that the wavelength-dependent skylight information can not be well represented using a scaling parameter. Most shadowed pixels have been detected as the correct material, resulting in largely decreased abundance estimation errors. The Fansky model confuses between materials with similar spectra, such as green fabric and grass, as well as red fabric and red metal sheets. In addition, it confuses blue materials with bitumen. Compared to other models, the ESMLM model achieves the best performance and can detect most ground targets with a total abundance estimation error of 5.233 pixels (corresponding to 5.68%). Specifically, the ESMLM model can better identify similar materials in shadowed pixels thanks to the advantageous and flexible modeling of nonlinear effects.

5.2.8 Real Dataset: HySpex/3K

5.2.8.1 Reconstruction Errors

The *RE* of each of the two subsets for all compared methods is depicted in Fig. 5.19. Separate results are shown for fully sunlit, (partly) shadowed pixels and the entire image, respectively. (Partly) shadowed pixels are identified using $Q > 0.1$, while fully sunlit pixels are identified using $Q \leq 0.1$, where Q values are computed using the proposed model. Results suggest that reconstruction errors highly depend on if and how the models consider the illumination

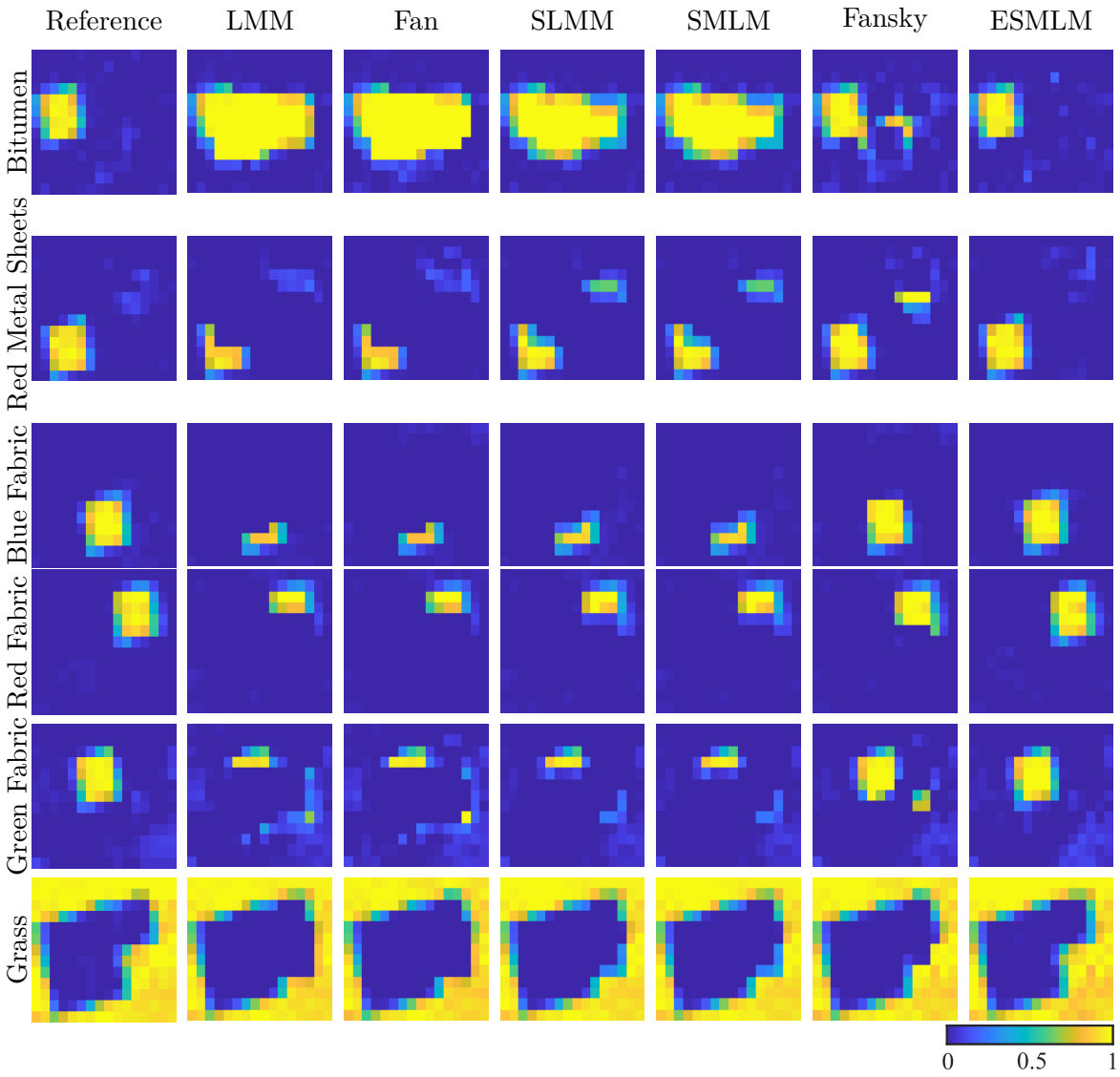


Figure 5.18: Abundance maps from the HySU dataset. Top to bottom: bitumen, red metal sheets, blue fabric, red fabric, green fabric, and grass. Left to right: reference, LMM, Fan model, SLMM, SMLM model, Fansky model, and ESMLM model. The reference abundance maps are computed using the shadow-free image through non-negative least squares.

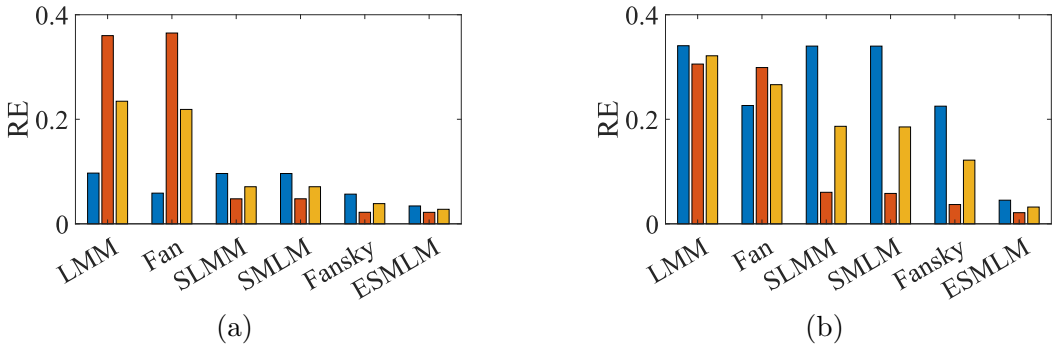


Figure 5.19: Mean reconstruction error (RE) of fully sunlit pixels (in blue), (partly) shadowed pixels (in orange), and the entire image (in yellow), for subset 1 in (a) and subset 2 in (b). Fully sunlit and (partly) shadowed pixels are identified as $Q \leq 0.1$ and $Q > 0.1$, respectively, where Q values are computed using the ESMLM model.

conditions and the nonlinearity. In (partly) shadowed pixels, the reconstruction errors largely decrease when using shadow-aware mixing models, i.e., the SLMM, SMLM, Fansky and ESMLM models. Among all shadow-aware models, the Fansky and ESMLM models consider the skylight information, and thus outperform other models. Compared to the Fansky model, ESMLM obtained the smallest reconstruction errors in (partly) shadowed pixels, especially when areas are shaded by vegetation, such as in subset 2. Since both models consider skylight information, this implies that the ESMLM model treats the nonlinearity better than the Fansky model in (partly) shadowed pixels. In sunlit regions, the reconstruction errors appear larger in subset 2, where the ground surface is covered mostly by vegetation. The Fan and Fansky models attained slightly lower errors than the linear models, but yielded higher errors than the ESMLM model regarding the neighborhood interactions. Overall, the ESMLM model attained the best pixel reconstruction and, in this respect, it produced a better representation of the ground mixtures.

Fig. 5.20 shows spectral reconstruction errors $SRE(\lambda)$ (Eq. (2.31)), which denote how well a spectral mixing model represent input pixels as a function of wavelength. When a mixing model obtains a good spectral representation, we expect the SRE values to be constant and small for all wavelengths. Instead, if SRE largely varies as a function of λ, the spectral unmixing method is not capable of dealing with specific wavelength-dependent effects. As both subsets contain large shadowed regions, the LMM and Fan models obtained the largest errors over the entire wavelength range. The SLMM and SMLM models obtained higher errors in the lower spectral range of $400 - 500nm$, because these assume the shadow effects to be wavelength-independent, and ignore the skylight which has the highest impact at shorter wavelengths. In subset 2, the SREs appear larger at longer wavelengths, and the spectral behavior of the errors shows vegetation characteristics. Compared to the Fansky model, the ESMLM model shows less vegetation characteristics, indicating that it provides a better spectral reconstruction performance for vegetation.

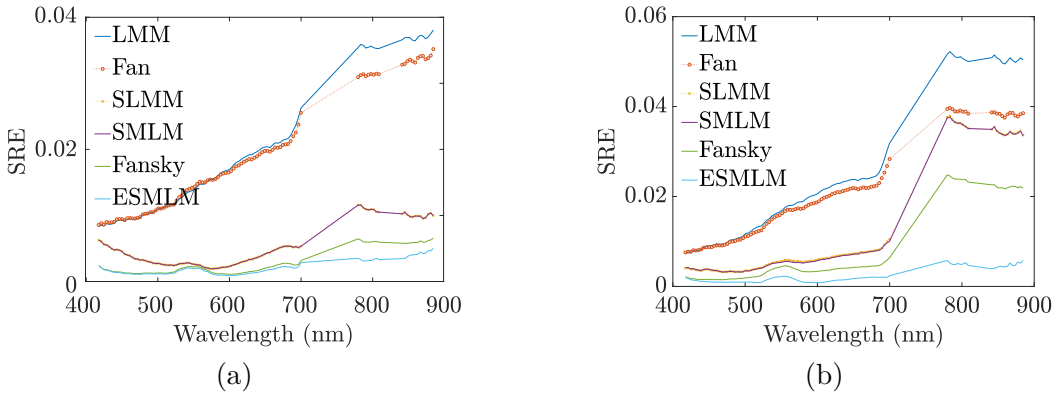


Figure 5.20: Spectral reconstruction errors (SRE) as a function of wavelength for subset 1 in (a) and subset 2 in (b).

Figure 5.21: Shadow-removed reconstructed images (true color composites) of subset 1 (first row) and subset 2 (second row).

5.2.8.2 Shadow-Removed Pixel Reconstruction

For spectral mixing models containing the shadow-related parameter Q, it is possible to perform shadow removal through pixel reconstruction. The idea is to "lighten up" the shadow fraction in a pixel, by replacing the illumination source for shadowed regions with the one for sunlit regions. In other words, the restoration process simulates that shadowed regions are exposed to direct solar illumination. Since SLMM and SMLM models do not contain the diffuse illumination source for shadowed regions, shadow removal can be performed by setting $Q = 0$ in the mixing models. For the Fansky and ESMLM models, we generate the restoration results by replacing $T(s_{0_{\text{diff}}})$ with $T(s_{0_{\text{g}}})$ in the mixing models. Fig. 5.21 shows the input and shadow-removed images for visual comparison. Since the shadow removal is performed by replacing the illumination sources in the mixing models, the shadow fraction Q has been naturally embedded in the restoration process, yielding physical-interpretable transitions at shadow boundaries in the shadow-removed images. As the values of Q are fractional in the range $[0, 1]$, a more realistic representation of shadows is provided.

Due to the lack of ground truth of the actual spectral reflectance and thus the actual pixel composition under the shadows, the shadow-removed images can only be qualitatively compared. For a more quantitative evaluation of the performance of shadow removal, we additionally designed an alternative test, by assuming that the region around a shadow

boundary should contain similar materials (see Fig. 5.22). In each subset we selected 7 regions, each consisting of sunlit pixels (located at the yellow markers) and (partly) shadowed pixels (located at the cyan markers) around a shadow boundary (Figs. 5.22 (a) and (e)). Spectra in sunlit pixels are selected from the input image, while spectra in (partly) shadowed pixels are selected from the restored image. In each region, the spectra of sunlit and (partly) shadowed pixels are individually averaged, resulting in 7 pairs of spectra, each consisting of a sunlit and a shadowed spectrum around a shadow boundary. The band-wise absolute differences between the sunlit and shadowed spectra are averaged over all 7 pairs, and plotted as the spectral error in Figs. 5.22 (b) and (f). In addition, in subset 1 we individually consider shadows caused by man-made objects, completely blocking direct sunlight, and shadows caused by vegetation that can partly block direct sunlight (Figs. 5.22 (c) and (d)).

The visual comparison of Fig. 5.21 can be interpreted, depending on the skylight information being taken into account or not. When excluding the skylight information (i.e., SLMM and SMLM models), results show less noise in shadowed restored regions. However, texture and spectral information can be lost in shadow areas, leading to a non-natural restoration result, and incorrect spectral information of the pixel composition under the shadow. In addition, the SMLM model has not removed all shadows, due to an inaccurate estimation of its parameters (see Section 5.2.8.3).

In contrast, the models that account for skylight show increased noise levels (Fig. 5.21), but on average perform better in restoring the spectral information in shadowed pixels (Fig. 5.22). The Fansky model behaves inconsistently in different types of shadows. Specifically, it performs worse than SLMM in the regions shadowed by vegetation (Fig. 5.22 (d) and (f)), probably due to its specific modeling of the nonlinearities. The proposed ESMLM model performs consistently better for different types of shadows, indicating the importance of the nonlinear modeling in shadow-aware mixing models.

5.2.8.3 Model Parameters

Some of the compared spectral mixing models generate physically interpretable pixelwise parameters providing valuable information. Fig. 5.23 shows the output parameter maps. All compared mixing models, except LMM and Fan, output Q, representing the fraction of shadows at sub-pixel level. Depending on the way the shadow effect is modeled, two categories of Q maps can be differentiated. The SLMM and the SMLM models treat shadow as a simple scaling effect without including skylight information. Despite Q being the spatial fraction of shadow in a pixel, it serves two functionalities here. One is to reduce the observed reflectance by scaling out the shadowed part Q of a pixel. The other is to use the remaining fractional value $1 - Q$ to "lighten up" the shadowed regions. From the figure, on can clearly observe that the Q values from SLMM and SMLM are consistently underestimated in shadowed pixels, because even in heavily shadowed regions the reflectance, although very small, is not equal to zero. Compared to the SLMM, the SMLM model can underestimate Q in (partly) shadowed pixels by overestimating the P values, such as in vegetation shadows in subset 2 and on the boundaries of the shadowed regions by the building in subset 1. The

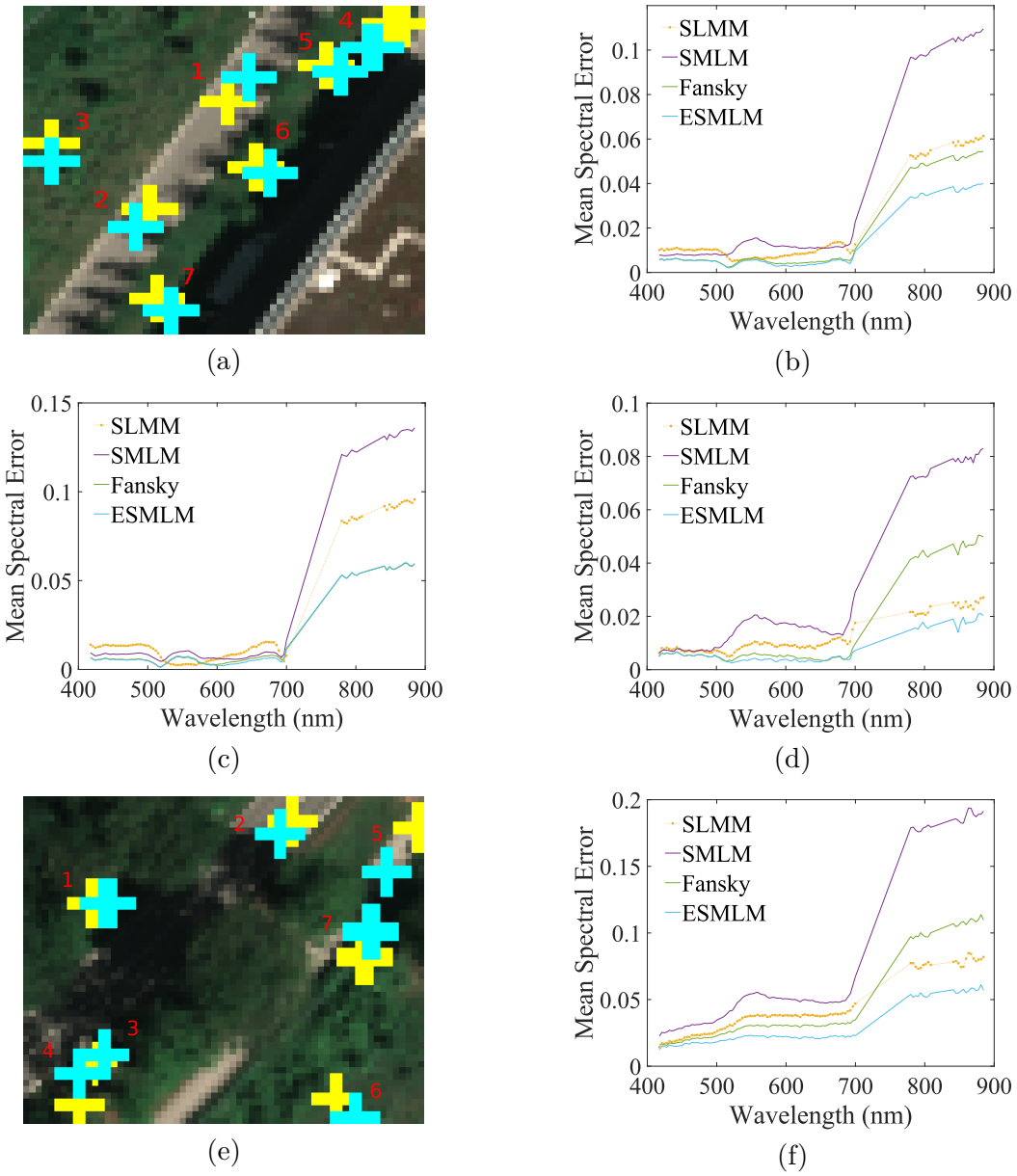
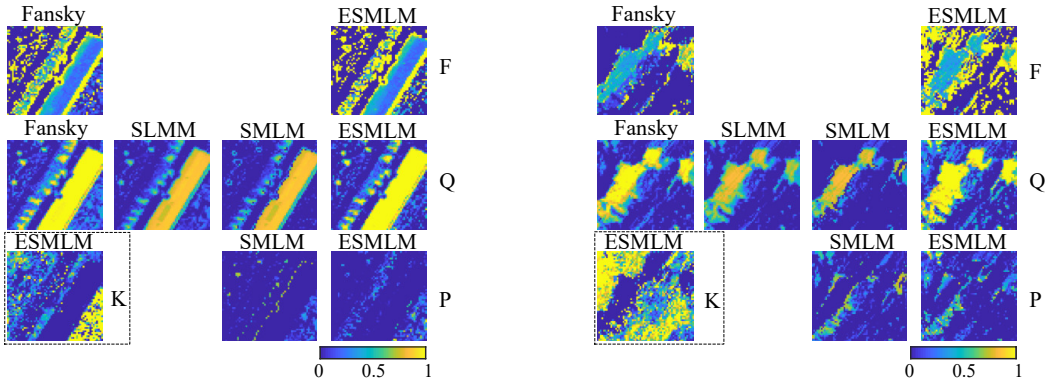


Figure 5.22: Spectral comparison between shadowed pixels in shadow-removed images and their corresponding sunlit pixels belonging to the same material in local neighborhoods. Subset 1: (a) locations of selected pairs of pixels, sunlit pixels are marked in yellow and (partly) shadowed pixels are marked in cyan, (b) spectral error (all regions), (c) spectral error in regions shadowed by man-made objects, (d) spectral error in regions shadowed by vegetation. Subset 2: (e) location of selected pairs of pixels, sunlit pixels are marked in yellow and (partly) shadowed pixels are marked in cyan, (f) spectral error over all regions (all regions are shadowed by vegetation).



(a) Subset 1 (b) Subset 2

Figure 5.23: Output parameter maps, from top to bottom: F (sky view factor), Q (spatial fraction of shadows), P (probability that a light ray undergoes additional interactions with endmembers), K (a strength factor of neighbor interactions, denoting the fraction of the scattered light from the neighborhood that is received by the pixel). Compared models, from left to right: Fansky, SLMM, SMLM, ESMLM.

inaccurate estimation of Q greatly decreases the performance for shadow removal (Fig. 5.21). In contrast, the ESMLM and Fansky models use the skylight to "lighten up" the shadowed areas, yielding a better estimation for Q. Compared to the Fansky model, the ESMLM model generates better Q maps, thanks to its superior nonlinear modeling. In subset 2, the ESMLM model estimates higher values of Q in fully shadowed pixels and can detect partial shadows in the bottom right area.

Beside shadows, the nonlinear behavior of the mixing models is also an important aspect. Fig. 5.23 shows the two parameters relevant to the nonlinearity, i.e., P and K, which describe within- and between-pixel optical interactions at the second order, respectively. The P parameter in the ESMLM model follows the definition from [151]. Although P can be negative in the SMLM model, we constrain $P \in [0, 1]$ in this work in order to keep its physical interpretation. The P maps of SMLM and ESMLM follow similar patterns, with increasing values in vegetated regions. In SMLM, P can have high values in (partly) shadowed pixels where Q is close to zero. This artifact can be observed on shadow boundaries of the building in subset 1, and tree-shadowed regions in both subsets. The ESMLM model shows a better estimation of both P and Q, due to the inclusion of skylight information.

The ESMLM model outputs high values of K, dominantly in (partly) sunlit vegetated areas, where pixels are expected to receive reflections from their neighborhood. Those regions show significantly lower reconstruction errors in the ESMLM model than in other models, indicating the advantage of modeling the neighbor interactions in the ESMLM model.

Finally, models accounting for skylight (ESMLM and Fansky) additionally output the topographic related parameter F, indicating the sky fraction that a ground pixel can "see". It is worth noting that F is only valid in (partly) shadowed pixels, as F is only involved in the skylight terms of the ESMLM and Fansky models. In this work, we set $F = 0$ in regions where $Q \leq 0.1$. Compared to the Fansky model, the ESMLM model is superior in estimating values of F in regions shadowed by vegetation. An example is shown in the upper-left

corner of subset 2, where the Fansky model obtained zero F, and thus merely reconstructs shadows by scaling sunlit pixels. Instead, the ESMLM model can balance the values of F by contributions of P. In regions where Fansky produces $F = 0$, the reconstruction errors of the ESMLM model are consistently lower by a value of 0.03 on average.

5.2.8.4 Abundances

Besides parameters, the proposed method outputs abundances that present the material components at sub-pixel level. Since we do not have ground truth data for this real dataset, we evaluate abundances qualitatively in this section. Fig. 5.24 and Fig. 5.25 show the abundance maps for all comparing models in two subsets. For each subset, we present two aggregate abundance maps of impervious surfaces and vegetation, by grouping materials with similar spectra.

Abundances in fully sunlit pixels are comparable among all models, while abundance maps show noticeably different patterns in (partly) shadowed pixels, depending on if and how the shadow and nonlinear effects are considered. In the LMM and Fan model, vegetation and impervious surface in the (partly) shadowed pixels are detected indiscriminately as impervious material. The reason is that the endmember library contains some impervious materials with low reflectances that are more similar to shadow spectra. Nevertheless, the LMM and Fan model show large RE values in (partly) shadowed pixels, indicating their unsatisfactory spectral representation in shadowed areas.

Results largely improve when considering shadow as a scaling effect, as is done by the SLMM and SMLM model. In the shadowed vegetation areas, SLMM and SMLM perform significantly better. Nevertheless, a small amount of impervious surface remains in shadowed vegetation areas, because it is used to compensate for the inaccuracy of the shadow modeling.

The Fansky and ESMLM models, that consider the skylight information, improve the performance in shadowed vegetation regions while introducing more noise. Compared to the Fansky model, the ESMLM model presents better results on vegetation shadows, thanks to the advantageous modeling of the nonlinear effects.

Some open questions remain for the deeply shadowed pixels caused by vegetation. For example, abundances in the shadowed impervious regions are not as large as expected in subset 2. One reason can be that the incoming light from the global illumination firstly interacts with trees before reaching the road, mixing vegetation features in the back-scatted signal in those regions. On the other hand, the LMM and Fan model estimate those regions as pure impervious materials, but it doesn't mean that they perform better, because they mainly confuse the shadowed pixels with impervious materials in the entire region.

5.2.8.5 Ablation Study

This section shows the results of an ablation study of the ESMLM model and analyzes the impact of the parameters P, Q, K by setting them to zero one at a time. Similar to the above experiments, we analyze the results in terms of reconstruction errors (Figs. 5.26 and

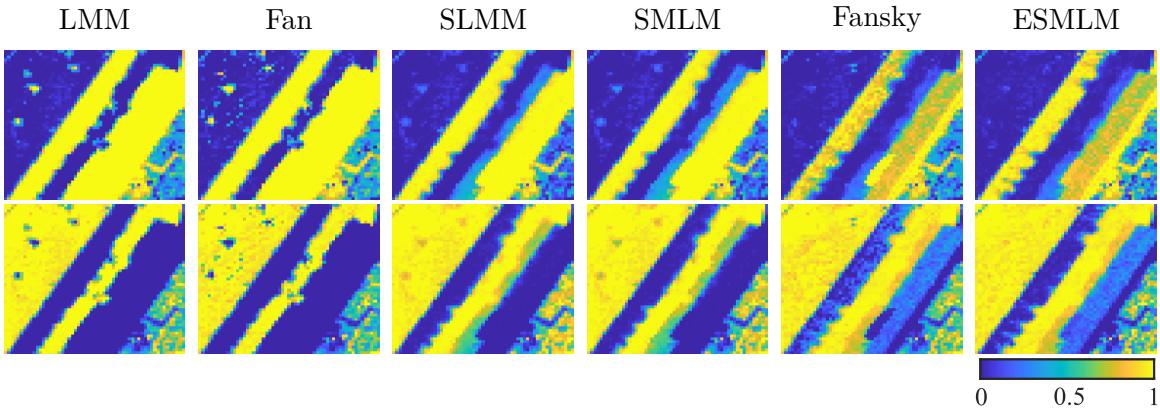


Figure 5.24: Abundance maps of the real hyperspectral imagery without ground truth for subset 1. First row: abundances of impervious materials; second row: abundances of vegetation.

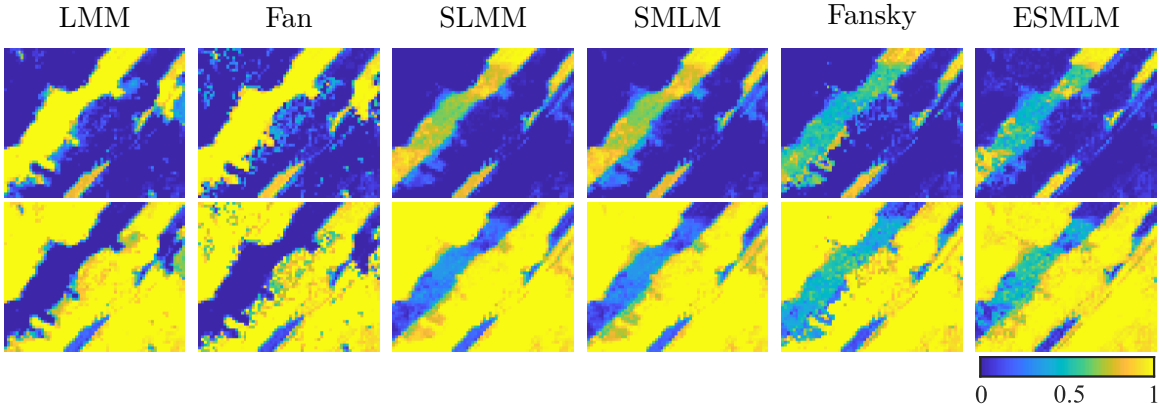


Figure 5.25: Abundance maps of the real hyperspectral imagery without ground truth for subset 2. First row: abundances of impervious materials; second row: abundances of vegetation.

5.27), shadow-removed reconstruction (Figs. 5.28 and 5.29), and output parameter maps (Fig. 5.30). We discuss the effect of each parameter in the following subsections.

- **The Role of Q**
  Q is the key parameter for modeling shadows. When $Q = 0$, the skylight-related terms become zero. Shadow removal cannot take place (Fig. 5.28), and mean spectral errors would become undesirably large in Fig. 5.29. As the shadow-related terms are removed, RE largely increases in shadowed regions in subset 1. In addition, SRE goes up for all wavelengths, and a significant increase is observed in the spectral range of 400 nm and 550 nm. This change is caused by the lack of skylight terms, which largely impacts on shorter wavelengths. In subset 2, RE does not considerably increase in (partly) shadowed pixels (Fig. 5.26 (b)), while SRE increases at shorter wavelengths in the spectral range of 400 nm and 500 nm (Fig. 5.27 (b)). The reason is that P replaces the role of Q to compensate for the reconstruction loss. In Fig. 5.30, it can be observed that the contribution of P increases. Despite a better reconstruction, this leads to an incorrect estimation of the parameters. This indicates that reconstruction errors can not be the only measure to evaluate the performance of mixing models.

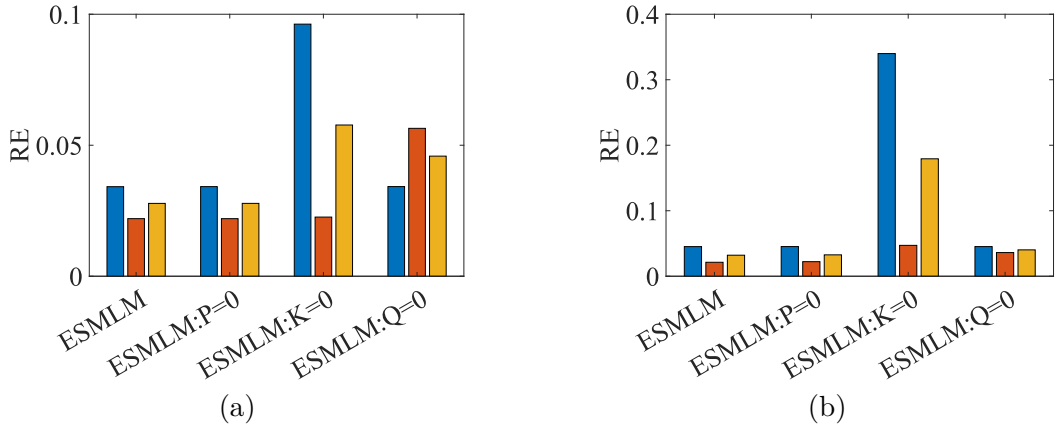


Figure 5.26: Mean reconstruction error (RE) in the ablation study for subset 1 in (a) and subset 2 in (b). Blue: fully sunlit pixels; Orange: (partly) shadowed pixels; Yellow: the entire image. Fully sunlit and (partly) shadowed pixels are identified as $Q \leq 0.1$ and $Q > 0.1$, respectively, where Q values are computed using the full model.

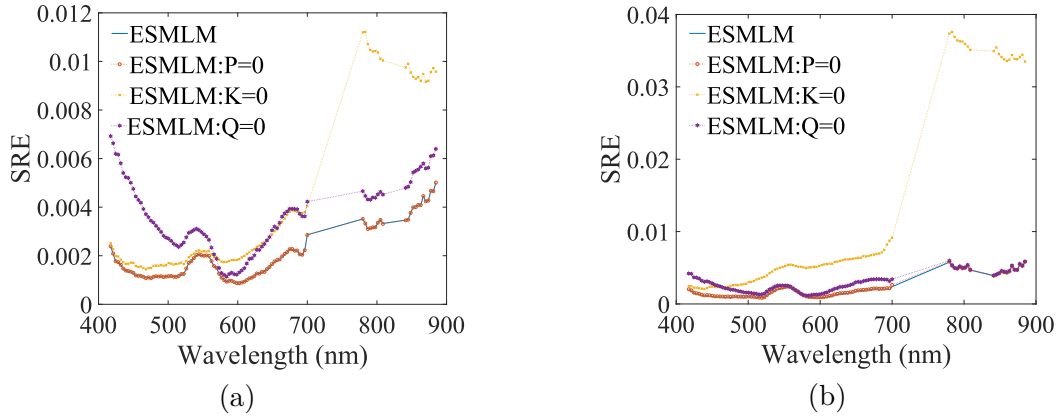


Figure 5.27: Spectral reconstruction error (SRE) as a function of wavelength in the ablation study for subset 1 in (a) and subset 2 in (b).

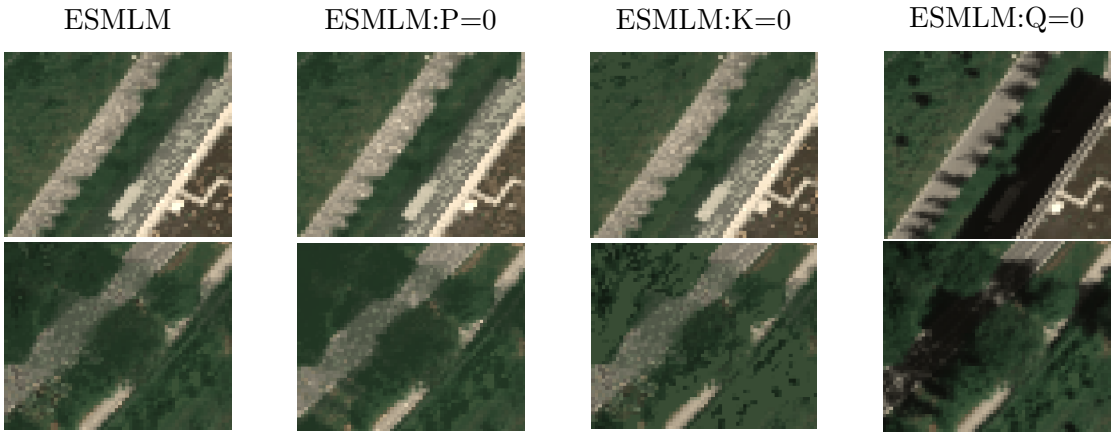


Figure 5.28: Shadow-removed reconstructed images (true color composites) in the ablation study for subset 1 (first row) and subset 2 (second row).

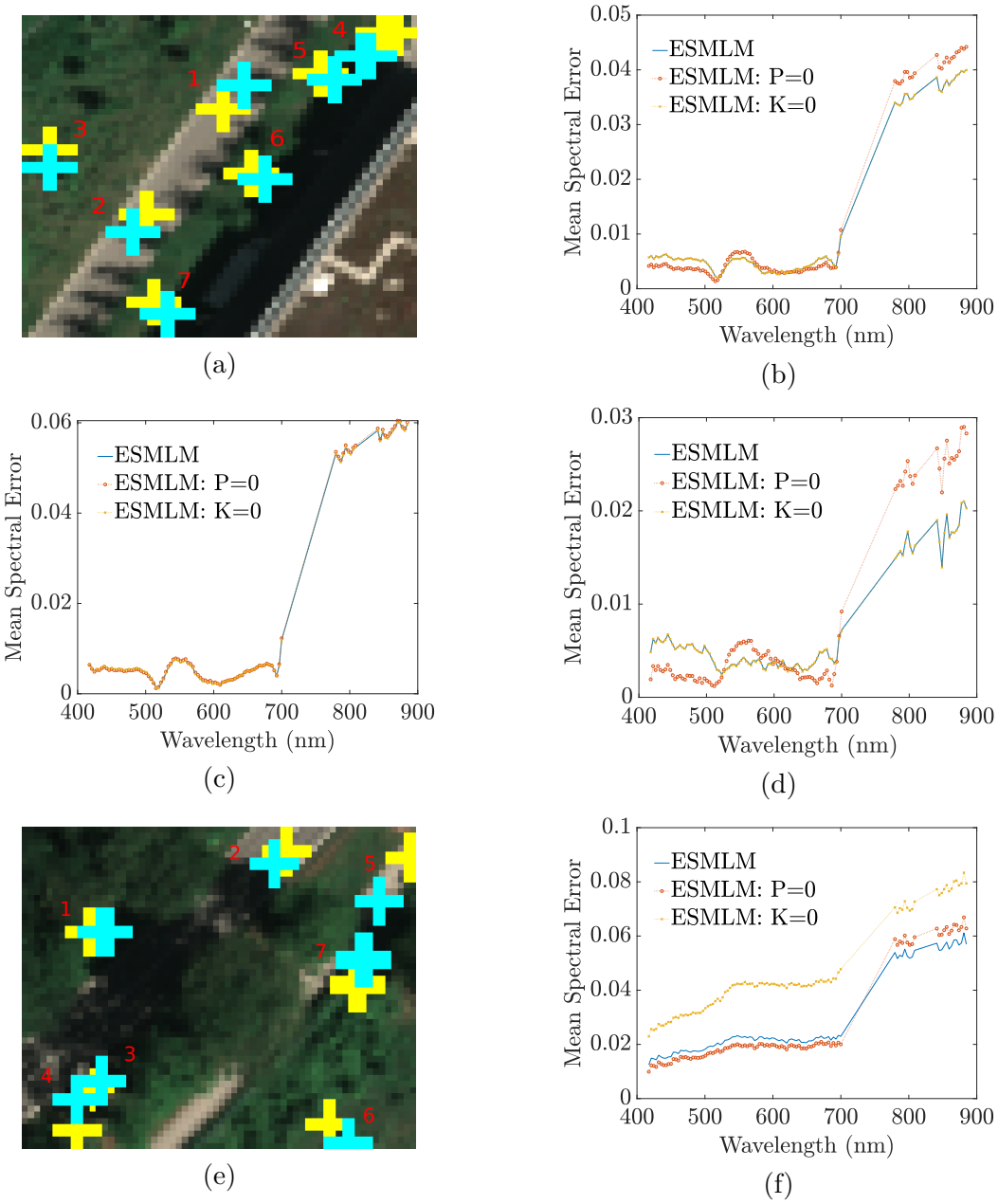
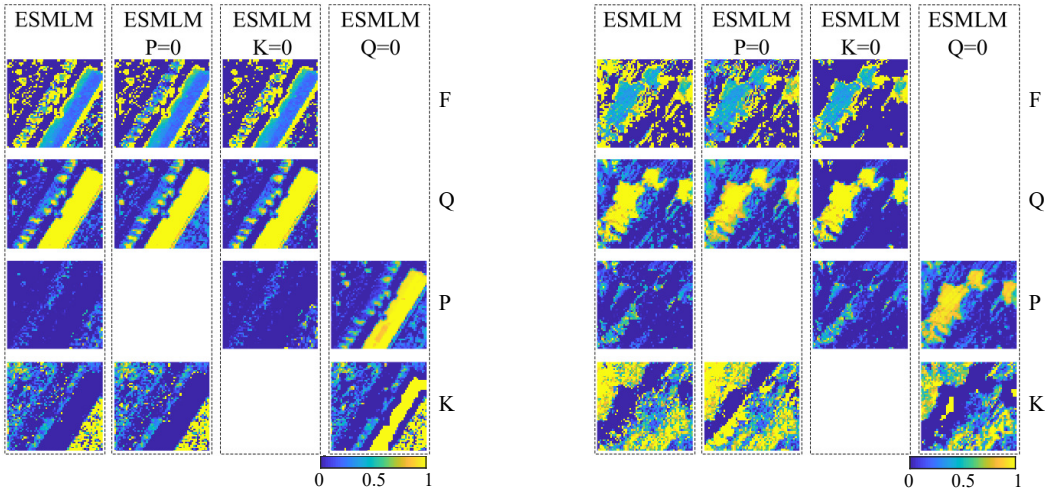


Figure 5.29: Ablation study: spectral comparison between shadowed pixels in shadow-removed images and their corresponding neighboring sunlit pixels, containing the same material. Subset 1: locations of selected pairs of pixels in (a), spectral errors in all regions in (b), in regions shadowed by man-made objects in (c), and in regions shadowed by vegetation in (d). Subset 2: location of selected pairs of pixels in (e), spectral errors in all regions (all regions are shadowed by vegetation) in (f). The ablated model with $Q = 0$ is not in the comparison, because it would exclude the shadow effect, causing undesirably large spectral errors.



(a) subset 1 (b) subset 2

Figure 5.30: Parameter maps in the ablation study. From top to bottom: F (the sky view factor), Q (the spatial fraction of shadow), P (the probability that a light ray undergoes additional interactions with endmembers), K (a strength factor of neighbor interactions, denoting the fraction of the scattered light from the neighborhood that is received by the pixel). From left to right: the ESMLM model, the ablated ESMLM models, with $P = 0$, $K = 0$, and $Q = 0$, respectively.

- **The Role of P**
  P is relevant for the within-pixel nonlinear behavior. Removing P does not affect reconstruction errors (Figs. 5.26 and 5.27), because Q and K compensate for the reconstruction loss. However, removing P impacts the estimation of other parameters (Fig. 5.30 (b)). When $P = 0$, the neighbor effect term $(1 - Q)(1 - P)K \sum_{i=1}^{p} a_i \boldsymbol{e}_i \odot \boldsymbol{\chi}$ becomes the only second-order reflection term in the model. Thus, the ablated model will estimate inaccurate values for Q, K, and F in vegetation shadows, where P is expected to have contributed. In addition, P plays an important role in removing shadow. When reconstructing the shadow-removed images, the sunlit regions should remain unchanged. However, we observed that the spectral distance using the L2-norm between input and restored images in sunlit regions increased consistently by a value of 0.025 on average in sunlit regions when $P = 0$, compared to the full model.

- **The Role of K**
  K is related to the between-pixel optical interactions that occur mainly in vegetated regions. When $K = 0$, the reconstruction errors largely increase in (partly) sunlit pixels (Fig. 5.26). In addition, SRE increases at longer wavelengths, which is caused by not accounting for the multiple interactions of vegetation in the local neighborhood. (Fig. 5.27). At shorter wavelengths, SRE only slightly increases, because the ablated model contains the key parameter Q for modeling shadows. The performance of shadow removal decreases when $K = 0$: shadow-removed images lose textural information (Fig. 5.28), and the mean spectral errors increase in subset 2, where vegetation dominates (Fig. 5.29 (f)).

5.2.8.6 Comparison Between Manual and Automatic Endmember Extraction Methods

In this section, we analyze the impact of endmember extraction methods on the result of the two subsets of the real hyperspectral image, and compare the unmixing results using manually extracted endmembers with those using automatically extracted endmembers. Considering the shadow issue, we designed a simple but effective method to select fully sunlit pixels before automatic endmember extraction. First, we carefully set an empirical threshold (set to 0.08 in this work), and then select pixels with mean reflectance larger than 0.08 as candidate sunlit pixels. However, candidate sunlit pixels may include partly sunlit pixels located at shadow boundaries. Thus, we additionally apply a Canny edge detector [209] to remove all boundary pixels from sunlit pixels candidates. In addition, considering endmember variability, we apply the method in [210] to extract endmember bundles based on Vertex Component Analysis (VCA) [71]. Finally, we merge endmembers with similar reflectances and show the selected endmembers in Fig. 5.31 (c).

We apply the automatically extracted endmembers to the two subsets of the real airborne hyperspectral imagery without ground truth. First, we compare the RE for subset 1 and 2 in Fig. 5.31 (f) and (g), respectively. Results depict that RE in (partly) shadowed pixels is comparable between automatic and manual endmember extraction methods. This is expected because (partly) shadowed pixels are excluded from the endmember extraction. In sunlit pixels, reconstruction errors decrease for all unmixing methods while their relative relationship remains, indicating that a better endmember library can be extracted using automatic endmember extraction. Moreover, we compare the output parameters (F, Q, P, and K) by the histogram of the parameter differences between the manually and automatically endmembers in Fig. 5.31 (a), (b), (d), (e). The differences between parameters estimated by two endmember libraries remain small, implying that results are not very sensitive to the endmember extraction method. We can conclude that manually extracted endmembers can be applied in our work.

5.2.8.7 Experimental Results and Discussion: the Entire Test Image

Finally, we applied the proposed model to the entire test image, resulting in RE of 0.04 in fully sunlit pixels, 0.03 in (partly) shadowed pixels, and 0.03 in the entire image. The low variation of RE values between different categories of pixels indicates that the ESMLM model provides a robust pixel representation over various illumination conditions and local structures. In addition, the output parameters of the entire image (Fig. 5.32 (b-e)) follow similar patterns as in the subset images. By interpreting the parameters, one can easily detect the different illumination conditions and local structures illustrated in Fig. 5.14. Specifically, (partly) shadowed pixels contain high Q values, while K mainly contributes in (partly) sunlit vegetated regions, and becomes typically higher when vegetation has larger height variations, such as trees. Compared to K, P mainly plays a role in (partly) shadowed pixels occluded by vegetation. Besides indicating the local structures, output parameters also play an important role in reconstructing the shadow-removed image shown in Fig. 5.32 (a).

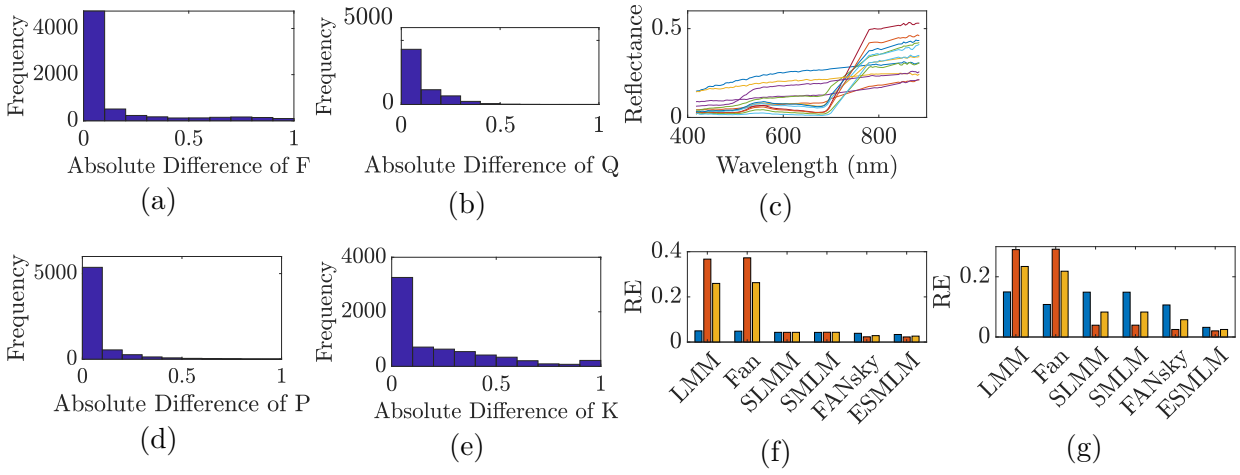


Figure 5.31: Comparison between manual and automatic endmember extraction methods. Automatically extracted endmembers in (c), RE computed by unmixing methods with automatically extracted endmembers for subset 1 in (f)) and subset 2 in (g), where blue, orange, and yellow colors represent fully sunlit pixels, (partly) shadowed pixels, and the entire image, respectively. Fully sunlit and (partly) shadowed pixels are identified as $Q \leq 0.1$ and $Q > 0.1$, respectively, where Q values are computed using the ESMLM model. Histogram of the absolute difference between manually and automatically extracted endmembers of parameter F in (a), Q in (b), P in (d), and K in (e).

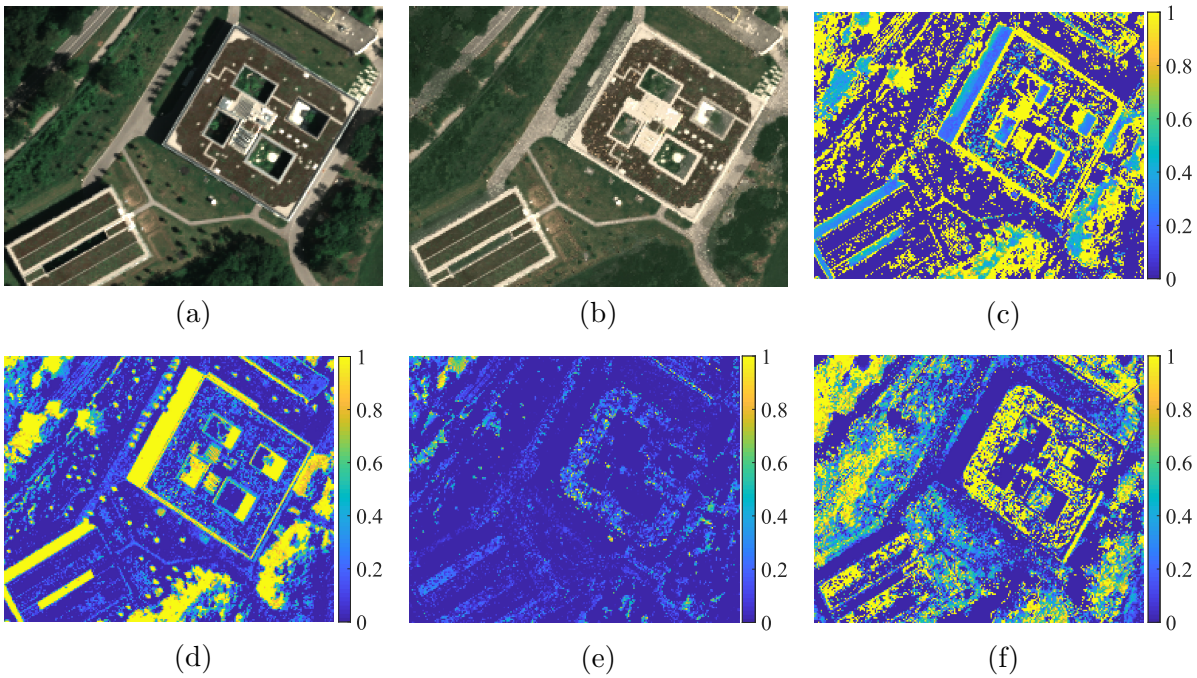


Figure 5.32: Results on the entire test image. (a) True composites of original image, (b) shadow-removed image. Parameter outputs F, Q, P, and K, are depicted in (c), (d), (e) and (f), respectively.

5.2.9 Summary

This section proposes an extended shadow multilinear (ESMLM) model for hyperspectral images based on radiative transfer theory, addressing shadow and nonlinear effects. The proposed model follows a graphical framework, and sums up all possible radiation paths initiated by the illumination sources. Three illumination sources are considered: direct, diffuse, and neighboring illuminations. The proposed model considers different shadow variants, i.e., fully shadowed pixels, partly shadowed and spatially separable pixels, and partly shadowed but spatially inseparable pixels. Additionally, two types of nonlinear interactions for different illumination conditions are modelled. Physical assumptions are made to simplify the ESMLM model, leading to four physically interpretable parameters: P (the probability that a light ray undergoes additional interactions with endmembers), Q (the spatial fraction of shadow), F (the sky view factor), and K (a strength factor of neighbor interactions, denoting the fraction of the scattered light from the neighborhood that is received by the pixel). Given these physically interpretable parameters as output, the proposed model characterizes the local structures of the ground surface and allows to reconstruct a shadow-removed image by simply "lighten up" the shadow-related terms.

The proposed model is compared with state-of-the-art mixing models on both synthetic dataset and real images with qualitative and quantitative measures. We first analyze reconstruction and abundance errors on simulated data with and without additional noise. After that, we simulate (partly) shadowed pixels in a real hyperspectral imagery with known abundance ground truth, and evaluate the performance of different mixing models. Furthermore, we analyze the unmixing models in airborne hyperspectral images with real shadowed pixels. Specifically, we discuss the reconstruction errors in spatial and spectral domains, and we compare shadow-removed images and model output parameters. Experimental results demonstrate that the proposed model performs consistently better in different ground scenarios with various illumination conditions. Moreover, we conduct an ablation study of the ESMLM model, in which we study the role and significance of each parameter separately. Experimental results demonstrate that the full model performs better than the ablated models.

Several open problems remain. Firstly, when including the skylight information, the shadow-removed images contain higher levels of noise, caused either by the low signal-to-noise ratio, or by strong nonlinear effects that take place in (partly) shadowed pixels. In addition, spectral errors of the proposed model, even though lower than in other models, remain large in some shadowed regions. Future work concerns including spatial information in the mixing model to promote spatial correlations among pixels.

5.3 Spatial-Spectral Shadow-Aware Mixing Model

Current shadow-aware hyperspectral unmixing methods often suffer from noisy abundance maps and inaccurate abundance estimation of shadowed pixels, as these are characterized by low reflectance values and signal-to-noise ratio. In order to achieve a shadow-insensitive abundance estimation, we propose a novel spatial-spectral shadow-aware mixing model (Spatial-spectral Shadow-aware Mixing (S3AM)). The approach models shadows by considering diffuse solar illumination and secondary illumination from neighbouring pixels. Besides, spatial regularization using shadow-aware weighted Total Variation is employed. Specifically, pixels in the local neighborhood of a target pixel take simultaneously into account spectral similarity measures derived from the imagery, elevation similarity measures derived from a Digital Surface Model, and the impact of shadows. The sky view factor F, needed as input for the model, is also derived from available Digital Surface Models (DSMs). The proposed approach is extensively validated and compared to state-of-the-art methods on two datasets. Results demonstrate that S3AM yields superior abundance estimation maps for real scenarios, by decreasing the noise in the results and achieving more accurate reconstructions in the presence of shadows.

5.3.1 Introduction

Spectral unmixing is a fundamental hyperspectral image analysis technique analyzing the composition of an image element at sub-pixel level [62, 66, 217]. A spectral mixing model describes how an incoming light ray from a given illumination source interacts with the targets on ground, before it is scattered back to the spectrometer [151]. In the past decades, numerous spectral mixing models with different physical assumptions have been proposed to tackle linear and nonlinear optical interactions [76, 86, 87, 89, 150, 161]. An introduction of spectral unmixing is detailed in Section 2.5.

One challenge for spectral unmixing is posed by shadow effects. Several methods have been developed to address this issue based on various strategies [149, 151, 162, 164, 165, 198, 199]. Their main idea is modeling spectra using endmembers and hyperparameters that can be relevant to shadow effects. Later, shadow-related parameters can be solved together with abundances, which are then used for computing shadow-removed images based on pixel reconstruction. Reviews on shadow-aware spectral unmixing methods are detailed in Chapter 3.

Nevertheless, several shadow-related challenges remain. Since the contribution from diffuse illumination is significantly smaller with respect to global illumination, modeling shadows with diffuse solar illumination terms may lead to over-fitting of the optimization problem. Moreover, the signal-to-noise ratio in shadowed pixels is much lower with respect to sunlit pixels. These factors lead to inaccurate abundance estimations and noisy abundance maps [199].

The above-mentioned problems may be solved by exploiting the spatial information in shadow-aware spectral unmixing methods. Spatial-spectral unmixing approaches consider the spatial dependence in local and non-local neighborhoods. On the one hand, in local

neighborhoods, abundances at a specific pixel are assumed to be strongly correlated to the ones of neighboring pixels [169, 175]; on the other hand, in non-local neighborhoods, similar patches in a larger region are assumed to share similar texture [176, 177].

In order to partly resolve the above challenges, we propose a spectral unmixing method with shadow-aware spatial constraints obtained from a hyperspectral image and a corresponding DSM generated by multi-view stereo imagery.

- Inspired by our previous work in Section 5.2, the spectral mixing model accounts for typical ground scenarios in the presence of shadows and nonlinear optical effects by considering multiple illumination sources: global solar illumination, diffuse solar illumination, and secondary illumination from neighbouring pixels. Specifically, global solar radiation is assumed to be the main illumination source for sunlit pixels, while diffuse solar radiation dominates in shadowed pixels. A ground pixel may receive light from some or all of the illumination sources, creating flexible spectral modeling for pixels with different illumination conditions.

- In order to alleviate the impact of shadows on the estimation of abundances, we propose a weighted TV constraint with shadow-insensitive weighting factors. Weighting factors are computed from the spectral angle and elevation differences between a target pixel and its neighboring pixels. In addition, a pre-computed shadow-related parameter is included in weighting factors in order to decrease the contribution of shadowed neighboring image elements of the target pixel.

- We inject elevation information from the DSM into the model in two ways. First, the elevation data provide illumination-insensitive TV weights, beneficial to the abundance estimation in shadowed pixels. Second, rather than being an additional model parameter, the sky view factor (F), required to calculate the contribution of diffuse solar illumination, is previously obtained from the elevation data, additionally decreasing the complexity of the spectral mixing model.

- We extensively validate the proposed method on two real hyperspectral images including shadows, both quantitatively and qualitatively. The proposed model significantly decreases the noise level in abundance maps, shows good robustness to shadow effects, and obtains more homogeneous abundance maps.

The remainder of this section is organized as follows. Section 5.3.2 introduces the proposed spatial-spectral shadow-aware mixing model with spatial constraints (S3AM). The proposed model is solved using the algorithm in Section 5.3.3. Section 5.3.4 describes the experimental setup, including datasets, compared methods, optimal parameter settings, and computational resources. Section 5.3.5 and Section 5.3.6 report and assess experimental results on two real datasets. Finally, we summarize our work and give prospects for future developments in Section 5.3.7.

5.3.2 Proposed Method

The proposed method is built upon the second contribution of this dissertation, namely the ESMLM model. We briefly remind the main concepts of the ESMLM model and its representation reported in Eq. (5.29). For further details, interested readers are referred to the extensive description in Section 5.2.

Denote a hyperspectral image with B spectral bands and N pixels as $\boldsymbol{X} \in \mathbb{R}^{B\times N}$, with $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N]$, where pixel $\boldsymbol{x}_j = (x_{j,1}, x_{j,2}, \cdots x_{j,B})^T \in \mathbb{R}^{B\times 1}$. An endmember library $\boldsymbol{E} \in \mathbb{R}^{B\times p}$ consists of p endmembers, where the i-th endmember is denoted as $\boldsymbol{e}_i \in \mathbb{R}^{B\times 1}$. The averaged spectrum in the first-order neighborhood of pixel $\boldsymbol{x}_j$ is denoted as $\boldsymbol{\chi}_j \in \mathbb{R}^{B\times 1}$. Denote an abundance matrix related to $\boldsymbol{E}$ as $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_N]$, with $\boldsymbol{a}_j \in \mathbb{R}^{p\times 1}$. In addition, four pixelwise parameters at pixel j are denotes as P_j, Q_j, K_j, F_j, and their corresponding vector forms are denoted as $\boldsymbol{P}$, $\boldsymbol{Q}$, $\boldsymbol{K}$, $\boldsymbol{F}$.

Spectral mixing models can be constructed by a ray-based description of the interaction of the incoming light with the ground materials [151, 199]. Following some physical assumptions, a light ray initiated from an illumination source interacts with ground materials with given probabilities before being scattered back to the sensor. The ESMLM model is based on ray-based descriptions. In order to account for shadows, the ESMLM model allows various illumination conditions in an image and accounts for typical scenarios related to the types and distribution of ground materials. Specifically, the ESMLM model considers three illumination sources: global solar illumination, diffuse solar illumination, and secondary illumination from neighbouring pixels. A light ray from each illumination source follows certain physical assumptions. For a given pixel $\boldsymbol{x}_j$, the ESMLM model sums up contributions of possible light rays initiated from all three illumination sources, as follows:

$$\begin{aligned}\boldsymbol{x}_j =&(1-Q_j)(1-P_j)\sum_{i=1}^{p} a_{j,i}\boldsymbol{e}_i + Q_j T(s_{0_{\text{diff}}}) \odot \sum_{i=1}^{p} a_{j,i}\boldsymbol{e}_i \\ &+ a_{j,i}\boldsymbol{e}_i P_j \sum_{i_1=1}^{p}\sum_{i_2=1}^{p} a_{j,i_1}a_{j,i_2}\boldsymbol{e}_{i_1} \odot \boldsymbol{e}_{i_2} \\ &+ (1-Q_j)(1-P_j)K_j \sum_{i=1}^{p} a_{j,i}\boldsymbol{e}_i \odot \boldsymbol{\chi}_j\end{aligned} \tag{5.29}$$

where $\sum_{i=1}^{p} a_{j,i} = 1$, $\forall i$: $a_{j,i} \geq 0$, and $P_j, Q_j, K_j \in [0,1]$.

$T(s_{0_{\text{diff}}})$ is computed as:

$$T(s_{0_{\text{diff}}}) = \frac{\boldsymbol{\tau}_{\text{diff}} \odot \boldsymbol{\mathcal{E}}_{\text{s}}}{\boldsymbol{\tau}_{\text{dir}} \odot \boldsymbol{\mathcal{E}}_{\text{l}} + \boldsymbol{\tau}_{\text{diff}} \odot \boldsymbol{\mathcal{E}}_{\text{s}}} \tag{5.30}$$

where $\frac{\tau_{\text{diff}}(\lambda)\mathcal{E}_{\text{s}}(\lambda)}{\tau_{\text{dir}}(\lambda)\mathcal{E}_{\text{l}}(\lambda)} = F(k_1\lambda^{-k_2} + k_3)$ with $k_1, k_2, k_3 > 0$.

The ESMLM model provides flexible nonlinear modeling with four parameters (P, Q, K and F) and accounts for different illumination conditions in an image element. Such flexibility brings challenges in solving the reverse problem due to the non-convexity of the objective

function. In particular, the ESMLM model becomes tri-convex, making it rather challenging to acquire a satisfying solution through the ADMM approach [97, 218].

In this section, we propose a spatial-spectral shadow-aware mixing (S3AM) model by embedding of spatial information. In order to make our problem bi-convex for an improved convergence ([97, 218]) we simplified the ESMLM model in two aspects. First, we set $P = 0$, because the inner-pixel second-order optical interactions have been observed to have minor impact on spectral unmixing results. Besides, we assume that the neighboring pixels contribute equally to a target pixel regardless of their illumination conditions, so that the neighbor illumination term is re-scaled solely according to parameter K.

The simplified model at pixel j is given by:

$$\begin{aligned}\boldsymbol{x}_j =&(1-Q_j)\boldsymbol{y}_j + Q_j\boldsymbol{y}_j \odot \widetilde{\boldsymbol{f}}_j + K_j\boldsymbol{y}_j \odot \boldsymbol{\chi}_j \\ =&(\mathbf{1}_B - \mathbf{1}_B Q_j) \odot \boldsymbol{y}_j + \mathbf{1}_B Q_j \odot \boldsymbol{y}_j \odot \widetilde{\boldsymbol{f}}_j \\ &+ \mathbf{1}_B K_j \boldsymbol{y}_j \odot \boldsymbol{\chi}_j \\ =&\boldsymbol{E} \odot (\mathbf{1}_B - \mathbf{1}_B Q_j + \mathbf{1}_B Q_j \odot \widetilde{\boldsymbol{f}}_j + \mathbf{1}_B K_j \odot \boldsymbol{\chi}_j)\mathbf{1}_p^T \boldsymbol{a}_j \\ =&\widetilde{\boldsymbol{E}}_j \boldsymbol{a}_j\end{aligned} \tag{5.31}$$

where

$$\widetilde{\boldsymbol{f}}_j = \frac{F_j \cdot (k_1 \boldsymbol{\lambda}^{-k_2} + k_3)}{1 + F_j \cdot (k_1 \boldsymbol{\lambda}^{-k_2} + k_3)} \tag{5.32}$$

$$\boldsymbol{y}_j = \boldsymbol{E}\boldsymbol{a}_j = \sum_{i=1}^{p} a_{j,i}\boldsymbol{e}_i \tag{5.33}$$

$$\widetilde{\boldsymbol{E}}_j = \boldsymbol{E} \odot ((\mathbf{1}_B - \mathbf{1}_B Q_j + \widetilde{\boldsymbol{f}}_j Q_j + \boldsymbol{\chi}_j K_j)\mathbf{1}_p^T) \tag{5.34}$$

We construct the optimization problem in vector form as:

$$\min_{\boldsymbol{a}_j, Q_j, K_j} \frac{1}{2} \sum_{j=1}^{N} \|\widetilde{\boldsymbol{E}}_j \boldsymbol{a}_j - \boldsymbol{x}_j\|_F^2 \tag{5.35}$$

The Abundance Non-negativity Constraint (ANC) and Abundance Sum-to-one Constraint (ASC) are applied on the abundances $\boldsymbol{a}_j$ [151, 199]. Additionally, we assume Q and $K \in [0, 1]$, in order to maintain their physical meanings:

$$\boldsymbol{a}_j \geq 0, \sum_{i=1}^{p} \boldsymbol{a}_{j,i} = 1, Q, K \in [0, 1] \tag{5.36}$$

Inspired by existing works on weighted total variation constraints for spectral unmixing ([84, 185]), the following spatial constraint on the abundances is proposed:

$$\sum_{j=1}^{N} \sum_{m \in \mathcal{N}(j)} R_{j,m} \|\boldsymbol{a}_j - \boldsymbol{a}_m\|_{1,1} \tag{5.37}$$

where $\mathcal{N}(j)$ denotes the first order neighborhood of the target pixel j. $R_{j,m}$ represents a

weighting factor describing the similarity between pixel j and m:

$$R_{j,m} = \frac{1}{Z_j}\left(R_{h_{j,m}} + R_{x_{j,m}}\right) \tag{5.38}$$

where $R_{h_{j,m}}$ and $R_{x_{j,m}}$ represent weighting factors related to height and spectral information, respectively. Z_j is the normalizing constant value and constraints the summation of weighting factors at pixel j to 1.

$R_{h_{j,m}}$ is defined as:

$$R_{h_{j,m}} = exp\left[-\frac{1}{\sigma_h^2}(1+\eta Q'_{j,m})T_{h_{j,m}}\right], \tag{5.39}$$

where σ_h^2 is a constant parameter controlling the weight range, and η reduces the influence of shadowed neighboring pixels on a target pixel using $Q'_{j,m}$, the shadow fraction at the neighboring pixel m for the target pixel j. $Q'_{j,m}$ is pre-calculated using the SLMM method [151]. Finally, $T_{h_{j,m}}$ is a height similarity measure, given by the normalized height difference between pixels j and m:

$$T_{h_{j,m}} = \frac{(h_j - h_m)^2}{(h_j + h_m)^2} \tag{5.40}$$

Normalized surface height above the ellipsoid h is provided by the DSM, which is illumination-insensitive and therefore robust to shadow effects. Thus, neighboring pixels with larger height similarities will have larger impact on the target pixel.

The weighting factor $R_{x_{j,m}}$ corresponds to spectral information:

$$R_{x_{j,m}} = exp\left[-\frac{1}{\sigma_x^2}(1+\eta Q'_{j,m})T_{x_{j,m}}\right], \tag{5.41}$$

where σ_x^2 is a constant parameter controlling the weight range, the shadow-related parameters η and $Q'_{j,m}$ are the same as in Eq. (5.39), and the spectral similarity measure $T_{x_{j,m}}$ is defined by the spectral angle [147, 153]:

$$T_{x_{j,m}} = \max(\arccos\frac{\boldsymbol{x}_j \cdot \boldsymbol{x}_m}{\|\boldsymbol{x}_j\|\|\boldsymbol{x}_m\|} - 0.1, 0) \tag{5.42}$$

Since shadow effects introduce spectral distortions [55, 103], the spectral angle between sunlit and shadowed pixels of the same material can be significantly larger than 0. We found this difference empirically to be around 0.1. In order to mitigate the impact of distortion in the spectral similarity measure, a value of 0.1 is then subtracted from the spectral angle in (5.42) up to a minimum value of 0.

Furthermore, as nonlinear effects typically do not depend on spectral, height, and shadow conditions, we apply a non-weighed total variation constraint on K:

$$\sum_{j=1}^{N}\sum_{m\in\mathcal{N}(j)}\|K_j - K_m\|_{1,1} \tag{5.43}$$

Hence, we define the optimization problem with spectral and spatial constraints as:

$$\begin{aligned}\min_{\boldsymbol{A},\boldsymbol{Q},\boldsymbol{K}} \frac{1}{2}\sum_{j=1}^{N}\|\widetilde{\boldsymbol{E}}_j\boldsymbol{a}_j-\boldsymbol{x}_j\|_F^2+\lambda_{\text{reg}}\|\boldsymbol{A}\boldsymbol{W}_1\|_{1,1}+\ell_{\mathcal{C}}(\boldsymbol{A})+\\ \ell_{\mathcal{S}}(\boldsymbol{A})+\ell_{\mathcal{M}}(\boldsymbol{Q})+\lambda_{\text{reg}}\|\boldsymbol{K}\boldsymbol{W}_2\|_{1,1}+\ell_{\mathcal{M}}(\boldsymbol{K})\end{aligned} \tag{5.44}$$

where $\ell_{\mathcal{C}}(\boldsymbol{A})=\{\boldsymbol{A}|\boldsymbol{A}\geq\boldsymbol{0}_{p\times N}\}$, $\ell_{\mathcal{S}}(\boldsymbol{A})=\{\boldsymbol{A}|\boldsymbol{1}_p^T\boldsymbol{A}=\boldsymbol{1}_N^T\}$, $\ell_{\mathcal{M}}(\boldsymbol{Q})=\{\boldsymbol{Q}|\boldsymbol{Q}\geq\boldsymbol{0}_{1\times N},\boldsymbol{Q}\leq\boldsymbol{1}_{1\times N}\}$ and $\ell_{\mathcal{M}}(\boldsymbol{K})=\{\boldsymbol{K}|\boldsymbol{K}\geq\boldsymbol{0}_{1\times N},\boldsymbol{K}\leq\boldsymbol{1}_{1\times N}\}$. The sparse matrix $\boldsymbol{W}_2=[\boldsymbol{W}_2^{\uparrow}\boldsymbol{W}_2^{\downarrow}\boldsymbol{W}_2^{\leftarrow}\boldsymbol{W}_2^{\rightarrow}]\in\mathbb{R}^{N\times 4N}$, where each element belongs to the set $\{-1,0,1\}$, consists of differential operators in four directions, i.e., up, down, left, and right. $\boldsymbol{K}\boldsymbol{W}_2$ computes the difference in K in each direction in the first-order neighborhood of each pixel. For instance, the difference in K in the upward direction at pixel j can be written as $K_{m^{\uparrow}}-K_j$, where $m^{\uparrow}$ denotes the index of the neighboring pixel in the upward direction of pixel j. Similarly, the sparse matrix $\boldsymbol{W}_1=[\boldsymbol{W}_1^{\uparrow}\boldsymbol{W}_1^{\downarrow}\boldsymbol{W}_1^{\leftarrow}\boldsymbol{W}_1^{\rightarrow}]\in\mathbb{R}^{N\times 4N}$ consists of differential operators in four directions weighted by the factor $R_{j,m}$. $\boldsymbol{A}\boldsymbol{W}_1$ computes the difference in A in each direction in the first order neighborhood of each pixel for each endmember, weighted by the factor $R_{j,m}$. For instance, the abundance difference in the upward direction at pixel j associated with endmember i can be written as $(a_{m^{\uparrow},i}-a_{j,i})R_{j,m^{\uparrow}}$, where $m^{\uparrow}$ denotes the index of the neighboring pixel in the upward direction of pixel j.

The above optimization is a bi-convex problem, and it is convex to $\boldsymbol{A}$ and $\{\boldsymbol{Q},\boldsymbol{K}\}$, respectively. Following [97, 218], we split the unknown variables into two groups and solve two convex problems sequentially using the ADMM approach. In the ADMM form, the optimization problem is given by:

$$\begin{aligned}&\min_{\boldsymbol{A},\boldsymbol{Q},\boldsymbol{K},\boldsymbol{G},\boldsymbol{H}} \frac{1}{2}\sum_{j=1}^{N}\|\widetilde{\boldsymbol{E}}_j\boldsymbol{a}_j-\boldsymbol{x}_j\|_F^2+\lambda_{\text{reg}}\|\boldsymbol{G}_2\|_{1,1}+\\ &\ell_{\mathcal{C}}(\boldsymbol{G}_3)+\ell_{\mathcal{S}}(\boldsymbol{G}_4)+\ell_{\mathcal{M}}(\boldsymbol{H}_1)+\lambda_{\text{reg}}\|\boldsymbol{H}_3\|_{1,1}+\ell_{\mathcal{M}}(\boldsymbol{H}_4)\\ &\begin{cases}\boldsymbol{G}_1=\boldsymbol{A}\\ \boldsymbol{G}_2=\boldsymbol{G}_1\boldsymbol{W}_1\\ \boldsymbol{G}_3=\boldsymbol{A}\\ \boldsymbol{G}_4=\boldsymbol{A}\end{cases}\quad\begin{cases}\boldsymbol{H}_1=\boldsymbol{Q}\\ \boldsymbol{H}_2=\boldsymbol{K}\\ \boldsymbol{H}_3=\boldsymbol{H}_2\boldsymbol{W}_2\\ \boldsymbol{H}_4=\boldsymbol{K}\end{cases}\end{aligned} \tag{5.45}$$

The solution of the optimization problem in Eq. (5.45) is reported in Algorithm 1 in Section 5.3.3.

5.3.3 The Solution of the Optimization Problem

This section presents the solution of the problem in Eq. (5.45) using Algorithm 1. Updating equations for primal and dual variables are given in the following subsections.

Algorithm 1: ADMM for the optimization problem in Eq. (5.45)

Input : $\boldsymbol{E}$, $\widetilde{\boldsymbol{f}}$, $\boldsymbol{X}$, $\boldsymbol{\chi}$, $\boldsymbol{\lambda}$, k_1, k_2, k_3, λ_{reg}, μ
Output : $\boldsymbol{A}$, $\boldsymbol{Q}$, $\boldsymbol{K}$
Initialize: $t = 0$, $\boldsymbol{A}^{(0)}$, $\boldsymbol{Q}^{(0)}$, $\boldsymbol{K}^{(0)}$, $\boldsymbol{G}^{(0)}$, $\boldsymbol{H}^{(0)}$, $\boldsymbol{U}^{(0)}$
while *the stopping criterion is not satisfied* **do**
  Given $\boldsymbol{Q}^{(t)}$, $\boldsymbol{K}^{(t)}$, $\boldsymbol{G}^{(t)}$, $\boldsymbol{U}^{(t)}$, update $\boldsymbol{A}^{(t+1)}$ with (5.48)
  Given $\boldsymbol{A}^{(t+1)}$, $\boldsymbol{U}^{(t)}$, update $\boldsymbol{G}^{(t+1)}$ with (5.49), (5.50), (5.51), (5.52)
  Given $\boldsymbol{A}^{(t+1)}$, $\boldsymbol{H}^{(t)}$, $\boldsymbol{U}^{(t)}$, update $\boldsymbol{Q}^{(t+1)}$ and $\boldsymbol{K}^{(t+1)}$ with (5.54), (5.55), (5.56)
  Given $\boldsymbol{Q}^{(t+1)}$, $\boldsymbol{K}^{(t+1)}$, $\boldsymbol{U}^{(t)}$, update $\boldsymbol{H}^{(t+1)}$ with (5.57), (5.58), (5.59), (5.60)
  Given $\boldsymbol{U}^{(t)}$, $\boldsymbol{A}^{(t+1)}$, $\boldsymbol{Q}^{(t+1)}$, $\boldsymbol{K}^{(t+1)}$, $\boldsymbol{G}^{(t+1)}$, $\boldsymbol{H}^{(t+1)}$, update $\boldsymbol{U}^{(t+1)}$ with (5.61)
  t = t + 1
end

5.3.3.1 Update $\boldsymbol{A}$ and $\boldsymbol{G}$

Given $\boldsymbol{Q}$, $\boldsymbol{K}$, and $\boldsymbol{H}$, the optimization problem in Eq. (5.45) can be rewritten as:

$$\begin{aligned}&\min_{\boldsymbol{A},\boldsymbol{G}} \frac{1}{2}\sum_{j=1}^{N} \|\widetilde{\boldsymbol{E}}_j\boldsymbol{a}_j - \boldsymbol{x}_j\|_F^2 + \lambda_{\text{reg}}\|\boldsymbol{G}_2\|_{1,1} + \ell_{\mathcal{C}}(\boldsymbol{G}_3) + \ell_{\mathcal{S}}(\boldsymbol{G}_4)\\ &\text{subject to} \begin{cases}\boldsymbol{G}_1 = \boldsymbol{A}\\ \boldsymbol{G}_2 = \boldsymbol{G}_1\boldsymbol{W}_1\\ \boldsymbol{G}_3 = \boldsymbol{A}\\ \boldsymbol{G}_4 = \boldsymbol{A}\end{cases}\end{aligned} \tag{5.46}$$

whose augmented Lagrangian is:

$$\begin{aligned}&\min_{\boldsymbol{A},\boldsymbol{G}} \frac{1}{2}\sum_{j=1}^{N} \|\widetilde{\boldsymbol{E}}_j\boldsymbol{a}_j - \boldsymbol{x}_j\|_F^2 + \lambda_{\text{reg}}\|\boldsymbol{G}_2\|_{1,1} + \ell_{\mathcal{C}}(\boldsymbol{G}_3) + \ell_{\mathcal{S}}(\boldsymbol{G}_4)\\ &+ \frac{\mu}{2}\|\boldsymbol{A} - \boldsymbol{G}_1 - \boldsymbol{U}_1\|_F^2 + \frac{\mu}{2}\|\boldsymbol{G}_1\boldsymbol{W}_1 - \boldsymbol{G}_2 - \boldsymbol{U}_2\|_F^2\\ &+ \frac{\mu}{2}\|\boldsymbol{A} - \boldsymbol{G}_3 - \boldsymbol{U}_3\|_F^2 + \frac{\mu}{2}\|\boldsymbol{A} - \boldsymbol{G}_4 - \boldsymbol{U}_4\|_F^2\end{aligned} \tag{5.47}$$

Thus, we can derive the optimizations with respect to $\boldsymbol{a}_j^{(t+1)}$:

$$\boldsymbol{a}_j^{(t+1)} = (\widetilde{\boldsymbol{E}}_j^T\widetilde{\boldsymbol{E}}_j + 3\mu\boldsymbol{I})^{-1}(\widetilde{\boldsymbol{E}}_j^T\boldsymbol{x}_j + \mu(\boldsymbol{J}_{1_j}^{(t)} + \boldsymbol{J}_{3_j}^{(t)} + \boldsymbol{J}_{4_j}^{(t)}) \tag{5.48}$$

where $\boldsymbol{J}_{1_j}^{(t)} = \boldsymbol{G}_{1_j}^{(t)} + \boldsymbol{U}_{1_j}^{(t)}$, $\boldsymbol{J}_{3_j}^{(t)} = \boldsymbol{G}_{3_j}^{(t)} + \boldsymbol{U}_{3_j}^{(t)}$, and $\boldsymbol{J}_{4_j}^{(t)} = \boldsymbol{G}_{4_j}^{(t)} + \boldsymbol{U}_{4_j}^{(t)}$.

Next, the optimizations with respect to $\boldsymbol{G}_1^{(t+1)}$, $\boldsymbol{G}_2^{(t+1)}$, $\boldsymbol{G}_3^{(t+1)}$, $\boldsymbol{G}_4^{(t+1)}$ are written as:

$$\boldsymbol{G}_1^{(t+1)} = \left[\boldsymbol{A}^{(t+1)} - \boldsymbol{U}_1^{(t)} + (\boldsymbol{G}_2^t + \boldsymbol{U}_2^{(t)})\boldsymbol{W}_1^T\right]\left[\boldsymbol{I} + \boldsymbol{W}_1\boldsymbol{W}_1^T\right]^{-1} \tag{5.49}$$

$$\boldsymbol{G}_2^{(t+1)} = soft\left(\boldsymbol{G}_1^{(t)}\boldsymbol{W}_1 - \boldsymbol{U}_2^{(t)}, \frac{\lambda_{\text{reg}}}{\mu}\right) \tag{5.50}$$

$$\boldsymbol{G}_3^{(t+1)} = max\left(\boldsymbol{A}^{(t+1)} - \boldsymbol{U}_3^{(t)}, 0\right) \tag{5.51}$$

$$\boldsymbol{G}_4^{(t+1)} = \left(\boldsymbol{A}^{(t+1)} - \boldsymbol{U}_4^{(t)}\right) + \frac{1}{p}\left[\mathbf{1}_N^T - \mathbf{1}_p^T\left(\boldsymbol{A}^{(t+1)} - \boldsymbol{U}_4^{(t)}\right)\right] \otimes \mathbf{1}_p \tag{5.52}$$

5.3.3.2 Update $\boldsymbol{Q}$, $\boldsymbol{K}$, and $\boldsymbol{H}$

Given $\boldsymbol{A}$ and $\boldsymbol{G}$, the optimization problem in Eq. (5.45) can be rewritten as:

$$\begin{aligned}
&\min_{\boldsymbol{Q},\boldsymbol{K},\boldsymbol{H}} \frac{1}{2}\sum_{j=1}^{N} \|\widetilde{\boldsymbol{E}_j}\boldsymbol{a}_j - \boldsymbol{x}_j\|_F^2 + \ell_{\mathcal{M}}(\boldsymbol{H}_1) + \lambda_{\text{reg}}\|\boldsymbol{H}_3\|_{1,1} + \ell_{\mathcal{M}}(\boldsymbol{H}_4) \\
&\text{subject to} \begin{cases} \boldsymbol{H}_1 = \boldsymbol{Q} \\ \boldsymbol{H}_2 = \boldsymbol{K} \\ \boldsymbol{H}_3 = \boldsymbol{H}_2\boldsymbol{W}_2 \\ \boldsymbol{H}_4 = \boldsymbol{K} \end{cases}
\end{aligned} \tag{5.53}$$

whose augmented Lagrangian is:

$$\begin{aligned}
&\min_{\boldsymbol{Q},\boldsymbol{K},\boldsymbol{H}} \frac{1}{2}\sum_{j=1}^{N} \|\widetilde{\boldsymbol{E}}_j\boldsymbol{a}_j - \boldsymbol{x}_j\|_F^2 + \ell_{\mathcal{M}}(\boldsymbol{H}_1) + \lambda_{\text{reg}}\|\boldsymbol{H}_3\|_{1,1} + \\
&\ell_{\mathcal{M}}(\boldsymbol{H}_4) + \frac{\mu}{2}\|\boldsymbol{Q} - \boldsymbol{H}_1 - \boldsymbol{U}_5\|_F^2 + \frac{\mu}{2}\|\boldsymbol{K}_1 - \boldsymbol{H}_2 - \boldsymbol{U}_6\|_F^2 \\
&+ \frac{\mu}{2}\|\boldsymbol{H}_2\boldsymbol{W}_2 - \boldsymbol{H}_3 - \boldsymbol{U}_7\|_F^2 + \frac{\mu}{2}\|\boldsymbol{K} - \boldsymbol{H}_4 - \boldsymbol{U}_8\|_F^2
\end{aligned}$$

The optimizations with respect to $\boldsymbol{Q}$ and $\boldsymbol{K}$ are solved pixelwise. For pixel j, we update $Q_j^{(t+1)}$ and $K_j^{(t+1)}$ using:

$$Q_j^{(t+1)} = \frac{C_3C_4 - C_2C_5}{C_1C_4 - C_2^2} \tag{5.54}$$

$$K_j^{(t+1)} = \frac{C_2C_3 - C_1C_5}{C_2^2 - C_1C_4} \tag{5.55}$$

where

$$\begin{cases} C_1 = & \sum_{i=1}^{B} (\widetilde{f}_{j,i} - 1)^2 y_{j,i}^2 + \mu \\ C_2 = & \sum_{i=1}^{B} (\widetilde{f}_{j,i} - 1) \chi_{j,i} y_{j,i}^2 \\ C_3 = & \sum_{i=1}^{B} (\widetilde{f}_{j,i} - 1) y_{j,i} (x_{j,i} - y_{j,i}) + \mu \boldsymbol{J}_{5_j}^{(t)} \\ C_4 = & \sum_{i=1}^{B} \chi_{j,i}^2 y_{j,i}^2 + 2\mu \\ C_5 = & \sum_{i=1}^{B} \chi_{j,i} y_{j,i} (x_{j,i} - y_{j,i}) + \mu \left(\boldsymbol{J}_{6_j}^{(t)} + \boldsymbol{J}_{8_j}^{(t)} \right) \end{cases} \tag{5.56}$$

with $\boldsymbol{y}_j = \boldsymbol{E}_j \boldsymbol{a}_j^{(t+1)}$, $\boldsymbol{J}_{5_j}^{(t)} = \boldsymbol{H}_{1_j}^{(t)} + \boldsymbol{U}_{5_j}^{(t)}$, $\boldsymbol{J}_{6_j}^{(t)} = \boldsymbol{H}_{2_j}^{(t)} + \boldsymbol{U}_{6_j}^{(t)}$, and $\boldsymbol{J}_{8_j}^{(t)} = \boldsymbol{H}_{4_j}^{(t)} + \boldsymbol{U}_{8_j}^{(t)}$.

Next, we optimize the objective function with respect to $\boldsymbol{H}_1^{(t+1)}$, $\boldsymbol{H}_2^{(t+1)}$, $\boldsymbol{H}_3^{(t+1)}$, and $\boldsymbol{H}_4^{(t+1)}$ using Eqs. (5.57), (5.58), (5.59), (5.60):

$$\boldsymbol{H}_1^{(t+1)} = \min \left(\max \left(\boldsymbol{Q}^{(t+1)} - \boldsymbol{U}_5^{(t)}, 0 \right), 1 \right) \tag{5.57}$$

$$\boldsymbol{H}_2^{(t+1)} = \left[\boldsymbol{K}^{(t+1)} - \boldsymbol{U}_6^{(t)} + (\boldsymbol{H}_3^t + \boldsymbol{U}_7^{(t)}) \boldsymbol{W}_2^T \right] \left[\boldsymbol{I} + \boldsymbol{W}_2 \boldsymbol{W}_2^T \right]^{-1} \tag{5.58}$$

$$\boldsymbol{H}_3^{(t+1)} = soft \left(\boldsymbol{H}_2^{(t)} \boldsymbol{W}_2 - \boldsymbol{U}_7^{(t)}, \frac{\lambda_{\text{reg}}}{\mu} \right) \tag{5.59}$$

$$\boldsymbol{H}_4^{(t+1)} = \min \left(\max \left(\boldsymbol{K}^{(t+1)} - \boldsymbol{U}_8^{(t)}, 0 \right), 1 \right) \tag{5.60}$$

5.3.3.3 Update $\boldsymbol{U}^{(t+1)}$

$$\begin{cases} \boldsymbol{U}_1^{(t+1)} = \boldsymbol{U}_1^{(t)} - \boldsymbol{A}^{(t+1)} + \boldsymbol{G}_1^{(t+1)} \\ \boldsymbol{U}_2^{(t+1)} = \boldsymbol{U}_2^{(t)} - \boldsymbol{G}_1^{(t+1)} W_1 + \boldsymbol{G}_2^{(t+1)} \\ \boldsymbol{U}_3^{(t+1)} = \boldsymbol{U}_3^{(t)} - \boldsymbol{A}^{(t+1)} + \boldsymbol{G}_3^{(t+1)} \\ \boldsymbol{U}_4^{(t+1)} = \boldsymbol{U}_4^{(t)} - \boldsymbol{A}^{(t+1)} + \boldsymbol{G}_4^{(t+1)} \\ \boldsymbol{U}_5^{(t+1)} = \boldsymbol{U}_5^{(t)} - \boldsymbol{Q}^{(t+1)} + \boldsymbol{H}_1^{(t+1)} \\ \boldsymbol{U}_6^{(t+1)} = \boldsymbol{U}_6^{(t)} - \boldsymbol{K}^{(t+1)} + \boldsymbol{H}_2^{(t+1)} \\ \boldsymbol{U}_7^{(t+1)} = \boldsymbol{U}_7^{(t)} - \boldsymbol{H}_2^{(t+1)} \boldsymbol{W}_2 + \boldsymbol{H}_3^{(t+1)} \\ \boldsymbol{U}_8^{(t+1)} = \boldsymbol{U}_8^{(t)} - \boldsymbol{K}^{(t+1)} + \boldsymbol{H}_4^{(t+1)} \end{cases} \tag{5.61}$$

5.3.4 Experimental Setup

5.3.4.1 Datasets

We validate the proposed method on the extended DLR HySU and HySpex/3K datasets. A detailed description of these datasets are presented in Chapter 4.

- **Extended DLR HySU**
  This experiment uses the hyperspectral images with simulated shadows, both with and without additional random noise ($SNR = 30$). In addition, the DSM is used as input for the proposed method and is also used to compute the sky view factor F based on [219].

- **HySpex/3K**
  A subregion is selected from the HySpex/3K dataset (Fig. 5.33). An airborne hyperspectral image is shown in Fig. 5.33 (a). Endmembers $\boldsymbol{E}$ have been extracted from the fully sunlit pixels using the method in [199, 210] (see Fig. 5.33 (b)). Moreover, multi-view stereo imagery acquired with the 3K camera system was employed to generate the DSM [170] in Fig. 5.33 (c), whose values represent surface height above the ellipsoid. In addition, height values were normalized within $[0, 1]$ to retain the relative height of the ground surface. After geometrical co-registration and re-sampling, the DSM and images share the same geo-coordinates and spatial resolution (i.e., 0.7 m). Given the height data, the sky view factor F was computed using the software System for Automated Geoscientific Analyses (SAGA) [219] (see Fig. 5.33 (d)).

5.3.4.2 Methods in Comparison

We compare the S3AM method to the following state-of-the-art spectral mixing models. A review of these methods can be found in Chapter 3.

- LMM [76]: a linear mixing model that does not take shadows into account.
- SLMM [149]: a linear mixing model accounting for shadows using a scaling factor while ignoring diffuse solar illumination.
- SMLM [151]: a nonlinear mixing model accounting for shadows using a scaling factor as in SLMM, along with nonlinear interactions using the multilinear mixing model [150].
- Shadow Removal Method based on Nonlinear Unmixing and Endmember Matching (NUEM) [155]: a shadow removal method based on nonlinear unmixing and endmember matching. This method requires a shadow mask as input, which has been computed by S3AM in the experiment.
- Fansky [198]: a nonlinear mixing model considering shadows based on both direct and diffuse solar illumination, along with nonlinear interactions using the Fan model [87].
- ESMLM [199]: an extended SMLM model incorporating direct and diffuse solar illumination as well as nonlinear interactions.

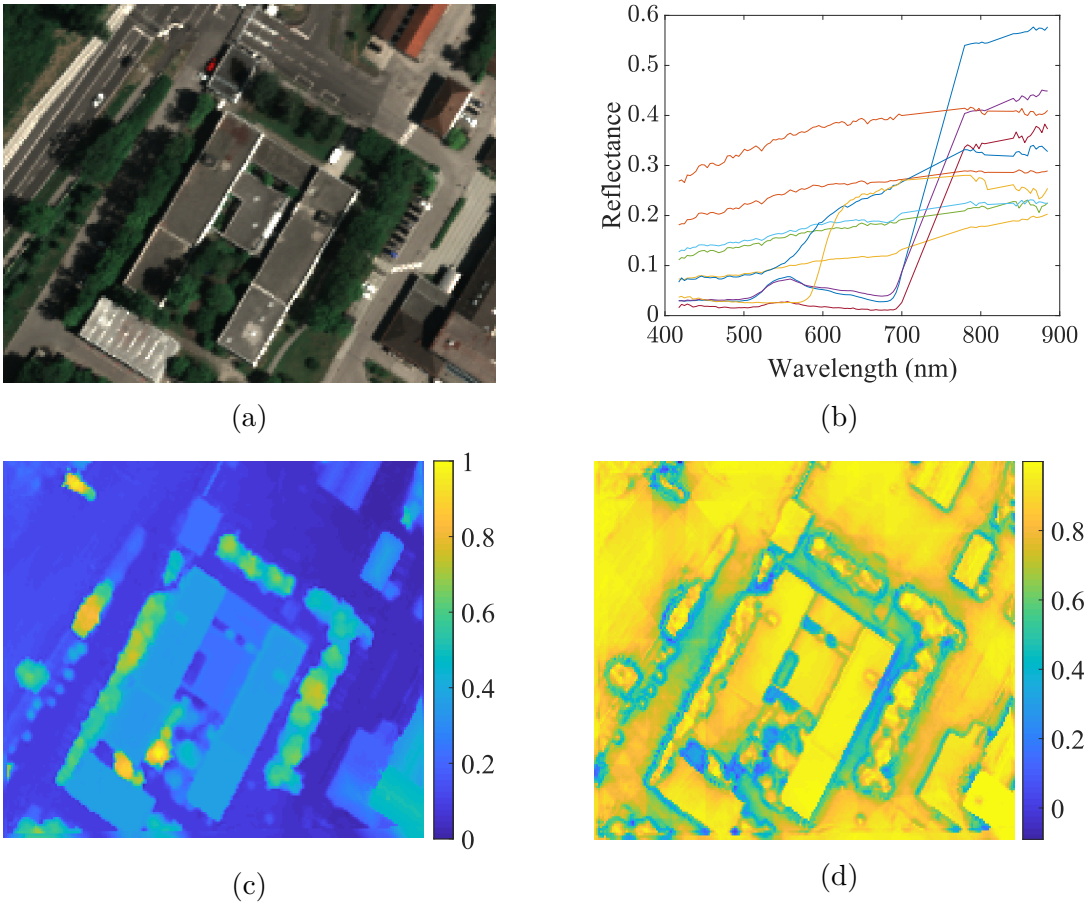


Figure 5.33: A subscene selected from the HySpex/3K dataset: (a) hyperspectral image as a true color composite acquired by the HySpex sensor in the study area of Oberpfaffenhofen, Bavaria, Germany; (b) endmember library, automatically extracted from (a); (c) normalized DSM; (d) sky view factor map derived from the DSM in (c).

- SCBMM [165]: a nonlinear mixing model considering shadows using both direct and diffuse solar illumination and nonlinear interactions using the bilinear mixing model [86].

Moreover, we investigate the impact of different variations of the spatial TV regularization in an ablation study.

5.3.4.3 Parameter Settings

- k_1, k_2, k_3
  Following our previous work in [199], k_1, k_2, k_3 are the parameters of a power function that models the ratio of diffuse to global solar irradiance on the ground surface. We assume that atmospheric conditions are consistent in the entire image region, so these parameters are constant in our experiments. In practice, we compute these parameters using ten pairs of pixels that are selected in the scene. Specifically, a pair of pixels contain a sunlit pixel and a fully shadowed pixel near a shadow boundary, assumed to consist of the same material. We avoid including vegetation during the selection in

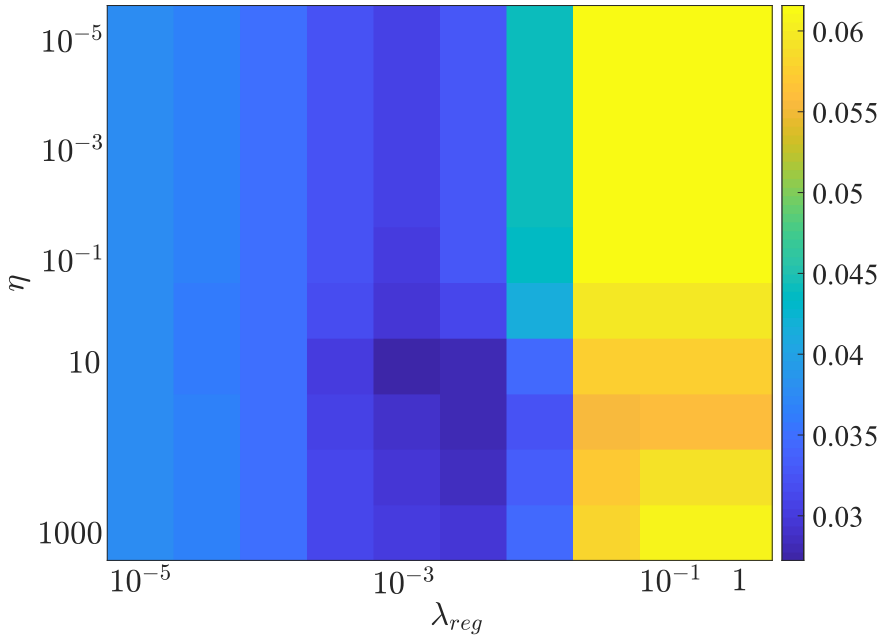


Figure 5.34: *AE* as a function of λ_{reg} and η for the DLR HySU dataset, where $\lambda_{reg} \in \{10^{-5}, 5\cdot10^{-5}, 10^{-4}, 5\cdot10^{-4}, 10^{-3}, 5\cdot10^{-3}, 0.01, 0.05, 0.1, 1\}$ and $\eta \in \{10^{-5}, 10^{-4}, 10^{-3}, 0.01, 0.1, 1, 10, 100, 500, 1000\}$.

order to avoid dealing with complex nonlinear effects. Then, k_1, k_2 and k_3 are solved by Eq. (5.31), with $K_j = 0$, $Q_j = 1$ and $p = 1$. The obtained values are: $k_1 = 0.579$; $k_2 = 6.974$; and $k_3 = 0.206$.

- λ_{reg} **and** η
  Fig. 5.34 presents *AE* as a function of λ_{reg} and η for the DLR HySU dataset. We calculate optimal values of these parameters from the values $\lambda_{\text{reg}} \in \{10^{-5}, 5\cdot10^{-5}, 10^{-4}, 5\cdot10^{-4}, 10^{-3}, 5\cdot10^{-3}, 0.01, 0.05, 0.1, 1\}$ and $\eta \in \{10^{-5}, 10^{-4}, 10^{-3}, 0.01, 0.1, 1, 10, 100, 500, 1000\}$ by minimizing *AE*, resulting in $\lambda_{\text{reg}} = 10^{-3}$ and $\eta = 10$ for S3AM. As ground truth abundances are not available for the HySpex dataset, we empirically determine λ_{reg} and η to be the same as for the DLR HySU dataset. Additionally, we analyze the impact of different values of λ_{reg} on the abundance maps in Section 5.3.6.

- σ_x^2 **and** σ_h^2
  These parameters represent the weight range in the exponential functions of the height-related (Eq. (5.39)) and spectral (Eq. (5.41)) weighting factors, respectively. In principle, one can optimize the values of σ_x^2 and σ_h^2 in a similar way as λ_{reg} and η, by minimizing the optimization error. However, too many free parameters can lead to over-fitting. In practice, more than 99.9% of T_x and T_h was found to lie within the range $[0, 0.5]$. Hence, we set the weighting ranges within the same span, and choose empirical values $\sigma_x^2 = \sigma_h^2 = 0.1$.

- μ
  The penalty parameter μ of the augmented Lagrangian (see Eq. (5.47)) was determined as in [82]. The initial value is set to $\mu = 0.001$, and is then updated iteratively by

keeping the ratio between primal and dual residual norms within a positive value of 10, as suggested in [97].

- **Initialization and Stopping Criteria**
  A fully constrained spectral unmixing method [196] based on the SLMM model has been applied to initialize $\boldsymbol{A}$ and $\boldsymbol{Q}$, while $\boldsymbol{K}$ and $\boldsymbol{U}$ are initialized to zero. In addition, the algorithm stops when the primal residual is less than $5 \cdot 10^{-4}$ or the maximum number of iterations, set as 100, is reached.

5.3.4.4 Computational Resources

All algorithms were developed in a MATLAB environment and run on an Intel Core i7 −8650 U CPU, 1.90 GHz machine with 4 Cores and 8 Logical Processors. We apply the MATLAB function FMINCON using the Sequential Quadratic Programming algorithm to perform the (non)linear optimization for LMM, SLMM, SMLM, NUEM, Fansky, and ESMLM. The function and constraint tolerance are set to 10^{-6} and 10^{-5}, respectively. In addition, for-loop iterations over all pixels were running in parallel on workers in a parallel pool.

The Fansky and SCBMM have the highest computational cost, depending on the number of pixels in the subset. Besides, the nonlinear optimization implemented by FMINCON in ESMLM, SMLM and SLMM requires more computational resources with respect to the linear optimization in LMM. The S3AM, solved by the ADMM approach has a short computation time, as MATLAB efficiently computes closed-form updates for this method.

Table 5.9: Running time of compared methods

Method	Running time (s)		
	HySU (208 pixel)	subset1 (1148 pixel)	subset 2 (1085 pixel)
LMM	0.47	2.56	2.27
SLMM	0.61	5.80	5.38
SMLM	0.70	9.42	8.33
NUEM	0.63	6.01	4.65
Fansky	7.42	112.75	109.20
ESMLM	2.59	26.66	25.47
SCBMM	11.43	42.25	40.43
S3AM	0.55	2.80	2.71

5.3.5 Results: DLR HySU

5.3.5.1 Pixel Reconstruction and Abundance Estimation

Fig. 5.35 shows the obtained *AE* and *RE* for all methods, from sunlit, (partly) shadowed and all pixels, respectively. The *AE* was obtained without taking grass into account. The reason for this is that reference abundance values were derived by LMM. In grass areas, nonlinear effects may be present, and this class contains non-negligible intraclass variations

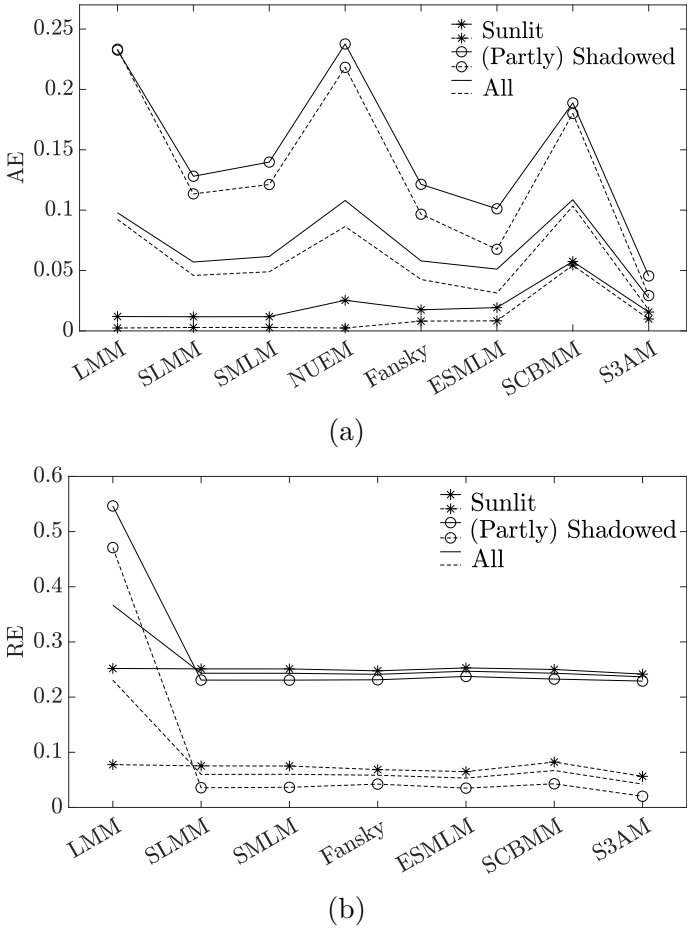


Figure 5.35: Comparison of mean abundance error AE in (a) and mean reconstruction error RE in (b) for the DLR HySU dataset. Solid and dashed lines represent results obtained using the input image with and without additional noise ($SNR = 30$), respectively. Sunlit and (partly) shadowed pixels are determined with $Q \leq 0.1$ and $Q > 0.1$, respectively.

with respect to the spectrum selected as endmember. Therefore, reference abundances for grass may be unreliable. In addition, NUEM is not included in the comparison of RE, as this method runs two unmixing processes followed by spectral matching. Furthermore, we show AE and RE on a degraded image with additive noise ($SNR = 30$), in order to evaluate the robustness to noise of the compared methods.

All compared methods show satisfactory REs in sunlit regions. In shadowed regions, the LMM obtains significantly higher RE compared to other methods, indicating the importance of shadow-aware modeling. Nevertheless, smaller REs do not necessarily imply a satisfactory abundance estimation. Compared to REs, we observe significantly larger differences of AEs among compared methods. In general, the better a model accounts for shadows, the better the abundance estimation. The LMM has the lowest performance, because shadow effects are simply ignored. The NUEM is also characterized by a large AE. Since shadow effects cause a wavelength-dependent spectral distortion, it is quite challenging to perform spectral matching between sunlit and shadowed pixels. Moreover, the SLMM and SMLM treat shadows as a scaling effect, and perform better at abundance estimation in shadowed

regions. The performance improves further when including the diffuse solar illumination as prior knowledge. Fansky, ESMLM, and S3AM select pairs of pixels from the input image as prior knowledge and estimate the ratio of diffuse and global solar illumination through a power function. The proposed approach, i.e., S3AM, clearly outperforms others thanks to the applied spatial constraints. In addition, the ratio of diffuse and global solar illuminations can be estimated along with abundance values in the unmixing process, i.e., SCBMM. Nevertheless, its *AE* appears higher than other methods. The reason is that the SCBMM does not assume stronger atmospheric scattering at shorter wavelengths [55]. Hence, its estimated diffuse radiation may not correspond to the spectral characteristics of shadows in practice. Furthermore, results in Fig. 5.35 show that both *RE*s and *AE*s increase considerably as the image is degraded by noise. Despite better abundance estimation, the Fansky and ESMLM are less robust to noise compared to LMM, SLMM, SMLM, NUEM, and SCBMM. Since the contribution of diffuse solar radiation is significantly smaller relative to global radiation, the optimization problem can be over-fitted and may lead to noisy abundance maps. This sensitivity to noise can be significantly alleviated by the proposed method S3AM, thanks to the spatial constraints.

A qualitative comparison of the abundance maps is displayed in Fig. 5.36. In order to assess the location of abundance errors, we overlay the *AE* map in grayscale as a semi-transparent layer over the optical image in Fig. 5.37. First of all, it can be observed that abundance errors are mainly located in shadowed regions. Compared to the LMM, the shadow-aware unmixing methods therein display improved abundance maps. While SLMM and SMLM treat shadow as a scaling effect, the SMLM further considers nonlinear reflections. Since the study region is a flat terrain with artificial materials, multiple reflections appear minor. Thus, the abundance maps of SLMM and SMLM are very similar. Unlike embedding a shadow-related parameter in the model, the NUEM matches spectra in shadowed areas with their corresponding ones in sunlit regions. However, accurate spectral matching is rather challenging due to spectral distortion caused by shadows. In particular, this dataset contains some materials with similar spectral information, making spectral matching more difficult. For example, the NUEM easily confuses red-painted metal sheets and red fabric.

Furthermore, we compare methods that take into account diffuse solar illumination, i.e., SCBMM, Fansky, ESMLM, and S3AM. The SCBMM may have the ability to estimate abundance values in shadowed regions, such as bitumen, while it seems challenging to estimate diffuse solar illumination without prior knowledge. Fansky, ESMLM, and S3AM, on the other hand, whose diffuse solar illumination is computed from manually selected pixels in the input image, perform better in abundance estimation.

However, without applying spatial constraints, abundance maps show higher noise levels and confusion between similar materials, such as red-painted metal sheets and red fabric. Such distortions can be alleviated by injecting spatial information into the analysis. By applying weighted TV constraints, S3AM considerably improves the abundance estimation step. Firstly, the noise level has been significantly reduced owing to the spatial constraints. In addition, the abundance estimation at the boundary pixels is significantly improved, due to the weighting of the spectral and height information in the spatial constraints.

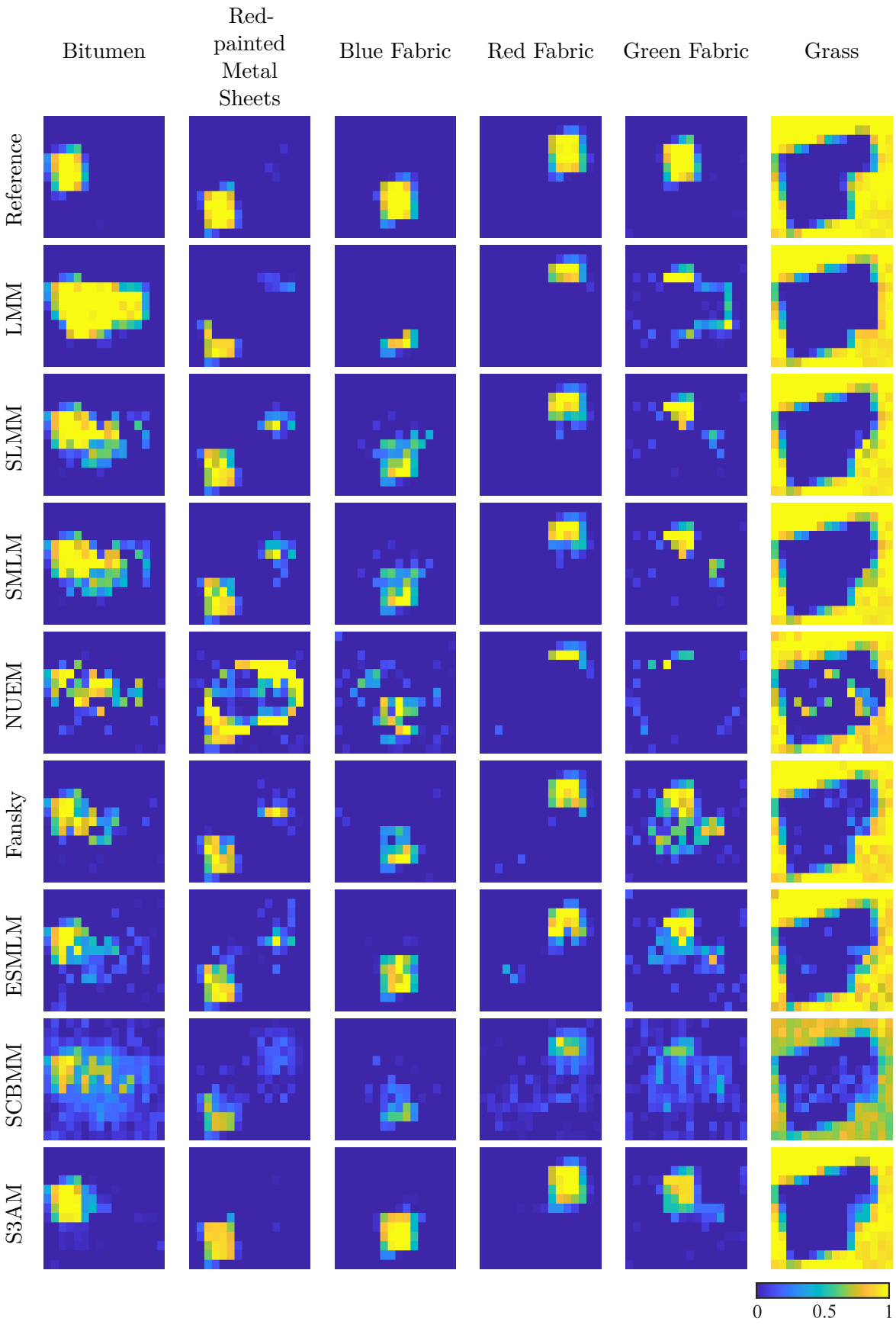


Figure 5.36: Abundance maps for the HySU dataset. Left to right: bitumen, red metal sheets, blue fabric, red fabric, green fabric, and grass. Top to bottom: reference, LMM, SLMM, SMLM, NUEM, Fansky, ESMLM, SCBMM, S3AM. The reference abundance maps are computed by applying fully constrained least squares unmixing using the library of known endmembers on the shadow-free image.

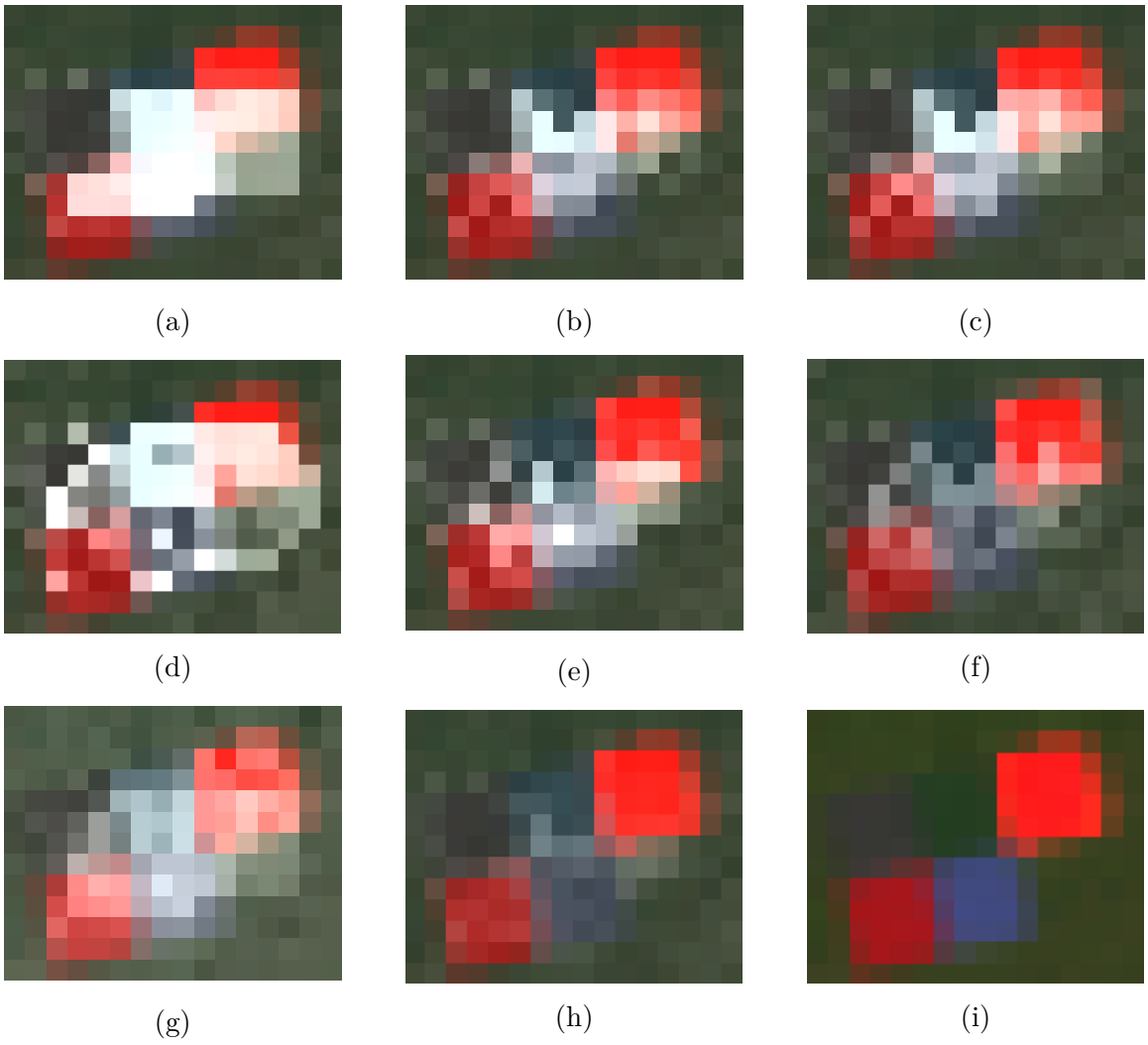


Figure 5.37: DLR HySU imagery overlaid with pixelwise mean abundance error maps in grayscale: (a) LMM, (b) SLMM, (c) SMLM, (d) NUEM, (e) Fansky, (f) ESMLM, (g) SCBMM, (h) S3AM, (i) reference image.

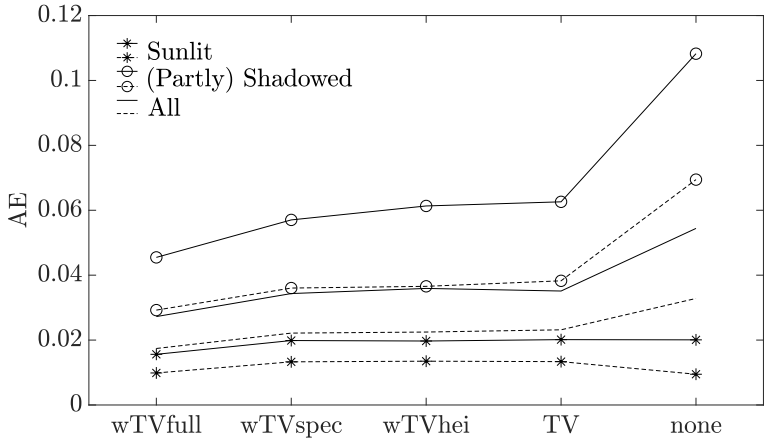


Figure 5.38: Comparison of mean abundance error AE of S3AM with ablated spatial constraints for the DLR HySU dataset. Solid and dashed lines represent results obtained on the input image with and without additional noise ($SNR = 30$), respectively.

5.3.5.2 Ablation Study

The S3AM method consists of a weighted TV constraint term, where the weights are formed by spectral and height features. In the ablation study, we investigate the individual contribution from each feature.

We refer to the weighted TV (Eq. (5.37)) in the S3AM method as wTVfull, where both height and spectral features are included and computed by Eqs. (5.39) and (5.41). In the ablation study, we regard height and spectral features one at a time by setting $R_{h_{j,m}} = 0$ and $R_{x_{j,m}} = 0$ in Eq. ((5.38)), respectively, resulting in the ablated TV forms wTVhei and wTVspec. In addition, we set the weights $R_{h_{j,m}} = R_{x_{j,m}} = 1$ in Eq. (5.38), resulting in a classic non-weighted TV, labeled as TV. Moreover, we ablate the spatial constraints entirely by setting $\lambda_{\text{reg}} = 0$ in Eq. (5.44), reducing the method to only the spectral mixing model, labeled as "none".

Fig. 5.38 compares AE of wTVfull, wTVspec, wTVhei, TV, and "none" in the ablation study. Specifically, we investigate AE in sunlit, (partly) shadowed, and all regions. Compared to sunlit regions, where the spatial constraints play a minor role, we observe considerable improvement in shadowed pixels by embedding spatial constraints, with wTVfull achieving the best abundance estimation, both with and without additional noise.

Figs. 5.39 and 5.40 show respectively abundance and abundance error maps. When no spatial constraints are applied, i.e., "none", resulting abundance maps are noisy. The ablation study shows that differences in abundances mainly appear on boundaries between different materials. Typically, TV oversmooths the boundaries between different materials, since it treats neighboring pixels equally. One example is visible at the transition from bitumen to green fabric. When applying weighted TV, the abundance estimation on the boundary pixels is largely improved and wTVspec better preserves the shape of the bitumen target with respect to wTVhei, because spectral information can better separate the two materials on the boundary between them. Since in our experiment the DSM is randomly generated, the height information does not fully correspond to the ground objects, leading to inaccurate

abundance estimations. In practice, spectral or height information alone might not be sufficient to distinguish ground materials. The configuration denoted as wTVfull, adopted by the proposed method S3AM, jointly considers spectral and height features and outperforms single-source weights both quantitatively and qualitatively. Specifically, wTVfull constraints visibly improve the abundance estimation on the boundary pixels, e.g., at the right border of bitumen and the bottom side of green fabric.

5.3.5.3 Shadow-Removed Pixel Reconstruction

Spectral mixing models considering shadows allow generating shadow-removed imagery through pixel reconstruction. Specifically, in SLMM, SMLM, ESMLM, and S3AM, the parameter Q represents the shadow fraction within a pixel. By setting $Q = 0$, shadows are removed during pixel reconstruction [199]. The NUEM, Fansky, and SCBMM methods compute abundance values separately in sunlit and shadowed groups. Then, shadow-removed pixels can be reconstructed using summed-up abundance values in the two groups and sunlit endmember spectra. Obviously, a better abundance estimation leads to better reconstruction and shadow removal. Fig. 5.41 compares the shadow-removed images computed by the different models. Despite confusion between bitumen and green fabric, the SLMM achieves a satisfactory restoration in shadowed areas, which is remarkable considering its simplicity. The SMLM shows a similar abundance estimation as the SLMM (see Fig. 5.36). However, some dark pixels appear in the reconstructed image of the SMLM, because of its incorrect estimation of parameters P and Q [198]. Large areas of red-painted metal sheets appear in the restored image of NUEM, because of the mismatch between endmembers extracted in sunlit and shadowed regions. For example, when minimizing the spectral angle, the spectrum of shadowed red fabric is associated with the red-painted metal sheets. In addition, the spectrum of shadowed green fabric is associated with bitumen. Moreover, Fansky, ESMLM, and SCBMM show material confusion between different red materials. The material boundaries in shadowed areas also appear reddish due to incorrect abundance estimation. Overall, the proposed method S3AM achieves the best qualitative shadow-removed image, thanks to its superior abundance estimation and spatial constraints.

5.3.6 Results: HySpex/3K

5.3.6.1 Abundance Estimation

Unlike for the DLR HySU dataset, we do not have ground-truth abundances for the HySpex/3K dataset. Therefore, in this section, we investigate and compare the methods qualitatively on two image subsets (Figs. 5.42 and 5.43). Fig. 5.42 compares the abundances of the roof material. The LMM seems to perform satisfactorily in shadowed pixels. However, it is worth noticing that LMM easily overestimates impervious surfaces, as can be observed in subset 2, where many vegetation pixels are incorrectly recognized as impervious surfaces (Fig. 5.43 (c)). Besides, the SLMM, SMLM, and NUEM show lower abundances in shadowed pixels on the roof (Fig. 5.42 (d)-(f)). Specifically, they confuse the spectra of the roof with other impervious materials (see Fig. 5.46), since these two materials

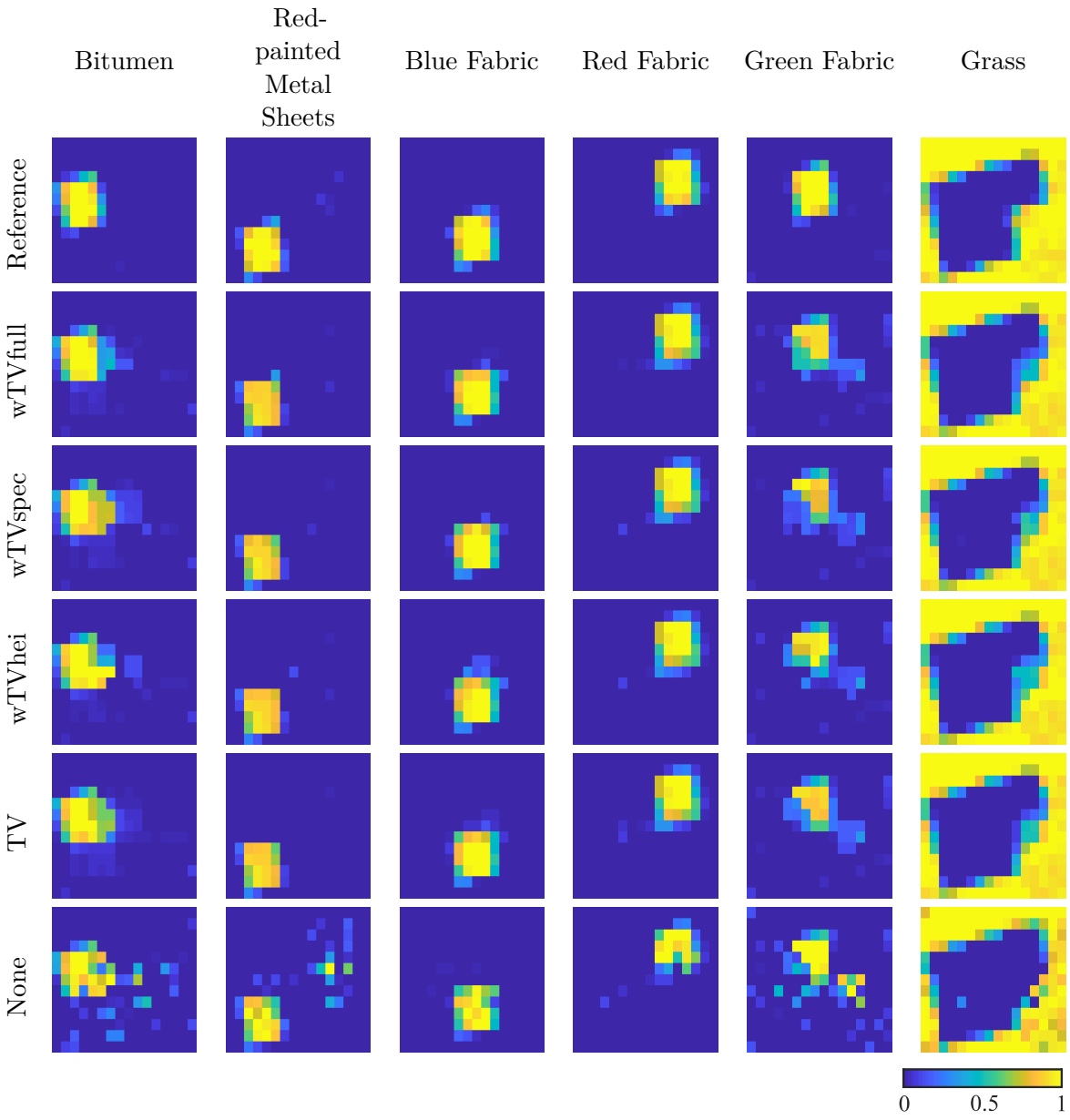


Figure 5.39: Abundance maps for the HySU dataset using S3AM with ablated spatial constraints. Left to right: bitumen, red metal sheets, blue fabric, red fabric, green fabric, and grass. Top to bottom: reference, wTVfull, wTVspec, wTVhei, TV, and "none". The reference abundance maps are computed by applying fully constrained least squares using the library of known endmembers on the shadow-free image.

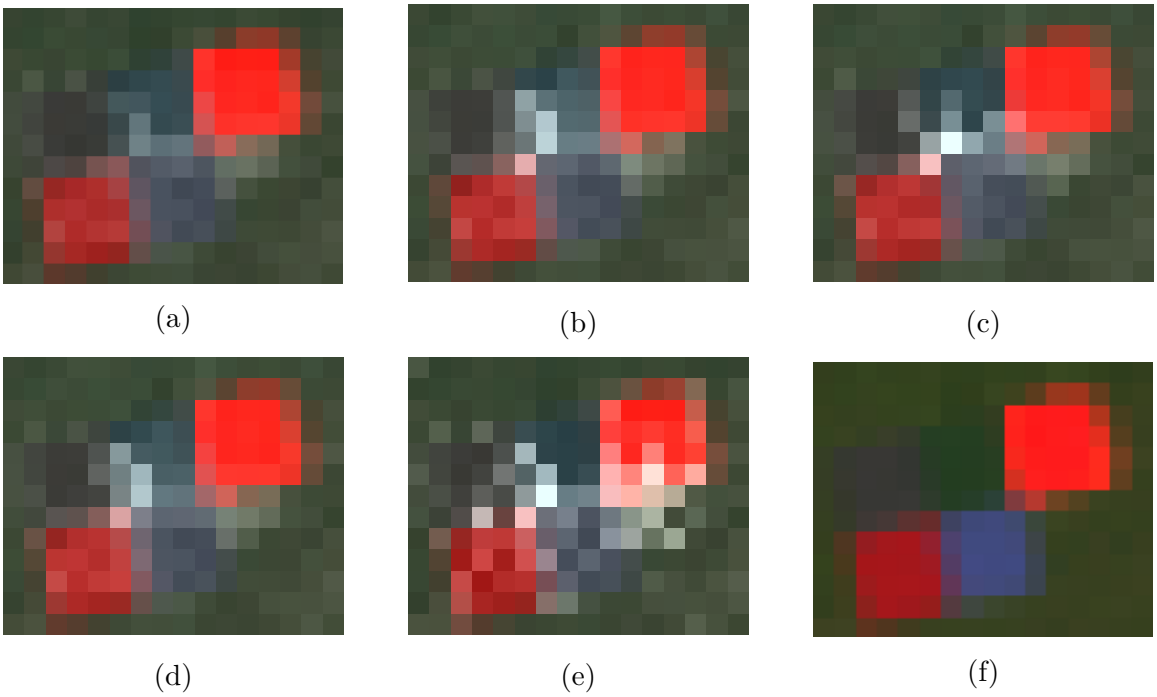


Figure 5.40: DLR HySU imagery overlaid with pixelwise mean abundance error maps in grayscale using S3AM with ablated spatial constraints: (a) wTVfull, (b) wTVspec, (c) wTVhei, (d) TV, (e) "none", (f) reference image.

contain similar spectral information. Therefore, in order to distinguish between similar spectra in the presence of shadows, it is essential to consider the diffuse solar radiation (see Figs. 5.42 (g)-(j)). In contrast, when determining materials with large spectral differences, such as vegetation and road in subset 2 (Fig. 5.43), SLMM and SMLM can also achieve satisfactory results. Moreover, by considering the diffuse solar illumination, Fansky, ESMLM, and SCBMM may achieve better abundance estimation at higher noise levels. In particular, the SCBMM shows a quite noisy abundance map in subset 1. Similar as in the DLR HySU dataset, the SCBMM may show decreased performance at some sunlit pixels (Fig. 5.43 (i)). In addition, the Fansky method appears noisier compared to ESMLM, and can only distinguish a part of the shadowed materials. Compared to Fansky and SCBMM, the ESMLM performs consistently better in both subsets.

The TV constraint further contributes to the abundance estimation in two aspects. First, given the large spectral variability in real hyperspectral imagery, the spectral unmixing methods without the TV constraint can easily confuse similar materials, thus producing considerably noisier abundance maps, while the spatial constraint promotes similar abundances in local neighborhoods, significantly reducing noise. Second, the abundance estimation is not as accurate in shadowed regions, where pixels contain a lower signal-to-noise ratio. The spatial constraint provides additional information to spectral models, thus achieving better abundance estimations.

An ablation study has been conducted in subset 1, using spatial constraints wTVfull, wTVspec, wTVhei, TV, and "none" in S3AM (see Fig. 5.44). As in the case of the HySU dataset, differences in TV constraints mainly affect mixed sunlit/shadowed pixels in the HySpex

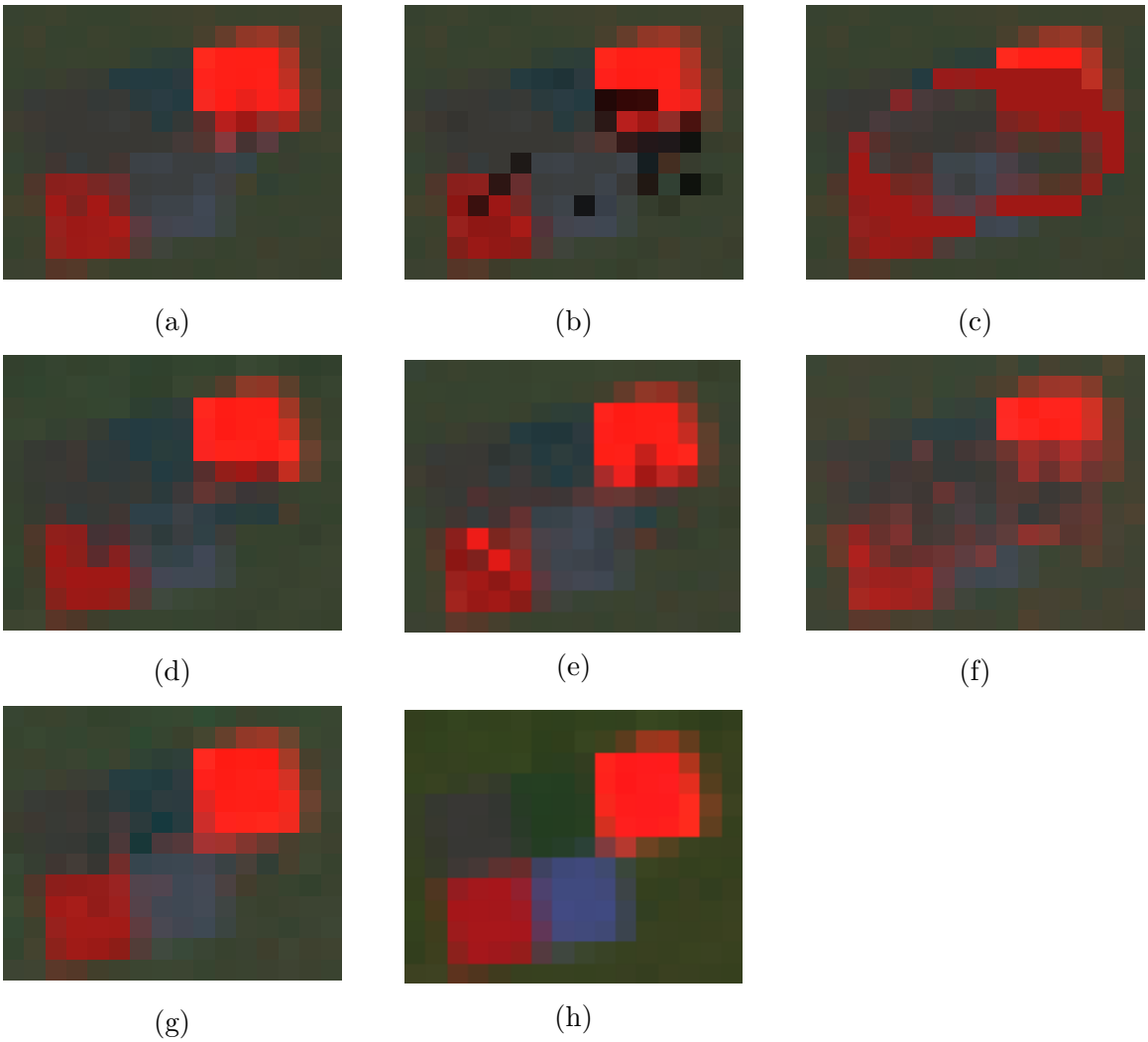


Figure 5.41: Restored images with removed shadows of the DLR HySU dataset generated by (a) SLMM, (b) SMLM, (c) NUEM, (d) Fansky, (e) ESMLM, (f) SCBMM, (g) S3AM, (h) reference image.

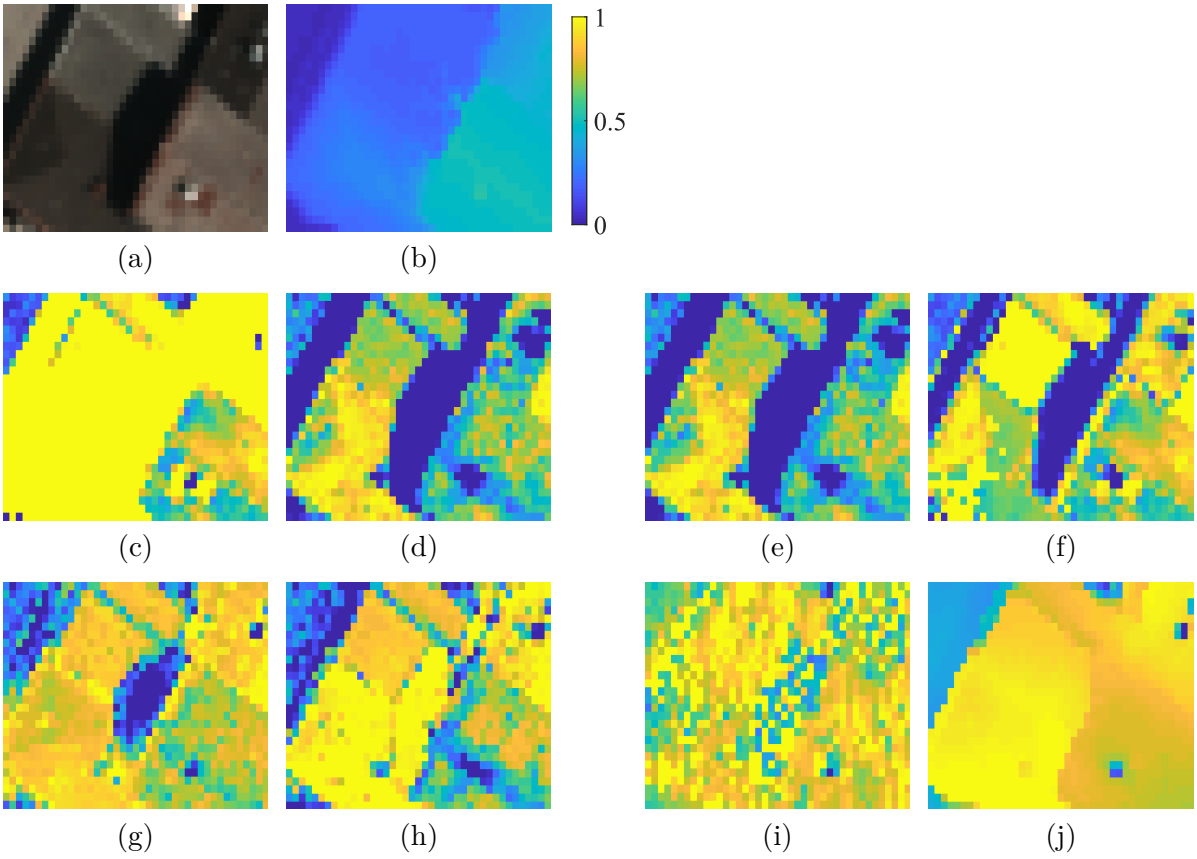


Figure 5.42: Subset 1 of the HySpex dataset: (a) true color composite, (b) DSM. Abundance maps of roof using: (c) LMM, (d) SLMM, (e) SMLM, (f) NUEM, (g) Fansky, (h) ESMLM, (i) SCBMM, (j) S3AM.

dataset. Specifically, the weighted TV methods, i.e., wTVspec, wTVhei, and wTVfull lead to sharp edges, while the classic TV method oversmooths transitions in boundary pixels. Examples can be spotted in the regions between the two roofs in subset 1.

5.3.6.2 The Impact of λ_{reg} on Abundance Estimation

Since it is very challenging to select an optimal λ_{reg} for the HySpex dataset, due to the lack of ground truth, the sensitivity of various TV constraints with respect to λ_{reg} is evaluated in Fig. 5.45. In our experiments so far, we used the optimized values of λ_{reg} from the DLR HySU dataset on the HySpex dataset. However, a similar dataset with ground truth may be not available at all in real-case scenarios. Fig. 5.45 compares abundances of roof in subset 1 for $\lambda_{\text{reg}} \in \{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$. As expected, the abundance maps become more homogeneous and less noisy as λ_{reg} increases. The abundance maps using $\lambda_{\text{reg}} = 10^{-3}$ present the best qualitative results. As discussed, the use of the weighted TV leads to better abundance estimation on boundary pixels, where the wTVfull achieves a balance between height and spectral information. This characteristic becomes more prominent with larger values of λ_{reg}. Specifically, the TV constraint using spectral information only, i.e., wTVspec, largely decreases the performance when $\lambda_{\text{reg}} \geq 10^{-2}$, as the computed weights do not match

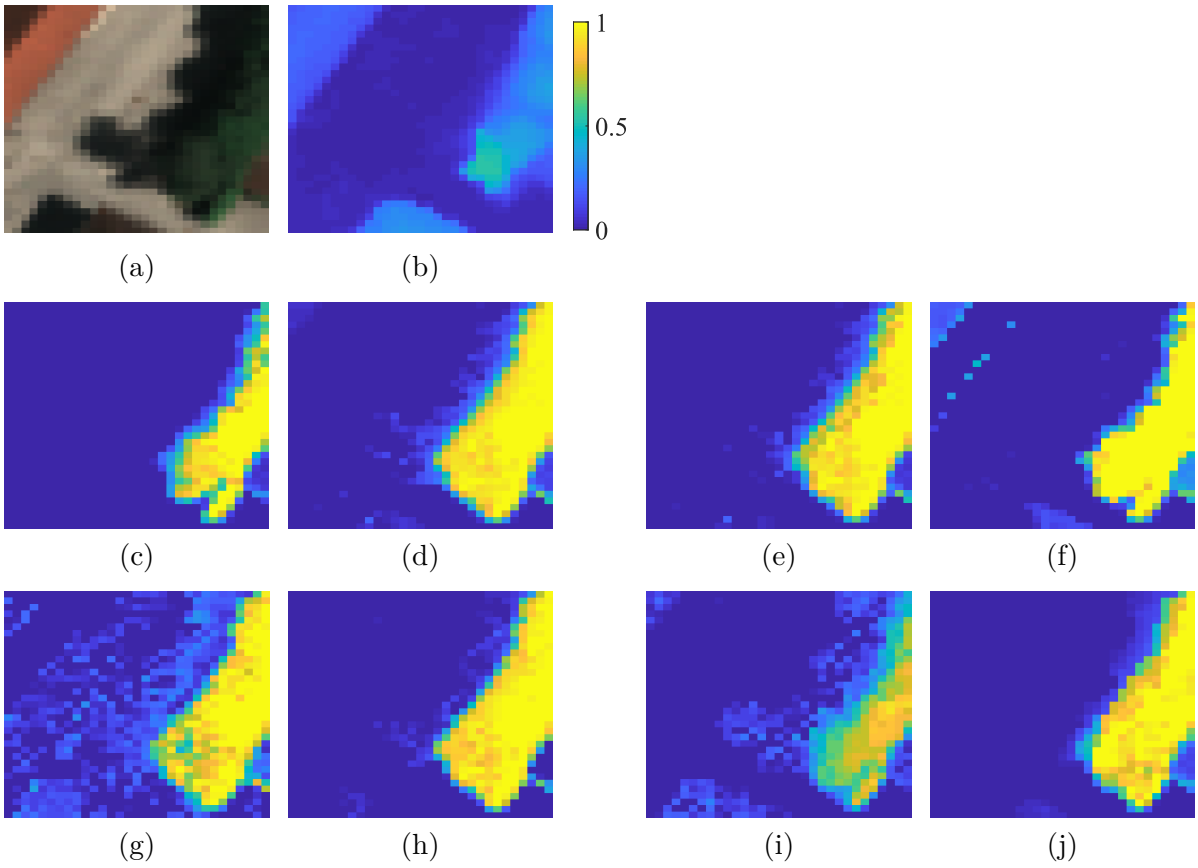


Figure 5.43: Subset 2 of the HySpex dataset: (a) true color composite, (b) DSM. Abundance maps of vegetation using: (c) LMM, (d) SLMM, (e) SMLM, (f) NUEM, (g) Fansky, (h) ESMLM, (i) SCBMM, (j) S3AM.

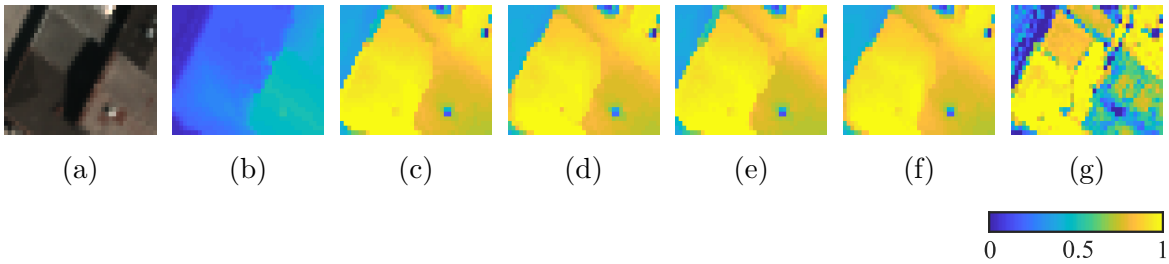


Figure 5.44: Subset 1 of the HySpex dataset: (a) true color composite, (b) DSM. Abundance maps of roof for ablation study using spatial constraints: (c) wTVfull, (d) wTVspec, (e) wTVhei, (f) TV, (g) none.

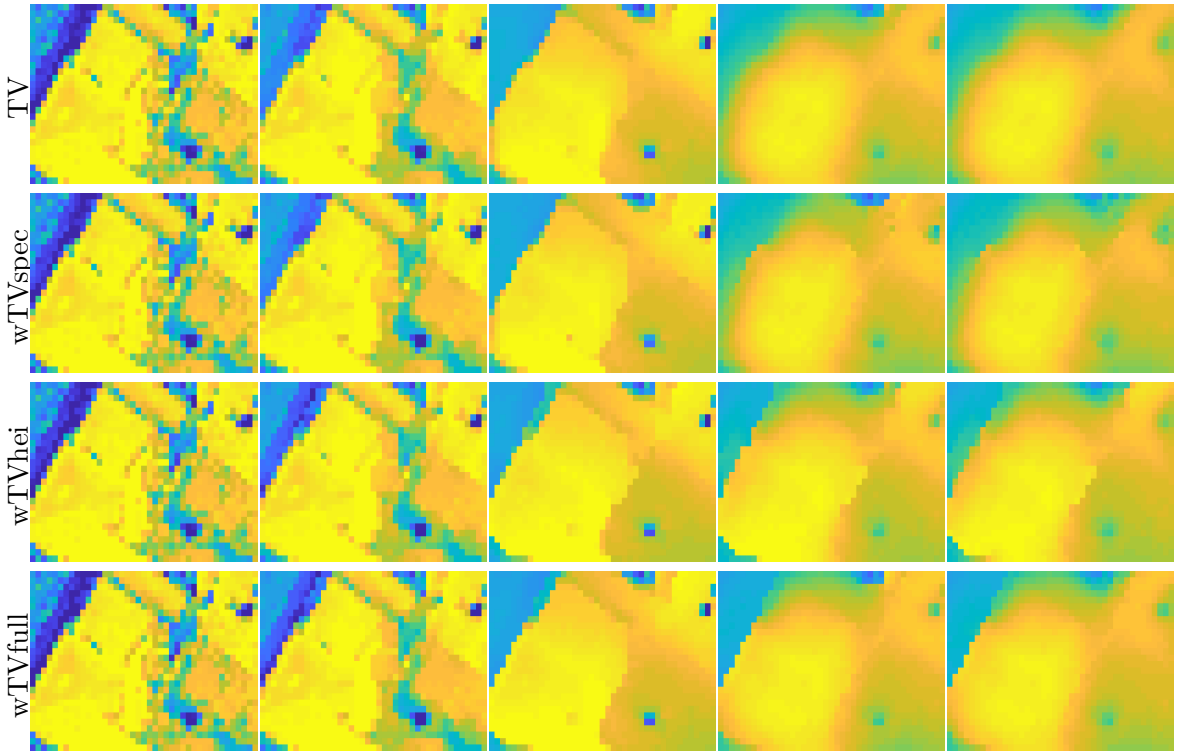


Figure 5.45: Abundance maps of roof as a function of λ_{reg} for subset 1 of the HySpex dataset using S3AM with ablated spatial constraints. Top to bottom: TV,wTVspec, wTVhei, wTVfull. Left to right: $\lambda_{\text{reg}} = 10^{-5}$, 10^{-4}, 10^{-3}, 10^{-2} and 10^{-1}.

well with the ground materials. Compared to TV and wTVspec, the wTVhei constraint leads to better results, but it can suffer from height inaccuracies. Despite some boundary pixels possibly being affected for larger values of λ_{reg}, the wTVfull constraint generally reaches a more robust abundance estimation in a broader range of values of λ_{reg}.

5.3.6.3 Shadow-Removed Pixel Reconstruction

Fig. 5.46 shows three examples of shadow removal results for the HySpex dataset. The restored images appear very bright in shadowed roof pixels when applying the SLMM, SMLM, and NUEM approaches. This corresponds to the poor abundance estimations in Fig. 5.42 (d)-(f). In addition, some dark pixels appear in the restored image by the SMLM, because of the incorrect estimation of the parameters P and Q. A similar artifact can be observed in the DLR HySU dataset (Fig. 5.41 (b)). Furthermore it is worth noticing that, when spectral angle distance successfully matches sunlit and shadowed pixels in a scene, the NUEM can achieve good results, such as in the second and third examples. However, this may work only in simple scenarios. In addition, the NUEM highly depends on the input shadow mask. Some artifacts appear when shadowed areas are over- or underestimated (see examples 1 and 3 in Fig. 5.46). Approaches considering the diffuse solar illumination achieve in all cases better shadow-removed images, while the results of SCBMM are much noisier. Despite the improvement, shadow removal by Fansky and ESMLM in the first example shows some

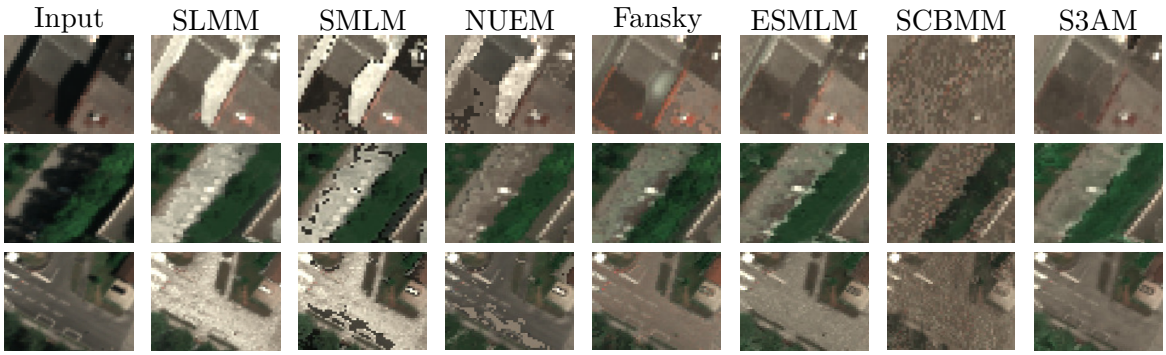


Figure 5.46: Three examples of shadow-removed results for the HySpex dataset. Left to right: input image, SLMM, SMLM, NUEM, Fansky, ESMLM, SCBMM, and S3AM.

spectral distortions. Finally, the reconstructed road computed by ESMLM and SCBMM in the third example appears noisy, with the line marking resulting almost invisible.

The advantages of S3AM compared to other methods appear evident. First, S3AM improves the restoration result in shadowed regions and retains spectral homogeneity. The shaded roof area in the first example of Fig. 5.46 shows that S3AM better reconstructs spectral features compared to ESMLM. Moreover, S3AM considerably reduces the noise in the shadow-removed imagery. This noise reduction can be spotted not only in shadowed regions, such as regions shaded by vegetation in the second example, but also in sunlit pixels, such as the impervious surface in the third example.

5.3.7 Summary

This section proposes a spatial-spectral shadow-aware mixing (S3AM) model. The spectral modeling accounts for shadows, following physical assumptions based on radiative transfer theory. Specifically, a light path initiates from an illumination source and interacts with endmembers before being scattered back to the observer. The model considers direct, diffuse, and neighboring illumination sources, where direct solar radiation is the dominant source for sunlit pixels, and diffuse solar radiation for shadowed pixels. A mixed pixel is then resolved by summing up the spectral contribution of all possible light paths.

S3AM embeds a DSM generated by multi-view stereo images. The sky view factor F, which is essential to estimate diffuse solar illumination in S3AM, can be conveniently computed using the DSM, reducing the model complexity. Moreover, we take into account the spatial relationship of abundances through weighted TV constraints, derived by spectral information from the hyperspectral imagery, height information from the DSM, and shadow information. The obtained optimization problem is bi-convex and is split into two convex problems, separately solved by the ADMM approach iteratively.

The proposed method has been extensively evaluated using two datasets, both quantitatively and qualitatively. Experiments demonstrate that the proposed method significantly reduces the noise level of abundance maps and improves the abundance estimation. Moreover, an ablation study is performed in which the proposed weighted TV constraint is compared

to different variations of the spatial TV regularization, or only considering the spectral information.

Several open questions remain. First of all, in order to quantitatively validate shadow-aware spectral unmixing methods in real scenarios, there is a lack of real datasets with shadows for which ground-truth abundances are known. Furthermore, although DSM derived by multi-view stereo imagery offers shadow-insensitive height information, the performance of S3AM can be degraded by noise and inaccurate values, especially on boundary regions with large height variations, causing imprecise TV weights and abundance estimation. In particular, urban areas represent the ideal application for this kind of analysis, as shadows are relevant and present across the image. However, DSMs can lead to inconsistencies due to occlusions, especially if the elevation model is derived by stereo images. Thus, future work should address spectral unmixing methods robust to inconsistencies and missing data in the DSM.

5.4 Discussion

5.4.1 Shadow Effects: a Field Study

This section investigates shadow effects in real scenarios using an imaging system mounted on a ground-based platform. The experiment measures spectra at sunlit and three typical shadow scenarios. Compared with airborne images, the primary objective of this field study is to mitigate atmospheric impacts on the measured spectra. In addition, in order to eliminate any inferring factors apart from shadow effects during the measurement, this experiment utilized a white Spectralon panel as target. The measured data can potentially be used to validate shadow removal approaches quantitatively.

The camera system, i.e., Cubert UHD-185 Firefly, manufactured by the company Cubert, was mounted on a ground-based platform. Please refer to Section 4.3 for a detailed description of the imaging system. The relative position between the panel and the sensor remained unchanged during the experiment. During the experiment the Spectralon panel was placed in the regions with different illumination conditions. Figs. 5.47 (a) - (d) presents four measured hyperspectral images as true color composites. Areas receiving direct solar radiation are fully sunlit (Fig. 5.47 (a)). The occluding objects forming shadows can be buildings and trees. Buildings are assumed to block direct solar radiation entirely, resulting in fully shadowed areas (Fig. 5.47 (b)). The imaging spectrometer measures a partly shadowed image when placing the panel on the border between fully shadowed and fully sunlit areas (5.47 (c)). In addition, direct solar radiation can transmit through and reflect between the leaves of trees, resulting in the second type of partly shadowed area (Fig. 5.47 (d)). Corresponding to Figs. 5.47 (a) - (d), Figs. 5.47 (e) and (f) illustrate their averaged reflectance spectra for each image without and with normalization, respectively. Compared to the reflectance in the fully sunlit region, shadow effects significantly decrease the spectral magnitude, whose strength depends on the amount of shadows. In addition, shadow effects demonstrate significant wavelength-dependent distortion on reflectance values, among which fully shadowed pixels show stronger distortion than partly shadowed pixels. Besides, pixels shaded by trees present significant vegetation patterns, namely a sharp red edge increase, especially in normalized reflectance.

5.4.2 Result Comparison: Investigation in a Larger Region

Three proposed methods in Sections 5.1, 5.2, and 5.3 incrementally improved one after the other in resolving shadow issues for hyperspectral imagery. In particular, the third work, i.e., S3AM, achieves the best performance regarding accuracy, robustness, and computational speed (Section 5.3). Nevertheless, S3AM was compared with state-of-the-art methods on small subsets with a total number of pixels between 200 and 1200. Besides, several recent methods aim to address shadow issues, but they were not included in Sections 5.1, 5.2, and 5.3. Hence, this section compares S3AM with state-of-the-art works in a larger study region using the HySpex/3K dataset (Section 4.1), which can benefit further development and deployment. Fig. 5.48 presents the study region with a size of 400-by-400 pixels.

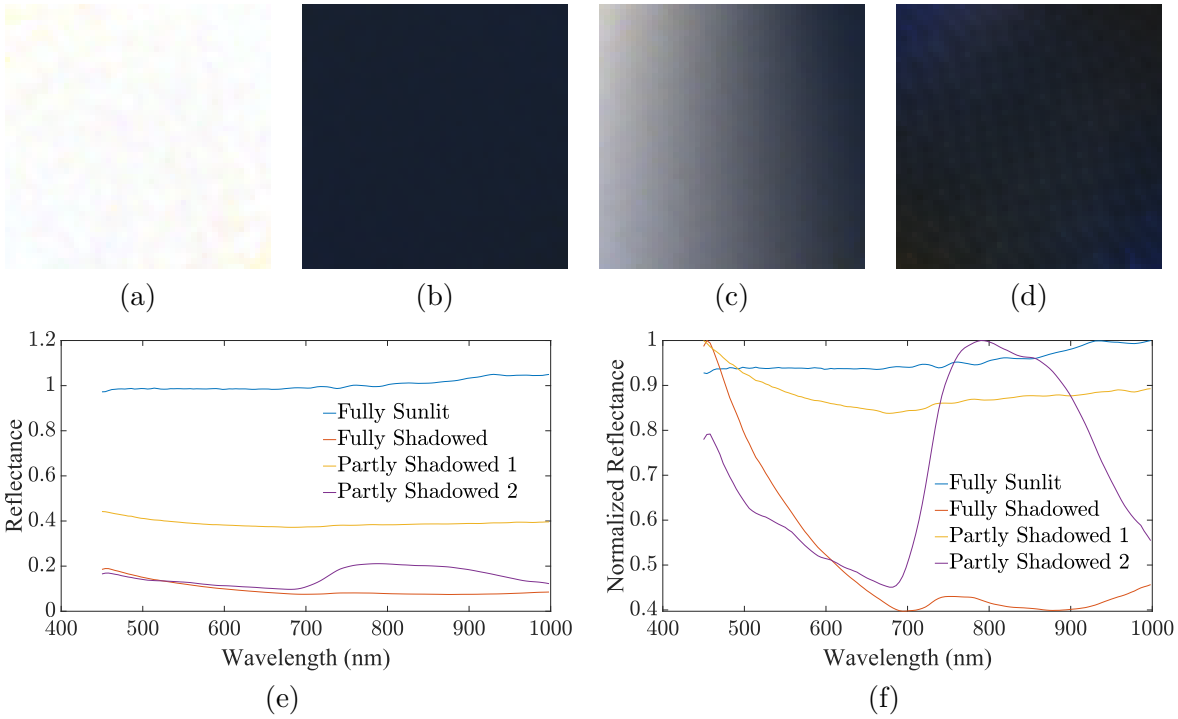


Figure 5.47: Four typical scenarios of spectral measurements using the camera system Cubert UHD-185 Firefly, placing a white Spectralon panel in regions with different illumination conditions: (a) fully sunlit area, (b) fully shadowed area formed by buildings, (c) partly shadowed area formed by buildings, (d) partly shadowed area formed by trees. Averaged reflectance spectra for each scenario without and with normalization in (e) and with normalization in (f), where the first and second partly shadowed cases correspond to (c) and (d), respectively.

It is worth noticing that few methods compared in Sections 5.1, 5.2, and 5.3 can be directly applied to large regions, as a result of their very high computational complexity, such as Fansky [198], ESMLM [199], and SCBMM [165]. Please refer to Section 5.3.4.4 for comparing their computational resources in smaller areas. Other methods with better computational performance in Sections 5.1, 5.2, and 5.3 were selected for comparison, namely SLMM and NUEM. In addition, ATCOR and Shadow Removal Method based on Multi-exposure Fusion (MEXPO), which were not included in previous experiments, are compared in this section. In summary, this section compares the following methods:

- SLMM, named after a shadow detection and removal method based on linear unmixing [149]. Please refer to Sections 3.4.1.3 and 3.4.2.3 for a review of this method.
- NUEM, named after a shadow removal method based on nonlinear unmixing and endmember matching [155]. Please refer to Section 3.4.2.1 for a review of this method.
- MEXPO, named after a shadow detection and removal method based on multi-exposure fusion [157]. Please refer to Section 3.4.2.2 for a review of this method.
- ATCOR, named after a atmospheric correction software that generates shadow detection result as a byproduct [63, 148]. Please refer to Section 3.4.1.3 for a review of this method.

Figure 5.48: A study region selected from the HySpex/3K dataset (Section 4.1). Hyperspectral imagery shown in true color composite with two zoomed in areas.

- S3AM, named after spatial-spectral shadow-aware mixing model for shadow compensation [200]. Please refer to the third contribution of this dissertation (Section 5.3).

Fig. 5.49 presents shadow masks and two zoomed-in areas generated by SLMM, MEXPO, ATCOR, and S3AM. MEXPO generated a binary shadow mask, while SLMM, ATCOR and S3AM produced non-binary results, i.e., soft shadow masks, where decimal values are allowed between 0 and 1. MEXPO tends to overdetect shadows compared to the other two methods, especially in the regions of dark vegetation (see the first example in Fig. 5.49 (b)). In addition, this method can also underdetect some areas, such as shadows occluded by trees in the second example. Since MEXPO detects shadows based on thresholding and morphological filters, selecting one threshold suitable for the whole image can be challenging, leading to over- and under-detection.

Compared to MEXPO, three physics-based methods, i.e., SLMM, ATCOR, and S3AM, produced significantly better results, indicating the advantages of considering physical assumptions in shadow detection tasks. Nevertheless, ATCOR can miss some regions, especially in light-shaded pixels. An example can be observed in shadows occluded by trees in the two zoomed-in areas (Fig. 5.49 (c)). SLMM and S3AM presents similar result and can detect most shadowed regions. For example, they successfully detect shadows occluded by trees in the second zoomed-in area. In addition, shadows occluded by the building in the second example present a smoother transition on shadow boundaries using S3AM. Nevertheless, SLMM and S3AM can be sensitive to spectral information. In the first zoomed-in example, SLMM overdetects some road pixels as shadows, and S3AM preserves road surface marking patterns in shadowed regions occluded by trees.

Table 5.10: Running time of shadow removal methods in the study region

Method	SLMM	NUEM	MEXPO	S3AM
Running Time (s)	31.42	35.72	18.85	813.74

Fig. 5.50 compares shadow removal results, and Table 5.10 compares their computational speeds. MEXPO removed shadows by fusing over- and under-exposure images. This approach may recover some partly shaded pixels, such as regions occluded by trees, while it cannot successfully recover fully shadowed pixels, such as regions occluded by buildings. In addition, some artifacts aroused by over-exposure images can be observed in the restored image. NUEM aims to match shadowed pixels with their corresponding sunlit ones by minimizing the spectral angle. Due to the characteristics of this method, all shadowed pixels can be "recovered" to some extent because all shadowed pixels will be matched to sunlit ones. However, the matching error seems large. In other words, many shadowed pixels have been matched with incorrect sunlit ones since shadow effects introduce large spectral distortion. Both SLMM and S3AM rely on spectral unmixing but with different physical assumptions: the former assumes shadows as a wavelength-independent scaling effect, while the latter regards shadows as wavelength-dependent spectral distortion. In Fig. 5.50 (a), SLMM recovered most shadowed pixels with very bright material, indicating that the scaling factor can not successfully model shadow effects. S3AM performs better than SLMM and the other two methods according to a qualitative interpretation. Nevertheless,

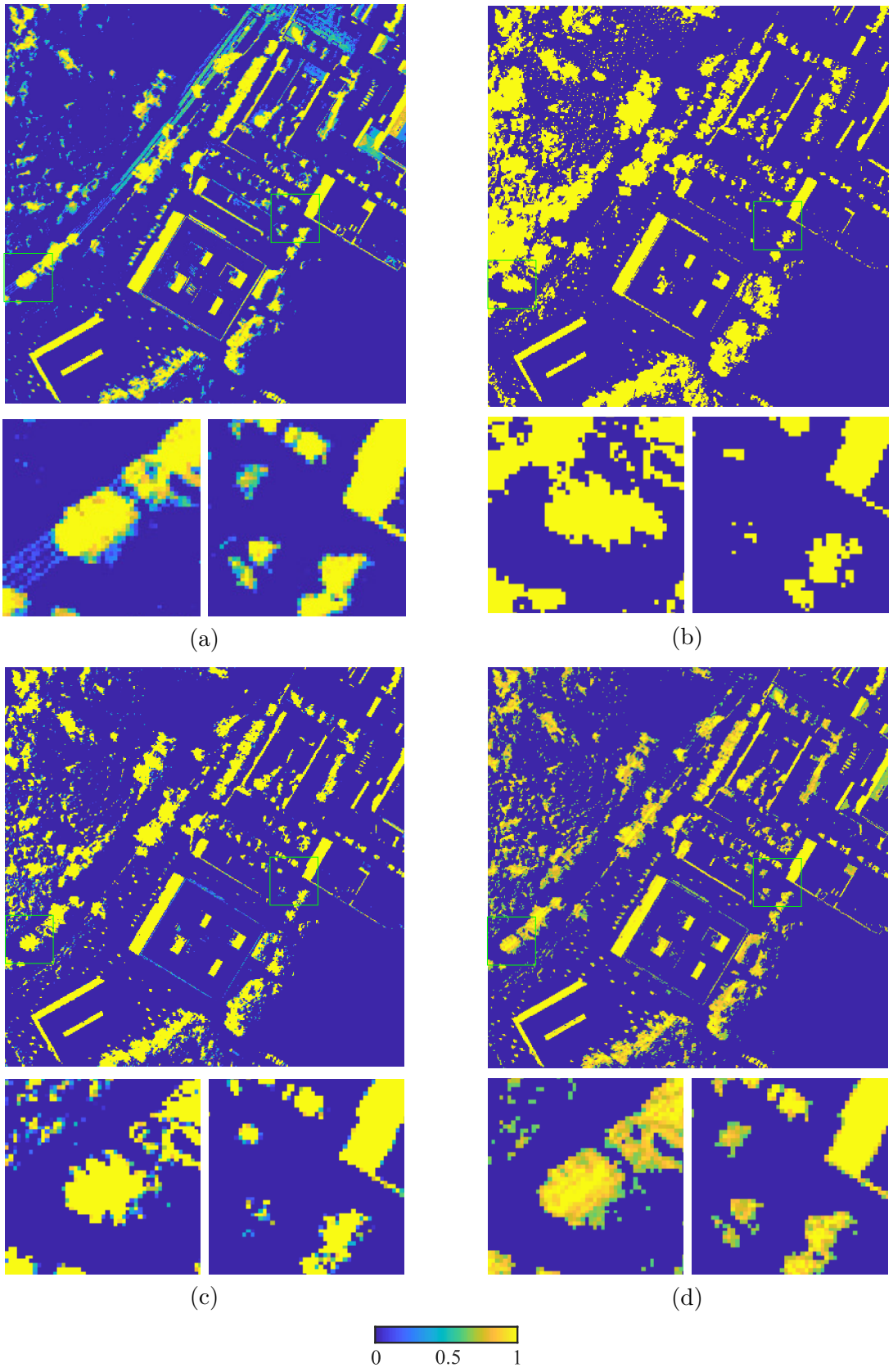
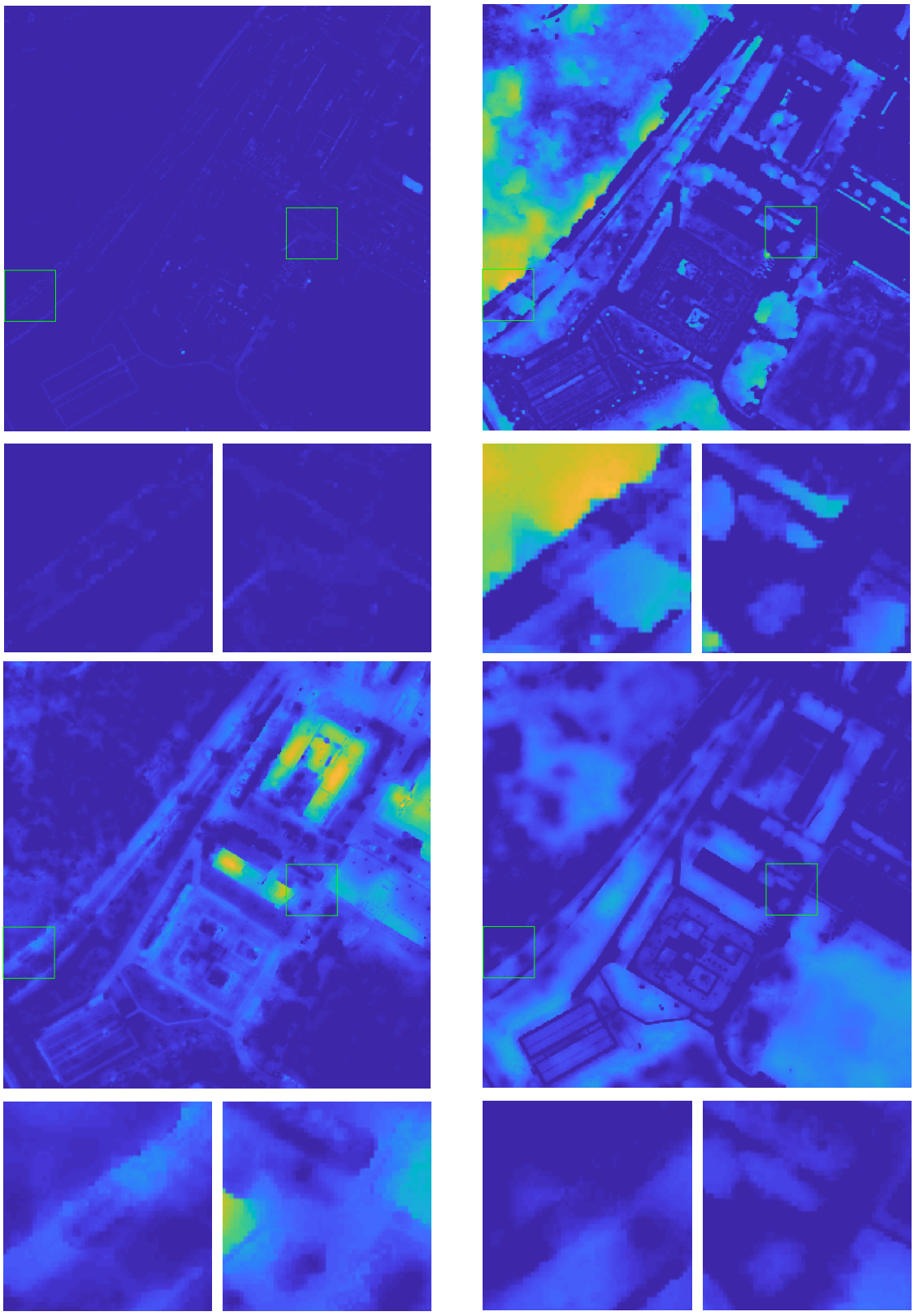


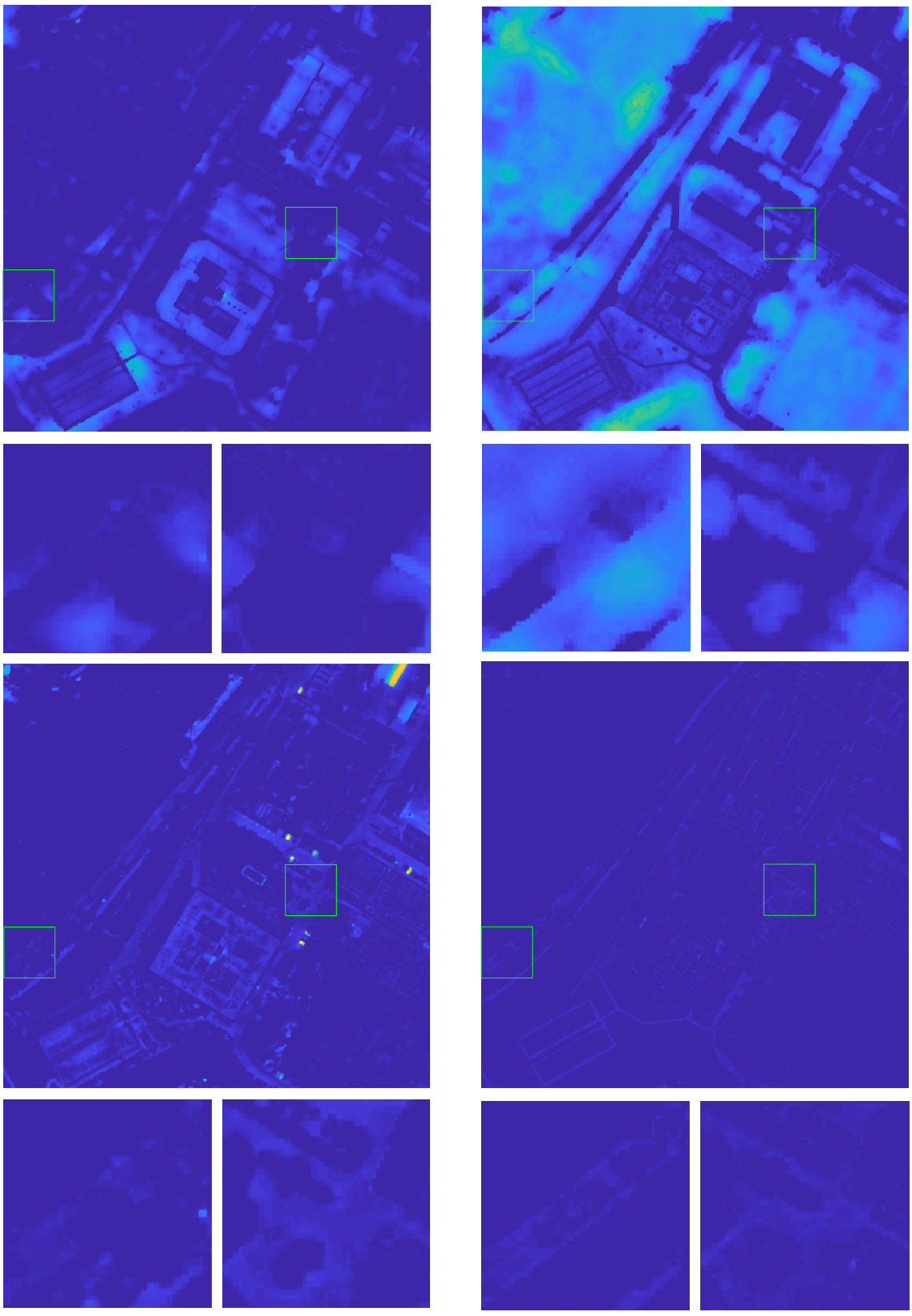
Figure 5.49: Comparison of shadow detection results in a study region with a size of 400-by-400 pixels. (a) SLMM, (b) MEXPO, (c) ATCOR, (d) S3AM.

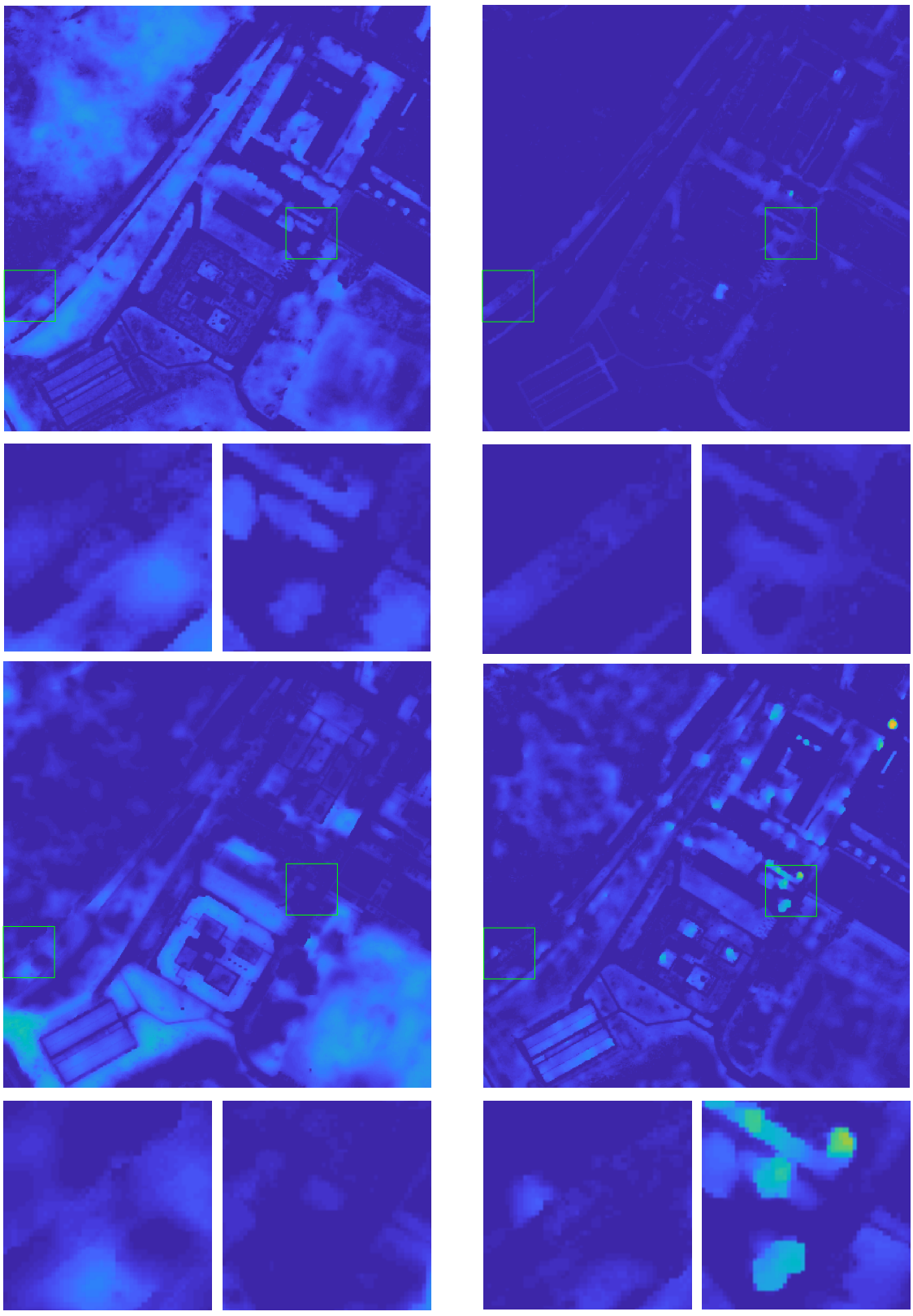
Figure 5.50: Comparison of shadow removal methods using hyperspectral imagery with a size of 400-by-400 pixels shown in true color composite. (a) SLMM (b) NUEM, (c) MEXPO, and (d) S3AM.

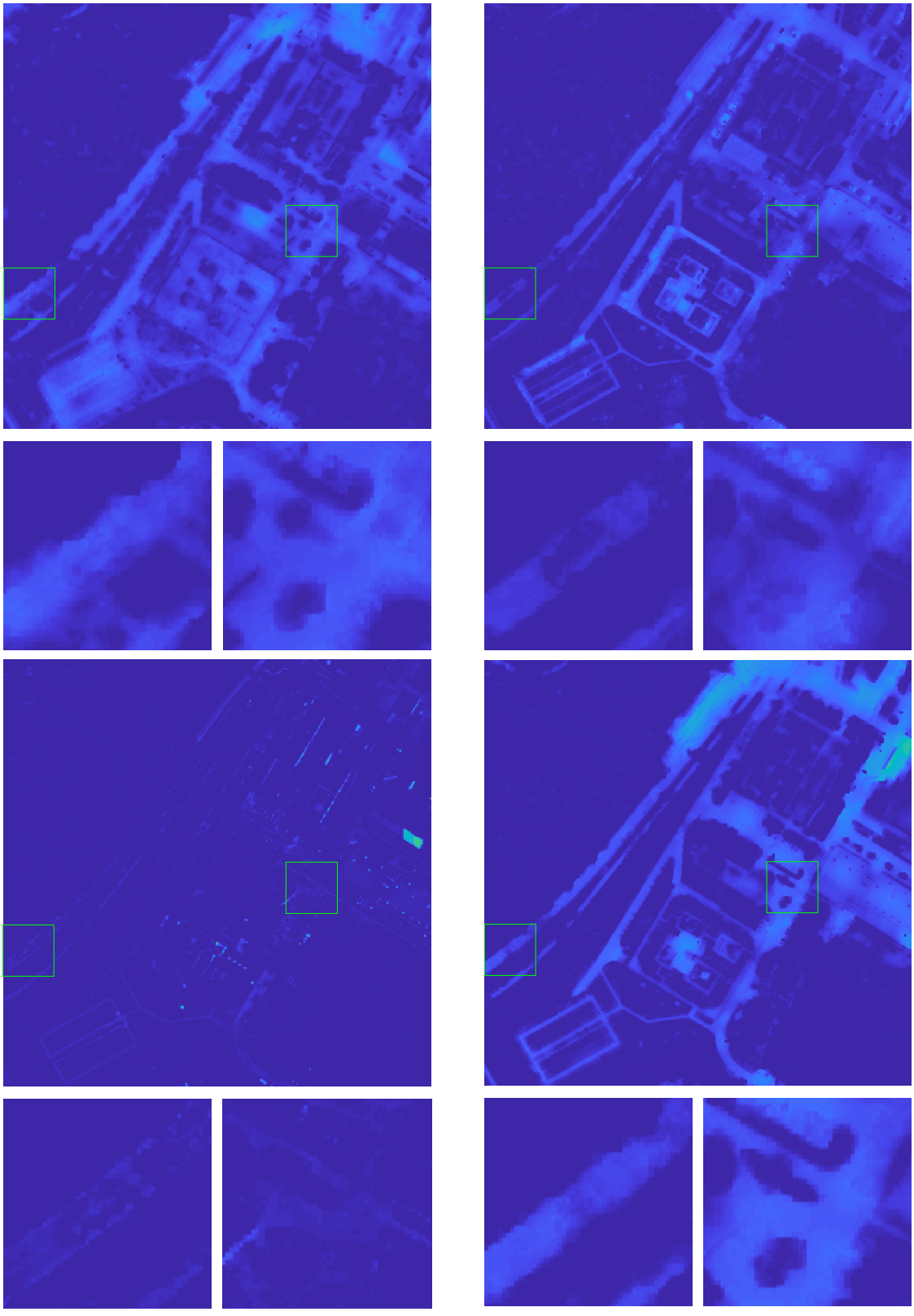
several open questions remain and require further investigation in future work. Specifically, the shadow-removed image shows the loss of textures, which is caused by two reasons. Firstly, the TV constraint introduces spatial relationships between pixels. More specifically, it promotes similar abundances in local neighborhoods, leading to blurred reconstructed images. Secondly, S3AM is based on spectral unmixing, where limited materials (i.e., endmembers) are extracted from the input image and involved in shadow restoration.

Finally, it is worth noticing that spectral unmixing-based methods yield not only shadow detection and removal results but also abundance maps that describe fractions of materials at each pixel. Fig. 5.51 shows an example of 20 abundance maps derived by S3AM, corresponding to 20 endmembers extracted using the method in 5.3. Specifically, since S3AM addresses shadow effects using other hyperparameters, its derived abundances are shadow-insensitive. These abundance maps can be used as input for other machine learning frameworks, such as classification and segmentation, and aid in deriving shadow-insensitive or shadow-invariant results.









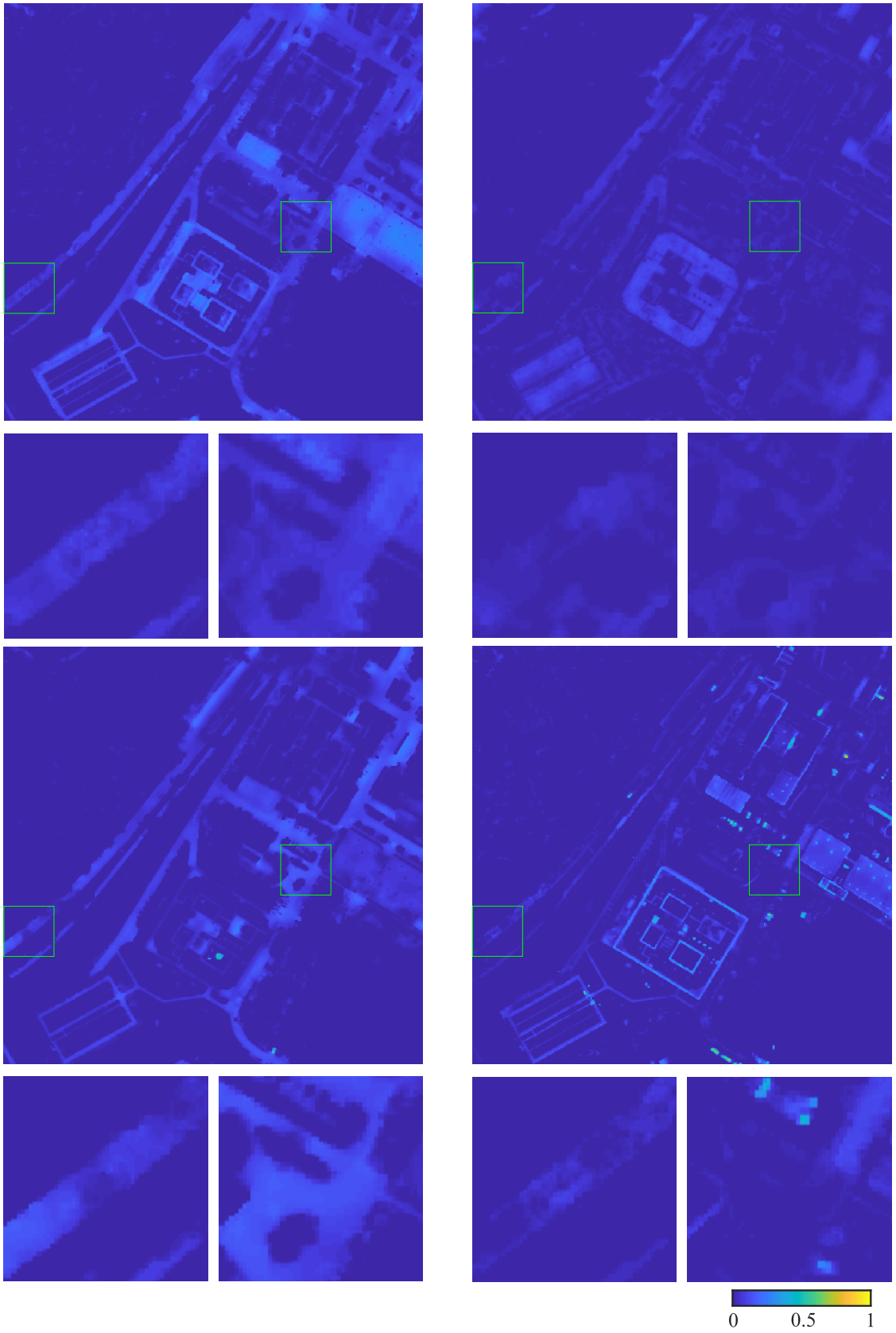


Figure 5.51: Abundance maps derived by S3AM, corresponding to 20 endmembers extracted using the method in Section 5.3.

6

Conclusion and Outlook

“ If the past could be changed,
it would not exist.
If the future could be stopped,
it would not survive.
If the present could be avoided,
it would not prevail.”

- Matshona Dhliwayo

Contents

6.1 Conclusion

Hyperspectral sensors have increased spectral resolution with respect to other optical imaging systems, allowing capturing additional information from targets on the ground surface using narrow spectral bands. Shadow effects caused by occlusion may pose a critical challenge for accurately analyzing hyperspectral imagery, with the problem not being investigated so far as for other types of optical images. In particular, the characteristics of hyperspectral imagery, e.g., high spectral resolution, coupled with other impacts from remote sensing, e.g., radiative propagation between the sensor and ground surface, and scarcity of training samples, lead to unique and unsolved challenges.

This dissertation presents three advanced and robust physics-aware algorithms, built on one another, for resolving shadow issues in hyperspectral imagery. In particular, physics-based assumptions derived from radiative transfer theory have demonstrated advantages in solving shadow problems.

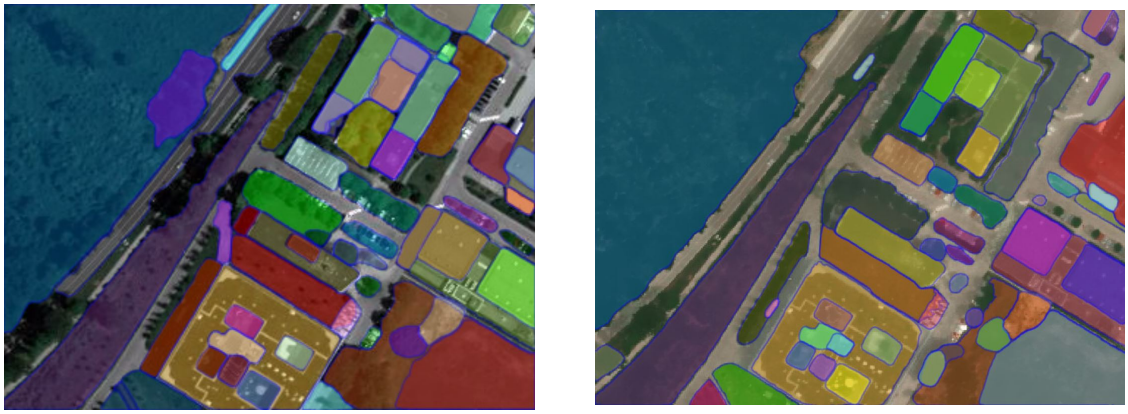
The first contribution in Section 5.1 proposes a novel framework to simultaneously detect and remove shadows based on nonlinear spectral unmixing. This requires an image and several atmosphere-related parameters as input, and yields shadow-removed images, soft shadow masks, and shadow-insensitive material abundances. Following several physical assumptions, this work proposes a spectral mixing model to compute sunlit and (partly) shadowed pixels. Later, unknown parameters were solved by spectral unmixing and used in pixel reconstruction for shadow removal. This work demonstrates the advantages of physics-based assumptions and spectral unmixing in shadow detection and removal but considering only simple ground scenarios.

The second contribution in Section 5.2 focuses on developing a spectral mixing model that can adapt to more complicated scenarios. In particular, it considers both shadow and nonlinear effects and addresses their intrinsic relationship. Besides, several physical assumptions are utilized to leverage the accuracy and simplicity of the proposed model. Compared to the first contribution, this framework yields a more generalized model and more accurate parameter estimation. Then, those estimated parameters were used to generate improved shadow-removed images.

Nevertheless, solving nonlinear unmixing, which can be regarded as a constrained optimization problem, is non-trivial. It can often be ill-posed and sensitive to noise. In order to achieve a more robust estimation, the third contribution of this dissertation embeds spatial and topographical information into spectral unmixing by incorporating Digital Surface Models (DSMs) in the analysis. This framework significantly improves the robustness of the optimization problem, yields more accurate estimations of physical parameters, and results in better performance on shadow-removed images. In addition, the computational speed of this method is comparable to linear spectral unmixing, thus making it possible to be applied to large images.

Referring to Fig. 1.1 in Chapter 1, Fig. 6.1 compares two segmentation results, where the input is represented by the original hyperspectral image and the shadow-removed image

derived by the approach described in Section 5.3, respectively. Results demonstrate that shadow-removed images have a large potential to improve segmentation performance.



(a) (b)

Figure 6.1: Comparison of image segmentation results generated by SAM [7] for the data selected from the HySpex/3K dataset (Section 4.1), acquired over Oberphaffenhofen, Bavaria, Germany. Input images: (a) airborne hyperspectral image shown as a true color composite acquired by the HySpex sensor (VNIR-1600), (b) shadow-removed image derived from (a) derived by S3AM in Section 5.3.

6.2 Outlook

This dissertation opens novel insights into tackling shadow issues in hyperspectral imagery while leaving several open questions for investigation in the future.

Several spectral mixing models are presented to simulate image pixels based on simplified radiative transfer theory. In particular, the ground surface is assumed to be Lambertian: however, radiative propagation is much more complicated in practice. For example, multiple optical interactions exist between the ground surface and the atmosphere, and the former is usually not perfectly Lambertian. These complications demand further investigations in spectral mixing models. Furthermore, it is vital to leverage the simplicity and accuracy of the models in order to solve spectral unmixing accurately and robustly. Thus, advanced optimization solvers are desired.

The third contribution has demonstrated that fusing hyperspectral imagery with DSMs improves the performance of shadow removal. Nevertheless, the inaccuracies in DSMs may be propagated to the obtained results, especially in some shadowed regions. An inaccurate DSM can lead to decreased performance in data fusion, demanding improved data fusion methods robust to the quality of DSMs.

Image time series can also provide valuable information. As the sun's position changes, illumination conditions in a region vary over time in a predictable manner. Thus, the fusion of multi-temporal images can help detect and remove shadows.

Moreover, physics-aware deep learning may be a research direction to tackle shadow issues in hyperspectral data. One of the main challenges of applying deep learning frameworks for shadow detection and removal in hyperspectral imagery is the lack of training datasets. Incorporating physics-prior can improve the explainable of the model while requiring fewer training samples.

Besides shadow detection and removal, there may be other solutions to tackle shadow effects in hyperspectral imagery. Depending on their applications, learning a shadow-invariant feature presentation from the image or shadow-aware classification and segmentation frameworks can be sufficient.

Finally, benchmark datasets considering shadows for airborne and spaceborne hyperspectral images are not publicly available. Thus, it is difficult to quantitatively evaluate the performance of shadow-aware methods in real scenarios. A possible solution for building a benchmark dataset can be manual labeling based on comprehensive field studies. Another possible solution is to obtain hyperspectral images with the same imaging system in an exact location at different acquisition times on the same day.

List of Symbols

a_i	Abundance of the endmember $\boldsymbol{e}_i$
a_{s_i}	Abundance of the shadowed endmember $\boldsymbol{e}_{s,i}$
B	Number of spectral bands
c_0, c_1	Parameters for radiometric calibration
dA	Area element on a surface
D	Euclidean distance
$\boldsymbol{e}_i$	Spectrum of the i-th endmember in the endmember library $\boldsymbol{E}$
$\boldsymbol{e}_{\mathrm{s}_i}$	Spectrum of the i-th shadowed endmember
$\boldsymbol{E}$	Endmember library
$\mathcal{E}$	Irradiance
$\mathcal{E}_{\mathrm{g}}$	Global solar irradiance at ground surface
$\mathcal{E}_{\mathrm{l}}$	Direct solar irradiance at ground surface
$\mathcal{E}_{\mathrm{s}}$	Diffuse solar irradiance at ground surface
F	Sky view factor
h	Surface elevation above the ellipsoid
$\hbar$	Reduced Planck's constant
I	Radiant intensity
k_1, k_2, k_3	Parameters describing the diffuse-to-direct solar irradiance
K	a strength factor of neighbor interactions, denoting the fraction of the scattered light from the neighborhood that is received by the pixel
L	Length of *path*
$\mathcal{L}$	Radiance
$\mathcal{L}_l$	At-sensor radiance with $\delta = 1$
$\mathcal{L}_p$	Path radiance
$\mathcal{L}_s$	At-sensor radiance with $\delta = 0$

m	Parameter in bilinear mixing models
M	Radiant exitance
n_i	Number of photons at the frequency index i
N	Number of pixels in an image
$\mathcal{N}(j)$	The first order neighborhood of the target pixel j in the spatial domain
$\boldsymbol{o}_j$	Residual vector at pixel j
$\boldsymbol{O}$	Residual matrix
p	Number of endmembers
$path$	A light path defined by the random variable $\{\Xi_i\}_{i\geq 0} \in S$
P	Probability that a light ray undergoes additional interactions with endmembers
$P(path)$	Probability of $path$
Q	Spatial fraction of shadow in a pixel
$\mathcal{Q}$	Total energy carried by photons
r	Material reflectance
$\hat{r}$	Estimated r by the atmospheric correction with $\delta = 1$
$\hat{r}_{\mathrm{s}}$	Estimated r by the atmospheric correction with $\delta = 0$
$R_{j,m}$	Weighting factor describing the similarity between pixel j and m
$R_{hj,m}$, $R_{xj,m}$	Weighting factors describing the similarity between pixel j and m, related to height and spectral information, respectively
S	A discrete set containing all possible interactions that a light ray can undergo from the illumination source(s) to the observer
t	Time
$T(s_i)$	Operator that acts on the light ray in state s_i
w	Single Scattering Albedo (SSA)
$\boldsymbol{W}_1$, $\boldsymbol{W}_2$	Sparse matrices consisting of differential operators in four directions, i.e., up, down, left, and right.

$\boldsymbol{x}_j$	Spectral signature of the j-th pixel in the hyperspectral image
$\hat{\boldsymbol{x}}_j$	Reconstructed spectral signature of $\boldsymbol{x}_j$
$\boldsymbol{X}$	Hyperspectral image
z	Unknown variable
Z	Normalizing constant value
α	Absorptance
β	Learning rate of Gradient Descent Algorithm (GDA)
γ	Parameter in bilinear mixing models
δ	Direct solar irradiance indicator
η	Parameter related to shadow effects
λ	Spectral wavelength
λ_{reg}	Regularization parameter
μ	Penalty parameter
ν_i	Frequency at index i
σ	Parameter controlling probability distribution
τ	Transmittance
τ_{dir}	Earth's atmospheric transmittance of the direct solar radiation
τ_{diff}	Earth's atmospheric transmittance of the diffuse solar radiation
Φ	Radiant flux
$\boldsymbol{\chi}$	Averaged spectrum of a pixel in the first-order neighborhood
ω	Weight factor according to inverse distance between neighboring and target pixels
$\mathrm{d}\Omega$	The element of solid angle
AE	Mean abundance error in spectral unmixing
DN	Digital number
RE	Mean reconstruction error in spectral unmixing

SNR	Signal-to-noise ratio
SRE	Spectral reconstruction error in spectral unmixing
$\odot$	Hadamard product
$\|\cdot\|_F$	Frobenius norm
$\|\cdot\|_{1,1}$	L1 norm of matrix
$\|\cdot\|_{2,1}$	L2 norm of matrix
$\|\cdot\|_0$	L0 norm

List of Abbreviations

2D two-dimensional

3D three-dimensional

ADMM Alternating Direction Method of Multipliers

ANC Abundance Nonnegative Constraint

ARGAN Attentive Recurrent Generative Adversarial Network

ASC Abundance Sum-to-one Constraint

ATCOR Atmospheric and Topographic Correction

AVIRIS Airborne Visible/Infrared Imaging Spectrometer

BRDF Bidirectional Reflectance Distribution Function

BRISK Binary Robust Invariant Scalable Keypoints

CASI Compact Airborne Spectrographic Imager

CEST Central European Summer Time

CGAN Conditional Generative Adversarial Network

CHRIS Compact High Resolution Imaging Spectrometer

ConvNet Convolutional Neural Network

DESIS Deutsches Zentrum für Luft- und Raumfahrt Earth Sensing Imaging Spectrometer

DLR Deutsches Zentrum für Luft- und Raumfahrt

DSM Digital Surface Model

EnMAP Environmental Mapping and Analysis Program

ESMLM Extended Shadow Multilinear Mixing

FCLS Fully Constrained Least Squares

GAN Generative Adversarial Network

GBM Generalized Bilinear Model

GDA Gradient Descent Algorithm

GPSO Global Particle Swarm Optimization

GSD Ground Sampling Distance

HCV Hue-Chroma-Value

HFC Harsanyi–Farrand–Chang

HIS Hue-Intensity-Saturation

HSV Hue-Saturation-Value

HSVDI Normalized Saturation Value Difference Index

HyMap Hyperspectral Mapper

HySime Hyperspectral Signal Identification by Minimum Error

HySU HyperSpectral Unmixing

IFOV Instantaneous Field of View

LAI Leaf Area Index

LASSO Least Absolute Shrinkage and Selection Operator

LiDAR Light Detection and Ranging

LMM Linear Mixing Model

LWIR Long-wave Infrared

MCMC Markov Chain Monte Carlo

MEXPO Shadow Removal Method based on Multi-exposure Fusion

MLM Multilinear Mixing

MODIS Moderate Resolution Imaging Spectroradiometer

NDVI Normalized Difference Vegetation Index

NIR Near-infrared

NNLS Non-negative Least Squares

NUEM Shadow Removal Method based on Nonlinear Unmixing and Endmember Matching

PPI Pure Pixel Index

PPNM Polynomial Post-nonlinear Model

RGB Red, Green, and Blue

RMS Root Mean Square

RNN Recurrent Neural Network

S3AM Spatial-spectral Shadow-aware Mixing

SAD Spectral Angle Distance

SAGA System for Automated Geoscientific Analyses

SAM Segment Anything Model

SAR Synthetic Aperture Radar

SCBMM Shadow-compensated Bilinear Mixing Model

SCGAN Shadow Conditional Generative Adversarial Network

SGA Simplex Growing Algorithm

SLMM Shadow Linear Mixing Model

SMLM Shadow Multilinear Mixing

SQP Sequential Quadratic Programming

SNR Signal-to-noise Ratio

SSA Single Scattering Albedo

STCGAN Stacked Conditional Generative Adversarial Network

SUnSAL Sparse Unmixing via Variable Splitting Augmented Lagrangian

SVM Support Vector Machine

SWIR Short-wave Infrared

TGV Total Generalized Variation

TIR Thermal Infrared

TV Total Variation

UAV Unmanned Aerial Vehicle

USGS United States Geological Survey

VCA Vertex Component Analysis

VNIR Visible and Near-infrared

List of Figures

List of Tables

Bibliography

[1] Hans Grahn and Paul Geladi. *Techniques and applications of hyperspectral image analysis*. John Wiley & Sons, 2007.

[2] Bing Lu, Phuong D Dao, Jiangui Liu, Yuhong He, and Jiali Shang. Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sensing*, 12(16):2659, 2020.

[3] Isabella Mariotto, Prasad S. Thenkabail, Alfredo Huete, E. Terrence Slonecker, and Alexander Platonov. Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the hyspiri mission. 139:291–305, 2013.

[4] Laura M. Dale, André Thewis, Christelle Boudry, Ioan Rotar, Pierre Dardenne, Vincent Baeten, and Juan A. Fernández Pierna. Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: A review. 48:142–159, 2013.

[5] Aliaksei Makarau, Rudolf Richter, Rupert Muller, and Peter Reinartz. Adaptive shadow detection using a blackbody radiator model. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):2049–2059, 2011.

[6] Wei Ji, Jingjing Li, Qi Bi, Wenbo Li, and Li Cheng. Segment anything is not always perfect: An investigation of sam on different real-world applications. *arXiv preprint arXiv:2304.05750*, 2023.

[7] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross Girshick. Segment anything, 2023.

[8] Graham D Finlayson, Mark S Drew, and Cheng Lu. Entropy minimization for shadow removal. *International Journal of Computer Vision*, 85(1):35–57, 2009.

[9] Liangqiong Qu, Jiandong Tian, Shengfeng He, Yandong Tang, and Rynson WH Lau. Deshadownet: A multi-context embedding deep network for shadow removal. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4067–4075, 2017.

[10] Zhihao Chen, Lei Zhu, Liang Wan, Song Wang, Wei Feng, and Pheng-Ann Heng. A multi-task mean teacher for semi-supervised shadow detection, 2020.

[11] David Ruiz Hidalgo, Bladimir Bacca Cortés, and Eduardo Caicedo Bravo. Dimensionality reduction of hyperspectral images of vegetation and crops based on self-organized maps. *Information Processing in Agriculture*, 8(2):310–327, 2021.

[12] Pedram Ghamisi, Naoto Yokoya, Jun Li, Wenzhi Liao, Sicong Liu, Javier Plaza, Behnood Rasti, and Antonio Plaza. Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 5(4):37–78, 2017.

[13] Juan M Jurado, Alfonso López, Luís Pádua, and Joaquim J Sousa. Remote sensing image fusion on 3d scenarios: A review of applications for agriculture and forestry. *International Journal of Applied Earth Observation and Geoinformation*, 112:102856, 2022.

[14] Fernando Sansò and Michael G Sideris. *Geoid determination: theory and methods*. Springer Science & Business Media, 2013.

[15] Maryam Imani and Hassan Ghassemian. An overview on spectral and spatial information fusion for hyperspectral image classification: Current trends and challenges. *Information fusion*, 59:59–83, 2020.

[16] Inès Meganem, Philippe Déliot, Xavier Briottet, Yannick Deville, and Shahram Hosseini. Linear–quadratic mixing model for reflectances in urban environments. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1):544–558, 2014.

[17] Christoph C. Borel, Siegfried A.W. Gerstl, and Bill J. Powers. The radiosity method in optical remote sensing of structured 3-d surfaces. *Remote Sensing of Environment*, 36(1):13–44, 1991.

[18] Xiaowei Hu, Chi-Wing Fu, Lei Zhu, Jing Qin, and Pheng-Ann Heng. Direction-aware spatial context features for shadow detection and removal. *IEEE transactions on pattern analysis and machine intelligence*, 42(11):2795–2808, 2019.

[19] Lin He, Jun Li, Chenying Liu, and Shutao Li. Recent advances on spectral–spatial hyperspectral image classification: An overview and new guidelines. *IEEE Transactions on Geoscience and Remote Sensing*, 56(3):1579–1597, 2017.

[20] Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. Image segmentation using deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(7):3523–3542, 2021.

[21] Shaohui Mei, Jingyu Ji, Yunhao Geng, Zhi Zhang, Xu Li, and Qian Du. Unsupervised spatial–spectral feature learning by 3d convolutional autoencoder for hyperspectral classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9):6808–6820, 2019.

[22] Chao Tao, Hongbo Pan, Yansheng Li, and Zhengrou Zou. Unsupervised spectral–spatial feature learning with stacked sparse autoencoder for hyperspectral imagery classification. *IEEE Geoscience and remote sensing letters*, 12(12):2438–2442, 2015.

[23] José M Bioucas-Dias, Antonio Plaza, Nicolas Dobigeon, Mario Parente, Qian Du, Paul Gader, and Jocelyn Chanussot. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE journal of selected topics in applied earth observations and remote sensing*, 5(2):354–379, 2012.

[24] James B Campbell and Randolph H Wynne. *Introduction to remote sensing*. Guilford press, 2011.

[25] Charles Toth and Grzegorz Jóźków. Remote sensing platforms and sensors: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115:22–36, 2016.

[26] Fawwaz T Ulaby, Richard K Moore, and Adrian K Fung. Microwave remote sensing: Active and passive. volume 3-from theory to applications. 1986.

[27] Hong Wang, Jianfeng Yang, and Bin Xue. A novel method for the definition of central wavelength and spectral bandwidth. *Results in Physics*, 18:103327, 2020.

[28] Robert Hirsch. *Exploring colour photography: a complete guide*. Laurence King Publishing, 2005.

[29] Gary A Shaw and Hsiaohua K Burke. Spectral imaging for remote sensing. *Lincoln laboratory journal*, 14(1):3–28, 2003.

[30] Nathalie Pettorelli, Jon Olav Vik, Atle Mysterud, Jean-Michel Gaillard, Compton J Tucker, and Nils Chr Stenseth. Using the satellite-derived ndvi to assess ecological responses to environmental change. *Trends in ecology & evolution*, 20(9):503–510, 2005.

[31] Lei Ji, Li Zhang, Bruce K Wylie, and Jennifer Rover. On the terminology of the spectral vegetation index (nir- swir)/(nir+ swir). *International journal of remote sensing*, 32(21):6901–6909, 2011.

[32] Teodoro Semeraro, Giovanni Mastroleo, Alessandro Pomes, Andrea Luvisi, Elena Gissi, and Roberta Aretano. Modelling fuzzy combination of remote sensing vegetation index for durum wheat crop analysis. *Computers and Electronics in Agriculture*, 156:684–692, 2019.

[33] Qihao Weng. Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. *ISPRS Journal of photogrammetry and remote sensing*, 64(4):335–344, 2009.

[34] Gerrit Polder and Aoife Gowen. The hype in spectral imaging. *Journal of Spectral Imaging*, 9, 2020.

[35] Megandhren Govender, Kershani Chetty, and Hartley Bulcock. A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water Sa*, 33(2):145–151, 2007.

[36] Aoife A Gowen, Colm P O'Donnell, Patrick J Cullen, Gérard Downey, and Jesus M Frias. Hyperspectral imaging–an emerging process analytical tool for food quality and safety control. *Trends in food science & technology*, 18(12):590–598, 2007.

[37] Hilda Deborah, Noël Richard, and Jon Yngve 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*, 8(6):3224–3234, 2015.

[38] Kyu-Sung Lee, Warren B Cohen, Robert E Kennedy, Thomas K Maiersperger, and Stith T Gower. Hyperspectral versus multispectral data for estimating leaf area index in four different biomes. *Remote Sensing of Environment*, 91(3-4):508–520, 2004.

[39] Yu Fenghua, Xu Tongyu, Du Wen, Ma Hang, Zhang Guosheng, and Chen Chunling. Radiative transfer models (rtms) for field phenotyping inversion of rice based on uav hyperspectral remote sensing. *International Journal of Agricultural and Biological Engineering*, 10(4):150–157, 2017.

[40] Michael Marshall and Prasad Thenkabail. Advantage of hyperspectral eo-1 hyperion over multispectral ikonos, geoeye-1, worldview-2, landsat etm+, and modis vegetation indices in crop biomass estimation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 108:205–218, 2015.

[41] Yen-Ben Cheng, Susan L Ustin, David Riaño, and Vern C Vanderbilt. Water content estimation from hyperspectral images and modis indexes in southeastern arizona. *Remote Sensing of Environment*, 112(2):363–374, 2008.

[42] Ariolfo Camacho Velasco, César Augusto Vargas García, and Henry Arguello Fuentes. A comparative study of target detection algorithms in hyperspectral imagery applied to agricultural crops in colombia. *Tecnura*, 20(49):86–99, 2016.

[43] Umberto Amato, Anestis Antoniadis, Maria Francesca Carfora, Paolo Colandrea, Vincenzo Cuomo, Monica Franzese, Stefano Pignatti, and Carmine Serio. Statistical classification for assessing prisma hyperspectral potential for agricultural land use. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2):615–625, 2013.

[44] Shutao Li, Weiwei Song, Leyuan Fang, Yushi Chen, Pedram Ghamisi, and Jon Atli Benediktsson. Deep learning for hyperspectral image classification: An overview. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9):6690–6709, 2019.

[45] Yanheng Wang, Danfeng Hong, Jianjun Sha, Lianru Gao, Lian Liu, Yonggang Zhang, and Xianhui Rong. Spectral–spatial–temporal transformers for hyperspectral image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14, 2022.

[46] Lianchong Zhang and Junshi Xia. Flood detection using multiple chinese satellite datasets during 2020 china summer floods. *Remote Sensing*, 14(1):51, 2021.

[47] Diana Krupnik and Shuhab Khan. Close-range, ground-based hyperspectral imaging for mining applications at various scales: Review and case studies. *Earth-science reviews*, 198:102952, 2019.

[48] Da-Wen Sun. *Hyperspectral imaging for food quality analysis and control*. Elsevier, 2010.

[49] Gamal ElMasry, Douglas F Barbin, Da-Wen Sun, and Paul Allen. Meat quality evaluation by hyperspectral imaging technique: an overview. *Critical reviews in food science and nutrition*, 52(8):689–711, 2012.

[50] Yidan Bao, Chunxiao Mi, Na Wu, Fei Liu, and Yong He. Rapid classification of wheat grain varieties using hyperspectral imaging and chemometrics. *Applied Sciences*, 9(19):4119, 2019.

[51] Zhi Liu, Hongjun Wang, and Qingli Li. Tongue tumor detection in medical hyperspectral images. *Sensors*, 12(1):162–174, 2011.

[52] Zhi Liu, Hongjun Wang, and Qingli Li. Tongue tumor detection in medical hyperspectral images. *Sensors*, 12(1):162–174, 2011.

[53] Silvia Serranti, Aldo Gargiulo, and Giuseppe Bonifazi. Characterization of post-consumer polyolefin wastes by hyperspectral imaging for quality control in recycling processes. *Waste Management*, 31(11):2217–2227, 2011.

[54] Silvia Serranti, Aldo Gargiulo, and Giuseppe Bonifazi. Hyperspectral imaging for process and quality control in recycling plants of polyolefin flakes. *Journal of Near Infrared Spectroscopy*, 20(5):573–581, 2012.

[55] John R Schott. *Remote sensing: the image chain approach*. Oxford University Press on Demand, 2007.

[56] L Frank, SJ Pedrotti, Leno M Pedrotti, and Leno S Pedrotti. Introduction to optics, 1993.

[57] Fred E Nicodemus. Directional reflectance and emissivity of an opaque surface. *Applied optics*, 4(7):767–775, 1965.

[58] Simon Niedenthal. Learning from the cornell box. *Leonardo*, 35(3):249–254, 2002.

[59] Huihui Song, Bo Huang, and Kaihua Zhang. Shadow detection and reconstruction in high-resolution satellite images via morphological filtering and example-based learning. *IEEE Transactions on geoscience and remote sensing*, 52(5):2545–2554, 2013.

[60] Lee Harrison, Joseph Michalsky, and Jerry Berndt. Automated multifilter rotating shadow-band radiometer: an instrument for optical depth and radiation measurements. *Applied Optics*, 33(22):5118–5125, 1994.

[61] lucas Silva. My Research Software, 2018.

[62] Nicolas Dobigeon, Jean-Yves Tourneret, Cédric Richard, José Carlos M Bermudez, Stephen McLaughlin, and Alfred O Hero. Nonlinear unmixing of hyperspectral images: Models and algorithms. *IEEE Signal Processing Magazine*, 31(1):82–94, 2013.

[63] R Richter and D Schläpfer. Atmospheric/topographic correction for airborne imagery. *ATCOR-4 user guide*, pages 565–02, 2011.

[64] Jasmine S Bartlett, Áurea M Ciotti, Richard F Davis, and John J Cullen. The spectral effects of clouds on solar irradiance. *Journal of Geophysical Research: Oceans*, 103(C13):31017–31031, 1998.

[65] Eric F Vermote and Svetlana Kotchenova. Atmospheric correction for the monitoring of land surfaces. *Journal of Geophysical Research: Atmospheres*, 113(D23), 2008.

[66] Rob Heylen, Mario Parente, and Paul Gader. A review of nonlinear hyperspectral unmixing methods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):1844–1868, 2014.

[67] José M Bioucas-Dias and José MP Nascimento. Hyperspectral subspace identification. *IEEE Transactions on Geoscience and Remote Sensing*, 46(8):2435–2445, 2008.

[68] JC Harsanyi, WH Farrand, and Chein-I Chang. Determining the number and identity of spectral endmembers: An integrated approach using neyman-pearson eigen-thresholding and iterative constrained rms error minimization. In *Proceedings of the Thematic Conference on Geologic Remote Sensing*, volume 1, pages 395–395. Environmental Research Institute of Michigan, 1993.

[69] Chein-I Chang and Qian Du. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. *IEEE Transactions on geoscience and remote sensing*, 42(3):608–619, 2004.

[70] Chein-I Chang and Antonio Plaza. A fast iterative algorithm for implementation of pixel purity index. *IEEE Geoscience and Remote Sensing Letters*, 3(1):63–67, 2006.

[71] José MP Nascimento and José MB Dias. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE transactions on Geoscience and Remote Sensing*, 43(4):898–910, 2005.

[72] C-I Chang, C-C Wu, Weimin Liu, and Y-C Ouyang. A new growing method for simplex-based endmember extraction algorithm. *IEEE transactions on geoscience and remote sensing*, 44(10):2804–2819, 2006.

[73] Dimitris G Manolakis, Ronald B Lockwood, and Thomas W Cooley. *Hyperspectral imaging remote sensing: physics, sensors, and algorithms*. Cambridge University Press, 2016.

[74] Christoph C Borel and Siegfried A.W Gerstl. Nonlinear spectral mixing models for vegetative and soil surfaces. *Remote Sensing of Environment*, 47(3):403–416, 1994.

[75] Qi Wei, José Bioucas-Dias, Nicolas Dobigeon, Jean-Yves Tourneret, Marcus Chen, and Simon Godsill. Multiband image fusion based on spectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12):7236–7249, 2016.

[76] Nirmal Keshava and John F Mustard. Spectral unmixing. *IEEE signal processing magazine*, 19(1):44–57, 2002.

[77] Marian-Daniel Iordache, José M Bioucas-Dias, and Antonio Plaza. Sparse unmixing of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):2014–2039, 2011.

[78] Jing Yao, Danfeng Hong, Lin Xu, Deyu Meng, Jocelyn Chanussot, and Zongben Xu. Sparsity-enhanced convolutional decomposition: A novel tensor-based paradigm for blind hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14, 2021.

[79] Nematollah Omidikia, Mahdiyeh Ghaffari, and Róbert Rajkó. Sparse non-negative multivariate curve resolution: L0, l1, or l2 norms? *Chemometrics and Intelligent Laboratory Systems*, 199:103969, 2020.

[80] Jonas Ranstam and JA Cook. Lasso regression. *Journal of British Surgery*, 105(10):1348–1348, 2018.

[81] Junmin Liu and Jiangshe Zhang. Spectral unmixing via compressive sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 52(11):7099–7110, 2014.

[82] Marian-Daniel Iordache, José M Bioucas-Dias, and Antonio Plaza. Total variation spatial regularization for sparse hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 50(11):4484–4502, 2012.

[83] Marian-Daniel Iordache, José M Bioucas-Dias, and Antonio Plaza. Sparse unmixing of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):2014–2039, 2011.

[84] Tatsumi Uezato, Mathieu Fauvel, and Nicolas Dobigeon. Hyperspectral image unmixing with lidar data-aided spatial regularization. *IEEE Transactions on Geoscience and Remote Sensing*, 56(7):4098–4108, 2018.

[85] Zina Mitraka, Fabio Del Frate, and Francesco Carbone. Nonlinear spectral unmixing of landsat imagery for urban surface cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(7):3340–3350, 2016.

[86] José MP Nascimento and José M Bioucas-Dias. Nonlinear mixture model for hyperspectral unmixing. In *Image and Signal Processing for Remote Sensing XV*, volume 7477, page 74770I. International Society for Optics and Photonics, 2009.

[87] Wenyi Fan, Baoxin Hu, John Miller, and Mingze Li. Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data. *International Journal of Remote Sensing*, 30(11):2951–2962, 2009.

[88] Yoann Altmann, Abderrahim Halimi, Nicolas Dobigeon, and Jean-Yves Tourneret. Supervised nonlinear spectral unmixing using a polynomial post nonlinear model for hyperspectral imagery. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1009–1012. IEEE, 2011.

[89] Abderrahim Halimi, Yoann Altmann, Nicolas Dobigeon, and Jean-Yves Tourneret. Nonlinear unmixing of hyperspectral images using a generalized bilinear model. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4153–4162, 2011.

[90] DB Nash and JE Conel. Spectral reflectance systematics for mixtures of powdered hypersthene, labradorite, and ilmenite. *Journal of Geophysical Research*, 79(11):1615–1621, 1974.

[91] Bruce Hapke. Bidirectional reflectance spectroscopy: 1. theory. *Journal of Geophysical Research: Solid Earth*, 86(B4):3039–3054, 1981.

[92] P Kubelka and F Munk. Reflection characteristics of paints. *Zeitschrift fur Technische Physik*, 12:593–601, 1931.

[93] Yurij Shkuratov, Larissa Starukhina, Harald Hoffmann, and Gabriele Arnold. A model of spectral albedo of particulate surfaces: Implications for optical properties of the moon. *Icarus*, 137(2):235–246, 1999.

[94] Yosio Edemir Shimabukuro and James A Smith. The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. *IEEE Transactions on Geoscience and Remote sensing*, 29(1):16–20, 1991.

[95] Rasmus Bro and Sijmen De Jong. A fast non-negativity-constrained least squares algorithm. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 11(5):393–401, 1997.

[96] Abderrahim Halimi, Yoann Altmann, Nicolas Dobigeon, and Jean-Yves Tourneret. Unmixing hyperspectral images using the generalized bilinear model. In *2011 IEEE International Geoscience and Remote Sensing Symposium*, pages 1886–1889. IEEE, 2011.

[97] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122, 2011.

[98] Manya V Afonso, José M Bioucas-Dias, and Mário AT Figueiredo. An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE transactions on image processing*, 20(3):681–695, 2010.

[99] Kezhi Li, Martin Sundin, Cristian R Rojas, Saikat Chatterjee, and Magnus Jansson. Alternating strategies with internal admm for low-rank matrix reconstruction. *Signal Processing*, 121:153–159, 2016.

[100] Yasser Mostafa. A review on various shadow detection and compensation techniques in remote sensing images. *Canadian journal of remote sensing*, 43(6):545–562, 2017.

[101] Paul M Dare. Shadow analysis in high-resolution satellite imagery of urban areas. *Photogrammetric Engineering & Remote Sensing*, 71(2):169–177, 2005.

[102] Edward A Ashton, Brian D Wemett, Robert A Leathers, and Trijntje V Downes. A novel method for illumination suppression in hyperspectral images. In *Algorithms and technologies for multispectral, hyperspectral, and ultraspectral imagery XIV*, volume 6966, page 69660C. International Society for Optics and Photonics, 2008.

[103] Fumio Yamazaki, Wen Liu, and Makiko Takasaki. Characteristics of shadow and removal of its effects for remote sensing imagery. In *2009 IEEE International Geoscience and Remote Sensing Symposium*, volume 4, pages IV–426. IEEE, 2009.

[104] Francois PS Luus, Frans Van den Bergh, and Bodhaswar Tikanath Jugpershad Maharaj. Adaptive threshold-based shadow masking for across-date settlement classification of panchromatic quickbird images. *IEEE Geoscience and Remote Sensing Letters*, 11(6):1153–1157, 2013.

[105] Nan Su, Ye Zhang, Shu Tian, Yiming Yan, and Xinyuan Miao. Shadow detection and removal for occluded object information recovery in urban high-resolution panchromatic satellite images. 9:2568–2582, 2016.

[106] Pierre Soille et al. *Morphological image analysis: principles and applications*, volume 2. Springer, 1999.

[107] Huifang Li, Liangpei Zhang, and Huanfeng Shen. An adaptive nonlocal regularized shadow removal method for aerial remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1):106–120, 2013.

[108] AmirReza Shahtahmassebi, Ning Yang, Ke Wang, Nathan Moore, and Zhangquan Shen. Review of shadow detection and de-shadowing methods in remote sensing. *Chinese Geographical Science*, 23(4):403–420, 2013.

[109] Yanfeng Wei, Zhongming Zhao, and Jianghong Song. Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection. In *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*, volume 3, pages 2008–2010. Ieee, 2004.

[110] Francois PS Luus, Frans Van den Bergh, and Bodhaswar TJ Maharaj. The effects of segmentation-based shadow removal on across-date settlement type classification of panchromatic quickbird images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(3):1274–1285, 2013.

[111] KA Divya, KI Roshna, and Shelmy Mathai. Shadow detection and removal by object-wise segmentation. In *2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pages 1–4. IEEE, 2015.

[112] Takashi Nakajima, Guo Tao, and Yoshifumi Yasuoka. Simulated recovery of information in shadow areas on ikonos image by combing als data. In *Proceeding of Asian conference on remote sensing (ACRS)*, 2002.

[113] Qingming Zhan, Wenzhong Shi, and Yinghui Xiao. Quantitative analysis of shadow effects in high-resolution images of urban areas. *International Archives of Photogrammetry and Remote Sensing*, 36(8/W27), 2005.

[114] Gustav Tolt, Michal Shimoni, and Jörgen Ahlberg. A shadow detection method for remote sensing images using vhr hyperspectral and lidar data. In *2011 IEEE international geoscience and remote sensing symposium*, pages 4423–4426. IEEE, 2011.

[115] Makoto Nagao, Takashi Matsuyama, and Yoshio Ikeda. Region extraction and shape analysis in aerial photographs. *Computer Graphics and Image Processing*, 10(3):195–223, 1979.

[116] Y. Chen, D. Wen, L. Jing, and P. Shi. Shadow information recovery in urban areas from very high resolution satellite imagery. 28:3249–3254, 2007.

[117] Karine RM Adeline, M Chen, X Briottet, SK Pang, and N Paparoditis. Shadow detection in very high spatial resolution aerial images: A comparative study. *ISPRS Journal of Photogrammetry and Remote Sensing*, 80:21–38, 2013.

[118] Xiaojun Qiao, Deshuai Yuan, and Hui Li. Urban shadow detection and classification using hyperspectral image. *Journal of the Indian Society of Remote Sensing*, 45(6):945–952, 2017.

[119] Victor JD Tsai. A comparative study on shadow compensation of color aerial images in invariant color models. *IEEE transactions on geoscience and remote sensing*, 44(6):1661–1671, 2006.

[120] Pooya Sarabandi, Fumio Yamazaki, Masashi Matsuoka, and Anne Kiremidjian. Shadow detection and radiometric restoration in satellite high resolution images. In *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*, volume 6, pages 3744–3747. IEEE, 2004.

[121] Haijian Ma, Qiming Qin, and Xinyi Shen. Shadow segmentation and compensation in high resolution satellite images. In *IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium*, volume 2, pages II–1036. IEEE, 2008.

[122] Dong Cai, Manchun Li, Zhiliang Bao, Zhenjie Chen, Wei Wei, and Hao Zhang. Study on shadow detection method on high resolution remote sensing image based on his space transformation and ndvi index. In *2010 18th International Conference on Geoinformatics*, pages 1–4. IEEE, 2010.

[123] Nicolas Martel-Brisson and Andre Zaccarin. Moving cast shadow detection from a gaussian mixture shadow model. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 643–648. IEEE, 2005.

[124] Martin D Levine and Jisnu Bhattacharyya. Removing shadows. *Pattern recognition letters*, 26(3):251–265, 2005.

[125] Salman Hameed Khan, Mohammed Bennamoun, Ferdous Sohel, and Roberto Togneri. Automatic feature learning for robust shadow detection. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1939–1946. IEEE, 2014.

[126] Hanna M Wallach. Conditional random fields: An introduction. *Technical Reports (CIS)*, page 22, 2004.

[127] Tomás F Yago Vicente, Le Hou, Chen-Ping Yu, Minh Hoai, and Dimitris Samaras. Large-scale training of shadow detectors with noisily-annotated shadow examples. In *European Conference on Computer Vision*, pages 816–832. Springer, 2016.

[128] Naoto Inoue and Toshihiko Yamasaki. Learning from synthetic shadows for shadow detection and removal. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(11):4187–4197, 2020.

[129] Vu Nguyen, Tomas F Yago Vicente, Maozheng Zhao, Minh Hoai, and Dimitris Samaras. Shadow detection with conditional generative adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4510–4518, 2017.

[130] Jifeng Wang, Xiang Li, and Jian Yang. Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal, 2018.

[131] Graham D Finlayson, Steven D Hordley, Cheng Lu, and Mark S Drew. On the removal of shadows from images. *IEEE transactions on pattern analysis and machine intelligence*, 28(1):59–68, 2005.

[132] Eli Arbel and Hagit Hel-Or. Shadow removal using intensity surfaces and texture anchor points. *IEEE transactions on pattern analysis and machine intelligence*, 33(6):1202–1216, 2010.

[133] Ling Zhang, Qing Zhang, and Chunxia Xiao. Shadow remover: Image shadow removal based on illumination recovering optimization. *IEEE Transactions on Image Processing*, 24(11):4623–4636, 2015.

[134] Yao Xiao, Efstratios Tsougenis, and Chi-Keung Tang. Shadow removal from single rgb-d images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3011–3018, 2014.

[135] Luca Lorenzi, Farid Melgani, and Grégoire Mercier. A complete processing chain for shadow detection and reconstruction in vhr images. *IEEE transactions on geoscience and remote sensing*, 50(9):3440–3452, 2012.

[136] Bin Ding, Chengjiang Long, Ling Zhang, and Chunxia Xiao. Argan: Attentive recurrent generative adversarial network for shadow detection and removal. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 10213–10222, 2019.

[137] Ruiqi Guo, Qieyun Dai, and Derek Hoiem. Paired regions for shadow detection and removal. *IEEE transactions on pattern analysis and machine intelligence*, 35(12):2956–2967, 2012.

[138] Jean-François Lalonde, Alexei A Efros, and Srinivasa G Narasimhan. Detecting ground shadows in outdoor consumer photographs. In *Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part II 11*, pages 322–335. Springer, 2010.

[139] Jiejie Zhu, Kegan GG Samuel, Syed Z Masood, and Marshall F Tappen. Learning to recognize shadows in monochromatic natural images. In *2010 IEEE Computer Society conference on computer vision and pattern recognition*, pages 223–230. IEEE, 2010.

[140] Hieu Le and Dimitris Samaras. Shadow removal via shadow image decomposition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8578–8587, 2019.

[141] Xiaowei Hu, Yitong Jiang, Chi-Wing Fu, and Pheng-Ann Heng. Mask-shadowgan: Learning to remove shadows from unpaired data. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 2472–2481, 2019.

[142] Florin-Alexandru Vasluianu, Andrés Romero, Luc Van Gool, and Radu Timofte. Shadow removal with paired and unpaired learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 826–835, 2021.

[143] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.

[144] Dominic Rüfenacht, Clément Fredembach, and Sabine Süsstrunk. Automatic and accurate shadow detection using near-infrared information. *IEEE transactions on pattern analysis and machine intelligence*, 36(8):1672–1678, 2013.

[145] Stuart Lloyd. Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137, 1982.

[146] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20:273–297, 1995.

[147] Jian Yang, Yuhong He, and John Caspersen. Fully constrained linear spectral unmixing based global shadow compensation for high resolution satellite imagery of urban areas. *International Journal of Applied Earth Observation and Geoinformation*, 38:88–98, 2015.

[148] R Richter and A Müller. De-shadowing of satellite/airborne imagery. *International Journal of Remote Sensing*, 26(15):3137–3148, 2005.

[149] Steven M Adler-Golden, Michael W Matthew, Gail P Anderson, Gerald W Felde, and James A Gardner. Algorithm for de-shadowing spectral imagery. In *Imaging Spectrometry VIII*, volume 4816, pages 203–210. International Society for Optics and Photonics, 2002.

[150] Rob Heylen and Paul Scheunders. A multilinear mixing model for nonlinear spectral unmixing. *IEEE transactions on geoscience and remote sensing*, 54(1):240–251, 2015.

[151] Rob Heylen, Vera Andrejchenko, Zohreh Zahiri, Mario Parente, and Paul Scheunders. Nonlinear hyperspectral unmixing with graphical models. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7):4844–4856, 2019.

[152] Mark Cameron and Lalit Kumar. Diffuse skylight as a surrogate for shadow detection in high-resolution imagery acquired under clear sky conditions. *Remote Sensing*, 10(8):1185, 2018.

[153] Fred A Kruse, AB Lefkoff, JW Boardman, KB Heidebrecht, AT Shapiro, PJ Barloon, and AFH Goetz. The spectral image processing system (sips)-interactive visualization and analysis of imaging spectrometer data. In *AIP Conference Proceedings*, volume 283, pages 192–201. American Institute of Physics, 1993.

[154] Fatih Omruuzun, Didem Ozisik Baskurt, Hazan Daglayan, and Yasemin Yardimci Cetin. Shadow removal from vnir hyperspectral remote sensing imagery with endmember signature analysis. In *Next-Generation Spectroscopic Technologies VIII*, volume 9482, page 94821F. International Society for Optics and Photonics, 2015.

[155] Min Zhao, Jie Chen, and Susanto Rahardja. Hyperspectral shadow removal via nonlinear unmixing. *IEEE Geoscience and Remote Sensing Letters*, 18(5):881–885, 2020.

[156] Guillaume Roussel, Christiane Weber, Xavier Ceamanos, and Xavier Briottet. A sun/shadow approach for the classification of hyperspectral data. In *2016 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–5. IEEE, 2016.

[157] Puhong Duan, Shangsong Hu, Xudong Kang, and Shutao Li. Shadow removal of hyperspectral remote sensing images with multiexposure fusion. 60:1–11, 2022.

[158] Yindan Zhang, Gang Chen, Jelena Vukomanovic, Kunwar K Singh, Yong Liu, Samuel Holden, and Ross K Meentemeyer. Recurrent shadow attention model (rsam) for shadow removal in high-resolution urban land-cover mapping. *Remote Sensing of Environment*, 247:111945, 2020.

[159] Min Zhao, Longbin Yan, and Jie Chen. Hyperspectral image shadow compensation via cycle-consistent adversarial networks. *Neurocomputing*, 450:61–69, 2021.

[160] Andrea Marinoni, Antonio Plaza, and Paolo Gamba. Harmonic mixture modeling for efficient nonlinear hyperspectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(9):4247–4256, 2016.

[161] Andrea Marinoni and Paolo Gamba. A novel approach for efficient p-linear hyperspectral unmixing. *IEEE Journal of Selected Topics in Signal Processing*, 9(6):1156–1168, 2015.

[162] Antonio Plaza, Pablo Martínez, Rosa Pérez, and Javier Plaza. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. *IEEE transactions on geoscience and remote sensing*, 42(3):650–663, 2004.

[163] Tomohiro Matsuki, Naoto Yokoya, and Akira Iwasaki. Hyperspectral tree species classification of japanese complex mixed forest with the aid of lidar data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(5):2177–2187, 2015.

[164] Tatsumi Uezato, Naoto Yokoya, and Wei He. Illumination invariant hyperspectral image unmixing based on a digital surface model. *IEEE Transactions on Image Processing*, 29:3652–3664, 2020.

[165] Bin Yang. Supervised nonlinear hyperspectral unmixing with automatic shadow compensation using multiswarm particle swarm optimization. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–18, 2022.

[166] Andries P. Engelbrecht. *Fundamentals of Computational Swarm Intelligence*. Wiley, 2005.

[167] Riccardo Poli, James Kennedy, and Tim Blackwell. Particle swarm optimization. *Swarm Intell.*, 1(1):33–57, 2007.

[168] Jiadi Yin, Jinwei Dong, Nicholas AS Hamm, Zhichao Li, Jianghao Wang, Hanfa Xing, and Ping Fu. Integrating remote sensing and geospatial big data for urban land use mapping: A review. *International Journal of Applied Earth Observation and Geoinformation*, 103:102514, 2021.

[169] Le Wang, Chen Shi, Chunyuan Diao, Wenjie Ji, and Dameng Yin. A survey of methods incorporating spatial information in image classification and spectral unmixing. *International Journal of Remote Sensing*, 37(16):3870–3910, 2016.

[170] P d'Angelo and F Kurz. Aircraft based real time bundle adjustment and digital surface model generation. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42:1643–1647, 2019.

[171] Ruisheng Wang, Jiju Peethambaran, and Dong Chen. Lidar point clouds to 3-d urban models : a review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(2):606–627, 2018.

[172] Xiaozhi Chen, Kaustav Kundu, Yukun Zhu, Huimin Ma, Sanja Fidler, and Raquel Urtasun. 3d object proposals using stereo imagery for accurate object class detection. *IEEE transactions on pattern analysis and machine intelligence*, 40(5):1259–1272, 2017.

[173] Paul L Basgall, Fred A Kruse, and Richard C Olsen. Comparison of lidar and stereo photogrammetric point clouds for change detection. In *Laser Radar Technology and Applications XIX; and Atmospheric Propagation XI*, volume 9080, pages 214–227. SPIE, 2014.

[174] Qiang Zhang, V Paúl Pauca, Robert J Plemmons, and D Dejan Nikic. Detecting objects under shadows by fusion of hyperspectral and lidar data: A physical model approach. In *2013 5th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–4. IEEE, 2013.

[175] Chen Shi and Le Wang. Incorporating spatial information in spectral unmixing: A review. *Remote Sensing of Environment*, 149:70–87, 2014.

[176] Stamatios Lefkimmiatis and Stanley Osher. Nonlocal structure tensor functionals for image regularization. *IEEE Transactions on Computational Imaging*, 1(1):16–29, 2015.

[177] Jing Yao, Deyu Meng, Qian Zhao, Wenfei Cao, and Zongben Xu. Nonconvex-sparsity and nonlocal-smoothness-based blind hyperspectral unmixing. *IEEE Transactions on Image Processing*, 28(6):2991–3006, 2019.

[178] F Van Der Meer. Iterative spectral unmixing (isu). *International Journal of Remote Sensing*, 20(17):3431–3436, 1999.

[179] Olivier Eches, Nicolas Dobigeon, and Jean-Yves Tourneret. Enhancing hyperspectral image unmixing with spatial correlations. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4239–4247, 2011.

[180] Persi Diaconis. The markov chain monte carlo revolution. *Bulletin of the American Mathematical Society*, 46(2):179–205, 2009.

[181] Danfeng Hong, Lianru Gao, Jing Yao, Naoto Yokoya, Jocelyn Chanussot, Uta Heiden, and Bing Zhang. Endmember-guided unmixing network (egu-net): A general deep learning framework for self-supervised hyperspectral unmixing. *IEEE Transactions on Neural Networks and Learning Systems*, 33(11):6518–6531, 2021.

[182] Xianfeng Song, Xiaoguang Jiang, and Xiaoping Rui. Spectral unmixing using linear unmixing under spatial autocorrelation constraints. In *2010 IEEE International Geoscience and Remote Sensing Symposium*, pages 975–978. IEEE, 2010.

[183] Alexey Castrodad, Zhengming Xing, John B Greer, Edward Bosch, Lawrence Carin, and Guillermo Sapiro. Learning discriminative sparse representations for modeling, source separation, and mapping of hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4263–4281, 2011.

[184] Pierre-Antoine Thouvenin, Nicolas Dobigeon, and Jean-Yves Tourneret. Hyperspectral unmixing with spectral variability using a perturbed linear mixing model. *IEEE Transactions on Signal Processing*, 64(2):525–538, 2015.

[185] Junmin Liu, Jiangshe Zhang, Yuelin Gao, Chunxia Zhang, and Zhihua Li. Enhancing spectral unmixing by local neighborhood weights. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(5):1545–1552, 2012.

[186] Yanfei Zhong, Ruyi Feng, and Liangpei Zhang. Non-local sparse unmixing for hyperspectral remote sensing imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):1889–1909, 2013.

[187] Maciel Zortea and Antonio Plaza. Spatial preprocessing for endmember extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 47(8):2679–2693, 2009.

[188] Daniele Cerra, Miguel Pato, Kevin Alonso, Claas Köhler, Mathias Schneider, Raquel de los Reyes, Emiliano Carmona, Rudolf Richter, Franz Kurz, Peter Reinartz, et al. Dlr hysu—a benchmark dataset for spectral unmixing. *Remote Sensing*, 13(13):2559, 2021.

[189] Karim Lenhard, Andreas Baumgartner, and Thomas Schwarzmaier. Independent laboratory characterization of neo hyspex imaging spectrometers vnir-1600 and swir-320m-e. *IEEE Transactions on Geoscience and Remote Sensing*, 53(4):1828–1841, 2014.

[190] Thomas Krauß, Pablo d'Angelo, Mathias Schneider, and Veronika Gstaiger. The fully automatic optical processing system catena at dlr. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40:177–183, 2013.

[191] Andreas Baumgartner and Claas Henning Köhler. Transformation of point spread functions on an individual pixel scale. *Optics Express*, 28(26):38682–38697, 2020.

[192] Peter Schwind, Mathias Schneider, and Rupert Müller. Improving hyspex sensor co-registration accuracy using brisk and sensor-model based ransac. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40:371–376, 2014.

[193] R Muller, Manfred Lehner, Rainer Muller, Peter Reinartz, Manfred Schroeder, and Beate Vollmer. A program for direct georeferencing of airborne and spaceborne line scanner images. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(1):148–153, 2002.

[194] Claas Henning Köhler. Airborne imaging spectrometer hyspex. *Journal of large-scale research facilities JLSRF*, 2(A93):1–6, 2016.

[195] Franz Kurz, Sebastian Tuermer, Oliver Meynberg, Dominik Rosenbaum, Hartmut Runge, Peter Reinartz, and Jens Leitloff. Low-cost optical camera systems for real-time mapping applications. *Photogrammetrie-Fernerkundung-Geoinformation*, pages 159–176, 2012.

[196] Daniel C Heinz et al. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE transactions on geoscience and remote sensing*, 39(3):529–545, 2001.

[197] Helge Aasen, Juliane Bendig, Andreas Bolten, Simon Bennertz, Max Willkomm, and Georg Bareth. Introduction and preliminary results of a calibration for full-frame hyperspectral cameras to monitor agricultural crops with uavs. In *ISPRS Technical Commission VII Symposium*, volume 40, pages 1–8. Copernicus GmbH, 2014.

[198] Guichen Zhang, Daniele Cerra, and Rupert Müller. Shadow detection and restoration for hyperspectral images based on nonlinear spectral unmixing. *Remote Sensing*, 12(23):3985, 2020.

[199] Guichen Zhang, Paul Scheunders, Daniele Cerra, and Rupert Müller. Shadow-aware nonlinear spectral unmixing for hyperspectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15:5514–5533, 2022.

[200] Guichen Zhang, Paul Scheunders, and Daniele Cerra. Shadow-aware nonlinear spectral unmixing with spatial regularization. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–16, 2023.

[201] José M Bioucas-Dias, Antonio Plaza, Gustavo Camps-Valls, Paul Scheunders, Nasser Nasrabadi, and Jocelyn Chanussot. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and remote sensing magazine*, 1(2):6–36, 2013.

[202] Ruiqi Guo, Qieyun Dai, and Derek Hoiem. Single-image shadow detection and removal using paired regions. In *CVPR 2011*, pages 2033–2040. IEEE, 2011.

[203] Nan Mo, Ruixi Zhu, Li Yan, and Zhan Zhao. Deshadowing of urban airborne imagery based on object-oriented automatic shadow detection and regional matching compensation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(2):585–605, 2018.

[204] Guichen Zhang, Daniele Cerra, and Rupert Mueller. Towards the spectral restoration of shadowed areas in hyperspectral images based on nonlinear unmixing. In *2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–5. IEEE, 2019.

[205] Guichen Zhang, Daniele Cerra, and Rupert Mueller. Improving the classification in shadowed areas using nonlinear spectral unmixing. IEEE. Manuscript submitted to 2020 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), in press.

[206] Philip N Slater, FJ Doyle, NL Fritz, and R Welch. Photographic systems for remote sensing. *Manual of remote sensing*, 1:231–291, 1983.

[207] Kristian Bredies, Karl Kunisch, and Thomas Pock. Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010.

[208] Ricardo Augusto Borsoi, Tales Imbiriba, José Carlos Moreira Bermudez, Cédric Richard, Jocelyn Chanussot, Lucas Drumetz, Jean-Yves Tourneret, Alina Zare, and Christian Jutten. Spectral variability in hyperspectral data unmixing: A comprehensive review. *arXiv preprint arXiv:2001.07307*, 2020.

[209] John Canny. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6):679–698, 1986.

[210] Ben Somers, Maciel Zortea, Antonio Plaza, and Gregory P Asner. Automated extraction of image-based endmember bundles for improved spectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2):396–408, 2012.

[211] Xiaodong Cun, Chi-Man Pun, and Cheng Shi. Towards ghost-free shadow removal via dual hierarchical aggregation network and shadow matting gan. In *AAAI*, pages 10680–10687, 2020.

[212] Lloyd Windrim, Rishi Ramakrishnan, Arman Melkumyan, and Richard J Murphy. A physics-based deep learning approach to shadow invariant representations of hyperspectral images. *IEEE Transactions on Image Processing*, 27(2):665–677, 2017.

[213] Salman H Khan, Mohammed Bennamoun, Ferdous Sohel, and Roberto Togneri. Automatic shadow detection and removal from a single image. *IEEE transactions on pattern analysis and machine intelligence*, 38(3):431–446, 2015.

[214] Marie K Svensson. Sky view factor analysis–implications for urban air temperature differences. *Meteorological applications*, 11(3):201–211, 2004.

[215] DG Kaskaoutis, HD Kambezidis, and Z Toth. Investigation about the dependence of spectral diffuse-to-direct-beam irradiance ratio on atmospheric turbidity and solar zenith angle. *Theoretical and applied climatology*, 89(3):245–256, 2007.

[216] Richard E Bird and Carol Riordan. Simple solar spectral model for direct and diffuse irradiance on horizontal and tilted planes at the earth's surface for cloudless atmospheres. *Journal of Applied Meteorology and Climatology*, 25(1):87–97, 1986.

[217] Muhammad Jaleed Khan, Hamid Saeed Khan, Adeel Yousaf, Khurram Khurshid, and Asad Abbas. Modern trends in hyperspectral image analysis: A review. *Ieee Access*, 6:14118–14129, 2018.

[218] Bin Yang and Bin Wang. Band-wise nonlinear unmixing for hyperspectral imagery using an extended multilinear mixing model. *IEEE Transactions on Geoscience and Remote Sensing*, 56(11):6747–6762, 2018.

[219] Olaf Conrad, Benjamin Bechtel, Michael Bock, Helge Dietrich, Elke Fischer, Lars Gerlitz, Jan Wehberg, Volker Wichmann, and Jürgen Böhner. System for automated geoscientific analyses (saga) v. 2.1. 4. *Geoscientific Model Development*, 8(7):1991–2007, 2015.