Close-Range Urban Material Classification with Hyperspectral and Textural Features

Zur Erlangung des akademischen Grades einer
Doktor-Ingenieurin

von der KIT-Fakultät
für Bauingenieur-, Geo- und Umweltwissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte
Dissertation

von

M.Sc. Rebecca Maria Ilehag
aus Stockholm (Schweden)

Tag der mündlichen Prüfung: 10.07.2020

Hauptreferent:	Prof. Dr. -Ing. habil. Stefan Hinz Institut für Photogrammetrie und Fernerkundung Karlsruher Institut für Technologie
Korreferent:	Prof. Dr. -Ing. Peter Reinartz Institut für Methodik der Fernerkundung Photogrammetrie und Bildanalyse Deutsches Zentrum für Luft- und Raumfahrt (DLR)

Karlsruhe 2020

Eidesstattliche Versicherung gemäß § 6 Abs. 1 Ziff. 4 der Promotionsordnung des Karlsruher Instituts für Technologie (KIT) für die KIT-Fakultät für Bauingenieur-, Geo- und Umweltwissenschaften

1. Bei der eingereichten Dissertation zu dem Thema

   Close-Range Urban Material Classification with Hyperspectral and Textural Features

   handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Ort und Datum Unterschrift

Abstract

With the recent advancements of unmanned aerial vehicles, close-range scene analysis of the urban environment is an expanding research field. This can either include assessment of buildings, such as thermal infrared analysis to determine the heating and cooling demand, or the classification of the exterior building materials as supplementary information for 3D city models and thermal 3D city simulations. The knowledge about the materials in the urban environment can likewise assist with localizing areas where the materials have low reflectance and are therefore vulnerable to the urban heat island effect. Material classification is commonly performed with high-dimensional hyperspectral data since the spectral characteristic features can easier be distinguished due to continuous data over a wide spectral range. Urban material spectral libraries contain such spectra that can be used for classification. The spectral libraries categorize their samples using an urban material taxonomy. However, spectral characteristic features can be very similar for different urban materials. A high inter-class spectral similarity (high inter-class correlation) is caused by similar chemical compositions of the materials. Misclassification frequently occurs since it is difficult to distinguish those material classes based on the spectral data. Furthermore, as no standardization exists concerning urban material taxonomy, the taxonomy of spectral libraries is usually not comparable.

Due to the aforementioned challenges, this thesis presents several analyzed aspects and created solutions concerning urban materials. It can be observed that different spectral libraries inconsistency define the taxonomy. To improve this, the taxonomy of such libraries should not be based on the local material diversity, but on global diversity. Consequently is the first spectral library KLUM developed which focuses on building facade materials. An intra-class spectral dissimilarity can be observed based on both the spectra from KLUM and on spectra from existing spectral libraries. Lastly, the relationship between spectral and textural features has been examined for the classification of urban materials. It can be concluded that spectral features are deemed more important to distinguish urban materials as the textural features are not improving the distinction. The classification results further indicate that spectral and textural features are not correlated with each other for certain materials.

Zusammenfassung

Mit der stetigen Weiterentwicklung von Drohnen ist die Nahbereichs-Szenenanalyse der urbanen Umgebung zu einem expandierenden Forschungsfeld geworden. Dies kann die Analyse von Gebäuden einschließen, wie z.B. im Rahmen einer Wärmeanalyse mit Wärmebildern zur Bestimmung des Heiz- und Kühlbedarfs. Dies kann ebenfalls die Klassifizierung der äußeren Baumaterialien als Zusatzinformation für 3D-Stadtmodelle und thermische 3D-Stadtsimulationen umfassen. Das Wissen über die Materialien in der urbanen Umgebung kann ebenfalls bei der Lokalisierung von Bereichen helfen, in denen Materialien ein niedriges Rückstrahlungsvermögen aufweisen und deswegen eine Gefährdung hinsichtlich urbaner Wärmeinseln besteht. Die Materialklassifizierung wird üblicherweise mit hochdimensionalen hyperspektralen Daten durchgeführt, weil die spektralen charakteristischen Merkmale aufgrund der kontinuierlichen Daten über einen breiten Spektralbereich leichter unterschieden werden können. Urbane spektrale Materialbibliotheken enthalten solche Spektren, die zur Klassifizierung verwendet werden können. Die Spektralbibliotheken kategorisieren ihre Proben anhand einer urbanen Materialtaxonomie. Die spektralen charakteristischen Merkmale können jedoch für verschiedene urbane Materialien sehr ähnlich sein. Eine hohe klassenübergreifende spektrale Ähnlichkeit (hohe klassenübergreifende Korrelation) wird durch ähnliche chemische Zusammensetzungen der Materialien verursacht. Eine Fehlklassifizierung geschieht deswegen häufig, weil es schwierig ist, diese Materialklassen auf der Grundlage der Spektraldaten zu unterscheiden. Weil zudem keine Standardisierung bezüglich der Taxonomie städtischer Materialien existiert, ist die Taxonomie von Spektralbibliotheken in der Regel nicht vergleichbar.

Aufgrund der oben genannten Herausforderungen werden im Rahmen dieser Dissertation verschiedene Aspekte analysiert und Lösungen für urbane Materialien erarbeitet. Es lässt sich beobachten, dass für verschiedene spektrale Bibliotheken die Taxonomie nicht einheitlich definiert ist. Um dies zu verbessern, sollte die Taxonomie von solchen Bibliotheken nicht auf der lokalen Materialvielfalt basieren, sondern auf der globalen Materialvielfalt. Folglich wurde mit KLUM die erste Spektralbibliothek entwickelt, die sich auf Materialien für Gebäudefassaden konzentriert. Eine spektrale Unähnlichkeit innerhalb der Klassen kann auf der Basis der Spektren aus KLUM und auf der Basis von Spektren aus bestehenden spektralen Bibliotheken beobachtet werden. Zuletzt wurde die Beziehung zwischen spektralen und texturalen Merkmalen für die Klassifizierung von urbanen Materialien untersucht. Es kann gefolgert werden, dass spektrale Merkmale für die Unterscheidung von urbanen Materialien für wichtiger erachtet werden, weil die texturalen Merkmale die Unterscheidung nicht verbessern. Die Klassifizierungsergebnisse zeigen auch, dass spektrale und texturale Merkmale bei bestimmten Materialien nicht miteinander korreliert sind.

Sammanfattning

På grund av den senaste utvecklingen av drönare är närbildsanalys inom urbana miljöer ett växande forskningsområde. Detta kan omfatta olika typer av byggnadsbedömningar, såsom värmeanalys för att bestämma värme- och kylbehovet med hjälp av värmebilder, men även klassificering av de yttre byggnadsmaterialen som kan användas som kompletterande information till tre-dimensionella stadsmodeller och simulering av värmeflödet inom stadsmiljön. Kunskapen om vilka material som finns inom den urbana miljön kan samtidigt hjälpa till att lokalisera områden där materialen har en låg reflektans och är därmed sårbara för uppkomsten av urbana värmeöar. Materialklassificering utförs vanligtvis med högdimensionell hyperspektral data eftersom de karakteristiska spektrala kännetecknen kan lättare urskiljas på grund av den kontinuerliga datan över ett brett spektralt intervall. Spektrala stadsmaterialsbibliotek innehåller sådan spektra som kan användas för klassificering. För att urskilja materialklasserna, kategoriserar de spektrala biblioteken med hjälp av en stadsmaterialstaxonomi. De karakteristiska spektrala kännetecknen för ett material kan dock vara vädligt likt de spektrala kännetecknen av att annat material. En stor spektral klasslikhet (en hög klasskorrelation) orsakas av materialens liknande kemiska kompositioner. Felaktig klassificering är därav en vanlig förekomst eftersom det är svårt att särskilja dessa materialklasser. Det finns även ingen generell standardisering av stadsmaterialstaxonomi och den taxonomin som de spektrala biblioteken använder är ofta ojämnförbar.

På grund av de ovannämnda utmaningarna presenterar denna avhandling flera analyserade aspekter och skapade lösningar angående materialen som finns i stadsmiljöer. Det går att observera att olika spektrala bibliotek inkonsekvent definierar taxonomin. För att förbättra detta bör taxonomi från sådana bibliotek inte vara baserad på den lokala materialmångfalden, utan på den globala materialmångfalden. Följaktligen är det första spektrala biblioteket KLUM skapad, vars fokus är fasadmaterial. En spektral skillnad inom en klass kan även observeras baserat på både spektra från KLUM och på spektra från de befintliga spektrala biblioteken. Slutligen har förhållandet mellan spektrala och texturella kännetecknen undersökts för att klassificera material i stadsmiljön. Slutsatsen som kan dras är att de spektrala kännetecknen anses vara viktigare för att urskilja materialklasserna eftersom de texturella kännetecknen inte bidrar med en förbättring. Klassificeringsresultaten indikerar även att de spektrala och texturella kännetecknen inte är korrelerade med varandra för vissa material.

Acknowledgments

I would firstly like to express my sincere gratitude to my supervisor Stefan Hinz for his support, for giving me the freedom to dive deep into the realm of material classification and most importantly, for asking me to pursuit a Ph.D. at IPF.

Secondly, a thank you to Peter Reinartz at DLR for being my co-referee and for the comments.

Furthermore, I would like to thank my everyday supervisor Andreas Schenk at IPF for the support, discussions and comments throughout the PhD.

My sincere thanks to my mentor Stefan Auer at DLR for his support, encouragement and for introducing me to Stefan Hinz back in 2016.

A thank you goes also to Clement Mallet at IGN for fruitful discussions and support.

I would like to thank the Graduate School of Climate and Environment (GRACE) at KIT for financially supporting this thesis and my stay abroad.

A thank you goes as well to Petra Helmholz and David Belton at the Department of Spatial Sciences at Curtin University for hosting me during my stay abroad in Perth, Australia. Furthermore, I would also like to thank Dimitri Bulatov, who stepped up as an additional mentor in Perth.

To all my colleagues at IPF, a big thank you. Without you all, my German would more or less be non-existing. To the best office mate in the world, Martin, a big thank you for always answering my random questions and for your unlimited support and encouragement. A special thank you as well to Uwe for helping me out with hyperspectral questions, to Jens L for his machine learning expertise and finally to Jens K for helping me out with the UAV and calibration of sensors (even if this thesis is covering nothing about that).

Last but not least, I would like to thank my old friends both back home in Sweden and my new friends in Germany. My family, especially my dad and mom, thank you for your support from abroad. Thank you for more or less only talking about things that have nothing to do with this thesis and encouraging me to do other stuff as well. To quote my dad the day after my Ph.D. defense: "22 years of studies are now finished and you should start doing something else". Finally, my German "gubbe", Joachim, thank you for always being there for me.

Contents

Acronyms

3D	Three-dimensional
AE	Autoencoder
ASD	Analytical Spectral Device
BRDF	Bidirectional Reflectance Distribution Function
CCD	Charged Coupled Device
CFS	Correlation-based Feature Selection
CNN	Convolutional Neural Network
CAE	Convolutional Autoencoder
DSM	Digital Surface Model
DTM	Digital Terrain Model
GAN	Generative Adversarial Network
GNSS	Global Navigation Satellite System
GSD	Ground Sample Distance
HGB	Histogram-based Gradient Boosting Classification Tree
KLUM	Karlsruhe Library of Urban Materials
LiDAR	Light Detection and Ranging
LWIR	Long-wave Infrared
MR8	Maximum Response 8 filter bank
MWIR	Mid-wave Infrared
nDSM	Normalized Digital Surface Model
NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared
PCA	Principle Component Analysis
RF	Random Forest
RGB	Red Green Blue
ReLU	Rectified Linear Unit

SLIC	Simple Linear Iterative Clustering
SA	Spectral Angle
SGA	Spectral Gradient Angle
SGA*	Modified Spectral Gradient Angle
SID	Spectral Information Divergence
SVM	Support Vector Machine
SWIR	Short-wave Infrared
SWIR1	FieldSpec-4 SWIR instrument (1001-1800 nm)
SWIR2	FieldSpec-4 SWIR instrument (1801-2500 nm)
TIR	Thermal Infrared
t-SNE	t-Distributed Stochastic Neighbor Embedding
UAV	Unmanned Aerial Vehicle
UHI	Urban Heat Island
VIS	Vegetation-impervious surface-soil
VLWIR	Very Long-wave Infrared
VNIR	Visible and Near-Infrared

1 Introduction

Knowledge about materials in the urban environment is used in several applications. Researchers and city planners who are working with high-level detailed city models or simulations need to know and localize the existing materials. Three-dimensional (3D) city models and the corresponding applications and toolboxes, such as CityGML [114, 121, 122, 175] and OpenStreetMap [182], can use the knowledge about the materials to establish urban inventories. Furthermore, thermal simulations of 3D city models are currently an expanding research field [72, 125, 256] and include the material information for the calculation of the albedo in the urban environment, for e.g. buildings and roads. Further knowledge about the materials in urban areas can likewise be an indicator on how to tackle and handle the Urban Heat Island (UHI) effect [134, 212, 227]. An UHI [178] is defined as an urban area with a significantly higher temperature than its surrounding areas which increasingly occurs in cities across the globe due to the decline of vegetation in cities [45, 241, 247]. Consequently, the overall albedo increases, making UHIs even more prominent [5]. Urban areas heat up during the day due to the incoming solar radiation and with the natural cooling system from the surrounding vegetation being drastically reduced or completely removed, the heat remains during the night. As the global temperature is increasing, urban areas will be even more vulnerable to the UHI effect if not addressed. By assessing urban materials, vulnerable areas can be localized. For the aforementioned examples, the material assessment is done on a larger scale, i.e. on city or district level, and is therefore performed with airborne or satellite acquired data. The material assessment is likewise done on a smaller scale, i.e. on building level, with the usage of close-range acquired data. This can include scene analysis with Unmanned Aerial Vehicle (UAV) to determine building materials [108, 112, 209]. Furthermore, by knowing the exterior building material, i.e. the facade materials, the energy demand of a building (heating and cooling) can be determined by combining the material information with Thermal Infrared (TIR) imagery [74, 83]. Additionally, to preserve and conserve cultural heritage objects where the materials are not well-documented, such as buildings or statues, the materials are determined through image classification [87, 209, 264].

Assessment of urban material can be done by utilizing the spectral information (features) gathered from hyperspectral and multispectral data [76, 233]. Hyperspectral data consists of numerous continuously acquired spectral bands (i.e. a spectral value at every wavelength) and broad spectral range (e.g. 350 - 2500 nm). Material assessment is therefore commonly done with hyperspectral data in selected parts of the spectrum where specific spectral characteristic features are more prominent since it eases the distinction [124, 139]. Multispectral data on the other hand consists of few spectral bands over spaced ranges which limits the possibility to distinct spectral characteristic features. If the access to specific parts of the spectral range is limited due to the available sensors, the urban material distinction can be more difficult to establish.

To classify materials with remote sensing data, spectral libraries are commonly used. A spectral library contains a collection of spectra and their spectral characteristic features. A library can contain e.g. different kinds of vegetation [50, 197, 272], soil properties [219] or corals [132]. Spectral libraries that contain

urban materials will in this thesis be focus. Such spectral libraries can be based on spectra acquired either *in-situ* [69, 172], in the laboratory [14, 49] or by combining airborne, *in-situ* and laboratory acquisitions [19, 98, 102, 124, 222]. How a spectral library describes and labels a material sample depends on the intended application. Furthermore, the majority of spectral libraries provide no imagery of the material samples. Therefore, it can be challenging to assure that the different spectral libraries contain the same material if it has been labeled in different manners [69].

Only relying on the spectral features to assess different materials can be a challenging task since the spectral characteristic features are very similar for certain material classes [61, 97, 136, 183]. This is known as high inter-class correlation, i.e. a high inter-class spectral similarity, due to similar chemical compositions. Therefore, misclassification frequently occurs. Figure 1.1 highlights the intra-class spectral similarity challenge by displaying the spectral features of three urban materials: *Asphalt*, *Conglomerate* and *Sandstone*. The spectral characteristic features of *Asphalt* and *Conglomerate* are similar, due to the high inter-class spectral similarity. The scalar vector differences of the two material classes are alike (i.e. the gradient) but an intensity difference of around 10% is likewise present. Depending on the chosen approach, it could be challenging to distinguish those two classes with only spectral data.

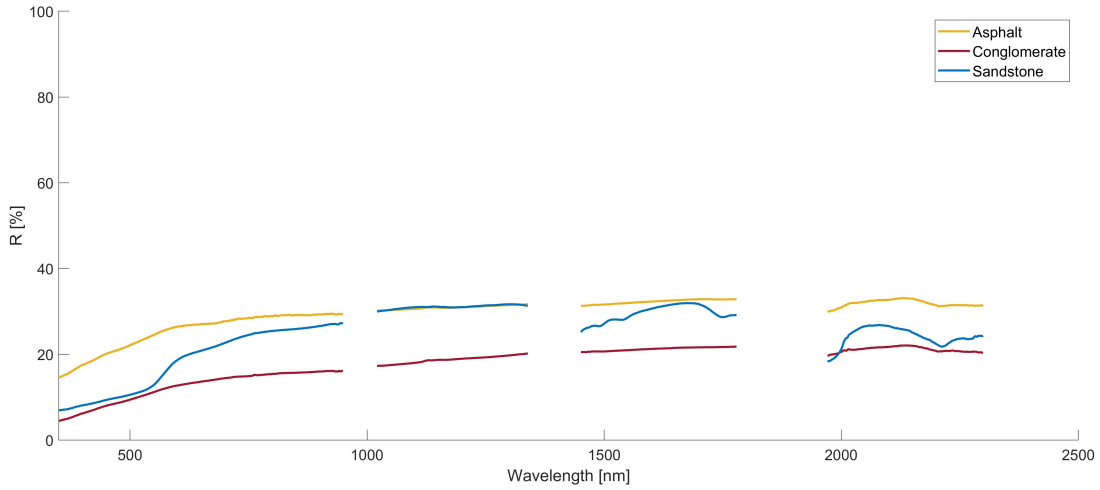


Figure 1.1: The spectral reflectance *R* of three samples representing the material classes *Asphalt* (yellow), *Conglomerate* (red) and *Sandstone* (blue). The gradients of the spectral characteristic features of *Asphalt* and *Conglomerate* are similar which could make the class distinction with only spectral features challenging. The spectral gaps are due to post-processing of the samples which is further explained in Chapter 5.

In short, material assessment is a complex topic. Different materials can possess similar spectral features due to the chemical composition and utilizing only spectral data is therefore not always sufficient.

1.1 Motivation

This thesis intends to tackle and present solutions to some of the challenges that material assessment faces. This is on the one hand motivated by the occurring inter-class spectral similarity which causes misclassification of urban materials. On the other hand, it is motivated by spectral libraries and the corresponding varied taxonomy (i.e. classification schemes and labeling systems) and limited additional data (e.g. imagery).

To improve material classification, the established inter-class spectral similarity has to be tackled first. As scene analysis with UAVs can acquire close-ranged data, low-level textural information can be added to the assessment in addition to the spectral features which could improve the distinction of urban materials. Combined spaceborne acquired spectral and textural data for assessment of e.g. tree species [73], agriculture [63, 168] and urban areas [263] have shown great potential. Textural features that describe the surface texture can be determined with standard and commonly used texture analysis methods relying on entropy [92], edge detection [36] and statistical filter banks [237]. They can likewise be determined with machine learning approaches such as Autoencoder (AE) [59, 126, 257]. Close-range acquired textural features contain low-level information and could contribute to improved material assessment if such features contain distinct information. Hence, this thesis also includes textural features to assess the material distinction.

As stated, spectral libraries tend not to provide imagery of the material samples. The majority of spectral libraries contain spectra of land cover and land use classes, as the common usage is to assess materials with airborne or spaceborne acquired data. Therefore, these libraries provide no imagery since the derived texture depends on the used sensor and its pixel-resolution. As close-range scene analysis with e.g. UAVs or Google Street View [176, 261] is an expanding research field, low-level textural features could be used as additional information for material assessment. The libraries that acquired the material samples close-ranged provide occasionally imagery. However, the imagery is rarely available for download as it is attached in technical reports (i.e. no separate image files). Therefore, it is not possible to further use the imagery for additional research. Furthermore, facade material assessment with oblique UAV-based sensors is not a common practice. Consequently, facade materials are not well represented in spectral libraries. There is therefore potential for the development of spectral libraries with new aspects that can improve and enhance close-range material assessment.

1.2 Contributions

Although research has been done within the realm of material assessment using remote sensing techniques, certain aspects should be examined more comprehensively. This thesis assesses, analyzes and develops approaches and datasets for material classification by focusing on two kinds of features: spectral and textural features.

The main contributions of this thesis are as follows:

- **Analysis of the existing urban material spectral libraries and assembly of a spectral library**
  Urban material spectral libraries contain spectra and the corresponding spectral characteristic features. The spectral characteristic features of a material depend on the influence of the surrounding environment and the chemical composition. As libraries contain samples that have acquired in different environments and regions, the spectral characteristic features that correspond to the same material are further analyzed to determine any dissimilarities. Furthermore, the material taxonomy that libraries use depends on the application, expertise and research field. The creators of a spectral library create their own taxonomy as no standardized definition exists. Therefore, the used material taxonomy in spectral libraries differs and it can be challenging to utilize different spectral libraries due to the various definitions. This thesis analyzes the used taxonomy by examining the existing spectral libraries and their various taxonomy to determine potential improvements. Consequently, the first spectral library with a focus on facade materials is created. Karlsruhe Library of Urban Materials (KLUM) consists mainly of facade materials to accommodate the future need for oblique UAV scene analysis which can include assessment of building facade materials. Furthermore, KLUM provides two kinds of imagery that can be further used for texture analysis. The utilized taxonomy describes the samples on a low level by providing information concerning the color, the surface structure, coating or texture and status.
- **Assessment of the existing approaches and tools used for urban material classification**
  There are numerous approaches and tools available for material classification. While handling spectral features and in particular those obtained from high-dimensional hyperspectral data, there are various approaches to either reduce the dimensionality or to select the most relevant features. In addition to feature extraction approaches, several classifiers are commonly used for high-dimensional data. These approaches have been thoroughly developed and evaluated for land cover and land-use classification. Assessment of the tools and approaches of classification of pure urban materials is sparse. Consequently, this thesis assesses those by utilizing two datasets, an existing spectral library and an *in-situ* acquired dataset. Furthermore, certain parts of the spectrum are more relevant for material classification than other [124, 139]. Access to specific parts of the spectral range can be limited due to the available sensors. It is therefore important to determine which parts of the spectrum are more suitable for material classification through analysis and evaluation of the information gain and loss in different spectral ranges.
- **Analysis of intra-class and inter-class spectral similarity**
  The inter-class spectral similarity of urban materials has in previous studies been established while intra-class spectral similarity has not. The intra-class spectral similarity is defined as the variation of the spectral features of one material. However, if the spectral features of one material represented by e.g. two samples do vary (they are dissimilar), misclassification can occur. If both inter-class spectral similarity and intra-class spectral dissimilarity are simultaneously present in a dataset, it can be challenging to distinguish the material classes as the spectral features are not

consistently distinct. Therefore, through extensive analysis and evaluation, both challenges are examined.

- **Assessment and evaluation of spectral and textural features importance for classification of urban materials**
  Due to the known inter-class spectral similarity, additional textural information can be added to enhance the urban material distinction. This is implemented through two case studies, one based on airborne acquired data and another based on close-range acquired data. The case studies express the textural features differently, one retrieving the features from standardized texture analysis approaches and the other from the machine learning approach AE [126]. Furthermore, based on the results from the two aforementioned case studies, a detailed evaluation of the feature importance is performed. This involves analysis of the limitations and challenges to summarize the potential of the two kinds of features. The two kinds of features are likewise individually analyzed by comparing different manners of expressing the features and how they consequently influence the final classification results.

1.3 Outline

This thesis is structured as follows. Chapter 2 presents the fundamentals, principles and definitions of urban materials which includes the physical definitions and characteristics of urban materials from the point of view of remote sensing. Chapter 3 presents an overview of the relevant and existing datasets that can be used for material assessment. This is supplemented by an analysis of the presented spectral libraries and their corresponding taxonomy. Consequently, the term taxonomy is discussed and the existing material taxonomy is examined. This is followed by Chapter 4 where the relevant classifiers, dimensionality reduction approaches, feature selection tools and texture analysis methods are presented. A case study with high-dimensional spectral data is likewise presented by assessing a few of the specified approaches and tools. The first published spectral library with the main focus on building facade materials (KLUM) is introduced in Chapter 5. The complete workflow is described, including the data acquisition, post-processing and both internal and external spectral screening with the acquired samples. Two case studies utilizing both spectral and textural features to classify materials are presented in Chapter 6. The case studies use either airborne acquired data to assess roof materials or close-range data to analyze a selection of samples from KLUM. Lastly, this thesis concludes with the summary, concluding remarks and future work in Chapter 7.

2 Fundamentals, Principles and Definitions

This chapter provides an overview from a remote sensing perspective of the principles, definitions and datasets concerning materials. This thesis focuses on urban materials, i.e. materials that are present in the urban environment. Firstly, an overview is given about the electromagnetic radiation and spectrum which includes the mathematical, physical and chemical definitions. Furthermore, the derivation of the spectral characteristic features is explained as these are dependent on materials due to the chemical compositions. Lastly, the concept of hyperspectral data is presented as spectral characteristics can be distinguished with such data due to it being continuous and covers a wide spectral range.

2.1 Electromagnetic Radiation and Spectrum

Electromagnetic radiation refers to waves that carry the energy and are represented as oscillations of electric and magnetic fields [223]. An electromagnetic wave can be defined and described by any of the following three physical properties: frequency f, wavelength λ, or photon energy Q [210]. The relationship between the three properties can be expressed through the following equation where c is the speed of light in vacuum (299792458 m/s) and h is the Planck's constant ($6.62606957 \cdot 10^{-34} Js$):

$$Q = \frac{hc}{\lambda} = fh \tag{2.1}$$

The energy of a photon Q is an electromagnetic particle with energy inversely proportional to its wavelength (as seen in Equation 2.1) and transfers the electromagnetic radiation [24]. Wave-particle duality describes the behavior of a photon, i.e. a photon can be described as either a particle or a wave. Photons can either be absorbed if the targeted object completely absorbs the photon, reflected if the photon bounces back from its targeted objects or transmitted if the photon passes through the targeted object [210].

Emissivity ϵ is a measure that expresses an object's ability of emitting thermal radiation. Emissivity ϵ is determined by calculating the ratio of the thermal radiation Φ between an object (Φ_o) and a perfect black body (Φ_b) according to the Stefan-Boltzmann law of radiation [3]. A black body absorbs all incident electromagnetic radiation throughout the entire electromagnetic spectrum, regardless of the incoming incidence angle. The ratio is determined within the range of $[0, 1]$ where 1 represents a black body. Emissivity ϵ is determined using the following equation:

$$\epsilon = \frac{\Phi_o}{\Phi_b} \tag{2.2}$$

The reflectance R of an object's surface is the measure of an object's ability to reflect electromagnetic radiation. It is expressed as the percentage within the range of $[0, 100]$ where 0% represents a black body

and 100% total reflectance. The reflectance R can be determined by either knowing the emissivity ϵ of the object or by measuring the reflecting thermal radiation Φ incident upon the object [3, 24]. It can be determined by calculating the difference between total reflection (100%) and the calculated emissivity ϵ of a targeted object using:

$$R = (1 - \epsilon) \cdot 100 \tag{2.3}$$

The reflectance R can also be determined if the thermal radiation Φ reflected from the object (Φ_r) and incident upon the object (Φ_i) have been measured, using:

$$R = \frac{\Phi_r}{\Phi_i} \cdot 100 \tag{2.4}$$

As the reflectance R depends on the targeted object due to its ability of emitting radiation, objects have distinct spectral characteristic features that depend on their material and the chemical compositions [105]. The reflectance R can be determined in selected parts of the electromagnetic spectrum, also known as spectral ranges. The electromagnetic spectrum describes the wavelength range of radiated electromagnetic waves which range from gamma rays to radio waves [3, 228].

Table 2.1: The defined spectral ranges in the electromagnetic spectrum that are used in this thesis. *Definition exclusively used by the remote sensing community.

Name	Wavelength
Visible	350 - 700 nm
Visible and Near-Infrared (VNIR)*	350 - 1000 nm
Near Infrared (NIR)	750 - 1400 nm
Short-wave Infrared (SWIR)*	1000 - 2500 nm
Mid-wave Infrared (MWIR)	3000 - 8000 nm
Long-wave Infrared (LWIR)	8000 - 12000 nm
Very Long-wave Infrared (VLWIR)	12000 - 30000 nm

The electromagnetic spectrum is divided into several spectral ranges and the defined ranges that are used in this thesis (the visible and the infrared spectra) are presented in Table 2.1. The remote sensing community uses another definition for the spectral ranges VNIR and SWIR than other research fields, as the range of NIR in the range of VNIR conflicts with the defined range of NIR [46]. These are thus denoted with a * in Table 2.1.

The visible spectrum is the part of the electromagnetic spectrum which is visible for the human eye. It can be divided into the pure spectral colors violet, blue, cyan, green, yellow, orange and red. Although they are defined as pure colors, the defined spectral ranges of those are continuous and overlapping. The colors blue, green and red are the three primary colors [3]. One primary color cannot be created by combining the other two primary colors, but can, when combined, create all other colors. The spectral ranges of the primary colors are used for Red Green Blue (RGB) cameras.

As stated, objects have different reflectivity that depends on the properties of the material and the chemical components. The characteristics of materials will in the following sub-chapter be further explained.

2.2 Characteristics of Materials in the Electromagnetic Spectrum

The spectral characteristic features of an object, referred to as the spectral signature of an object, are represented as the variation of the spectral reflectance R within the range of $[0, 100\%]$ in relation to the wavelength. The spectral characteristic features of an object depend on its physical (environmental) conditions, chemical compositions and spectral behavior at different wavelengths [19, 105, 193]. These features can be characterized by the absorption bands (local spectral minima), the reflectance peaks (local spectral maxima), an increase or a decrease of reflectance in specific spectral ranges, the brightness (the mean of the spectra) of a material in a specific part of the spectrum and the continuity (the shape) of the spectral reflectance curve [98]. The absorption bands and the reflectance peaks can be further explained based on their depth, width, position and symmetry (see Figure 2.1 for an illustration) [46, 48, 235]. These features can be exploited for further analysis since they occur in specific parts of the spectrum which depends on the targeted objects [98, 229, 235, 240].

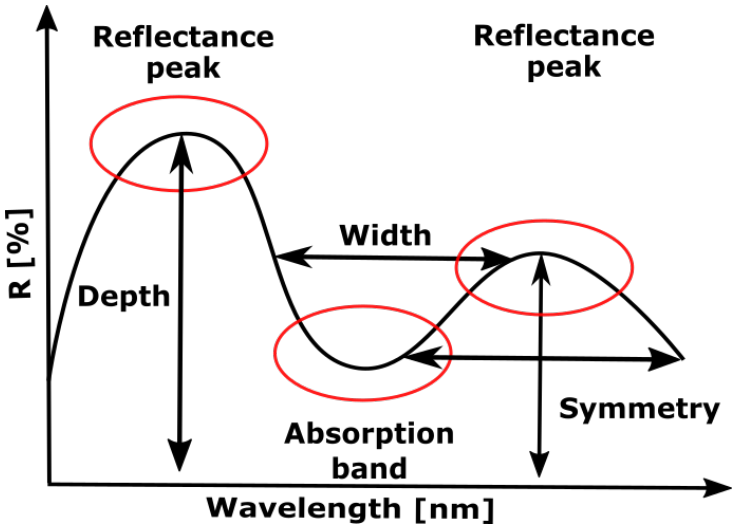


Figure 2.1: An artificial spectrum that displays an absorption band and two reflectance peaks in addition to how the depth, width and symmetry of these bands and peaks can be measured. The shown reflectance peaks are symmetrical as the measured depths are located in the middle of the measured width. If the peaks are located more to the right or left, the reflectance peaks would have been asymmetrical right or left respectively.

2.2.1 Physical Conditions

The physical (environmental) conditions affect the albedo (which expresses the diffuse reflection of a surface) and influence the spectra in the visible spectral range. This includes shadowing effects (change of illumination or solar radiation), particle-size distribution (the relative amount of particles according to size), the geometry of the object (which influences the surface reflection angle) and the refraction index (how fast light travels through the material in vacuum) [19]. Some of the physical conditions can be

regulated in a controlled environment, e.g. in a laboratory, where the incoming illumination angle and intensity can be manually adjusted. However, the chemical properties affect the spectral characteristic features more, such as the absorption bands and the reflectance peaks, due to the chemical response at specific wavelengths [19, 105].

Lambertian reflectance is a property that defines how a surface ideally reflects incoming radiance (diffuse reflection/scattering). A surface is Lambertian if it follows Lambert's cosine law and the radiance is the same from any viewing angle [186]. Urban materials usually reveal a non-Lambertian reflectance behavior [141, 165] that depends on the incoming solar radiance angle for both vertical and horizontal surfaces [102]. Depending on the material and its surface structure, some materials tend to display forward-scattering behaviors (e.g. concrete), some back-scattering behaviors while others are more Lambertian (e.g. asphalt) [102, 135].

The Bidirectional Reflectance Distribution Function (BRDF) [174] is a function that defines how the radiance is reflected against a surface and is calculated based on the direction of the incoming and outgoing radiance. Radiance accounts for the incoming or reflected direction of the energy of the electromagnetic radiation. Radiance is the radiant power (radiant energy per unit time) which is either emitted, reflected, transmitted or received by a surface per unit surface per unit steradian (solid angle). This is an important aspect for observation of an object at a specific position but with varied incoming illumination angles from a light source (e.g. the sun). Irradiance is the radiant power incident on a surface per unit area. The BRDF $f_r(\omega_i, \omega_r)$ can be determined by using the following equation, where L is the radiance, E the irradiance, i the incoming radiance, r the reflected radiance, ω_i the incoming radiance direction, ω_r the outgoing radiance direction and θ_i the angle between ω_i and the surface normal:

$$f_r(\omega_i, \omega_r) = \frac{dL_r(\omega_r)}{dE_i(\omega_r)} = \frac{dL_r(\omega_r)}{L_i(\omega_i cos\theta_i d\omega_i)} \tag{2.5}$$

The BRDF can likewise influence the measured spectra of an urban material and should, if possible, be accounted for [135, 202, 216]. Figure 2.2 illustrates calculated BRDF for the land cover classes pine shrubs and pure sand at 10°, 30° and 50°.

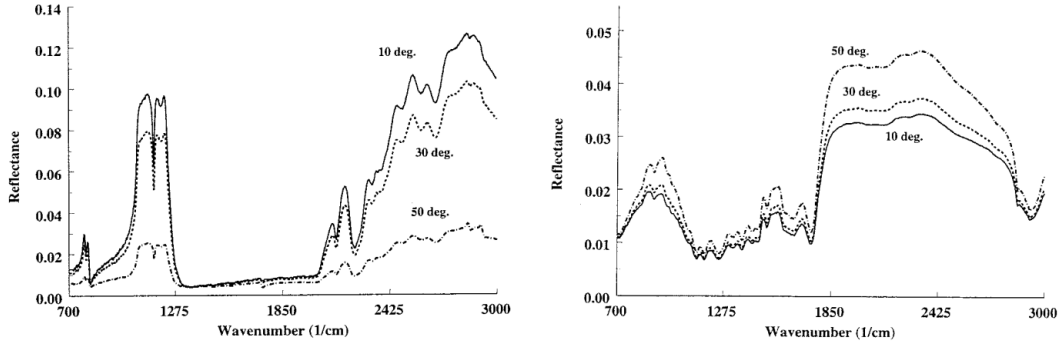


Figure 2.2: An example of calculated BRDF at 10°, 30° and 50° for pine shrubs (left subfigure) and pure sand (right subfigure) [220]. A spectral difference can be observed as the reflectance intensity varies across the spectral range. Nonetheless, the locations of the spectral characteristic feature remain the same (e.g. the absorption bands and the reflectance peaks). ©[1998]IEEE. Reprinted with permission from [220].

2.2.2 Physicochemical Processes

Two kinds of physicochemical processes affect the spectral absorption (the spectral characteristic features) of materials in the visible and infrared range due to its chemical properties: electronic transitions [34, 35] and vibrational transitions [70].

In solid-state chemistry, electrons can be shared between individual atoms or ions. Absorption of a specific wavelength occurs when an electron shifts from one energy state to a higher one. The transition metal ions, such as iron (Fe), cobalt (Co), nickel (Ni), copper (Cu) and zinc (Zn), have partially filled electron levels and can cause the electronic transitions in the range of VNIR [34, 35, 86, 160]. The nature of electronic transitions can be either crystal field effects, charge transfers, color centers or conduction bands [34, 35].

Spectral characteristic features in the range of SWIR are dominantly influenced by vibrational transitions [70]. These vibrational transitions can be observed in the range of SWIR as small measurable absorptions bands. This can be observed for molecules such as OH^- (hydroxide ion), H_2O (water) and CO_3^{2-} (carbonate ion) [47, 105]. OH can be bound to magnesium (Mg) and aluminum (Al), and its vibrational transition can be observed around 1400 nm and to around either 2200 nm (if aluminum is present) or 2300 nm (if magnesium is present). An example of such a vibrational transition can be seen in Figure 2.3 for four *Muscovite* samples. The vibrational transition attributed to the overtone combination of OH-bearing metal and OH can be observed around 2200 - 2300 nm. The vibrational transition can be observed for H_2O-bearing minerals around 942, 1135, 1450 and 1875 nm [4]. The vibrational transition caused by CO_3^{2-} can be observed within the range of 1600 - 2500 nm [46, 47, 105, 235].

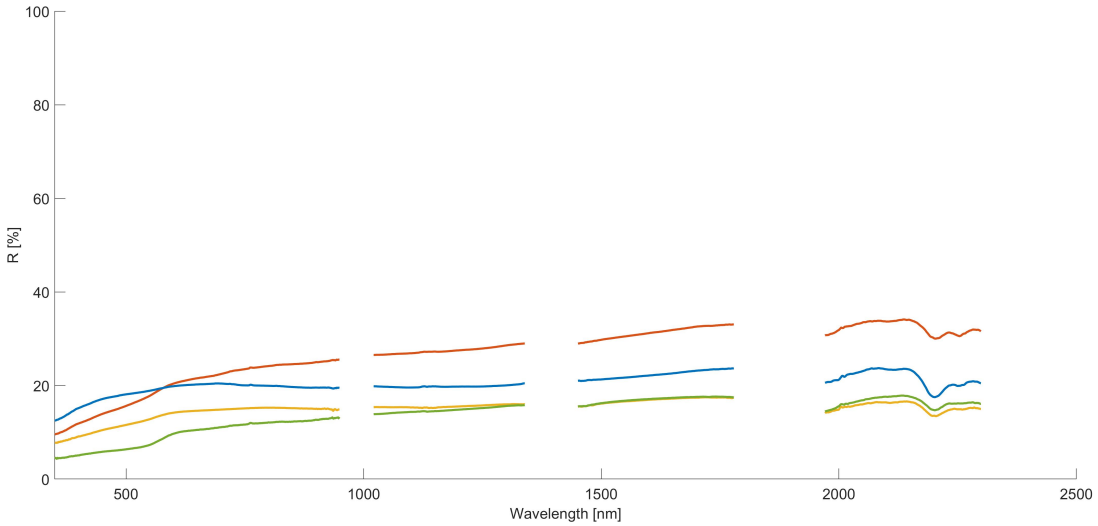


Figure 2.3: An example of observable vibrational transitions for three *Muscovite* samples. The vibrational transitions can be observed as smaller absorption bands around 2200 nm. The spectral gaps are due to post-processing which is further explained in Chapter 5.

An overview of the spectral ranges where the electronic and vibrational features can be observed for selected minerals can be seen in Figure 2.4 [105]. The smaller figure presents vibrational features and the larger the electronic features.

Materials in the urban environment are however complex and can consist of a mixture of material components and few materials in the urban areas are pure (i.e. consisting of only one chemical component).

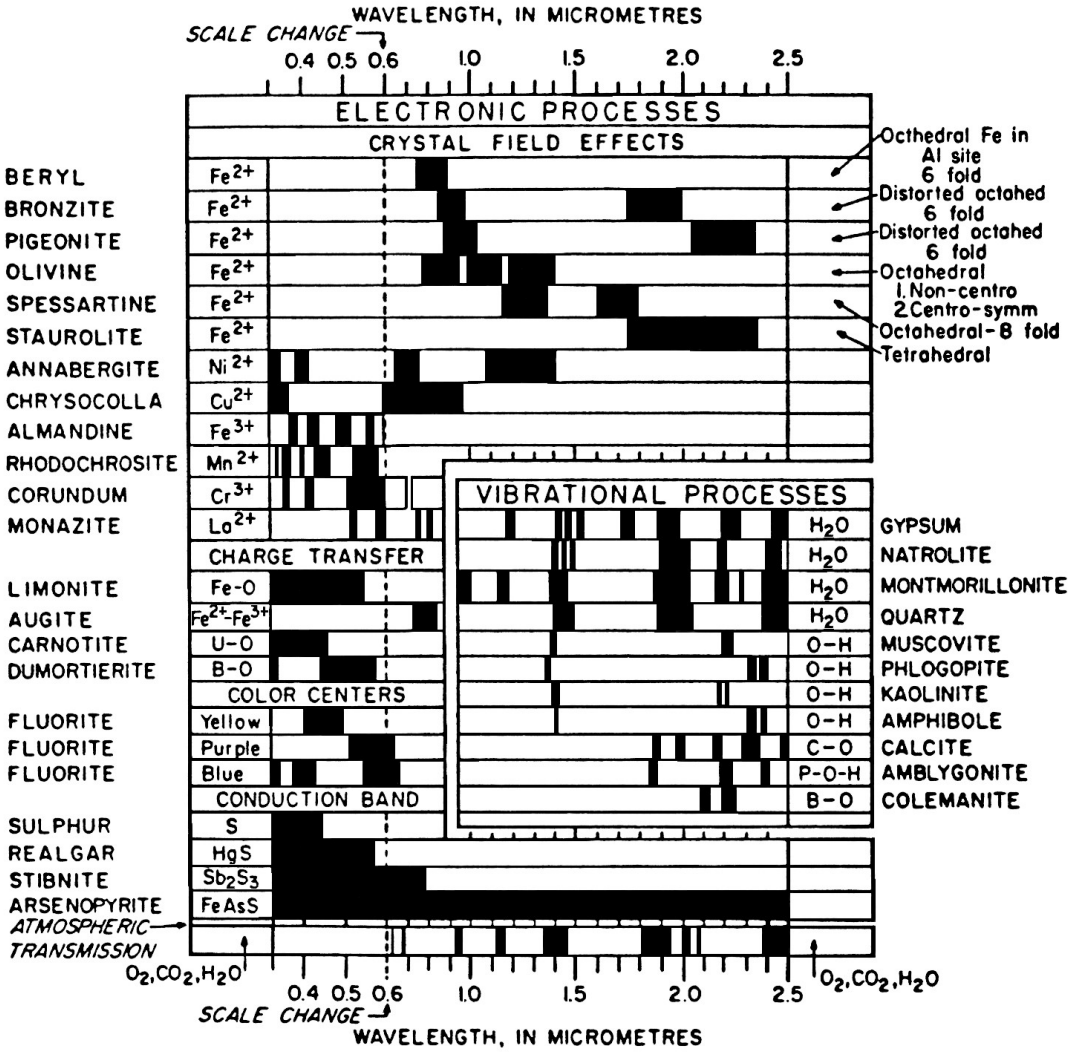


Figure 2.4: Overview of the spectral ranges where the electronic and vibrational transitions can be observed for a selection of minerals. The spectral ranges which are affected by the transitions are marked in black [105]. Reprinted with permission granted through the STM permissions guidelines[1].

Contrary, they contain several chemical components with various ratios. For instance, *Asphalt* is a mixture of a variety of hydrocarbons with the elemental composition of carbon (C, 82 - 88%) and hydrogen (H, 8 - 11%) in addition to one or more of nitrogen (N), sulfur (S) and oxygen (O) (all of which are important for the polarity of the binder) [187, 225]. The base element composition is similar as it consists of carbon (C) and hydrogen (H) but the ratio varies. Similar to *Asphalt*, *Concrete* consists of a mixture of chemical components. This includes silicon dioxide (SiO_2), aluminum oxide (Al_2O_3), calcium oxide (CaO) and ferric oxide (Fe_2O_3) and the ratio between these four in addition to smaller binders does likewise vary [15, 163, 171, 195]. This mixture of chemical components can be seen across various kinds of building materials [39]. Not only man-made materials consist of different mixtures of chemical components, natural materials do as well. The chemical composition of natural materials does vary naturally on a regional level [78, 113]. The ratio between potassium oxide (K_2O) and sodium oxide (Na_2O) is an important factor for classifying *Sandstone* into subclasses [167, 215]. The elemental composition of *Wood* varies depending on the species. *Wood* consists mainly of cellulose, hemicelluloses, lignin and extractives in varying proportions [188, 189, 213].

2.3 Hyperspectral Sensing

Hyperspectral imaging, also known as imaging spectroscopy, provides imagery data that has for each pixel a vector representation corresponding to hundreds of contiguous spectral bands [82]. This can be illustrated as a hyperspectral data cube (hypercube), a stack of images where each layer represents the measured radiation at the corresponding band (wavelength) [21]. Hyperspectral imaging can be acquired with spaceborne, airborne or ground-based sensors. Hyperspectral data can likewise be acquired with spectroscopy where only a spectral vector is retrieved [48]. Such data is often acquired close-ranged with hand-held sensors. For an arbitrary scene, hyperspectral imaging can express the scene as an image where each pixel contains spectral data. Through spectroscopy, the same scene would be expressed as one single spectrum which depends on where the spectral data was acquired (e.g. the spectrum for one chosen pixel). See Figure 2.5 for an overview of the different approaches for hyperspectral acquisition. Hyperspectral data is nowadays used for food safety and quality control [149, 150, 196, 255] and precision farming [153, 154, 170]. Applications may also lie beyond these two fields [6, 23, 90].

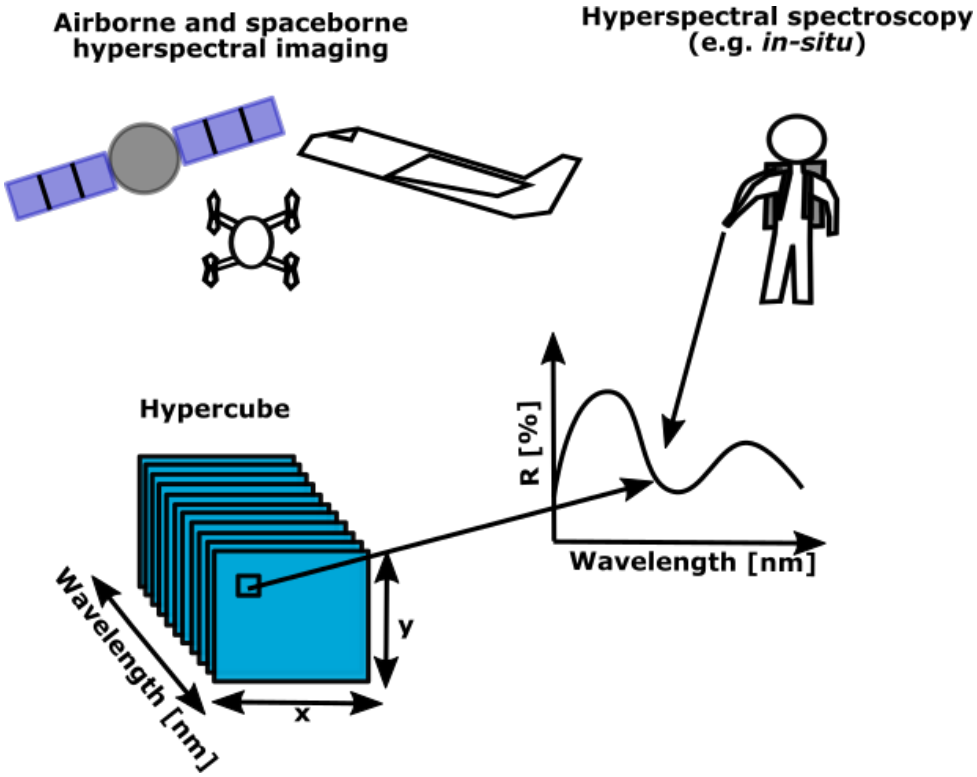


Figure 2.5: Overview of hyperspectral imaging and spectroscopy. Hyperspectral imaging can be acquired with e.g. spaceborne and airborne sensors where each layer of a hypercube represents the measured radiation at the corresponding band. Through spectroscopy, a spectral vector that represents the measured radiation at the corresponding band is acquired.

While working with hyperspectral data, dimensionality refers to the number of spectral bands, meaning, that hyperspectral data is high-dimensional compared to RGB (three bands) and multispectral data (e.g. five bands) as they consist of much fewer bands. As hyperspectral data is high-dimensional with hundreds of spectral bands, the phenomenon referred to as "curse of dimensionality" is present [18]. This refers to the challenges of analyzing and organizing high-dimensional data that do not occur with

[1] https://www.stm-assoc.org/2020_05_07_STM_Permissions_Guidelines_2014.pdf

low-dimensional data. The organization of data relies on finding similarities. However, the organization of high-dimensional data can be challenging as the data can appear both similar and dissimilar depending on the viewing point. Expressed from a hyperspectral point of view, certain spectral characteristic features can be similar in one part of the spectrum and dissimilar in another.

As hyperspectral data is high-dimensional, adjacent bands tend to be strongly correlated and thus, there is a high degree of redundancy. Consequently, the classification of hyperspectral data deals with relevant, irrelevant and redundant features. The Hughes phenomenon [104] (a known "curse of dimensionality" problem [18]) therefore commonly occurs, which happens when classification for a small fixed number of training samples firstly increases the accuracy but as the dimension increases, the classification accuracy decreases [166]. Therefore, common approaches to address this are either to reduce the dimensionality or to select a subset of important features. This is further discussed in Chapter 4.2.

This thesis is not dealing with hyperspectral imagery where additional information can be exploited by extracting the spatial-spectral information for classification tasks [71, 147, 164], where the spatial relation between neighboring pixels is likewise considered.

2.4 Summary

This chapter presented the essential foundation of materials from a remote sensing point of view. This included firstly a technical explanation of electromagnetic radiation and spectrum. Secondly, an overview was given of how the spectral characteristic features depend on the chemical compounds and their interference with the electromagnetic spectrum. Lastly, the definition of hyperspectral sensing was likewise explained.

3 The Taxonomy of Urban Materials

Relevant urban material datasets from the remote sensing and computer vision communities, which either contain spectra (spectral libraries) or imagery (texture databases), are presented in this chapter. Datasets tend to utilize different taxonomies as no standardization has been enforced and implemented. In this thesis, the taxonomy of a dataset consists of two parts, a classification scheme and a labeling system. The classification scheme is the hierarchical structure that a dataset uses to split its samples into different ranks (e.g. the number of the ranks). The labeling system is how the dataset assigns labels to its samples and which labels it uses (e.g. low-level or high-level description). The labels of various datasets are not always easily comparable. This chapter compares and analyzes the labels the existing datasets use. This chapter is further dedicated to the different employed material taxonomies used by the existing spectral libraries as it is both a complex and comprehensive topic. This includes an overview of the principles of taxonomy and the various classification schemes and labeling systems.

For example, a material sample can be described on a high level (less detailed) in one dataset (e.g. just the material label, such as *Asphalt*) and in another dataset on a low level (very detailed), where the surface conditions and the color can be described (e.g. *Grey old asphalt*). Hence, it can be challenging to assure that samples from different datasets are comparable matches purely based on the assigned label due to the varied taxonomies. Spectra corresponding to the same material but defined through different labeling systems are therefore compared to assess if the samples should have been labeled as the same material. Furthermore, the classification schemes and labeling systems are further analyzed to determine the similarities.

3.1 Urban Material Datasets

The currently existing and publicly available datasets that contain material information, either as spectra or imagery (texture), will in this sub-chapter be presented. The described datasets may not exclusively contain urban materials since the datasets have different purposes. Datasets that consist of urban material spectra, namely spectral libraries, are firstly presented. Secondly, spectra for selected materials will be extracted from several libraries for spectral comparison and analysis. By comparing the same material's spectral characteristics, it is possible to determine any spectral similarities and dissimilarities. Thirdly, datasets that contain imagery of urban materials, so-called texture databases, are presented. A comparison between spectral libraries and texture databases is lastly presented.

3.1.1 Spectral Libraries

A spectral library is a database that consists of a collection of acquired and labeled material samples in addition to the corresponding spectral characteristic features. The spectral reflectance R covers in the majority of the presented spectral libraries only the spectral ranges VNIR and SWIR, but sometimes additional ranges. One of the first recorded spectral libraries was acquired by the American federal space agency NASA in 1985 that recorded 156 samples and the corresponding spectral characteristic features [26]. The samples were acquired either in the laboratory, in the field or from an aircraft covering the VNIR and SWIR spectral ranges. The spectral library consists of six land cover categories: agriculture, trees, shrubs and grasses, rocks and soils, water, snow and clouds, and miscellaneous. By assembling a collection of spectra, this library provided further information to the design of satellite sensors by determining important spectral ranges where the land cover types show distinct spectral characteristic features.

Due to the advances of hyperspectral imaging, such as the airborne sensors AVIRIS [151] and HYDICE [199], reference data was needed to evaluate the acquired spectra. To the best of the author's knowledge, the first publicly known spectral library that contains urban materials is the Price's 1995 spectra collection, a collection that is based on spectra from several other datasets [194]. The collection contains spectra of soil, agriculture, vegetation, water, rocks, natural and man-made materials and minerals. Since the Price's 1995 spectra collection along with the further advances of hyperspectral imaging, several urban material spectral libraries have emerged. These can contain land cover and land-use classes or only urban materials. The spectra can be based on spaceborne, airborne or close-ranged (e.g. *in-situ* or in the laboratory) acquired data.

Selected urban material spectral libraries will in the following text be presented. This includes an overview of the data acquisition, the post-processing, the material classes and the used classification schemes and labeling system. The classification schemes are here simplified since the libraries are using different hierarchy structures which occasionally generate many sub-classes. The simplified classification schemes contain only the labeling of the two material classes *Asphalt* and *Concrete* in addition to one further selected urban material class. Furthermore, a summary and an overview of the presented spectral libraries can be seen in Table 3.12. If available, the website of a spectral library is provided.

PUSL and CASILIB Spectral Libraries of Tel-Aviv, Israel

Two of the first spectral libraries that focus on urban materials were created in Tel-Aviv, Israel, namely PUSL and CASILIB [19]. This was done by acquiring spectra in the spectral range of 400 - 1100 nm (VNIR) by combining spectral information acquired during one day simultaneously from an airborne system with the push-broom sensor CASI [10] and *in-situ* using an Analytical Spectral Device (ASD) field spectroradiometer. The field spectra were resampled to be comparable with the airborne sensor. Post-processing included geometric distortion correction due to the aircraft's movement and atmospheric correction as the data was acquired outdoors.

Two spectral libraries were created; PUSL with a focus on urban materials and CASILIB with representative objects in the city of Tel-Aviv. PUSL was based on selected samples from Price's 1995 spectra collection [194] which were chosen to represent urban materials in Tel-Aviv. CASILIB on the other hand

Table 3.1: Simplified classification scheme that the spectral library CASILIB uses [19]. Three ranks are used to describe the samples and the material compositions are defined on rank 3: *Material*.

Rank 1: Groups	**Rank 2: Sub-group**	**Rank 3: Material**
Natural bodies (liquid and solid)	Liquid	-
	Solid	-
Urban surfaces	Urban surfaces 1	Schoolyard concrete
	Urban surfaces 2	Asphalt (dark)
		Asphalt (light)
		School roof (asphalt)
		Marble pavement
	Urban surfaces 3	-
Vegetation	Vegetation	-
Pigment spectra	Pure color pigment	-

was generated with the acquired data to represent existing objects in selected areas in Tel-Aviv by combining the a priori knowledge from the *in-situ* acquired spectra with the airborne CASI data. The sample labeling and descriptions are defined inconsistently. The description includes occasionally the color and usage of the sample (e.g. roof or schoolyard), but not consistently across all samples. Nonetheless, the CASILIB library utilizes a land cover based classification scheme that consists of four major groups. The simplification can be seen in Table 3.1 where the labeling of the material classes *Asphalt*, *Concrete* and *Marble* are presented.

HYSL and FISL Spectral Libraries of Dresden and Potsdam, Germany

The spectral libraries HYSL and FISL consist of spectra acquired in selected areas in the two German cities Dresden and Potsdam [96, 98]. The spectra were acquired with the airborne hyperspectral imaging sensor HyMap [51] and the ASD FieldSpec Pro FR spectroradiometer in the spectral range of 400 - 2500 nm (VNIR and SWIR). The imagery of Dresden was acquired during three different years (1999, 2000 and 2003) with acquisition dates between May and August to account for annual and seasonal differences. The post-processing consisted of radiometric and spectral calibration of the imagery, transformation of the radiance data into reflectance, geometrical correction of the imagery and atmospheric correction and resampling of the field spectra to be comparable with the imagery.

HYSL consists of the image spectra from 21,670 pixels (samples) while FISL consists of the field spectra. FISL consists of fewer roof materials, as field measurements of roofs were not always possible. Both libraries provide additional information about the color and relevant chemical compounds for each material class. The libraries utilize a classification scheme with four ranks to categorize urban surface cover types into a land cover based hierarchical system [204] (seen in Table 3.2). The labeling of the material classes *Asphalt*, *Concrete* and *Roofing tile* is shown. These two libraries are the only ones that provide additional information about the chemical compounds. HYSL and FISL utilize the classification

Table 3.2: Simplified classification scheme of the spectral libraries HYSL and FISL [96, 98]. Three ranks are used to describe the samples and the material compositions are defined on rank 4: *Material*.

Rank 1: I	**Rank 2: II**	**Rank 3: III**	**Rank 4: IV**	**Rank 5: Material**	**Rank 6: Color**
Impervious	Overbuilt	Buildings/roofs	Mineral	Concrete	Grey
					Red
				Roofing tile	Red (new)
				Roofing tile	Red (old)
			Hydrocarbons	Bitumen	Green/grey
					Red
					Grey
	Open areas	Fully impervious	Mineral	Concrete	Bright/grey
				Pavement of concrete	Bright/grey
					Red
			Hydrocarbons	Asphalt	Dark/grey
		Partially impervious	-	-	-
Pervious	Vegetated	Green spaces	-	-	-
		Trees/bushes	-	-	-
	Non-vegetated	Bare soil	-	-	-
		Water	-	-	-

scheme consistently, with one exception where one class is split based on the sample age (*Red roofing tiles (old)* and *Red roofing tiles (new)*).

Santa Barbara Spectral Library, US

The Santa Barbara spectral library[1] contains urban material samples acquired in Santa Barbara and Goleta, CA, US [100, 102]. The library was assembled using spectra acquired from the airborne sensor AVIRIS [151] and *in-situ* using the hand-held spectroradiometer ASD FR in the spectral range of 350 - 2400 nm (VNIR and SWIR). The post-processing included band reduction of the airborne spectra due to the atmospheric influence and resampling of the field spectra to ensure comparability with the airborne acquired spectra.

The library consists of close to 1000 individual spectra, categorized into 108 urban land cover classes. The land cover based hierarchical classification scheme, as seen in Table 3.3, is split into four ranks and follows a previously defined scheme [8]. Rank 1 represents the main land cover classes, rank 2 the usage or functionality, rank 3 the material properties and rank 4 material-specific characteristics, such as the color. Thus, the labeling is consistently used and properly assigned. The labeling of the material classes *Asphalt*, *Concrete* and *Metal* is shown, where the labeling does not specify the material composition of the *Metal* samples.

[1] `http://www.geo-informatie.nl/Projects/Santa_Barbara_Urban_Spectral_Library/urbanspec/sburbspec.htm`

Table 3.3: Simplified classification scheme that the spectral library Santa Barbara uses [100, 102]. Four ranks are used to describe the samples and the material compositions are defined on rank 3: *Level 3*.

Rank 1: **Level 1**	**Rank 2:** **Level 2**	**Rank 3:** **Level 3**	**Rank 4:** **Level 4**
Built-up	Buildings/roofs	Asphalt roof	Light grey asphalt
	Transportation areas	Asphalt roads	Light asphalt (old)
			Dark asphalt (new)
		Metal roof	Brown metal
			Light grey metal
			Green metal
		Concrete roads	Light concrete
	Sport infrastructure	-	-
Vegetation	Green vegetation	-	-
	Non-photosynthetic vegetation	-	-
Non-urban bare surfaces	Bare soil	-	-
Water bodies	Natural/quasi-natural water bodies	-	-
	Swimming pools	-	-

Table 3.4: The used classification scheme of the spectral library ASTER (version 2.0) [14]. Four ranks are used to describe the samples and the material compositions are defined on rank 3: *Sub-class*.

Rank 1: **Spectral type**	**Rank 2:** **Class**	**Rank 3:** **Sub-class**	**Rank 4:** **Name**
Lunar	-	-	-
Man-made	Road	Paving asphalt	Construction asphalt
	Roofing material	Roofing shingle	Asphalt shingle
		Shingle	Reddish asphalt shingle
			Reddish asphalt roofing shingle
	Concrete	Paving concrete	Construction concrete
			Asphaltic concrete
	General construction material	Cement cinderblock	Construction concrete
		Brick	Bare red brick
			Weathered red brick
			Red smooth-faced brick
Meteorites	-	-	-
Minerals	-	-	-
Non-photosynthetic vegetation	-	-	-
Rock	-	-	-
Soil	-	-	-
Vegetation	-	-	-
Water	-	-	-

ASTER Spectral Library, US

The ASTER spectral library (version 2.0)[2] consists of both natural and man-made materials that have been acquired at various locations in the US [14]. The library is a collection of spectra acquired in the laboratory using hand-held spectroradiometers from three external sources, the Jet Propulsion Laboratory, Johns Hopkins University and the United States Geological Survey. The spectra cover the large spectral range of 400 - 15400 nm (from VNIR to VLWIR). The ASTER library is a part of the ECOSTRESS library, an extension of the ASTER library with a focus on vegetation [161].

The ASTER library contains 3447 spectra which are split into eight classes as seen in Table 3.4, where the majority of the spectra are natural materials (only around 2% are man-made materials). The labeling of the material classes *Asphalt*, *Concrete* and *Brick* is shown. The library is utilizing a hierarchical scheme where each type is split into classes and sub-classes. The library provides a sample description in terms of particle size, color and appearance. The sample naming is however not consistently used, the color is occasionally included (e.g. *Reddish asphalt shingle* compared to *Asphalt shingle*).

DESIREX Spectral Library of Madrid, Spain

The DESIREX spectral library consists of spectra acquired in selected areas in Madrid, Spain [221, 222] to understand the occurrence of UHIs. The spectra were collected using an airborne imaging system and several field instruments in the spectral ranges of 445 - 1015 nm, 1550 - 1750 nm, 2000 - 2553 nm, 3450 - 5250 nm, and 8400 - 12450 nm (VNIR, SWIR, MWIR and LWIR). Additionally, the data was complemented with satellite imagery from several satellites and spectra from the spectral library ASTER [14]. The post-processing included radiometric and geometric calibration, atmospheric correction and georeferencing. As DESIREX also contains spectra in the further ranges of the electromagnetic spectrum, additional processing steps were required.

13 land cover classes were defined from the acquired dataset (see Table 3.5). The labeling of the material classes *Asphalt*, *Concrete* and *Brick* is shown. No hierarchical classification scheme was used as the samples were not split into sub-classes. More detailed information about some of the samples can however be retrieved from the labeling of the individual samples [222]. The samples are not always labeled in such a manner that makes the material classes obvious (e.g. *Parking lot*). This indicates that no external use of the library was intended.

UPMSpeclib Spectral Library of Serdang, Malaysia

The spectral library UPMSpeclib was acquired *in-situ* on the main campus of Universiti Putra Malaysia in Serdang, Malaysia in the spectral range of 350 - 2500 nm (VNIR and SWIR) [172]. The spectra were acquired *in-situ* in sets of five for each sample by using the ASD FieldSpec-3 spectroradiometer. The locations on the campus were pre-selected. No data processing was conducted even if the library was acquired *in-situ* and was therefore affected by the atmospheric water vapor absorption. This can be seen in the provided spectral figures, e.g. as noise in the SWIR spectral range or as a dip of the spectral reflectance R around 1850 nm.

[2] `https://speclib.jpl.nasa.gov/`

Table 3.5: The used classification scheme of the spectral library DESIREX [221, 222]. Two ranks are used to describe the samples and the material compositions are defined on rank 1: *Class*.

Rank 1: **Class**	**Rank 2:** **Sub-class**
Water (lakes)	-
Water (swimming pools)	-
Trees	-
Green grass	-
Roads with asphalt	Parking lot
	Asphalt
Other roads and pavements	-
Roofs with asphalt	Composition shingle
Roofs with red bricks/tiles	Red brick
Roofs with concrete	Concrete
Roofs with metal	-
Shadows	

22 acquired land cover samples were split into two categories: man-made materials and non-built-up surfaces (as seen in Table 3.6). The labeling of the classes *Asphalt*, *Concrete* and *Interlocking block* (unknown material) is shown. The sample description is rather consistent, the man-made materials are all described on a higher level without mentioning the color. However, the sub-class *Road* is the exception, as the *Interlocking* and *Block* describe the usage and not the material. For materials with several samples, the samples were further categorized into either old or new, or rough or smooth.

Table 3.6: The used classification scheme of the spectral library UPMSpeclib [172]. Three ranks are used to describe the samples (two ranks share the same label, *Urban materials*) and the material compositions are defined on rank 3: *Types*.

Rank 1: **Urban materials 1**	**Rank 2:** **Urban materials 2**	**Rank 3:** **Types**
Man-made materials	Roofing	New concrete tile
		Old concrete tile
		Concrete slab
	Road	Asphalt
		New interlocking block
		Old interlocking block
Non-built-up surfaces	Vegetation	-
	Soil	-

LUMA Spectral Library of London, UK

The LUMA spectral library was acquired in the laboratory in addition to complimentary *in-situ* measurements of urban materials present in London, UK [124]. The data acquisition was performed in several spectral ranges: 350 - 2500 nm (VNIR and SWIR) in addition to 8 - 14 μm (LWIR), using the HR-1024 field spectroradiometer and the M2000 Fourier Transform Infrared spectrometer respectively. The data processing consisted of interpolating overlapping data in the spectral ranges SWIR and LWIR by performing a radiometric calibration, correcting background noise from the surrounding environment in addition to separating emissivity and temperature.

LUMA consists of 74 samples that were acquired from construction sites in London and are categorized into ten classes based on appearance and usage (seen in Table 3.7). The labeling of the classes *Asphalt*, *Concrete* and *Metal* is shown. Furthermore, metadata that describes the material, the color and the status (age and surface condition) of the samples in addition to imagery are likewise included and available for download[3]. However, the metadata is not provided in the available downloadable files. The utilized sample description is consistent with a few exceptions, where the material of a class or sub-class is occasionally described on a low level (e.g. sub-class *Metal painted*, material *Metal with paint*) and occasionally on a high level (e.g. sub-class *Metal*, material *Aluminum plus zinc*). However, LUMA is one of the few spectral libraries that does provide imagery in addition to low-level sample description.

Berlin-Urban Gradient Spectral Library of Berlin, Germany

The Berlin-Urban Gradient library[4] consists of image spectra that have been combined from two HyMap images and simulated hyperspectral EnMap data in the spectral range of 450 - 2450 nm (VNIR and SWIR) [179, 180]. The imagery acquired with the airborne hyperspectral sensor HyMap [51] has a Ground Sample Distance (GSD) of 3.6 and 9 m respectively while the simulated EnMap data has a GSD of 30 m. The EnMap program is a hyperspectral satellite mission [89] with a scheduled launch during 2020[5]. As preparation for the mission, EnMap data is simulated using existing hyperspectral data, as done here with HyMap imagery. The HyMap imagery was processed in terms of encompassed system correction, geocoding, atmospheric correction and noisy band removal.

The library contains 75 samples that are hierarchically categorized into major land cover classes according to the Vegetation-impervious surface-soil (VIS) model [200], as seen in Table 3.8. The labeling of the material classes *Asphalt*, *Concrete* and *Unknown* is shown. While focusing on the labels of the impervious urban material samples, most of the samples do include color and material description. Few of the exceptions are e.g. *White roof material (unknown)* and *Concrete*, thus making the sample description slightly inconsistent.

[3] https://urban-meteorology-reading.github.io/SLUM
[4] http://dataservices.gfz-potsdam.de/enmap/showshort.php?id=escidoc:1480925
[5] https://directory.eoportal.org/web/eoportal/satellite-missions/e/enmap

Table 3.7: The used classification scheme of the spectral library LUMA [124]. Five ranks are used to describe the samples and the material compositions are defined on rank 1: *Material class*.

Rank 1: Material class	**Rank 2: Sub-class**	**Rank 3: Material**	**Rank 4: Color**	**Rank 5: Status**
Quartzite	-	-	-	-
Stone	-	-	-	-
Granite	-	-	-	-
Asphalt	Road asphalt	Asphalt with stone aggregate	Black/grey	Weathered
		Tarmac	Black	Weathered
	Asphalt roofing shingle	Tarmac roofing paper	Grey	New
Concrete/cement				
Brick	-	-	-	-
Roofing shingle	-	-	-	-
Roofing tiles	-	-	-	-
Metal	Metal	Aluminum plus zinc	Grey, dull	New
		Aluminum, stucco	Grey, shiny	New
		Aluminum	Grey	Weathered
		Lead	Grey	Weathered
		Iron	Black	Weathered
	Metal painted	Metal with paint	Green	New
			Copper patina	New
			Slate grey	New
PVC	-	-	-	-

Table 3.8: The classification scheme of the spectral library Berlin-Urban Gradient [179, 180]. Three ranks are used to describe the samples and the material compositions are defined on rank 3: *Urban materials*.

Rank 1: VIS	**Rank 2: Extended VIS**	**Rank 3: Urban materials**
Impervious	Roof	Bitumen
		White roof material (unknown)
		Greenish roof material (unknown)
		Light gray roof material (unknown)
	Pavement	Asphalt
		Concrete
Vegetation	Low vegetation	-
	Tree	-
Soil	-	-
Other	-	-

USGS Spectral Library, US

The USGS spectral library version 7[6] consists of both man-made and natural materials. It was acquired in various locations in the US over 20 years [120]. All samples were acquired in the spectral range of 350 - 5000 nm (VNIR and SWIR) while some samples were additionally acquired in the spectral range of 5 - 25 μm (LWIR and VLWIR). The library was assembled by combining spectra from hyperspectral imagery from the airborne sensor AVIRIS [236] and three types of hand-held sensors for *in-situ* and laboratory acquisitions. Rock, soil and mineral spectra were acquired both *in-situ* and in the laboratory.

Table 3.9: The classification scheme of the spectral library USGS [120]. Three ranks are used to describe the samples and the material compositions are defined on rank 2: *Material type*.

Rank 1: Category	**Rank 2: Class**	**Rank 3: Material type**	**Rank 4: Material**
Minerals	-	-	-
Soils	-	-	-
Coatings	-	-	-
Liquids	-	-	-
Organics	-	-	-
Artificial	Man-made	Asphalt	Road asphalt
			Asphalt shingle
			Asphalt roof
		Concrete	Concrete
		Cement + gypsum	Cement + gypsum
		Wood	Wood
			Pine plywood
Vegetation	-	-	-

The library consists of around 1200 samples that are split into seven classes (see also Table 3.9). The labeling of the material classes *Asphalt*, *Concrete* and *Wood* is shown. The sample descriptions are written as longer sentences and state the appearance of the samples and a small discussion about the spectra (e.g. the color). Imagery is available for some selected samples. The sample description for the man-made materials varies, occasionally with more or less information.

SLyRUM Spectral Library of Barcelona, Spain

The SlyRUM spectral library[7] is a dataset containing spectra of roofing materials which exist in the Mediterranean region (Barcelona, Spain) [265]. The spectra were acquired in a controlled laboratory environment in the spectral range of 400 - 1000 nm (VNIR) by using the sensor AISA-EAGLE 2 from four different acquisition angles (0°, 5°,10° and 15°). The post-processing consisted of creating vector layer masks.

[6] `https://crustal.usgs.gov/speclab/QueryAll07a.php`
[7] `https://ddd.uab.cat/record/196065`

Table 3.10: The classification scheme of the spectral library SLyRUM [265]. Four ranks are used to describe the samples and the material compositions are defined on rank 1: *Material class*.

Rank 1: **Material class**	**Rank 2:** **Material**	**Rank 3:** **Color**	*Rank 4:* **Status**
Ceramic	-	-	-
Concrete/cement	Concrete tile with texture	Grey, dull	New, porous
	Concrete brick	Grey, dull	New, porous
Metal	-	-	-
Plastic	-	-	-
Stone	-	-	-
Wood	Wood shingle	Light brown, dull	New
	Wood shingle with blue paint	Blue, synthetic enamel	New
	Wood shingle with red paint	Red, synthetic enamel	New

SLyRUM contains 39 samples that represent six different roofing materials, shown in Table 3.10. The labeling of the material classes *Concrete* and *Wood* is shown. In addition to the class and the specific material, the color, the status (age and surface condition) and imagery of the samples are included. The classification scheme is used consistently but the color description does occasionally include information about the surface coating (e.g. dull, shiny, etc.) which should have been categorized in another rank.

Brussels Spectral Library, Belgium

The land cover spectral library of Brussels [60] was acquired using the imaging hyperspectral sensor APEX [214] in the spectral range of 372 - 2540 nm (VNIR and SWIR). The imagery had a GSD of 2 m and the image spectra were randomly selected in groups of three to 12 pixels and manually labeled. The post-processing consisted of geometric correction, re-projection of the coordinate system and atmospheric correction.

The library contains 752 spectra (pixels) split into 38 land cover classes and the hierarchical classification scheme uses the land cover class labeling according to Table 3.11. The labeling of the material classes *Asphalt*, *Concrete* and *Ceramic* is shown. The classification scheme utilizes two ranks and is not consistent concerning the labeling of the materials. Few of the samples are labeled with a color or color hue description (e.g. *Red ceramic tile* and *Dark ceramic tile*) while others are labeled without it (e.g. *Asphalt* and *Concrete*).

3.1.2 Comparison of Spectral Libraries

The presented spectral libraries are labeling the spectra in different manners with various classification schemes (different hierarchical systems) and taxonomies. Depending on the focus of the spectral library, e.g. land cover classes or urban materials, the classification scheme is structured in different manners. Spectral libraries with a focus on land cover classes (which is the majority of the presented libraries), such

Table 3.11: The classification scheme of the spectral library Brussels [60]. Two ranks are used to describe the samples and the material compositions are defined on rank 2: *Material/type*.

Rank 1: Land cover class	**Rank 2: Material/type**
Roof	Bitumen
	Dark ceramic tile
	Red ceramic tile
Pavement	Asphalt
	Concrete
	Red concrete pavers
Woody vegetation	-
Non-woody vegetation	-
Soil	-
Water	-

as HYSL and FISL [96, 98], Santa Barbara [100, 101] and USGS [120], are using a more general and generic labeling on the highest rank (e.g. vegetation and man-made materials) since they cover a larger sample variation (see e.g. the corresponding Tables 3.2, 3.3 and 3.9). This can likewise be observed based on how these libraries acquired their spectra. As these libraries are completely or partly based on airborne imagery spectra, they tend to use a higher level to describe the samples since the spectra is probably used as training data for the classification of spaceborne or airborne imagery. Low-level details are not necessary as they may not be distinct in the utilized imagery.

On the other hand, libraries with a focus on only urban materials, such as LUMA [124] and SLyRUM [265], include the material type on the highest rank (e.g. ceramic, concrete/cement). Thus, through the classification scheme, they can describe the spectra on a low level which includes the color and the surface conditions. Likewise, these have acquired their spectra in the laboratory or *in-situ* and can therefore perform a detailed study of the samples. The schemes reflect this closer observation. The two libraries that are exceptions of this statement are ASTER [14] and UPMSpeclib [172]. Although the library ASTER was created in the laboratory, it was assembled as a supplement to the airborne sensor ASTER [258] and its projects. The UPMSpeclib library consists of land cover classes even if it was acquired *in-situ* since its purpose was to train a classifier for classification of a satellite image [172].

Table 3.12: An overview of the presented urban material spectral libraries. The table contains information about the acquisition area, the content, the covering spectral ranges and how the data was acquired. *Libraries available for download.

Spectral Library	**Acquisition Area**	**Content**	**Spectral Range**	**Data Acquisition**
PUSL and CASILIB [19]	Tel-Aviv, Israel	55 samples, of which: 11 facade 16 ground 2 roof	VNIR	Airborne *In-situ*
HYSL and FISL [96, 98]	Dresden and Potsdam, Germany	32 samples, of which: 8 ground 13 roof	VNIR SWIR	Airborne *In-situ*
Santa Barbara* [100, 102]	Santa Barbara, CA, US	26 samples, of which: 9 ground 3 roof	VNIR SWIR	Airborne *In-situ*
ASTER* [14]	Various locations, US	3420 samples, of which: 28 facade 10 ground 18 roof	VNIR SWIR TIR	Laboratory
DESIREX [221, 222]	Madrid, Spain	27 samples, of which: 3 facade 18 ground 1 roof	VNIR SWIR MWIR LWIR	Airborne *In-situ*
UPMSpeclib [172]	Serdang, Malaysia	22 samples, of which: 15 roof 3 ground	VNIR SWIR	*In-situ*
LUMA* [124]	London, UK	74 samples, of which: 48 facade 9 ground 17 roof	VNIR SWIR LWIR	Laboratory *In-situ*
Berlin-Urban Gradient* [179, 180]	Berlin, Germany	75 samples, of which: 7 ground 23 roof	VNIR SWIR	Airborne Simulated
USGS* [120]	Various locations, US	2468 samples, of which: 9 facades 15 ground 16 roof	VNIR SWIR LWIR VLWIR	Airborne Laboratory *In-situ*
SLyRUM* [265]	Barcelona, Spain	39 samples, of which: 39 roof	VNIR	Laboratory
Brussels [60]	Brussels, Belgium	752 samples, of which: 221 ground 304 roof	VNIR SWIR	Airborne

Due to the different approaches of labeling a material, it can be challenging to find samples with the same corresponding material labels in the various spectral libraries. A spectral comparison between samples representing the same material class is only possible by using the assigned material labels. However, as stated earlier, a material's spectral characteristic features are due to the chemical composition. If these are not provided, it is only possible to rely on the assigned labels. Spectra that have been labeled as *Asphalt* and *Concrete* are extracted from the libraries Santa Barbara [100, 101], ASTER [14], LUMA [124] and USGS [120]. The selected samples that represent the two materials can be seen in Figure 3.1 and are the following:

Asphalt:

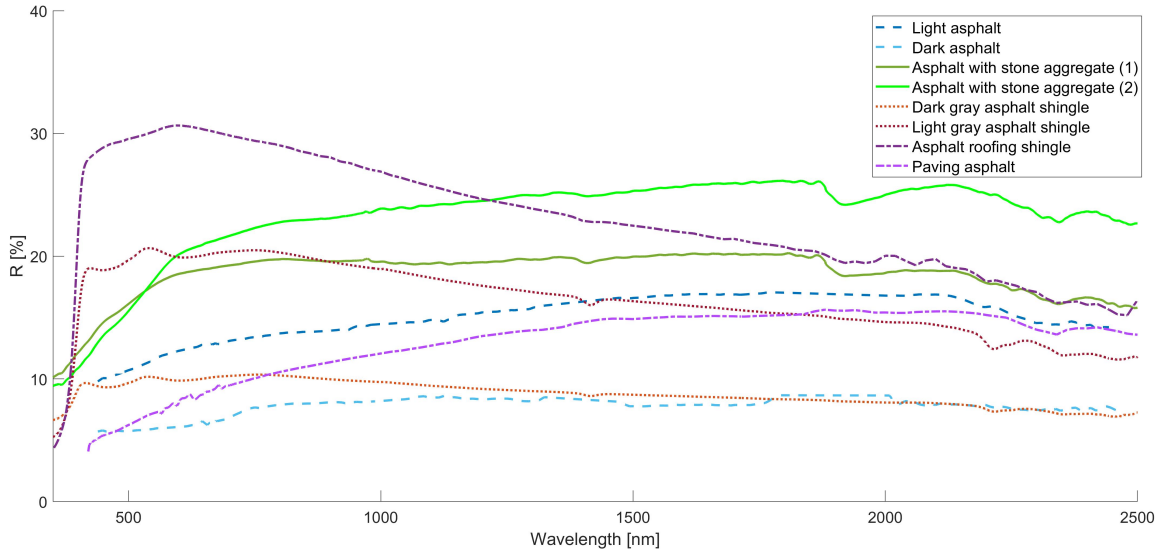
- *Light asphalt* (Santa Barbara)
- *Dark asphalt* (Santa Barbara)
- *Asphalt with stone aggregate (1)* (LUMA)
- *Asphalt with stone aggregate (2)* (LUMA)
- *Dark gray asphalt shingle* (USGS)
- *Light gray asphalt shingle* (USGS)
- *Asphalt roofing shingle* (ASTER)
- *Paving asphalt* (ASTER)

Concrete:

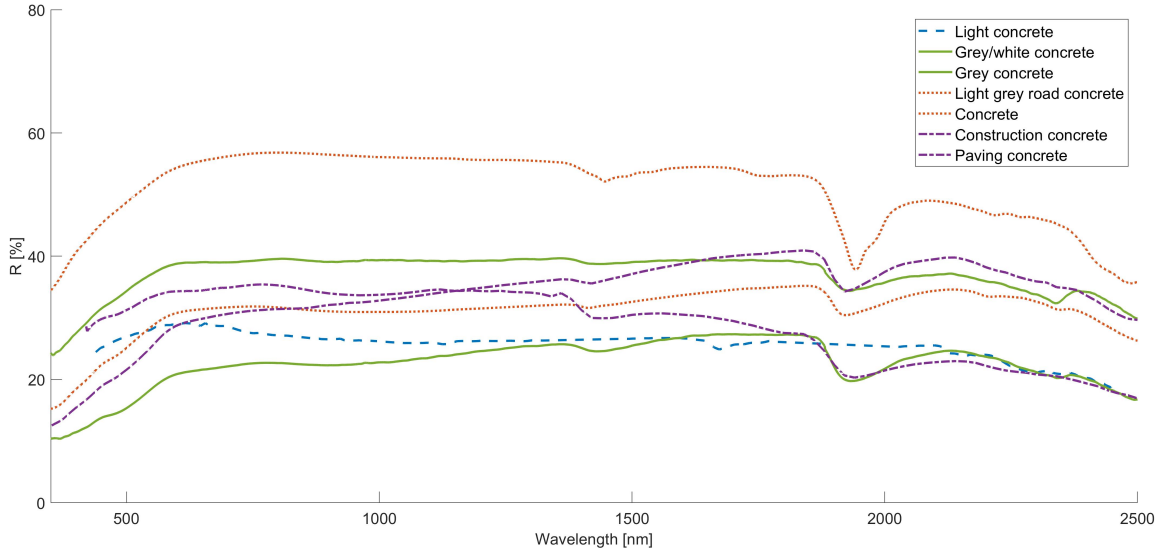
- *Light concrete* (Santa Barbara)
- *Gray/white concrete* (LUMA)
- *Gray concrete* (LUMA)
- *Light gray road concrete road* (USGS)
- *Concrete* (USGS)
- *Construction concrete* (ASTER)
- *Paving concrete* (ASTER)

Figure 3.1 displays a high intra-class spectral variation as the spectral characteristic features of the same material varies prominently. Some of the *Asphalt* samples in Figure 3.1a have different color coatings and a spectral difference can be observed in the visible part of the spectrum. However, a spectral difference can be observed across the entire spectrum. As previously discussed in Subsection 2.2.2, this is due to the elemental composition of *Asphalt* (mainly carbon and hydrogen) and the regional dependency [187]. Likewise, concrete consists of a mixture of element components with a varying ratio (e.g. silicon dioxide and aluminum oxide). This can be observed in Figure 3.1b as the spectral characteristic features vary. Although there is a spectral variation among the samples, all have been labeled as either *Asphalt* or *Concrete*. However, samples with similar labels do not ensure similar spectral characteristic features, as seen e.g. *Light asphalt* and *Light gray asphalt shingle* in addition to *Light concrete* and *Light gray road concrete*. On the other hand, samples with dissimilar labels can have similar spectral features, such as *Light asphalt* and *Paving asphalt* in addition to *Construction concrete* and *Light gray road concrete*. This further indicates the challenge of finding samples with the same material label and similar spectral characteristic features in the different spectral libraries. The labels are not assigned based on the chemical compounds, they are based on the appearance or usage which are defined by the creator of the library. The labeling, i.e. the taxonomy, is as observed both complex and comprehensive and will therefore be further discussed in Chapter 3.3.

Finally, the majority of the spectral libraries do not provide imagery of the samples as they are focusing on land cover classification. The spaceborne or airborne imagery would consist of few pixels and would not provide any further gained information about the sample. LUMA [124], USGS (for selected samples) [120] and SLyRUM [265] are the exceptions as they acquired all or part of the spectra close-ranged. However, USGS is the only library that offers direct download of the imagery while the other two provide the imagery in form of a technical document. If imagery is provided, the comparison of material samples can be performed even further by exploiting and extracting textural information about the material samples.



(a) *Asphalt*



(b) *Concrete*

Figure 3.1: (a-b): A spectral comparison of selected samples that represent the two material classes *Asphalt* and *Concrete*. An intra-class spectral variation between the samples can be observed in both (a) and (b), indicating an intra-class spectral dissimilarity. The library Santa Barbara is represented as blue dashed lines, the library LUMA as green solid lines, the library USGS as red dotted lines and the library ASTER as purple dash-dotted lines.

3.1.3 Texture Databases

Texture can be described and defined as visual information about the physical shape of an object [158] or as a function that represents the spatial arrangement and variation of pixel intensity values in an image [44, 232]. Texture analysis can be used to extract textural features, discriminate texture in an image into different regions, classification or reconstructing surface geometry onto a 3D model [158]. A texture database is a collection of imagery that can be used for further textural analysis. A texture database can for instance be used for classification tasks by determining distinct textural features with a chosen textural measure (described in more detail in Chapter 4.3).

While spectral libraries that have a strong footing in the remote sensing community, texture databases are not as well-established and utilized. These datasets have their footing in the computer vision community, which is reflected by the different manner of labeling the samples (i.e. the taxonomy) and by the lack of classification schemes. Additionally, this is likewise reflected in the distance between the object and the camera as these datasets contain imagery that were taken from a close range and not from airborne nor spaceborne platforms. The imagery represent objects from near to determine detailed textural features.

The following three types of texture databases need to be considered. Firstly, several texture databases consist of RGB imagery of the same sample acquired under different conditions concerning scale, illumination and viewing angle in a controlled environment (in a laboratory). By capturing the sample under different conditions, the imagery provide a wide textural diversity that contributes to a solid foundation for classification tasks. Such databases include the CUReT Database [58], the Outex Texture Dataset [177], the UBO2014 Database [245] and the KTH-TIPS Database [94]. To account for intra-class variation, the extension KTH-TIPS2 Database [37] was generated where further material classes and more samples of the same material class were acquired again under different conditions. However, these imagery have been acquired in a controlled environment and are therefore not taking into consideration the natural varied conditions.

Secondly, several databases have acquired their imagery in an uncontrolled environment (not in a laboratory) but likewise under different conditions, such as the viewing angle, illumination and scale. By acquiring the data in an uncontrolled environment, the true and varied complexity of one material class can be captured (e.g. changes of the illumination influences the observed surface color and the different scales the observed textural details) and thus, accounting for a larger intra-class variation. Such databases include the Flickr Material Database [218], the OpenSurfaces dataset [16] and the Materials in Context (MINC) Database [17].

Lastly, there are databases that have either collected the imagery from several sources or acquired the imagery in an uncontrolled environment. These datasets do not contain samples acquired under different conditions and do not provide a large textural variation of one material. These datasets rather aim to provide a large number of samples. This includes the Texture library [75], the ETHZ Synthesizeability Dataset (which also contains synthetic images) [57], the OpenStreetMap Texture Library [181] and the Salzburg texture image database (STex) [133].

Texture databases tend to contain a large variety of materials and focus rarely purely on urban materials. A detailed comparison between the texture databases will not be performed. As classification schemes are rarely used, the chosen taxonomy reflects this. Urban materials that are reoccurring present in these datasets are *Concrete*, *Granite*, *Metal* and *Wood*. A commonly named material class is *Stone* [17, 57,

75, 133, 138]. However, *Stone* is a widely defined material class as *Stone* occasionally includes samples representing materials such as *Asphalt* and *Sandstone*. *Brick* is likewise defined as a material class due to its unique textural structure but without a specific material assignment [17, 57, 58, 138, 181]. Furthermore, classes such as *Door* [75], *Building* [133], *Floor* [57, 133], *Tile* [17, 57, 133, 177, 181] and *Wall* [57, 133, 138, 181] may be important for texture analysis due to the texture but do not provide information about the actual used material. Thus, contrary to spectral libraries, texture databases focus on the appearance of the samples to categorize them into different classes since texture analysis relies on visual information.

3.1.4 Differences Between Spectral Libraries and Texture Databases

As the two types of datasets have their strong footing in two different research fields, remote sensing and computer vision respectively, the purpose differs. Spectral libraries utilize more comprehensive classification schemes and labeling as the purpose is to distinguish many complex classes on a low level. Texture databases utilize a simplified labeling system as the need to distinguish e.g. different types of land cover classes or urban materials is a rare task. These rather serve the purpose of distinguishing classes on a high level (e.g. *Stone* and *Wood*) or based on the usage (e.g. *Door* and *Wall*).

The measured spectra provide information about the chemical compounds of the material while imagery provide visual information about the material's appearance. A material can visually appear as another material that cannot be detected with imagery but would be detected with measured spectra. On the other hand, the spectral characteristic features can vary due to the used chemical compounds which would indicate a dissimilarity with measured spectra but could be indicated as visually similar. Thus, both kinds of datasets have strengths and weaknesses. By simultaneously using information from both kinds of datasets to i.e. distinguish material classes, additional information can be gained (in more detail in Chapter 6 where this idea is applied).

3.2 Taxonomy Approaches

Taxonomy is the science of classification which involves how objects are categorized or arranged into groups or classes based on similarities [68, 159, 198]. Taxonomy is originally rooted in biology and is in modern times regarded to have been founded by the Swedish botanist Carl Linnaeus [146]. The Linnaean taxonomy consists of classifying living things into three domains (classes) by using a rank-based classification scheme. The three classes animals, plants and minerals have each in turn several underlying orders, such as family, genus and species. Taxonomy is in modern times established in other fields as well, such as education (e.g. Bloom's taxonomy) [127] and classification of soil [250].

Taxonomy can be hierarchically described where the relationship between the ranks is explained as "IS-A" [28]. A "IS-A" relationship describes one object's inheritance of another object. The object of a "IS-A" relationship is the subset of the higher object and is therefore "a kind of" (the consecutive lower rank). For instance, "animal" is defined as the highest rank in a scheme and "dog" is the consecutive rank. Hence, "a dog is a kind of animal". The highest rank in a classification scheme defines the restrictions and roles of the subsets (the lower ranks). Spectral libraries tend to use a hierarchically based classification scheme with "IS-A" relationships as observed by the libraries Santa Barbara [100, 102] and UPMSpecLib

[172]. They describe the lower ranks in the schemes with a "IS-A" relationship as seen in Tables 3.3 and 3.6, where e.g. "Transportation areas are built-up areas" and "Road is a man-made material".

In ecology, the diversity of a dataset can be split into α, β and γ diversities [249]. Firstly, α diversity refers to the average species diversity in a local area (in one limited area). This can in general be described as recognizing, describing and naming objects [68]. Secondly, β diversity is defined as the ratio between the local and regional species diversity. This can likewise be described as comparing the relationship between objects and classifying those based on the observations [68]. Lastly, γ diversity is the total diversity of species. γ diversity can generally be described as studying the variation among the objects (the intra-class diversity) [68]. Based on the existing species diversity, a dataset can define its taxonomy to reflect the needed diversity level. A dataset that covers the local diversity may define its labels using α diversity.

α, β and γ diversities can likewise be used to describe how spectral libraries have defined their labeling system, based on the sample diversity. α diversity is used when the labeling is based on the locally available and acquired urban material samples. The libraries tend to acquire urban material samples that are locally available and therefore describe the local average diversity. They describe likewise the samples to fit the only intended use (i.e. the local use). For example, spectral libraries that contain land cover classes describe their samples on a high level as low-level details are not recognizable from airborne or spaceborne sensors. Contrary, spectral libraries that describe their samples on a low level acquired their spectra close-ranged, which makes it possible to recognize and describe the corresponding low-level details. β diversity is on the other hand referred to regional or global labeling by accounting the regional or global variation. The majority of the presented spectral libraries have acquired their samples in a small local area and label therefore their samples based on the local diversity. The two libraries ASTER [14] and USGS [120] acquired the spectra at several locations and those libraries had to establish a β diversity to cover the regional variation. β diversity can be extended by comparing the local labels of one material class of different spectral libraries to establish the regional variation. γ diversity is referred to as intra-class assessment. This is covered by the spectral libraries that contain several samples of the same material to provide variation and diversity. This labeling is nonetheless partly based on the α diversity, as the libraries label the samples based on the local intra-class variation. To conclude, the α and γ diversities are defined differently in spectral libraries (as likewise observed in the various classification schemes in the tables in Chapter 3.1). To define a general β diversity for all libraries, extensive analysis of the existing libraries is needed to establish a global variation. However, this is challenging to determine.

3.3 Existing Approaches for Taxonomy of Urban Materials

To the best of the author's knowledge, no standardized classification scheme for urban material spectral libraries has been established so far. A land cover and land-use classification scheme was nonetheless defined in 1971 [7] and further extended in 1973 [8]. These articles suggest using 10 criteria to create a scheme. For instance, "The minimum level of interpretation accuracy in the identification of land-use and land cover categories from remote sensor data should be at least 85%", "The accuracy of interpretation for the several categories should be about equal", "The classification system should be applicable over extensive areas" and "Aggregation of categories must be possible". A systematic structure can be established while following these proposed criteria, but the articles suggest labeling and describing urban

material classes based on the usage (e.g. *Commercial and services* and *Residential*) and not on the actual material (i.e. the chemical components). Among the presented libraries that contain land cover classes, the Santa Barbara spectral library [100, 102] is the only library that intentionally follows this proposed classification scheme. While all libraries use hierarchy-structured classification schemes, the labeling systems differ.

The spectral characteristic features of the material classes *Asphalt* and *Concrete* were compared in Figure 3.1. For the sake of consistency, these two material classes are again used to compare how the libraries structure the classification schemes and define the labeling systems. Tables 3.13 and 3.14 present respectively how the libraries label the two material classes *Asphalt* and *Concrete* (the libraries that contain at least one sample of the material class).

Table 3.13: Taxonomy comparison of how the different spectral libraries have categorized the material class *Asphalt*. The comparison scheme is split into four ranks, corresponding to the most commonly used ranks. Ranks 4 and 5 of the library LUMA are merged into one while ranks 1 and 2 of the libraries HYSL and FISL are excluded.

Spectral Library	**Rank 1**	**Rank 2**	**Rank 3**	**Rank 4**
CASILIB [19]	-	Urban surfaces	Urban surfaces 2	Asphalt (dark) Asphalt (light) School roof
HYSL and FISL [96, 98]	Roof materials	Hydrocarbons	Bitumen	Green/gray Red Gray
	Impervious open surfaces	Hydrocarbons	Asphalt	Dark/gray
Santa Barbara [100, 102]	Built-up	Buildings/roof Transportation areas	Asphalt roof Asphalt roads	Light gray asphalt Light asphalt (old) Dark asphalt (new)
ASTER [14]	Man-made	Road Roofing material	Paving asphalt Roofing shingle Shingle	Construction asphalt Asphalt shingle Reddish asphalt shingle Reddish asphalt roofing shingle Asphalt roofing shingle
DESIREX [221, 222]	-	-	Road with asphalt	Asphalt Parking lot
			Roof with asphalt	Composition shingle
UPMSpecLib [172]	-	Man-made materials	Road	Asphalt
LUMA [124]	Asphalt	Road asphalt	Asphalt with stone aggregate	Black/gray, weathered
			Tarmac	Black, weathered
		Asphalt roofing paper	Tarmac roofing paper	Gray, new
Berlin-Urban Gradient [179, 180]	-	Impervious	Pavement	Asphalt
			Roof	Bitumen
USGS [120]	Artificial	Man-made	Asphalt	Road asphalt Asphalt shingle Asphalt roof
Brussels [60]	-	-	Roof	Bitumen
			Pavement	Asphalt

Table 3.14: Taxonomy comparison of how the different spectral libraries have categorized the material class *Concrete*. The comparison scheme is split into four ranks, corresponding to the most commonly used ranks. Ranks 4 and 5 of the library LUMA are merged into one while ranks 1 and 2 of the libraries HYSL and FISL are excluded.

Spectral Library	Rank 1	Rank 2	Rank 3	Rank 4
CASILIB [19]	-	Urban surfaces	Urban surfaces 1	Schoolyard concrete
HYSL and FISL [96, 98]	Roof materials	Mineral	Concrete	Grey
				Red
	Impervious open surfaces	Mineral	Concrete	Bright/gray
	Partially impervious open surfaces	Mineral	Pavement of concrete	Bright/gray
				Red
Santa Barbara [100, 102]	Built-up	Transportation areas	Concrete roads	Light concrete
ASTER [14]	Man-made	General construction material	Cement cinderblock	Construction concrete
		Concrete	Paving concrete	Asphaltic concrete
				Construction concrete
DESIREX [221, 222]	-	-	Roofs with concrete	Concrete
UPMSpecLib [172]	-	Man-made materials	Roofing	New concrete tile
	-			Old concrete tile
	-			Concrete slab
LUMA [124]	Concrete/cement	Concrete	Concrete	Gray/white, new
				Gray, weathered, rough
Berlin-Urban Gradient [179, 180]	-	Impervious	Pavement	Concrete
USGS [120]	Artificial	Man-made	Concrete	Concrete
	-		Cement+Gypsum	Cement+Gypsum
SLyRUM [265]	Concrete/cement	Concrete tile with texture	Gray, dull	New, porous
		Concrete brick	Gray, dull	New, porous
Brussels [60]	-	-	Pavement	Concrete
				Red concrete pavers

As observed, the libraries define and label the material classes *Asphalt* and *Concrete* differently. Furthermore, the ranks the libraries use to define the two material classes vary likewise. As previously discussed, the two close-range acquired libraries LUMA [124] and SLyRUM [265] define the materials at rank 1 while land cover libraries such as HYSL and FISL [96, 98], USGS [120] and Santa Barbara [100, 101] define the materials at the lower ranks (ranks 3 or 4). Furthermore, given the information in the different ranks differs (e.g. the usage, color and age of the sample). The information is occasionally given for only certain samples in a library, e.g. two out of three *Asphalt* samples in the Santa Barbara library describe the age, two out of three *Asphalt* samples in the CASILIB library describe the color and two out of three *Asphalt* samples in the USGS library describe the usage. The majority of libraries do however distinguish the material classes *Asphalt* and *Concrete* as either road or roof material, indicating that those should be split into two different land-use classes even if the chemical composition is not provided. The general chemical composition is only given by the library FISL [96, 98]. Even if FISL splits up the samples of material class *Concrete* into road and roof, the description of the chemical composition remains the same.

As the libraries are describing the same materials differently, it can be challenging to assess if samples labeled as *Asphalt* through e.g. *Built-up - buildings/roof - asphalt roof - light gray asphalt*, *Man-made materials - roofing materials - roofing shingle - asphalt shingle*, *Asphalt - asphalt roofing paper - tarmac roofing paper - gray* and *Man-made - asphalt - asphalt roof* are truly describing the same type of *Asphalt*. While *Asphalt* is described relatively consistently with the various labeling systems, more liberties have

been taken to label the material class *Concrete*. The labeling of *Concrete* varies likewise, e.g. *Impervious open surfaces - mineral - concrete - bright/gray*, *Built-up - transportation areas - concrete roads - light concrete*, *Man-made materials - concrete - paving concrete - construction concrete* and *Pavement - red concrete pavers*. Again, it is not possible to ensure that these samples are describing the same type of *Concrete*. As the spectral characteristic features likewise can differ (as seen in Figure 3.1), imagery of the samples could further indicate if they are representing the same material.

Nonetheless, as no standardized labeling systems and classification schemes are used, the creator of a library can label the material classes according to her own definitions. As the libraries are mostly based on spectra acquired in one region, they present the α diversity and label the samples accordingly. They may label and split the samples of one material based on a color difference since they acquired two samples of one material with two different colors, such as *Asphalt (light)* and *Asphalt (dark)*. Contrary, libraries that contain only one sample of one material occasionally label this sample without any color, such as *Asphalt*. To ensure that libraries use a similar labeling system, libraries should label their samples with a β diversity in mind. They should fit the labeling for regional or global diversity and not a local diversity (α diversity). As spectral libraries are not necessarily used for the regions where the spectra were acquired [67, 190, 260], the labeling should likewise reflect this.

3.4 Summary

This chapter has presented the relevant datasets that contain either spectral or textural urban material information. This included a detailed overview of the existing spectral libraries, an analysis of the corresponding classification schemes and a comparison of the spectral characteristic features for the two material classes *Asphalt* and *Concrete*. A brief overview of relevant texture databases was likewise presented and the two types of datasets were compared to determine the differences. To further analyze the datasets to understand how they have created their classification schemes and labeling systems, the term taxonomy was introduced. Taxonomy involves how objects are categorized or arranged into groups or classes. The labeling system can be based on α, β or γ diversities. The urban material classification schemes tend to use a hierarchical structure. The taxonomy of the presented spectral libraries is analyzed and compared. The two material classes *Asphalt* and *Concrete* are again chosen to compare how the corresponding labels have been defined. It can be concluded that libraries create their labeling based on α diversity. To improve the comparability of libraries, a standardized classification scheme with an accompanied labeling system should be developed based on β diversity. Based on the observations made about the taxonomy of the existing spectral libraries, the knowledge about material taxonomy is applied to the spectral library KLUM, presented in Chapter 5.

4 Approaches and Tools for Material Classification

This chapter gives an overview of the tools and approaches that can be used for assessment and enhancement of hyperspectral data for material classification. The presented tools and approaches are used in this thesis to assess materials utilizing spectral and textural features (data). Firstly, Chapter 4.1 describes the different classifiers which can be used to assign material class labels. This includes both supervised and unsupervised classifiers. Secondly, as this thesis mostly utilizes high-dimensional data, Chapter 4.2 explains the methods used for selecting important features and for reducing the data dimensionality. Thirdly, texture analysis approaches that are used to extract textural features (information) are presented in Chapter 4.3. Subsequently, a material classification case study is presented in Chapter 4.4 which evaluates a selection of the presented tools and approaches. The case study utilizes two spectral datasets and analyzes, in addition to the various methods, the spectral range importance. Lastly, this chapter is summarized in Chapter 4.5.

4.1 Classifiers

Machine learning-based classification relies on data observations to determine and assign labels to a new set of data observations. Classification can consist of assigning class labels to pixels or segments in an image, points in a 3D point cloud or spectra in a spectral dataset. Classification can be done with a generative or discriminative approach [25]. Given a joint probability of the input x and the corresponding label y as $p(x,y)$, a generative classifier tries to learn a model that describes the same joint probability as $p(x,y)$. A discriminative classifier tries to learn a model that expresses a decision boundary that depends on the conditional density $p(y|x)$ without any assumption about the distribution of the input x. The classifiers used in this thesis are all discriminative since they base their classification on decision boundaries.

Supervised classification involves training a classifier using known and pre-labeled data to map a learned function onto unlabeled data [123]. Given a dataset with corresponding known labels, the data can be split into training, testing and validation data. The ratio between the three kinds of data commonly depends on the size of the data. A common practice is to split the data into 80% training and 20% testing. If validation data is likewise used, the split could be e.g. 80% training, 10% testing and 10% validation or 70% training, 20% testing and 10% validation. The training data consists of data, such as vectors representing hyperspectral data, and the corresponding known assigned labels that the classifier uses to learn a function which describes the relationship between the data and the labels. The learned function is mapped upon the testing data (which likewise contains data and known assigned labels) and predicts the labels. The performance of the function the classifier used can be evaluated using a confusion matrix based on the predicted labels and the known assigned labels of the testing data. Instead of using the testing data to evaluate the performance of a classifier, independent validation data can likewise be used.

As the testing data is commonly used to tune the parameters of a supervised classifier, the classifier can be over-fitted. The validation data can therefore be used to independently evaluate the performance as the data has not been used for the parameter tuning. From the confusion matrix, several useful statistical parameters can be derived.

A confusion matrix, an error matrix, visualizes the performance of a supervised classifier by presenting the predicted class labels (the mapped testing data) across the rows and the known assigned labels across the columns [191]. The matrix visualizes the number of false positive (FP, type I error), false negative (FN, type II error), true positive (TP), and true negative (TN) predicted labels for object X [191]. Table 4.1 illustrates a confusion matrix for a simple two-class problem.

Table 4.1: A simple example of a confusion matrix and how the different outcomes (e.g. FP and TP) are distributed for a two-class problem.

		Actual label	
		X	Not X
Predicted label	X	TP	FP
	Not X	FN	TN

Based on the confusion matrix, five useful parameters can be calculated to evaluate the classifier. Firstly, the overall accuracy O indicates the overall performance within the range of $[0, 1]$, where 1 is the highest accuracy. For a multi-class problem, the overall accuracy O can be calculated with one of the following expressions. It can be calculated with either c_{ii} and c_{ij}, where c represents the samples while i and j are the indexes of the rows and columns respectively (i.e. ii is the position of the diagonal) or the class individual TP_i, TN_i, FP_i and FN_i:

$$O = \frac{\sum_i c_{ii}}{\sum_i \sum_j c_{ij}} = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \tag{4.1}$$

Secondly, the κ-value indicates how well the classes can be separated within the range of $[0, 1]$, where 1 corresponds to perfect separation:

$$\kappa = \frac{O - p_{e,i}}{1 - p_{e,i}} \tag{4.2}$$

$$p_{e,i} = \frac{(TP_i + FP_i)(TP_i + FN_i) + (FN_i + TN_i)(FP_i + TN_i)}{TP_i + TN_i + FP_i + FN_i} \tag{4.3}$$

Thirdly, recall R indicates the ratio of relevant observations that were retrieved within the range of $[0, 1]$, where 1 indicates perfect recall R:

$$R_i = \frac{TP_i}{TP_i + FP_i} \tag{4.4}$$

Fourthly, precision P indicates the ratio of relevant observations among the retrieved observations within the range of $[0, 1]$, where 1 indicates perfect precision P:

$$P = \frac{TP_i}{TP_i + FN_i} \tag{4.5}$$

Finally, the F_1-score is the harmonic mean of recall R and precision P within the range of $[0, 1]$, where 1 is the highest score:

$$F_{1,i} = 2 \cdot \left(\frac{R_i \cdot P_i}{R_i + P_i} \right) \tag{4.6}$$

Table 4.2 contains an artificial confusion matrix of three classes which be used to further demonstrate how the five parameters are calculated. By assuming a dataset of 20 animals such as 10 *Cats*, 6 *Dogs* and 4 *Birds*, the confusion matrix could look like Table 4.2.

Table 4.2: An artificial example of a confusion matrix and how the different outcomes are distributed for a multi-class problem.

		Actual label		
		Cat	Dog	Bird
Predicted label	Cat	7	1	0
	Dog	2	4	1
	Bird	1	1	3

The overall accuracy O of this multi-class problem is $(7+4+3)/20 = 0.7$ and the κ-value is $(0.7 - p_e)/(1 - p_e) = 0.53$, where $p_e = 0.2 + 0.105 + 0.05 = 0.355$. For the class *Cat*, the recall R is $7/(7 + (1 + 0)) = 0.875$, the precision P $7/(7 + (2 + 1)) = 0.7$ and the F_1-score $2 \cdot (0.7 \cdot 0.875)/(0.7 + 0.875) = 0.78$. These parameters indicate that the class *Cat* can be rather well separated from the other two classes but the general class separation is poor as seen by the low κ-value. The ideal confusion matrix would have all samples in the diagonal (i.e. 10, 6 and 4), which would indicate that no samples have been incorrectly predicted. The worst-case scenario would be if the diagonal contains only zero, which would indicate that all samples have been incorrectly predicted.

Unsupervised classification does require neither training, testing nor validation data (i.e. no labeled data). The chosen algorithm clusters the data into suitable classes without any interference from the user. The clusters are based on the algorithm's analysis and prediction of the presented data without needing any assigned labels [111]. Unsupervised classification can be used when no prior information is available or when the aim is to find a structure in the dataset. Semi-supervised classification is another approach that combines a small labeled dataset with a large unlabeled dataset for the training of a classifier [271].

The supervised and unsupervised classifiers that have been used throughout this thesis will be further explained in the following text. Figure 4.1 illustrates the decision boundaries of the different discriminative supervised classifiers which will be presented in the following text using randomly generated datasets. The code to generate the illustration is available online from the toolbox *scikit-learn*[1] and has been adapted to illustrate the decision boundaries for the classifiers presented in this chapter. The points in solid colors represent the training dataset and semi-transparent points the testing dataset.

[1] `https://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html`

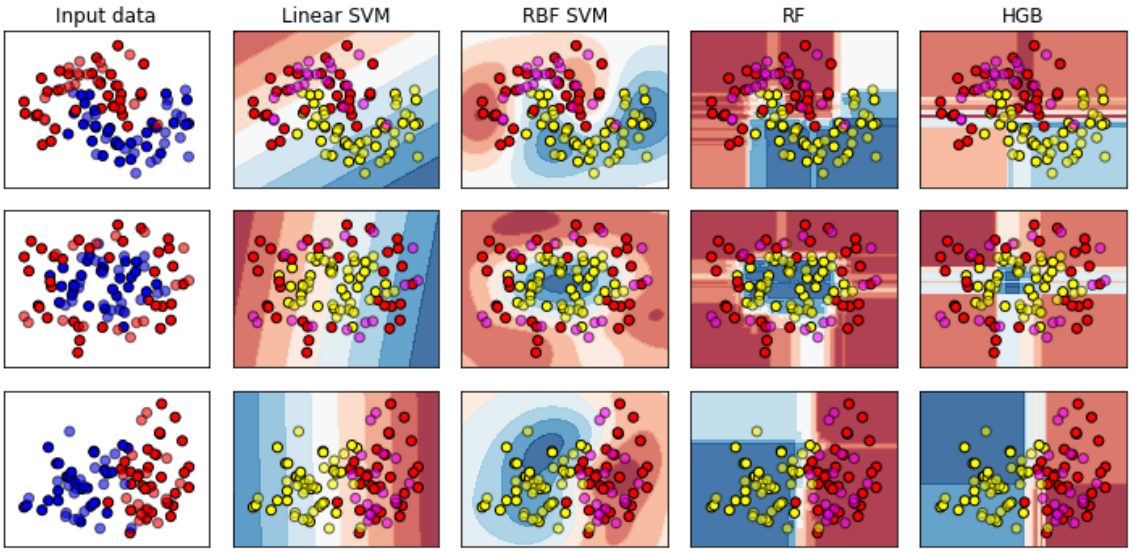


Figure 4.1: Visual comparison of the decision boundaries of the classifiers SVM, RF and HGB.

Support Vector Machine

The Support Vector Machine (SVM) is a supervised classifier that originally utilized a binary approach to linearly separate two classes (one-vs-one scheme) [53, 253]. The separation of the classes is done within a high-dimensional feature space by constructing one or several hyperplanes (i.e. decision boundaries). The classes can be separated both linearly and non-linearly. For non-linear separation, a kernel function $k(x,y)$ is needed which can be based on the Gaussian radial basis function $k(x,y) = exp(-\gamma\|x_i - x_j\|^2)$, polynomial functions $k(x,y) = (x_i \cdot x_j)^2, k(x,y) = (x_i \cdot x_j + 1)^2$ or the hyperbolic tangent $k(x,y) = tanh(kx_i \cdot x_j + c)$. For more complex classification tasks, the classes are commonly not linearly separable and a kernel function has to be utilized to separate and map the data onto a new higher-dimensional space.

The SVM is a max-margin learning approach and aims therefore to maximize the margin d (i.e. the width around the decision boundaries) between the classes in the new feature space. Generally, the SVM can utilize a hard-margin approach, where the aim is to correctly separate all training data into their assigned classes. This can result in complex hyperplanes as some classes are closely located in the feature space. Furthermore, a hard-margin approach tends to over-fit the data. On the other hand, a soft-margin approach utilizes a less complex hyperplane and allows for some misclassification of the training data [244]. A soft-margin approach has likewise a better generalization of new data.

The SVM can be defined as following for a linear separation. Given are a training dataset as $X = x_1, ..., x_n$, the corresponding binary labels as $Y = y_1, ..., y_n$ (where $y \in -1, 1$) and the hyperplanes H_1 and H_2 which correspond to the linear separation of $y_i = 1$ and $y_i = -1$ respectively. With w being the weight assigned to x (coefficients in the primal problem) and b the bias, the linear separation can be defined with the following equation:

$$H_1 : w \cdot x_1 + b = 1$$
$$H_2 : w \cdot x_1 + b = -1$$

The distance between the two hyperplanes H_1 and H_2 is $\frac{2}{\|w\|}$ and the median plane H_0, which is located between the two hyperplanes H_1 and H_2 (with a margin of $+d$ and $-d$ respectively) can be defined as:

$$H_0 : w^t x + b = 0 \tag{4.7}$$

To maximize the margin d, $\|w\|$ can be minimized as:

$$w \cdot x_1 + b \geq 1, \text{ for } y = 1$$
$$w \cdot x_1 + b \leq -1, \text{ for } y = -1$$

This is now a constrained optimization problem which has to be solved with quadratic programming with primal or dual form (for more details, refer to [53]). If margin d is hard-margin, the two hyperplanes H_1 and H_2 would be expressed as:

$$H_1 : w \cdot x_1 + b = 0$$
$$H_2 : w \cdot x_1 + b = 0$$

For a soft-margin approach, the two hyperplanes H_1 and H_2 would be expressed as:

$$H_1 : max(0, 1 - y_i(w \cdot x_1 + b - 1))$$
$$H_2 : max(0, 1 - y_i(w \cdot x_1 + b + 1))$$

For multi-class classification tasks, the SVM classifier uses a combination of several binary SVMs which allows performing of multiple binary classifications instead of one multi-class classification [40]. The one-vs-one scheme is thus used to train several SVMs against each other to distinguish samples of one class from samples of another class. This improves the distinction of classes that are closely located in the new high-dimensional space.

To initiate the SVM classifier, certain parameters can be defined. This includes the kernel function $k(x, y)$ and the class weights w.

Random Forest

An ensemble approach combines several learning algorithms to build a predictive model [205]. Ensemble learning is done by generating sets of weak learners and combining them to assemble one strong learner to decrease the variance (bagging) [29] and the bias (boosting) while improving the predictions (stacking). Bagging consists of selecting a predefined number of weak learners which are all independently trained on randomly chosen subsets of the training data. By making each weak learner different,

the correlation is decreased which improves the robustness and generalization [55]. Boosting on the other hand is a method that improves the performance of weak learners by training the weak learner several times. This is done by adding more decision trees in addition to the previous predictions [205]. Stacking involves training several learners and determining which of the weak learners are reliable and which are not [252].

The Random Forest (RF) classifier is an ensemble bagging approach [32]. The RF classifier trains several decision trees (weak learners) on randomly chosen subsets of the training data. These are then combined into strong trees (learners). The general algorithm can be explained as follows. Given a training dataset $X = x_1, ..., x_n$ and the corresponding labels $Y = y_1, ..., y_n$, bagging is applied B times by randomly selecting a subset of X and the corresponding Y as X_b, Y_b with replacement to train an individual decision tree f_b. Selecting a subset with replacement involves sampling the training data while keeping its size (e.g. the original training data [1,2,3,4,5,6] can be sampled as [1,1,2,2,4,5]). After each individual decision tree f_b has been trained with X_b, Y_b, the samples x' that were not used for training that individual decision tree is predicated by averaging the predictions of all the individual decision trees with x' as:

$$\hat{f}_b = \frac{1}{B} \sum_{b=1}^{B} f_b(x') \tag{4.8}$$

Each decision tree f_b can be further split and trained with a randomly selected subset from X, Y, called feature bagging. This is to ensure that the decision trees are not further trained on the corresponding X_b, Y_b since strong features would otherwise be over-represented in the trees. If all trees are using strong features, they are all correlated. Hence, by randomly choosing subsets from X, Y, the correlation between the trees is reduced.

To initiate the RF classifier, certain parameters can be defined. This includes the number of decision trees B, the maximum depth of a decision tree f_b and the number of features to consider for splits.

Histogram-based Gradient Boosting Classification Tree

Gradient decision tree boosting is likewise an ensemble approach that generates an additive model by using decision trees that are trained in a sequence [77]. Decision trees are added after each iteration in a forward-wise manner to the additive model to improve its predictions. The additive model consists of weak learners while new learners are added to compensate for the weaknesses, here as gradients. This iterative manner creates a strong learner which reduces the bias and variance [30]. The additive model is fitted (minimizing) with residual errors.

The general algorithm can be explained as follows. Given are a training dataset $X = x_1, ..., x_n$, the corresponding labels $Y = y_1, ..., y_n$, a chosen loss function $L(y(f(x)))$ and a model $F(x)$ that should be minimized with respect to the squared error. A decision tree $h(x_1)$ is added to the model $F(x)$ as:

$$F(x_1) + h(x_1) = y_1, ..., F(x_n) + h(x_n) = y_n \tag{4.9}$$

The additive model can be fitted where h is used to compensate the weakness of $F(x)$ as:

$$(x_1, y_1 - F(x_1)), ..., (x_n, y_n - F(x_n)) \tag{4.10}$$

For the minimized loss function J, the negative gradient $-g(x_i)$ can be used. This is calculated as:

$$J = \sum_i L(y_i(f(x_i))) \tag{4.11}$$

$$-g(x_i) = y_i - F(x_i) = -\frac{\delta J}{\delta F(x_i)} \tag{4.12}$$

As gradient boosting is an iterative process, the initial model $F(x) = \frac{\sum_{i=1}^{n} y_i}{n}$ is iterated until convergence when the decision tree h fits the negative gradients $-g(x_i)$. To determine the decision trees, the splits can e.g. be pre-sorted based on the optimal split points [162, 217] or be based on histograms by determining the potential gain and splits therefore the samples into histogram bins during the training of the classifier [116].

The classifier Histogram-based Gradient Boosting Classification Tree (HGB) is an approach based on histogram splits as it splits the data into integer-valued bins during training and reduces therefore the number of splitting points. Parameters that can be chosen are e.g. the maximum tree depth of each decision tree, the loss function and the maximum number of iterations. The HGB classifier is an experimental approach available from the toolbox *scikit-learn* [185]. It is based on the classifier LightGBM which splits its decision trees on a histogram basis by establishing the potential gain [116]. It splits its trees on a best-first basis instead of a level-wise basis. To initiate the classifier HGB, several parameters can be selected to further modify the model. This includes the loss function $L(y(f(x)))$, the shrinkage (a learning rate that determines how the weights are adjusted concerning minimizing the loss function $L(y(f(x))))$, the maximum number of iterations and the maximum depth of a decision tree h.

k-means Clustering

The only unsupervised classifier used in this thesis is the k-means clustering approach [148, 152]. k-means clusters data by discovering natural clusters of a dataset which is defined as finding k clusters which are "based on a measure of similarity such that the similarities between objects in the same group are high while the similarities between objects in different groups are low" [111]. Given a set of observations $x_1, ..., x_n$, k-means clusters iteratively n observations into k clusters (i.e. groups/classes) where k is pre-defined by the user. Each observation x_n is assigned to the cluster k with the nearest cluster center (centroid) μ_k. A chosen distance metric, e.g. the squared Euclidean distance or sum of the absolute differences, determines which cluster center μ_k is nearest to each observation $x_1, ..., x_n$. If the distance metric is the Euclidean distance, k-means can be determined as follows, where c_k is the number of observations in cluster k:

$$J(\mu_k) = \sum_{j=1}^{k} \sum_{i=1}^{c_k} \|x_i - \mu_j\|^2 \tag{4.13}$$

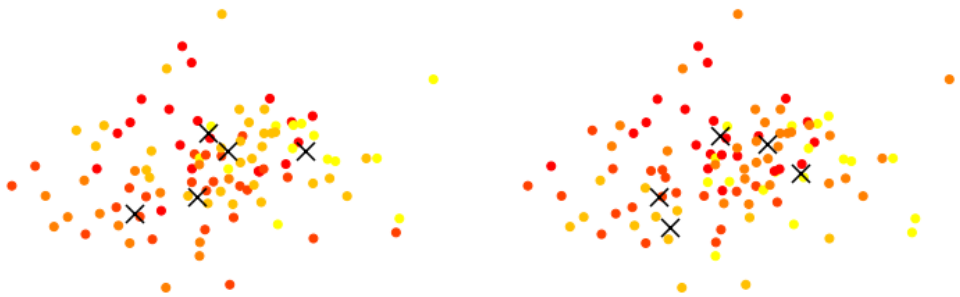
A chosen cluster initialization determines how the first clusters k are formed. The initial clusters k can be e.g. randomly chosen, based on the distance, based on the variance or through seeding [13, 38]. Depending on the chosen cluster initialization, the final clusters can differ as they converge to the local minima [111]. After initializing the first clusters k, the aforementioned equation is used to calculate the distance between each cluster center μ_k and observation x_n. Each observation x_n is assigned to the nearest cluster center μ_k (with the minimized distance). New clusters are determined as:

$$\mu_k = \frac{1}{c_k} \sum_{i=1}^{c_k} x_i \tag{4.14}$$

The distance is recalculated between each cluster center μ_k and observation x_n until the iteration has reached its maximum number of iterations. Figure 4.2 illustrates five cluster center with k-means clustering where the cluster initialization is either randomly chosen or center seeding (i.e. k-means++). The cluster initialization k-means++ uses center seeds [13] by firstly randomly selecting a uniform observation from x_n and initiates it as the first cluster center μ_1. The distances between the observations x_n and μ_1 are determined, denoted as $d(x_n, \mu_j)$. The next cluster center μ_2 is randomly selected with a probability as:

$$\frac{d^2(x_n, \mu_1)}{\sum_{j=1}^{n} d^2(x_j, \mu_1)} \tag{4.15}$$

This is repeated until all clusters k have been assigned a cluster center μ_k.



(a) Random initialization. (b) k-means++ initialization

Figure 4.2: An illustration of k-means cluster centers from a randomly generated dataset where two different cluster initializations are used. (a): Random initialization. (b): k-means++ initialization. The clusters in (a) and (b) are not the same as the five cluster centers (marked with X) are located at different locations.

4.2 Dimensionality Reduction and Feature Selection Methods

As discussed in Chapter 2.3, hyperspectral data is high-dimensional due to its number of spectral bands. To counter the previously explained phenomena "curse of dimensionality" [18] and Hughes phenomenon [104] which are reoccurring challenges with hyperspectral data, approaches that perform dimensionality reduction and feature selection are here presented.

Dimensionality reduction can be performed using algorithms such as Principle Component Analysis (PCA) [184], Independent Component Analysis (ICA) [52] that separates multivariate signals into sub-components, non-negative matrix factorization which factorizes the data into easier interpretative vectors [140], partial least-squares regression that determines a linear regression model of the variables and transforms those into a new space [79] or t-Distributed Stochastic Neighbor Embedding (t-SNE) [234]. Such algorithms transform the high-dimensional data into a new space of a lower dimensionality where the original data is no longer recognizable. Feature selection methods on the other hand extract the subsets with the most relevant and important features. This can include the algorithms Correlation-based Feature Selection (CFS) [91], SPEC which is a filter model that determines the best eigenvectors in a system [267], Fast Correlation based-filter (FCBF) that measures the feature class correlation using a chosen threshold [262] or AE [126]. The selected features can likewise be transformed into another space, depending on the approach. These approaches are not exclusively used for high-dimensional data, e.g. the feature selection approach AE can be used for extracting features in imagery as well [59, 128, 270].

Principal Component Analysis

Aa statistical approach used to reduce the dimensionality of a dataset can be performed with PCA [184]. It uses an orthogonal transformation to convert possibly correlated observed data into a new representation relying on a set of linearly uncorrelated variables, which are referred to as principal components. These components are organized such that the first principal component covers the highest variability of the data (it contains as much of the data variation as possible) and that each following component covers the highest possible variability under the constraint that it is orthogonal to all previous components. The most relevant information of the observed data is contained by the first few principal components. The PCA encoding consists of the first few principal components that contain the most relevant information since the assumption is that no significant loss of information is discarded in the remaining principal components.

The general approach is the following. Given a set of normalized observations $X = x_1, ..., x_n$ (where $X \in 0, 1$) with a dimensionality of m, the average value $\overline{x}_{nm}$ is calculated across each dimension m. The covariance c_{nm} is then determined for each x_n across its dimension m as:

$$c_{nm} = \frac{1}{n-1} \sum_{n=1}^{m} \left(x_{nm} - \overline{x}_n \right) \left(x_{nm} - \overline{x}_m \right) \tag{4.16}$$

Based on the covariance c_{nm}, the eigenvalues λ and the corresponding eigenvectors v can be determined. By ranking the eigenvectors v based on the corresponding eigenvalues λ (highest to lowest), the first principal component is the eigenvector with the highest eigenvalue λ. Through the ranking, it is possible to retrieve the components in the order of relevance and likewise retrieve the k components which correspond to the wanted reduced dimensionality k. To reduce the dimensionality of the original observation X, the eigenvectors v_k that correspond to the k principal components are retrieved and the original observation X is transformed as:

$$X' = v_k^T X \tag{4.17}$$

The probabilistic PCA uses a maximum-likelihood approach and is based on a Gaussian latent variable model [230]. The probabilistic PCA is computed by representing each x_n as latent variables z_n with a lower dimension than m. The distribution of z_n is $z_n \in WW^T, \sigma^2 I$ where W is the weight matrix and σ^2 the noise variance.

t-distributed Stochastic Neighbor Embedding

t-SNE is a non-linear approach used for dimensionality reduction of high-dimensional data [234]. It can likewise be used for visualization of high-dimensional data as it transforms the data into a two- or three-dimensional space. The algorithm consists of minimizing the divergence from a high-dimensional space and to a low-dimensional space, where the distribution which measures a corresponding pairwise similarity between the two spaces should be minimized.

Firstly, to determine the similarity of the pair in the high-dimensional space, a distance measure is calculated, such as the Euclidean distance. The pairwise similarity is determined using "the conditional probability $p_{j|i}$ that point x_i would select point x_j as its neighbor if neighbors were picked in proportion to their probability density under a Gaussian centered at x_i" [234]. This is calculated as follows, where σ_i is the Gaussian kernel, N the set of high-dimensional objects $x_i, ..., x_n$ and the conditional probability $p_{i|i} = 0$ (since the focus is on pairwise similarities):

$$p_{j|i} = \frac{exp(-d(x_i, x_j)^2/2\sigma_i^2)}{\sum_{k \neq i} exp(-d(x_i, x_k)^2/2\sigma_i^2)} \tag{4.18}$$

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N} \tag{4.19}$$

The conditional probability $p_{j|i}$ is relatively high for regions with densely located data points and small for regions with widely separated points due to the variation of the Gaussian kernel σ_i. The algorithm tries to learn an embedding to represent each object in the two- or three-dimensional space that reflects the conditional probability $p_{j|i}$ as well as possible. The conditional probability $q_{j|i}$ is likewise determined in the low-dimensional space with the low-dimensionally located objects $y_i, ..., y_n$ and $q_{i|i} = 0$:

$$q_{j|i} = \frac{exp(-||y_i - y_j||^2)}{\sum_{k \neq i} exp(-||y_i - y_k||^2)} \tag{4.20}$$

The location of the objects $y_i, ..., y_n$ in the low-dimensional space is determined by minimizing the Kullback-Leibler (KL) divergence across all objects using a gradient descent method. The KL divergence is a measure that describes the difference between one probability distribution from a second probability distribution [130]. Due to its asymmetry, the measure preserves the data structure in the low-dimensional space of the high-dimensional space. The divergence C can be expressed as the following equation, where P_i and Q_i represent the joint probability distribution:

$$C = KL(P_i||Q_i) = \sum_{i \neq j} p_{j|i} log \frac{p_{i|j}}{q_{i|j}} \tag{4.21}$$

The divergence is finally minimized in the low-dimensional space by using:

$$\frac{\delta C}{\delta y_i} = 2 \sum_j (p_{j|i} - q_{j|i} + p_{i|j} + q_{i|j})(y_i - y_j) \tag{4.22}$$

This results in an optimized low-dimensional space (either two- or three-dimensional) that maps the high-dimensional data while keeping the pairwise similarity.

Correlation-based Feature Selection

CFS is an approach for feature selection which accounts for the correlation between features (samples) and the corresponding classes by following the statement "A good feature subset is one that contains features highly correlated (predictive of) the class, yet uncorrelated with (not predictive of) each other" [91]. A feature is defined as redundant if it is highly correlated with one or more other features. CFS bases its approach on reducing the feature redundancy by evaluating subsets of features to determine the subset with the least redundancy correlated according to a heuristic evaluation function. For each feature subset S that contains k features, a heuristic merit M_S is calculated which is used to rank all possible feature subsets S to determine the one with the least redundancy. To calculate the merit M_S, the average feature-class correlation $\overline{r_{cf}}$ ($f \in S$) and the average feature-feature correlation $\overline{r_{ff}}$ need to firstly be defined. There are three variations of CFS and these are characterized by which correlation estimation they use. $\overline{r_{cf}}$ and $\overline{r_{ff}}$ are therefore determined with either a symmetrical uncertainty [192], a normalized symmetrical minimum description length [31] or a symmetrical relief [201]. The heuristic merit M_S is calculated with the following equation:

$$M_S = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{ff}}}}. \tag{4.23}$$

A suitable feature subset S is derived iteratively by selecting the M_S with the highest merit by maximizing M_S over the set of all possible feature subsets S (the highest-ranked M_S). The iteration can be done with three types of search strategies, either forward selection (adds a feature k to the feature subset S), backward elimination (removes a feature k from the feature subset S) or best first (starts with either all features or no features). The strategy of best first consists of either searching forward or backward by either adding or removing single features with a stopping criterion to hinder the search throughout the entire feature subset. The best first search is performed until the merit M_S converges to a stable value and five consecutive subsets S have not improved the current M_S. The suitable feature subset S is thus used to reduce the dimensionality of the data.

Autoencoder

AE is a data compression method that can be used to learn important features in an unsupervised manner using artificial neural networks [126]. An AE learns a compressed representation (also referred to as latent representation) of a dataset through an encoder by reducing its dimensionality while extracting the most important features. The compressed representation is reconstructed using a decoder where the AE attempts to restore the compressed representation into a reconstruction similar to the original form. To

compress and reconstruct the dataset, connected hidden layers, which describe how the data should be compressed and reconstructed, are used (see Figure 4.3 for the general AE architecture).

The input layer (the input data) is compressed using the encoder by extracting the most important features according to applied hidden layers. The intermediate outcome is the compressed representation, a dimensionality-reduced representation of the input layer. The size of the compressed representation depends on the architecture, the type of input data and the design of the hidden layers. To reconstruct the compressed representation, the hidden layers of the decoder produce an output layer with the same size as the original input layer. AEs are lossy and the reconstructed data is therefore degraded compared to the original data [84]. An AE is designed and trained for specific sets of data and is therefore only able to compress and reconstruct datasets comparable to what it was learned for. Thus, if it is learned to compress and reconstruct images of animals, it would perform poorly on images of cars [84]. An AE requires both training and testing data to assess the quality of the reconstruction. This is assessed with a loss function (e.g. mean squared error).

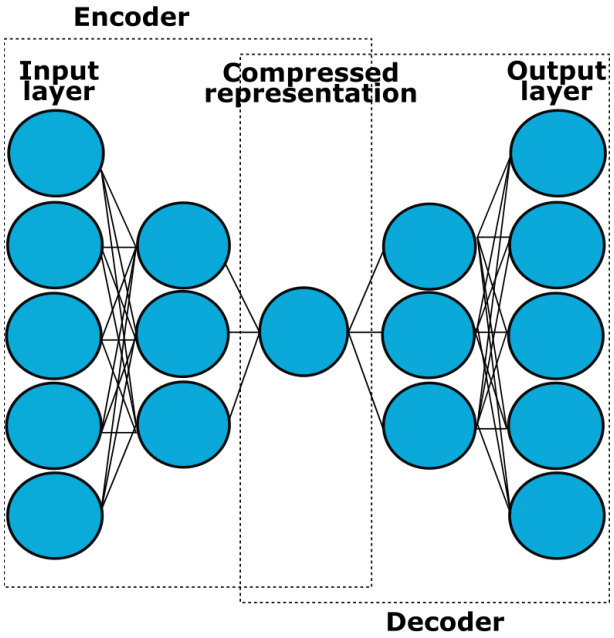


Figure 4.3: The general architecture of an AE. An AE consists of an encoder that compresses the input data through the hidden layer into a compressed representation. The decoder decompresses the compressed representation with hidden layers. The design of the hidden layers is based on the intended application and the type of input data.

There are various kinds of AEs that utilize different approaches to extract features. A sparse AE [173] learns the compressed representation through sparsity by using a few hidden layers and responds to the unique statistical features instead of the feature importance [84]. A variational AE [64] behaves similarly to Generative Adversarial Network (GAN) [85]. A GAN is a machine learning approach that consists of two neural networks that are competing against each other to train and generate new data. This new data consists of the same statistics as the initial training dataset as it aims to appear as real data and not generated data. Similarly, a variational AE generates new generative data by ensuring that the encoder is regularized to avoid overfitting and that the compressed representation consists of suitable distributed properties which enable the generative process (the casual relationship). A denoising AE uti-

lizes corrupted and uncorrupted data to learn how to denoise data, such as corrupted signals [238, 239]. A denoising AE is trained on both the corrupted and uncorrupted data. The uncorrupted data receives intentionally randomly added noise, making it likewise corrupted. By adding this noise, it forces the denoising AE to extract the most important features during the encoding which makes the model more robust [84, 239]. A deep AE [84, 239] has several connected hidden layers in the encoder and decoder. The output from one hidden layer is the input of the next consecutive hidden layer. Thus, the hidden layers are connected and learn either the compression or the reconstruction from each other. Additionally, a Convolutional Autoencoder (CAE) [157] utilizes convolutional layers (in a similar manner as a Convolutional Neural Network (CNN)) as hidden layers. A convolutional layer has parameters consisting of kernels or filters that learn high-level features. They convolve across the input data and generate during the forward pass dot products between the entries of the filter or kernels and the input data for every position. As they convolve across the input data, an activation map is generated which corresponds to the computed dot product at each position. These activation maps can be forwarded to the next hidden layer. Different kinds of AE can be combined, e.g. deep and convolutional [80, 103], sparse and denoising [137, 226] or deep and sparse [257, 266].

The compressed representation can be used as an input for classification tasks as it contains the relevant features, e.g. land cover classification [142] or SAR image classification [80]. As the compressed representation is dimensionality-reduced, it can be used for high-dimensional data such as hyperspectral data [137, 145, 269] or for extracting important features from images [59, 257].

4.3 Approaches for Texture Analysis

As stated in Chapter 3.1.3, texture can be described in various ways and can explain the properties of the 3D shape of an object. These properties can describe the characteristic of a pattern or sub-pattern based on e.g. roughness, smoothness, uniformity or randomness [158]. To express the textural properties of an object, texture analysis can be used. Texture analysis can, as stated in Chapter 3.1.3, be used for different applications, such as image classification [42, 242] and image segmentation [99, 155]. The approaches for texture analysis are categorized into four groups: structural methods, statistical methods, model-based methods and transform-based methods [12, 20, 158, 254].

Structural methods describe the texture as a set of defined primitives or elements (on a micro-level) with a hierarchical spatial arrangement (on a macro-level), such as regularly spaced lines that are located parallel to each other. The texture is defined by set rules that define the attributes and placement of the elements. Structural methods are limited for the description of texture as they can only describe regular textures [20]. Nonetheless, the mathematical morphological operators erosion, dilation, opening and closing are considered as powerful tools [93, 158]. These were originally developed for binary image classification, i.e. the image contained only pixel values as $\in 0, 1$. Given an image A and a structuring element B (a shape with a chosen size that is used for interaction with A, e.g. a square with the size of $3x3$), the operator erosion uses B for interaction with A and reduces the shape of the objects in A. The operator dilation uses B for interaction with A and expands the shape of the objects in A. The operator opening performs first erosion and then dilation while the operator closing performs first dilation and then erosion.

Statistical methods describe the texture as a spatial pixel gray value distribution in an image. First-order approaches only account for the pixels' gray values and not the spatial relationship. This generates simple statistics that can be visualized as histograms. Entropy is a first-order approach which is a measure of information content [92]. It expresses the randomness of an intensity distribution through the following equation, where G is the number of intensity levels in the image, z_i the random pixels with a certain intensity value and $p(z_i)$ the number of pixels that have a gray value equal to z_i:

$$Entropy = -\sum_{i=0}^{G-1} p(z_i) log_2 p(z_i) \tag{4.24}$$

Second-order approaches account for the spatial relationship as well where the gray-level co-occurrence matrix (GLCM) is among one of the most famous approaches [92]. Given a displacement vector, the GLCM allows the deriving of statistical textural features of neighboring pixels, such as contrast (the local variation) and correlation (the joint probability occurrence of a specified combination).

Model-based methods express the texture as an empirical (fractal or stochastic) model that describes the image's synthesis (pixel structure). The model's parameters that describe the qualitative attributes of the texture are calculated and used as texture features for further analysis. Such methods include Markov random fields [56], autoregressive models [156] and fractal models [117].

Transform-based methods, also known as filter-based methods, apply filter banks on images to transform them into a space where new features are distinguishable. The applied filter calculates energy based on the filter response from the frequency or/and spatial domain. This includes the Fourier [27], Gabor [246] and wavelet [11] transformations. Lines, edges or dots can be detected in the spatial domain by using gradient filters by measuring the filter density response [254]. This includes the Canny edge detection approach [36] which firstly applies a Gaussian filter to smooth the image (reducing noise) and detects the intensity gradients using the following equations. G_x is the first derivative in the horizontal direction, G_y is the first derivative in the vertical direction and θ is the angle which is rounded to represent the four angles $0°, 45°, 90°, 135°$:

$$G = \sqrt{G_x^2 + G_y^2} \tag{4.25}$$

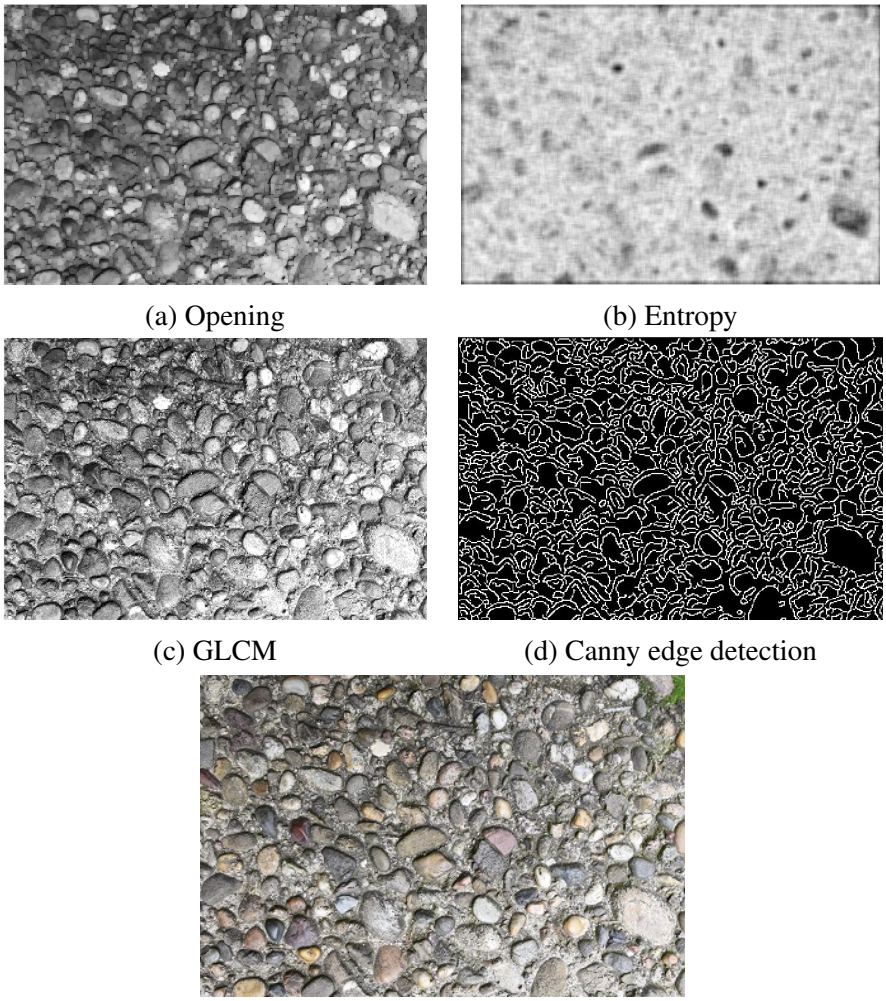
$$\theta = tan^{-1} \frac{G_y}{G_x} \tag{4.26}$$

This is followed by the elimination of any noise by comparing the edge strength of one pixel to the neighboring gradient direction (called non-maximum suppression). Then, a double threshold is applied to detect potential edges. The edges are determined by eliminating either individual weak edges or edges that are not connected to strong edges.

Another transform-based method is the Maximum Response 8 filter bank (MR8) and consists of 38 invariant filters and eight filter responses [237]. MR8 is derived by extracting the maximum responses across all orientations which results in eight filter responses. The filter bank consists of one Gaussian filter and one Laplacian of a Gaussian filter, both with a filter size σ of 10 pixels. Additionally, the bank contains an edge filter (first derivative) and a bar filter (second derivative). These two filters are

applied at three different scales with six different orientations $(\sigma_x, \sigma_y) = (1,3),(2,6),(4,12)$. The eight filters correspond to the direct response of the two isotropic filters (Gaussian and Laplacian) and the response of six maximum filters across all orientations from the oriented filters (edge and bar). These are derived by merging the responses at each scale by determining the maximum response.

Figure 4.4 illustrates the structural method operator opening (a square with the size $3x3$), the statistical methods entropy and GLCM (with the offsets $[0,1;-1,1;-1,0;-1,-1]$) in addition to the transform-based method Canny edge detection.



(a) Opening (b) Entropy

(c) GLCM (d) Canny edge detection

(e) Original image

Figure 4.4: (a-d): Different texture analysis approaches performed with the image in (e).

4.4 Case Study – Assessing the Approaches and Tools with Hyperspectral Data

This sub-chapter presents a case study that assessed a selection of the aforementioned classifiers and feature selection tools using two different datasets. To improve the readability within the context of this thesis, many paragraphs have been rephrased. Most content of this chapter has been published in the article:

R. Ilehag, M. Weinmann, A. Schenk, S. Keller, B. Jutzi and S. Hinz. Revisiting Existing Classification Approaches for Building Materials Based on Hyperspectral Data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII(3), pp. 65 - 71, 2017

To assess and evaluate the different approaches available for material classification with high-dimensional data, it is necessary to perform a comprehensive analysis. This sub-chapter utilizes therefore two urban materials datasets, the spectral library Santa Barbara [102] and the *in-situ* acquired dataset URBAN-SAT. Firstly, both datasets are individually used to compare and evaluate the performance of two supervised classifiers represented by RF [32] and SVM [253]. Secondly, they are used to assess the added value of the dimensionality reduction approach PCA [184, 230] and the feature selection CFS [91]. Lastly, the final assessment consists of limiting the classification of urban materials to certain spectral bands. The selected datasets, the proposed methodology and the retrieved results are described in the following sub-chapters.

4.4.1 Datasets

The first dataset is the spectral library Santa Barbara [102]. As explained in Chapter 3.1, the Santa Barbara spectral library (acquired in Santa Barbara, CA, US) contains close to 1000 spectra in the spectral range of VNIR and SWIR. The Santa Barbara spectral library contains 26 labeled land cover and land-use classes, such as various roof and road materials. The selected classes used for the assessment contain a sufficient amount of samples and exist in the built-up area (mostly ground and roof materials).

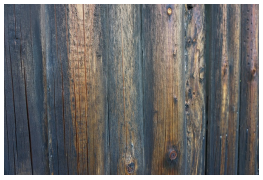
The second dataset, referred to as URBAN-SAT, was acquired *in-situ* around the city of Karlsruhe, Germany. It was acquired in August of 2017 with sunny conditions and with a few clouds passing by. The dataset was acquired *in-situ* using the spectroradiometer ASD FieldSpec 4 Hi-Res[2]. The spectroradiometer acquires spectra in the spectral range of 350 – 2500 nm and consists of three sensors: a VNIR sensor and two SWIR sensors. The VNIR sensor acquires spectra in the range of 350 – 1000 nm while the two SWIR sensors acquire spectra in the range of 1001 – 1800 nm and 1801 - 2500 nm respectively. The two SWIR sensors will here on be referred to as FieldSpec-4 SWIR instrument (1001-1800 nm) (SWIR1) and FieldSpec-4 SWIR instrument (1801-2500 nm) (SWIR2) respectively. FieldSpec-4 has a spectral sampling of 1.4 nm and a spectral resolution of 3 nm in the spectral range of the VNIR sensor. In the spectral range of the two sensors SWIR1 and SWIR2, the spectral sampling is 1.1 nm and the spectral resolution 8 nm. There are in total 2151 channels which have a wavelength accuracy of 0.5 nm. The optical fiber probe has a field of view of 25°.

[2] www.malvernpanalytical.com

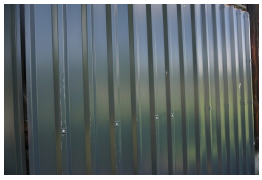
To reduce the likelihood of acquiring a sample that contains several urban materials, the acquisition distance was kept small. The spectra were collected from a distance of 20 cm from the targeted sample with a coverage area of around 9 cm in diameter. For each sample, 10 consecutively acquired spectra L_t were collected within one minute. Such a set will from hereon be referred to as an intra-set L. The measurement setup remained the same throughout the data acquisition. The spectroradiometer was firstly radiometrically calibrated using a 95% reflectance plate (representing total reflectance) and a black environment that the sensor achieves by closing its optical fiber probe (representing a black body) for correction of the incoming solar irradiance $S(\lambda)$. The reference spectrum E was acquired before the acquisition of an intra-set L. Both were acquired under the same conditions by facing the reference plate in the same direction as the sample surface. Thus, the solar incident angle θ was kept the same for the reference spectrum E and the intra-set L.

URBAN-SAT consists of several urban materials but mainly building materials. 12 material classes were acquired: *Asphalt*, *Brick*, *Concrete*, *Fibre cement*, *Glass*, *Granite*, *Metal (painted/coated or plain)*, *Plaster*, *Roofing felt*, *Roof tile (ceramic)*, *Sandstone* and *Wood (painted)*. Each material sample was acquired as one intra-set L, comprising of 470 spectra in total. Selected samples of URBAN-SAT can be seen in Figure 4.5. Neither the surface structure, the coating, the texture nor the color was considered while labeling the samples. For instance, the samples shown in 4.5(a) and 4.5(b) are labeled as the same material class *Wood* even if the surface coating is different.

(a) *Wood (painted)*

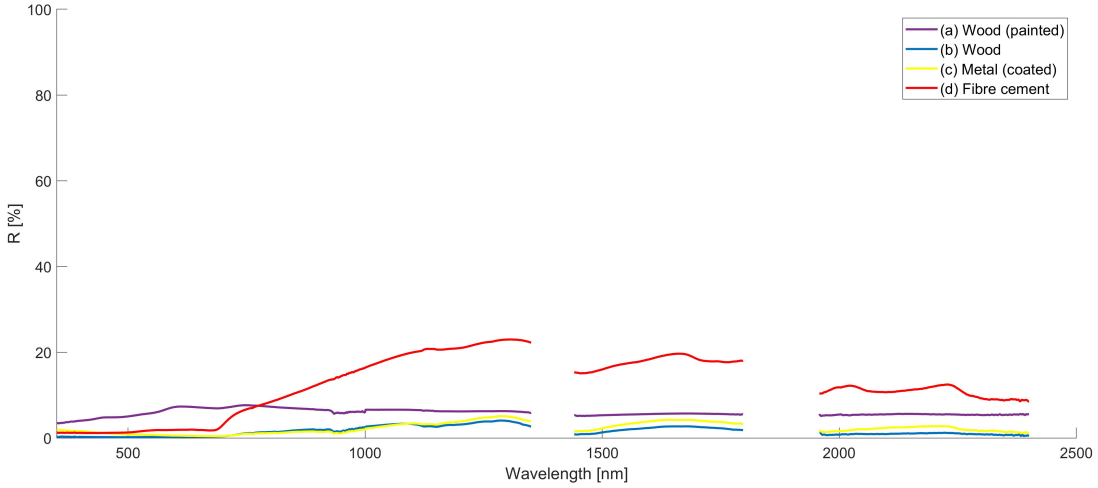


(b) *Wood*



(c) *Metal (coated)*

(d) *Fibre cement*



(e) Spectra of samples (a-d)

Figure 4.5: (a-d): Imagery of selected samples from the dataset URBAN-SAT. (e): The corresponding average spectral reflectance R of the selected samples. The spectral gaps are due to the post-processing which are described later in this chapter.

To compare the two datasets, selected classes that contain a sufficient amount of samples are chosen for the assessment. These are listed below and consist of 21 URBAN - SAT samples (corresponding to 210 spectra) and 44 samples from Santa Barbara samples (corresponding to 440 spectra). Highlighted

are the urban materials which are present in both datasets, enabling the possibility of comparing the spectral characteristic features. The two datasets are using different taxonomies. The Santa Barbara spectral library focuses on land cover and land-use classes whereas URBAN-SAT focuses on only urban materials and adds, if necessary, additional information about the sample's appearance.

Santa Barbara :

- *Asphalt (road)*
- *Composite (roof)*
- ***Concrete (roof)***
- *Gravel (road)*
- ***Metal (roof)***
- *Tar (roof)*
- *Tile (roof)*
- ***Wood (roof)***

URBAN-SAT :

- ***Concrete***
- *Fibre cement*
- *Glass*
- ***Metal (painted/coated or plain)***
- ***Wood (painted)***

4.4.2 Proposed Method

To assess the classification capability using the spectral features from the two datasets, three steps are conducted. Firstly, some minor post-processing procedures are performed on the two datasets. Secondly, the spectral features are expressed as either CFS encodings or PCA encodings. Lastly, classification is performed with the classifiers RF or SVM.

As the URBAN-SAT dataset was acquired *in-situ* in an uncontrolled environment, some measures are needed. The pre-processing consists firstly of removing the spectral bands that are sensitive to noise. The spectral range of 1351-1440 nm, 1798-1957 nm and >2400 nm are removed as they are affected by the water vapor absorption [81, 228]. To keep the spectral ranges consistent, these spectral bands are removed for both datasets. Secondly, the spectral reflectance R is calculated for every measured spectrum L_t in an intra-set L for the URBAN-SAT dataset. This is done by using the acquired reference spectrum E with the following equation:

$$R = \frac{L_t}{E} \tag{4.27}$$

Lastly, to determine which sub-bands are sufficient for classification of urban materials, the following four sub-bands are selected for the two datasets:

- Full spectral range: 400 - 2400 nm
- Visible range: 400 - 700 nm
- VNIR range: 450 - 900 nm
- SWIR range: 1000 - 2400 nm

The full spectral range is chosen as the reference and baseline for the assessment of the other selected sub-bands. The visible spectral range is chosen since it corresponds to the spectral range of RGB. This spectral range can only contribute with information about the color of the material samples which

is insufficient for urban material classification [139, 207]. The VNIR spectral range is chosen since it represents the spectral range of selected light-weighted hyperspectral sensors for UAV applications, such as Cubert Firefly S185[3], HSC-2 Hyperspectral Camera[4] and Nano-Hyperspec[5]. Additionally, this spectral range is commonly used for studies related to vegetation, water and land cover classification [143, 251]. If this spectral range is suitable for urban material classification as well, one hyperspectral sensor would be sufficient for both urban material and vegetation/water classification. The SWIR spectral range is chosen since it contributes the most for urban material classification unless the LWIR spectral range is available [124, 139].

Following the minor post-processing measures and limiting the spectral range of four sub-bands, PCA and CFS are employed onto the two datasets. The standard PCA is used to reduce the dimensionality by acquiring the first principal components that cover 99.99% of the variability of the data [184]. The relevant features are extracted using the CFS implementation provided by [268]. Finally, the two supervised classifiers, RF [32] and SVM [40, 53, 253] are employed where the RF implementation by [144] is used. The classification parameters are kept simple. The maximum depth of RF is set to five and SVM is set to use the Gaussian radial basis function as the kernel function. 10 randomly selected samples per class are retrieved to obtain a training set for both classifiers. The remaining samples are used for testing. 76% (160 spectra) of the URBAN-SAT spectra are used for training and 24% (50 spectra) for testing, while 82% (364 spectra) of the Santa Barbara spectra are used for training and 18% (80 spectra) for testing.

The workflow for the processing of the URBAN-SAT dataset can be seen in Figure 4.6. A detailed overview that displays the sub-band selection, the derivation of the spectral expressions and the classification can be seen in Figure 4.7.

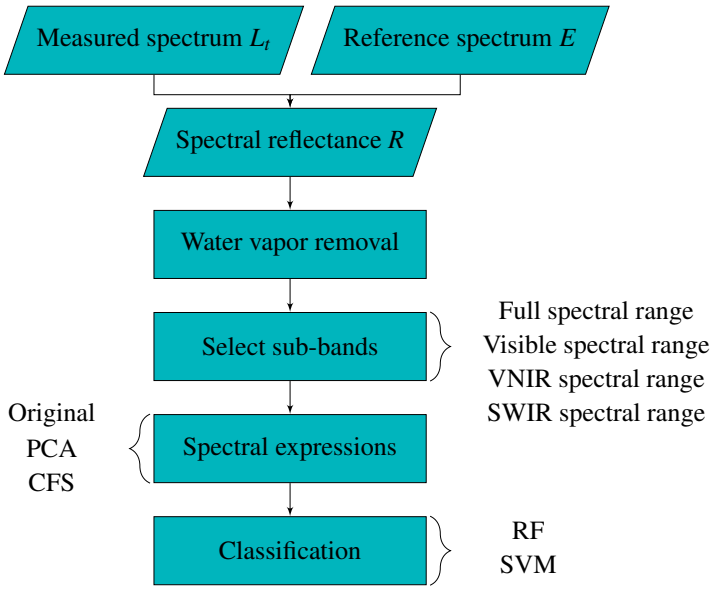


Figure 4.6: The proposed workflow for the processing of the URBAN-SAT dataset. The sub-band selection, the feature selection and the classification can be seen in Figure 4.7. The processing steps are likewise applied to the Santa Barbara dataset.

[3] www.cubert-gmbh.com
[4] www.senop.fi
[5] www.headwallphotonics.com

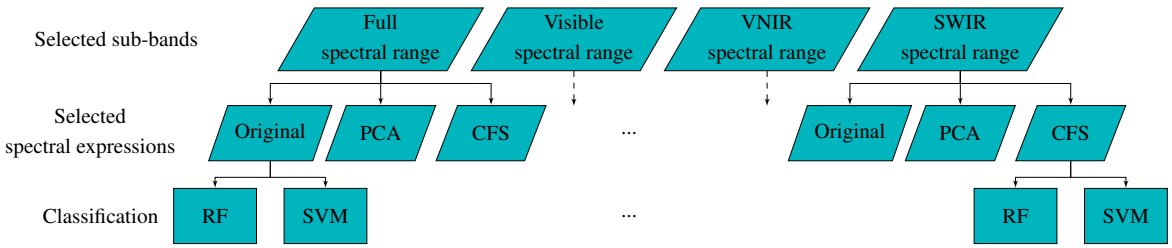


Figure 4.7: A detailed workflow of the processing steps involving sub-band selection, feature selection and classification. The workflow is here simplified to illustrate the workflow of only four scenarios (e.g. the original spectral features of the full spectral range are used as input to the classifiers RF and SVM).

In the end, 48 different scenarios are generated from the two datasets, consisting of four different sub-band selections, three spectral expressions and two classifiers. To evaluate the performance of the 48 scenarios, an evaluation matrix is provided that contains the overall accuracy O, the κ-value, the average recall $\overline{R}$, the average precision $\overline{P}$ and the average $\overline{F_1}$-score across all classes.

4.4.3 Results and Discussion

The classification results for the 48 scenarios can be seen in Table 4.3. The RF classifier using the PCA encoding in the full spectral range achieves the highest overall accuracy O across the 48 scenarios, regardless of the used dataset. This is further supported by the remaining evaluation metrics.

By comparing the two classifiers regardless of the spectral expression, it is apparent that the RF classifier outperforms the SVM classifier based on the evaluation metrics. The two spectral expressions from the encodings of PCA and CFS are compared with the baseline (the original spectrum). It can be observed that the PCA encoding achieves the highest evaluation scores. The evaluation scores further indicate that it is preferable to use the baseline than the CFS encoding. This indicates that the feature selection approach CFS has retrieved features less relevant for urban material classification since those are not distinct.

The evaluation scores indicate that the full spectral range is preferable for the material distinction, followed by the spectral range of SWIR. The spectral ranges of visible and VNIR achieve the lowest scores, indicating that the visible range is inferior. Materials are not bound to a specific color and by only utilizing this spectral range, the spectral features provide mostly only information about the material color. Nonetheless, the overall accuracy O is still above 80% which implies that materials in the datasets hcan be distinguished based on their color. The color of the material classes in the datasets is likely consistent, e.g. *Asphalt* is consistent grey and *Wood* consistent brown. Hence, for such datasets, it might not be necessary to utilize the SWIR spectral range.

Figure 4.8 displays a spectral comparison between the three material classes that were present in both datasets: *Metal*, *Wood* and *Concrete*. As observed, the spectral reflectance R differs, even if the samples have been labeled as the same material class. This can be due to how the materials have been defined through the taxonomy or due to difference chemical compounds. While both datasets contain the material class *Metal*, the spectral dissimilarity is likely caused by different chemical compounds, as *Metal* is a high-level description and could be e.g. *Aluminum*, *Iron* or *Steel*. As the Santa Barbara spectral library

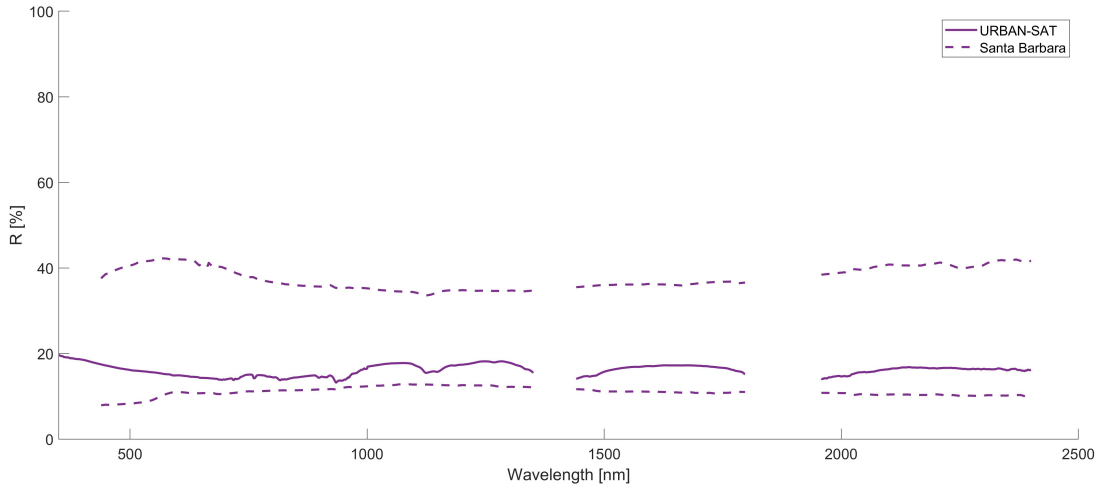
Table 4.3: The achieved classification results for the two datasets. The evaluation scores indicate that the best scores are retrieved with PCA encoding with the classifier RF. The higher evaluation scores are achieved when the full spectral range is used, shortly followed by the spectral range of SWIR. Highlighted are the best scores of each sub-band and evaluation score pair.

		Full spectral range			Visible spectral range			VNIR spectral range			SWIR spectral range		
		orig	PCA	CFS	orig	PCA	CFS	orig	PCA	CFS	orig	PCA	CFS
O [%]	$\text{URBANSAT}_{\text{RF}}$	85.0	87.5	69.8	71.9	**81.3**	63.1	80.6	**83.8**	65.0	76.9	75.6	71.3
	$\text{URBANSAT}_{\text{SVM}}$	77.5	76.9	71.3	68.1	68.1	60.0	70.6	69.4	64.4	66.9	65.0	65.6
	$\text{SantaBarbara}_{\text{RF}}$	86.0	**90.4**	85.4	70.3	**81.6**	67.6	70.6	81.9	71.4	69.0	**85.2**	62.9
	$\text{SantaBarbara}_{\text{SVM}}$	82.4	82.1	83.5	65.6	65.6	65.9	70.1	70.1	71.4	74.5	72.8	69.5
κ [%]	$\text{URBANSAT}_{\text{RF}}$	81.1	84.1	61.8	64.5	76.4	53.2	75.6	**79.5**	55.4	70.7	69.1	63.5
	$\text{URBANSAT}_{\text{SVM}}$	71.6	70.8	63.3	59.9	59.9	49.5	62.8	61.2	54.9	57.8	55.4	55.9
	$\text{SantaBarbara}_{\text{RF}}$	83.5	**88.7**	82.9	65.7	**78.5**	62.4	66.1	**78.8**	67.0	63.9	**82.5**	57.1
	$\text{SantaBarbara}_{\text{SVM}}$	79.4	79.1	80.6	60.8	60.8	61.0	65.6	65.6	67.1	70.2	68.3	64.6
$\overline{R}$ [%]	$\text{URBANSAT}_{\text{RF}}$	85.8	88.3	70.8	71.0	**81.7**	61.2	81.7	**84.7**	63.8	75.0	77.2	69.0
	$\text{URBANSAT}_{\text{SVM}}$	77.8	77.3	69.8	67.8	67.8	58.8	71.5	70.2	64.3	64.3	62.5	62.8
	$\text{SantaBarbara}_{\text{RF}}$	86.9	**92.4**	85.8	74.5	**80.8**	68.5	74.2	**85.8**	74.5	73.0	**87.9**	68.9
	$\text{SantaBarbara}_{\text{SVM}}$	85.3	85.0	85.2	70.7	70.7	69.7	75.3	75.3	76.4	77.4	75.7	73.5
$\overline{P}$ [%]	$\text{URBANSAT}_{\text{RF}}$	86.0	**90.2**	71.2	74.1	**82.9**	64.1	81.3	**85.0**	63.9	76.0	79.1	70.6
	$\text{URBANSAT}_{\text{SVM}}$	78.1	77.5	72.1	72.9	72.9	59.4	71.5	70.7	63.9	67.5	65.7	69.0
	$\text{SantaBarbara}_{\text{RF}}$	85.3	**90.8**	80.7	67.2	77.1	64.2	70.4	79.3	72.1	67.2	**84.8**	60.9
	$\text{SantaBarbara}_{\text{SVM}}$	81.8	81.5	81.1	62.3	62.6	66.0	72.9	72.9	74.3	72.1	70.5	65.7
$\overline{F_1}$ [%]	$\text{URBANSAT}_{\text{RF}}$	84.8	88.9	68.3	70.5	**80.5**	60.2	80.3	**83.8**	62.7	75.1	76.2	69.3
	$\text{URBANSAT}_{\text{SVM}}$	77.3	76.7	70.4	66.7	66.7	58.1	70.6	69.2	63.2	65.1	63.2	64.2
	$\text{SantaBarbara}_{\text{RF}}$	84.8	**91.0**	81.3	66.5	77.6	63.2	68.3	80.5	69.5	67.3	**85.1**	61.6
	$\text{SantaBarbara}_{\text{SVM}}$	82.2	82.0	81.9	60.2	60.1	61.8	69.7	69.7	71.1	71.1	69.8	66.0

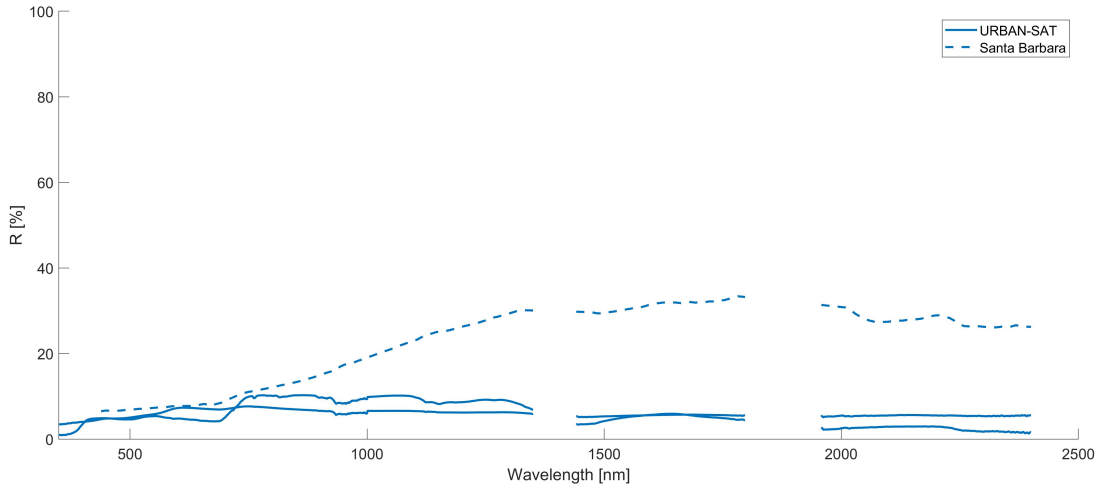
does not provide imagery and only a label that describes the material sample, it is not possible to further examine the difference.

The data acquisition and measurement setup were kept simple which leaves space for potential improvements. The post-processing consisted of two minor steps, by removing noise caused by water vapor in the atmosphere and by calculating the spectral reflectance R for each individual measured spectrum L_t. The reference spectrum E was however acquired before the acquisition of an intra-set L and the incoming solar irradiance $S(\lambda)$ can change throughout the acquisition. A further processing step should have been taken to remove the artificial spectral features observed around 950 nm. These are most likely caused by either insufficient instrument calibration or instrument malfunction.

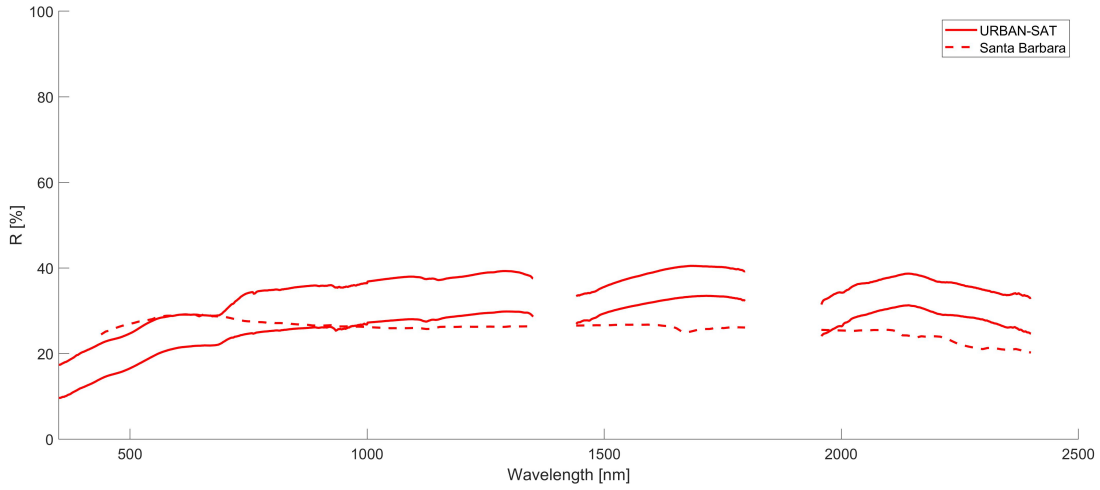
The URBAN-SAT dataset is small and it contains few samples and a small variation. The labeling system is simple and can be improved by adding further low-level information about the samples since it was acquired close-range.



(a) *Metal*



(b) *Wood*



(c) *Concrete*

Figure 4.8: (a-c): The spectral reflectance R of selected samples from the material classes *Metal*, *Wood* and *Concrete*. The samples from the URBAN-SAT dataset are represented as full lines and those from the Santa Barbara spectral library as dashed lines. An intra-class dissimilarity can be observed across all three material classes as the spectral reflectance R varies.

4.5 Summary

In this chapter, the relevant tools and approaches that are utilized in this thesis are introduced. This includes explanations and comparisons of the classifiers, the dimensionality reduction and feature selection tools in addition to an overview of texture analysis measures. Finally, a case study that used some of the presented classifiers and feature selection tools for material classification using two spectral datasets was presented. The case study showed that the RF classifier with PCA-encoded spectra achieved the highest overall accuracy among the 48 scenarios. It also showcased that the spectral range of SWIR contains the most important spectral features for material classification, indicated by the classification evaluation scores. However, the case study highlighted the challenges that occur while handling different spectral datasets due to intra-set spectral dissimilarities. This is caused by differently used taxonomies and as no imagery was available for the Santa Barbara spectral library, a visible comparison was not possible.

5 Spectral Library KLUM

This chapter presents the close-range acquired spectral library Karlsruhe Library of Urban Materials (KLUM), the first published spectral library with a focus on building facade materials. To improve the readability within the context of this thesis, many paragraphs have been rephrased. Most content of this chapter has been published in the article:

R. Ilehag, A. Schenk, Y. Huang and S. Hinz. KLUM: An Urban VNIR and SWIR Spectral Library Consisting of Building Materials. *Remote Sensing*, 11(18), pp. 2149:1 - 28, 2019

The improvements, limitations and potentials that spectral libraries possess have been observed and analyzed in the previous chapters. Firstly, additional information regarding the samples, such as detailed metadata and imagery, is often lacking. Due to limited additional information about the available samples, it can be challenging to either validate, compare or match acquired samples with existing spectral libraries purely based on the labels [69]. Secondly, as observed in the various spectral libraries, specifically apparent in Figures 3.1 and 4.8, the characteristic spectral features of one material can vary. This is likely due to either incorrect labeling or different variations of the surface structures, texture and coating in addition to colors and conditions one material can possess. This variation is neither well represented in libraries nor studied as the variation of one material is difficult to completely cover. Lastly, as presented in Chapter 3.1 and summarized in Table 3.12, the majority of existing spectral libraries contain few building facade samples. As the usage of UAVs is increasing and consequently, close-ranged scene analysis likewise, there is an increasing need for material classification of building facades [108, 112, 209, 264]. Based on the knowledge and experience gathered about existing spectral libraries and their corresponding taxonomy in addition to the case study presented in the previous chapter, this chapter is dedicated to the developed spectral library KLUM.

The chapter is structured as follows. Firstly, data acquisition and measurement setup are described in Chapter 5.1. Since KLUM was acquired *in-situ* and in an uncontrolled environment, several post-processing procedures were necessary which are explained in Chapter 5.2. These include the incorporation of a measure for solar irradiance adjustment and recalculation of the measured reference spectrum E due to an unexpected instrumental malfunction that caused signal clipping. Chapter 5.3 is dedicated to spectral screening. Three spectral metrics are used to assess the spectral similarity of samples with the same material composition to group samples into clusters. Furthermore, they are used to compare and assess the taxonomy the labeling of the clusters with two existing spectral libraries. The finalized spectral library is presented in Chapter 5.4. The chapter presents likewise a discussion regarding spectral libraries and observed intra- and inter-class spectral similarities. Additionally, a discussion about the methodology is presented. Finally, the chapter is concluded with a summary in Chapter 5.5. KLUM is publicly available at `https://github.com/rebeccailehag/KLUM_library`, which includes the samples, the corresponding metadata and imagery. The extended imagery is available at `https://github.com/rebeccailehag/KLUM_extended_imagery`.

5.1 Data Acquisition

The KLUM spectral library utilizes a similar data acquisition setup as the URBAN-SAT dataset, whose measurement setup is described in Chapter 4.4.1. KLUM is likewise acquired with the spectroradiometer ASD FieldSpec4 Hi-Res and acquires a set of 10 consecutively acquired spectra L_t of each sample, i.e. an intra-set L. Chapter 4.4.1 is referred to for additional information about the ASD FieldSpec 4 Hi-Res, the reflectance plate and its acquisition position in reference to the acquired sample and the acquisition of an intra-set L.

The measurement setup of KLUM is extended and improved by measuring the incoming solar irradiance $S(\lambda)$ throughout the acquisition. Furthermore, additional datasets are acquired as supplement: imagery and additional metadata. Additionally, the geolocation is recorded with a Global Navigation Satellite System (GNSS) receiver for possible future revisits.

5.1.1 Spectra and Solar Irradiance

In-situ collected spectra should be acquired during cloud-free conditions to ensure quality [208]. However, as *in-situ* acquisitions are performed in uncontrolled environments, the solar irradiance $S(\lambda)$ can alter during acquisition due to occasionally passing clouds. Therefore, the measurement setup is extended by introducing the Qmini Wide VIS spectrometer[1], an upward-looking spectral reference. It is faced towards the sun to account for any potential alternations by continuously collecting the incoming solar irradiance $S(\lambda)$. The spectrometer Qmini uses a Charged Coupled Device (CCD) detector with a spectral resolution of 1.5 nm within the spectral range of 200 - 1025 nm. This spectrometer enables therefore adjustment and correction of intra-sets L. The setup is visualized in Figure 5.1, which displays the setup of the reflectance plate in reference to the sample of interest. The reflectance plate is facing the same direction as the sample surface and the incoming solar incidence angle θ is the same for the two surfaces. If the the incoming solar incidence angle θ is not the same, the acquired sample's spectrum will be of poor quality and contain spectral abnormalities.

KLUM was acquired in the city of Karlsruhe between July 2 and August 8 2018 during days with sunny weather but with occasional passing clouds. All samples were acquired directly in the sun and their locations were selected to reduce the impact from opposite surfaces, such as facades and windows. The majority of the samples were acquired directly *in-situ* in the city from e.g. buildings, roads and pavements. Selected samples were additionally acquired outdoors at a local building supplier. These additional samples were acquired to increase the number of roofing materials. In short, facade, ground and roof materials were acquired to produce a diverse spectral library. The acquired samples are in various states of weathering and they also range in age. In total, 326 samples were acquired (around 3260 spectra).

[1] www.rgb-photonics.com

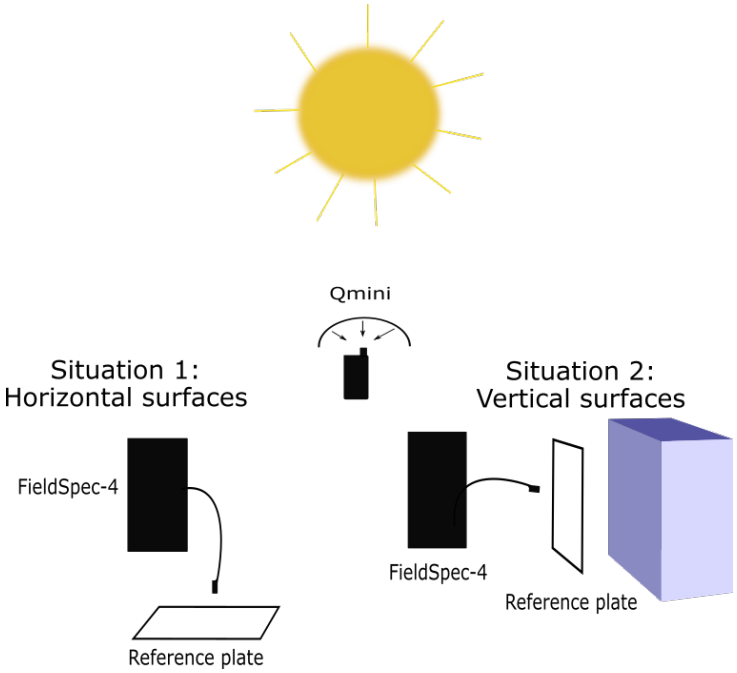


Figure 5.1: The measurement setup used to acquire the samples of the KLUM spectral library. The reflectance plate is placed to face the same direction as the samples, which are either horizontal or vertical surfaces. The Qmini spectrometer is faced towards the sun.

5.1.2 Imagery

Two kinds of imagery were acquired. The first kind of imagery contains an image of each acquired sample which is used as visual reference and control during post-processing. The second kind of imagery is more sophisticated and contains three to four images of selected samples at different scales. This imagery was acquired in October and November 2019 with an acquisition distance of 0.2 - 1.2 m. The extended imagery provides imagery at different scales and provides therefore a larger textural variation on both a finer and rougher detail level (see Figure 5.2 for an example). The extended imagery was acquired with a Nikon D810, a single-lens digital camera with 36.6 megapixels and a sensor size of 35.9 mm x 24 mm. The images have a pixel resolution of 6144 x 4080 pixels. The extended imagery consists of 116 selected samples for 61 samples (corresponding to the samples presented in the case study in Chapter 6.2).

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

Figure 5.2: (a-c): Three images of the sample A003 (*Asphalt*) from the extended KLUM imagery. The images illustrate the three scales, with an acquisition distance of 0.2 - 1.2 m.

5.1.3 Additional Metadata

The solar incidence angle θ is collected post-acquisition for potential future research (e.g. to calculate the BRDF [174]) and statistical purposes. The solar incidence angle θ of each sample is determined post-acquisition by utilizing the geolocation and time of acquisition. The solar incidence angle θ of each sample is available for download at `https://github.com/rebeccailehag/KLUM_library`.

The solar incidence angle θ upon a surface is calculated for the samples with the following equation, where L is the latitude, δ the declination, β the horizontal surface tile angle, Z_s the surface azimuth angle and h the hour angle [115]:

$$\begin{aligned} cos(\theta) = sin(L)sin(\delta)cos(\beta) - cos(L)sin(\delta)sin(\beta)cos(Z_s) + cos(L)cos(\delta)cos(h)cos(\beta) \\ + sin(L)cos(\delta)cos(h)sin(\beta)cos(Z_s) + cos(\delta)sin(h)sin(\beta)sin(Z_s) \end{aligned} \tag{5.1}$$

The solar declination δ is the angular distance of the sun's beam of the equator (either north or south) [115]. The solar declination δ varies throughout the year and is determined using the following equation, where N is the day of the year:

$$\delta = 23.45 sin\Big[\frac{360}{365}(284 + N)\Big] \tag{5.2}$$

The hour angle h is the angle between the two planes, where both planes contain the earth's axis but one contains the zenith and the other a location on earth. The hour angle h is at local solar noon zero, which is positive in the afternoon and negative in the morning. The hour angle h is determined with the following equation, where AST is the corrected local solar time (noon):

$$h = 15(AST - 12) \tag{5.3}$$

For horizontal surfaces (i.e ground), where $\beta = 0°$ and $\theta = \Phi$, Equation 5.1 can be reduced to:

$$cos(\theta) = sin(L)sin(\delta) + cos(L)cos(\delta)cos(h) \tag{5.4}$$

For vertical surfaces (i.e. facades), where $\beta = 90°$, Equation 5.1 can be reduced to:

$$cos(\theta) = -cos(L)sin(\delta)cos(Z_s) + sin(L)cos(\delta)cos(h)cos(Z_s) + cos(\delta)sin(h)sin(Z_s) \tag{5.5}$$

For vertical surfaces, additional information is needed regarding the direction the surface is facing. To keep it simple, the surface directions are either south ($Z_s = 0°$), south-east ($Z_s = -45°$) or east ($Z_s = -90°$).

5.2 Post-Processing

Three refinement steps are firstly employed before the actual post-processing workflow to ease the workflow. 25 samples have a spectral reflectance R of above or close to 100% which were caused by incorrect calibration setup. Those samples were completely excluded from the dataset. Due to the poor calibration of the spectroradiometer, noise mostly in the spectral range of sensor SWIR2 across all L_t in an intra-set L is at times present. The sensor SWIR2 suffered also from a measuring malfunction which caused the reference spectrum E to appear as a completely flat signal. Due to these two issues with sensor SWIR2, this spectral range is excluded for the affected samples.

The post-processing workflow is implemented to improve and adjust the remaining samples and the corresponding spectra. This workflow covers the elimination of noise and outliers in an intra-set L, intra-set adjustment using the acquired solar irradiance $S(\lambda)$ and recalculation of the reference spectrum E caused by an instrumental malfunction. Parts of the post-processing are either done for just one of the FieldSpec sensors or all three sensors (as defined in Chapter 4.4.1). Consequently, some samples have one part of the spectral range completely removed.

The utilized post-processing workflow can be seen in Figure 5.3. The workflow starts with acquiring an intra-set L and the corresponding reference spectrum E for one sample. It receives as output one fully processed sample, representing the average spectral reflectance R of one intra-set L. Each step will be explained in the order of appearance. Firstly, as the samples are acquired *in-situ*, water vapor in the atmosphere absorbs radiation corresponding to certain spectral ranges which cannot be controlled [81, 228]. The spectral range of 1340 - 1450 nm, 1780 - 1970 nm, and 2300 - 2500 nm are therefore removed.

The three following processing steps involve the measured individual samples L_t in an intra-set L. The first step involves the removal of outliers in L. As an intra-set L is acquired in a set of 10, any obvious outliers caused by e.g. unexpected movement of the sensor carrier during the acquisition can easily be manually detected and excluded. Around 2% of the acquired spectra are labeled as outliers due to the spectral features being distinctly different from the remaining L_t in L. The second step concerns the adjustment of L using the acquired $S(\lambda)$ which is further explained in Chapter 5.2.1. The following step involves the removal of noise which is caused by either instrumental malfunction or irregularities in the spectral range of sensor SWIR2. An individual samples L_t is flagged as noisy if its spectral variation is above the threshold of 10%. The average spectral reflectance R of one intra-set L is calculated with:

$$R = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{L_t}{E} \right) \tag{5.6}$$

Due to an unforeseen instrumental malfunction during the radiometric calibration of the spectroradiometer FieldSpec-4, the reference spectrum E suffered from signal clipping. The defect is known and has been confirmed by the producers. It occurs during the radiometric calibration when it is performing the black body calibration. The sensor should close its optical fiber probe to achieve a black environment where no light is present (resembling a black body). However, the malfunction is causing some light to enter since the fiber probe is not completely closing. The reference spectrum E is therefore cut off at specific spectral ranges which occurred for 61 samples (19%). To counter this defect, a post-processing step of recalculating the reference spectrum E is implemented which is explained in Chapter 5.2.2. Due

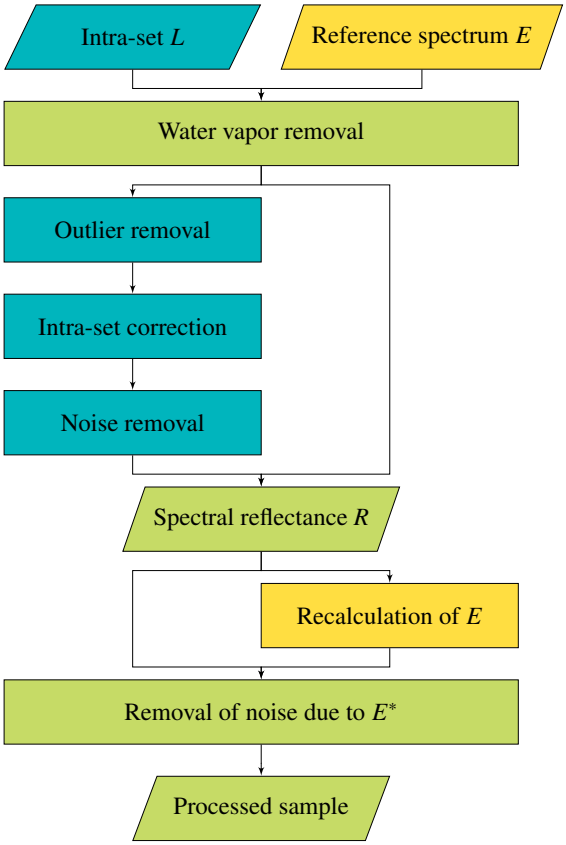


Figure 5.3: The implemented post-processing workflow. Blue represents the processing steps that involve the measured spectra L, yellow the steps that involve the reference spectra E and green the steps that involve both kinds of spectra.

to this recalculation, additional noise is introduced in the spectral range of 950 - 1020 nm which is therefore excluded. This noise is caused by utilizing the spectrometer Qmini since it contains noise in this spectral range. In total, approximately 20% of the full spectral range is excluded due to various kinds of noise.

This concludes the post-processing workflow and the output is a sample that can be further used for spectral screening. To summarize, 44% of the 326 samples experienced one of the aforementioned issues in either one or several parts of the spectral range. 25% of the samples had to be completely excluded after the post-processing, whereas 245 samples remain for the spectral screening. The distribution of the samples that required further post-processing and the samples that were excluded due to the aforementioned issues is:

- Noise: 59 samples (18%)
  - Completely excluded: 47 samples (14%)
- Spectral reflectance R close or above 100%: 25 samples (8%)
  - Completely excluded: 25 samples (8%)
- Sensor SWIR2 malfunction: 8 samples (2%)
- Signal clipping: 61 samples (19%)
  - Completely excluded: 9 samples (3%)

5.2.1 Adjustment of Intra-sets Using Solar Irradiance

As only one reference spectrum E is acquired before the acquisition of an intra-set L, the solar irradiance $S(\lambda)$ can alter throughout the acquisition due to cloud coverage changes. This affects each individual measured spectrum L_t since the reference spectrum E would have been measured differently. By using the the solar irradiance $S(\lambda)$ which the spectrometer Qmini collected during the acquisition, it is possible to adjust an intra-set L by making the individual spectra L_t more homogeneous. By appointing the first measured spectrum in one intra-set L as the initial spectrum L_t^0 and extracting the solar irradiance $S(\lambda)$ acquired at the corresponding time stamp $S_t^0(\lambda)$, it is possible to determine if $S_t(\lambda)$ has been altered for the remaining L_t by computing a ratio corresponding to the alternation. Firstly, to determine if this post-processing step is necessary, the standard deviation σ is computed. If the maximum σ is more than 2% over the complete spectral range, it is an indication that there has been a noticeable alternation. If so, the solar irradiance $S(\lambda)$ is extracted and matched with the corresponding individual spectrum L_t utilizing the synchronized GNSS-time stamps from the two instruments. The adjusted intra-set L^* is calculated using the average ratio between $S_t^0(\lambda)$ and $S_t(\lambda)$ over the full spectral range using

$$L^* = L \cdot \frac{1}{n} \sum_{j=1}^{n} \left(\frac{S_t(\lambda_j)}{S_t^0(\lambda_j)} \right) \tag{5.7}$$

To assure that adjusted intra-set L^* is more homogeneous than the original intra-set L, the maximum standard deviation σ for L^* and L are determined and compared. If the maximum σ of L^* is more than the σ of L, L^* is rejected and the original L is kept. This can be expressed as:

$$L = \begin{cases} L^*, \text{ if } max(\sigma_L) \geq max(\sigma_L^*) \\ L, \text{ otherwise} \end{cases} \tag{5.8}$$

To summarize, this post-processing step enables the possibility to adjust an L using a calculated solar alternation ratio. Figure 5.4 showcases this post-processing step with a comparison between the original intra-set L and the adjusted intra-set L^*, where the adjusted intra-set L^* is more homogeneous.

5.2.2 Recalculation of Reference Spectrum

As the unforeseen instrumental malfunction with the spectroradiometer's optical fiber probe occurred, the measured reference spectrum E is cut off for approximately 19% of all acquired samples. This occurs if the acquired signal is restricted to a certain data range and if reached, the signal is cut off and appears as a flat signal [1]. This phenomenon can be observed for certain parts of the spectral range and at different thresholds. No defined threshold can be utilized for all samples but the spectral ranges can nonetheless be defined. To recover 19% of the samples, a post-processing step is implemented to recalculate clipped reference spectra, defined as E^*. This post-processing step is done separately for the three spectroradiometers' sensors as they require different approaches.

Due to the clipped signal occurring at different thresholds and spectral ranges, the first step consists of detecting flat signals in the reference spectrum E. This consists of firstly finding signal peaks that occur

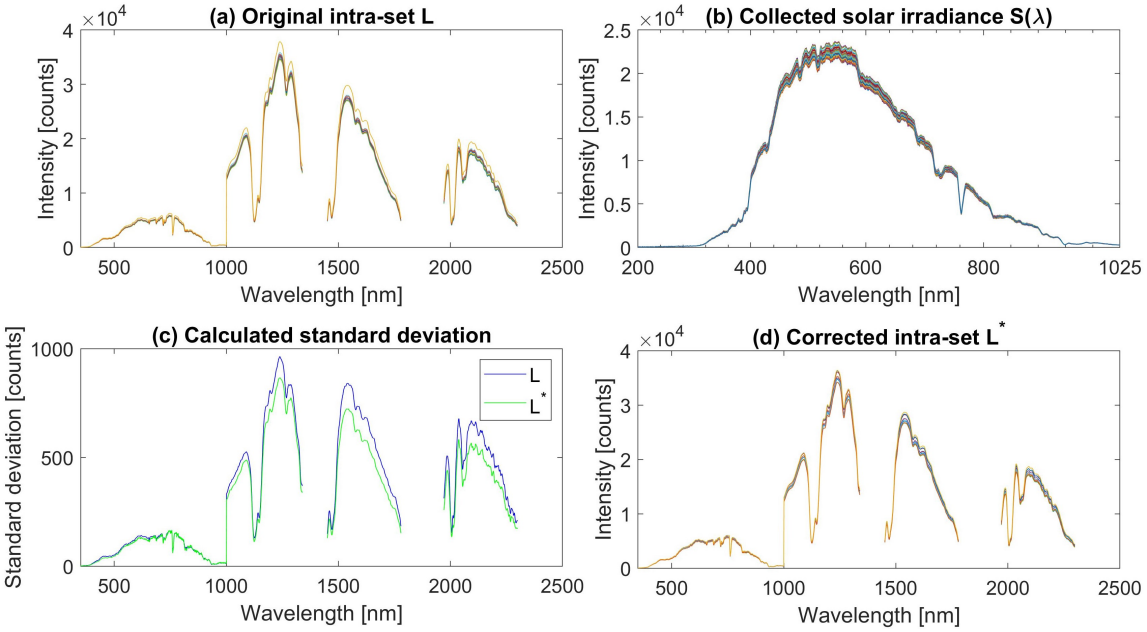


Figure 5.4: The workflow of adjusting an intra-set L by using collected solar irradiance $S(\lambda)$. (a): The original acquired intra-set L. (b): The corresponding collected solar irradiance $S(\lambda)$. (c): A comparison of the standard deviation σ across the spectral range of the original intra-set L and adjusted intra-set L^*. The maximum σ across the spectral range is extracted for both sets and compared. (d): The adjusted intra-set L^*.

in certain parts of the spectral range. As the spectral ranges that are affected by signal clipping are always the same (e.g. 1500 - 1700 nm), it is possible to limit the search to these parts. Secondly, once the signal peaks are found, it is necessary to determine if those peaks are representing a flat signal or not. This is determined by computing the intensity difference between the peaks. If the computed difference represents approximately an intensity difference of ±10, a flat signal has been found. A reference spectrum E with a clipped signal is defined as E^s. Once determined, the consecutively acquired reference spectrum E which did not suffer from signal clipping, E^+, in the same part of the spectrum is chosen. To retrieve a E^+ suitable for E^s, the surrounding environment should not have changed much in addition to the incoming solar irradiance $S(\lambda)$ and its incidence angle θ. If these parameters are have been altered too much, E^* cannot be properly recalculated. A ratio c which represents the average difference in the spectral range which did not suffer from signal clipping between E^s and E^+, E^s_{nc} and E^+_{nc} respectively, is calculated using:

$$c = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{E^s_{nc}(\lambda_i)}{E^+_{nc}(\lambda_i)}\right) \tag{5.9}$$

In addition to this ratio, a second ratio is needed to recalculate the reference spectrum E in the spectral range of VNIR. If the reference spectrum E is clipped off in the VNIR spectral range, around 53% of the spectral range is affected and the ratio c is insufficient. As the spectrometer likewise covers this spectral range, it enables the possibility to determine a vector ratio based on the average $S(\lambda)$ alternation between the two acquisitions of E^+ and E^s using

$$s(\lambda) = \frac{1}{m} \sum_{j=1}^{m} \left(\frac{S^{s}_{nc}(\lambda_j)}{S^{+}_{nc}(\lambda_j)} \right) \tag{5.10}$$

Depending on the spectral range, E^* is determined using:

$$E^* = \begin{cases} E^+ \cdot c \cdot s(\lambda), \text{ in the spectral range of VNIR} \\ E^+ \cdot c, \text{ in the spectral range of SWIR} \end{cases} \tag{5.11}$$

To evaluate the quality of E^* and to determine if E^* has been properly recalculated, the quality measure Q is calculated. Q utilizes the parts of the spectral range that are not clipped off in E^s_{nc} and E^*_{nc}, in addition to the previously determined ratio c and $s(\lambda)$ (for the spectral range of VNIR) to retrieve a value ranging between 0 and 1. 0 represents a perfectly recalculated E^* and a threshold is set to 0.05 to eliminate E^* of poor quality. Q is determined in the spectral range of SWIR using:

$$Q = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\frac{E^s_{nc}(\lambda_i) \cdot c}{E^*_{nc}(\lambda_i)} - 1 \right)^2} \tag{5.12}$$

Figure 5.5 displays E^s, E^+ and E^* for one selected sample. Here, the recalculated reference spectrum E^* passes the quality control in the spectral range of sensor SWIR1 while the recalculated reference spectrum E^* in the spectral range of SWIR2 does not. In the end, a spectral range representing one of the three sensors for 46% of the 61 samples has to be excluded and 15% of the 61 samples have to be completely excluded. The spectral range of VNIR had to be excluded for 26% of the samples and the spectral range of sensor SWIR2 for 20% of the samples. The parameters that are needed to receive a satisfactory E^* are for the majority of the samples (61%) not fulfilled.

5.3 Spectral Screening

To assign appropriate labels to the samples and their classes and sub-classes, the samples are firstly clustered into groups that consist of at least two samples with a similar chemical composition. This is executed to provide several samples of similar characteristic spectral features. Spectral screening is the chosen methodology to determine the spectral similarity with different spectral metrics [203]. The spectral similarity of the samples is determined by using three kinds of spectral metrics by clustering the samples with spectral similarities into groups. This is followed by comparing the clusters with samples of similar chemical compositions from existing spectral libraries to validate and compare the spectra and the corresponding labels. Each class and sub-class is lastly evaluated to ensure high intra-class and intra-sub-class spectral similarity.

5.3.1 Metrics for Spectral Screening

To determine the material clusters (samples with similar chemical composition) and the corresponding sub-classes (samples of the same material but with e.g. different surface characteristics), three spectral

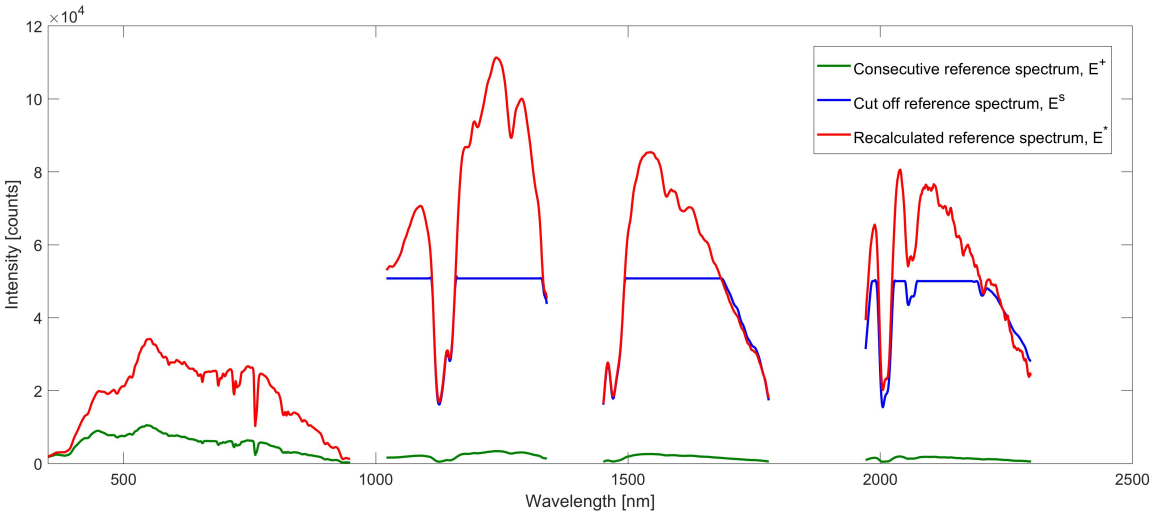


Figure 5.5: A visualization of how the reference spectrum is recalculated. The reference spectrum E^s shown in blue suffered from a clipped signal in the spectral range of the sensors SWIR1 and SWIR2. The consecutively acquired reference spectrum E^+ shown in green did not suffer from signal clipping. In the spectral range of sensor SWIR1, the recalculated reference spectrum E^* shown in red passes the quality control as $Q < 0.05$ while E^* in the spectral range of the sensor SWIR2 does not.

metrics are used. The first chosen metric is the most commonly used metric among the three metrics, the Spectral Angle (SA) [118, 129]. The SA is suitable for continuous data such as hyperspectral data as it determines the angle between two spectral vectors as a measure for the spectral similarity. The smaller the angle, the more similar the two spectra are. This metric is suitable for spectra acquired during different down-welling irradiance and albedo conditions, as the metric is relatively robust for such alternations since it accounts for the vector direction and not the vector length. For two spectral vectors x and y, the SA can be described according to the following formula, where $\langle x,y\rangle$ represents the dot product of the two spectral vectors and $\|.\|_2$ represents the Euclidean norm:

$$SA(x,y) = arccos\left(\frac{\langle x,y\rangle}{\|x\|_2\|y\|_2}\right) \tag{5.13}$$

The metrics Spectral Information Divergence (SID) [41] and Spectral Gradient Angle (SGA) [9] are likewise used. The SID determines a divergence metric between two spectral vectors. Again, the smaller the divergence metric, the more similar the two spectra are. The quality of spectral similarity has been proven to be higher with the SID than with the SA [203]. Given two n-dimensional vectors x and y, the SID is defined as:

$$SID(x,y) = \sum_{i=1}^{n}\left(\frac{x_i}{\sum_{j=1}^{n}x_j} - \frac{y_i}{\sum_{j=1}^{n}y_j}\right)\left(log\frac{x_i}{\sum_{j=1}^{n}x_j} - log\frac{y_i}{\sum_{j=1}^{n}y_j}\right) \tag{5.14}$$

The SGA on the other hand calculates the SA between two absolute spectral gradients. The SGA utilizes the absolute vector gradient instead of the vector direction. The slope change is hence considered which increases the robustness against static offsets. It is therefore likewise invariant to geometry and incident

illumination. The gradients SG of two n-dimensional vectors x and y are firstly determined which are used as input to calculate the SA, as defined in Equation 5.13. The SGA is defined as:

$$SG(x) = (x_2 - x_1, x_3 - x_2, x_4 - x_3, ..., x_n - x_{n-1}) \tag{5.15}$$

$$SGA(x,y) = SA(abs(SG(x)), abs(SG(y))) \tag{5.16}$$

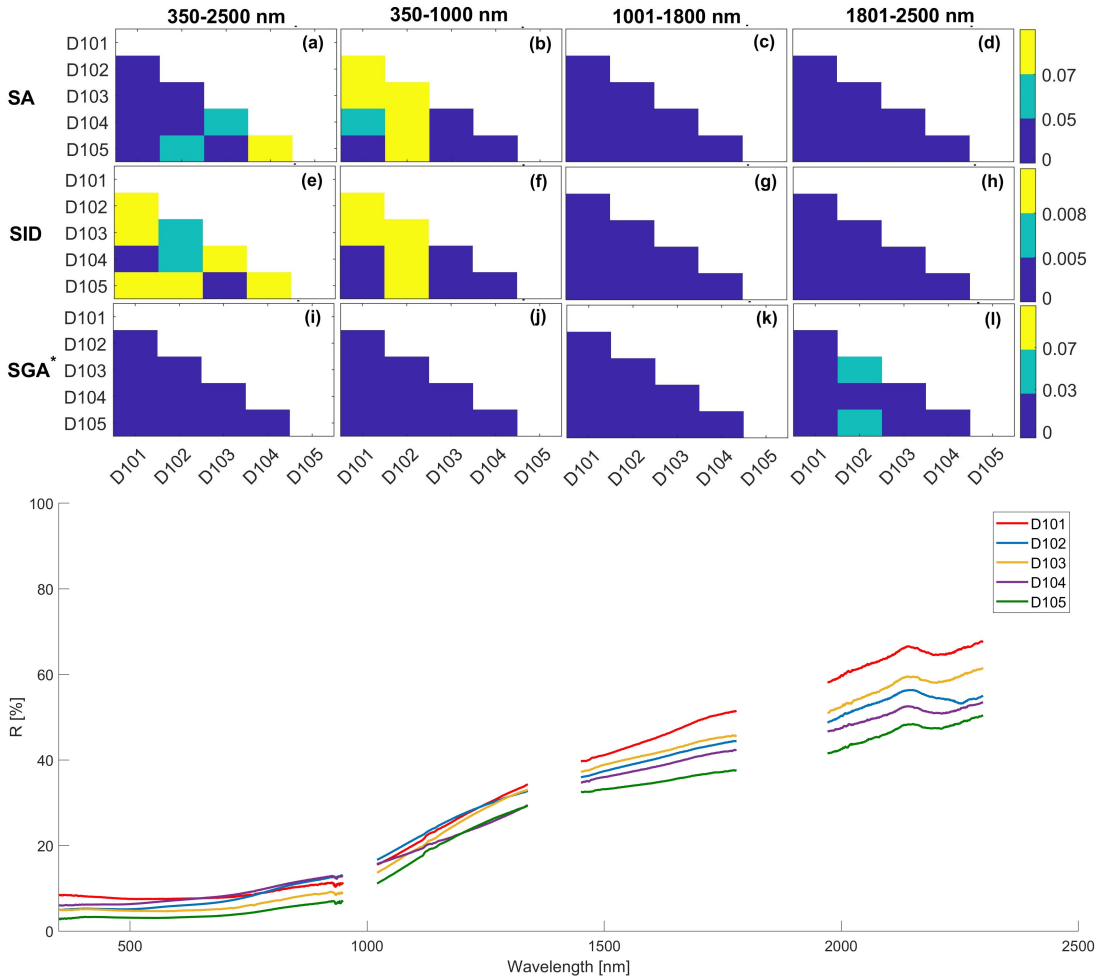
However, the SGA is calculated using the absolute gradient. Two spectral gradients will therefore appear the same even if one spectrum has a negative slope and the other spectrum a positive slope as the absolute derivative will be the same. A modified version of the SGA, denoted as Modified Spectral Gradient Angle (SGA*), is thus implemented. The SGA* adds 1 (but could be any other fixed number) to the calculated vector gradients instead of determining the absolute gradient. This allows a distinction between spectra with the same absolute negative and positive derivative. This can be expressed as

$$SGA^*(x,y) = SA(SG(x) + 1, SG(y) + 1) \tag{5.17}$$

5.3.2 Spectral Screening of KLUM

245 samples remain after the post-processing and are, with the use of the aforementioned spectral metrics, assigned into material clusters in an iterative procedure. The initial iteration is based on the observations made in the field. By iterating the spectral screening, the larger material classes can be split into more refined sub-classes. The three spectral metrics SA, SID and SGA* are calculated for every sample pair using the entire spectral range and each of the three spectral ranges that correspond to the spectroradiometer sensors. This allows a spectral similarity analysis of the samples at different parts of the spectrum. If one part of the spectrum has been excluded due the post-processing, the corresponding sample pair has likewise this part of the spectrum excluded. The three spectral metrics are given different threshold values to determine suitable pairings through spectral similarity scores, resembling those used in [203]. 12 matrices, three for the entire spectral range and nine for each spectroradiometer sensor are determined which display the spectral similarity score between each sample pair. Figure 5.6 visualizes the spectral similarity scores between four samples from one sub-class. The different colors represent three levels of similarity: very similar (dark blue), rather similar (blue) and not similar (yellow) (similar to [203]). The spectral similarity is very similar in almost all parts of the spectrum since the four samples were clustered into the same sub-class. Pairings where all three metrics indicate poor similarity in several parts of the spectrum are rejected. However, an indication of a poor similarity in the spectral range of VNIR, likewise seen in Figure 5.6, is reoccurring this spectral range contains spectral information from mostly the visible spectrum (i.e. the material color).

In the end, 26% of the 245 samples have to be excluded from KLUM after this procedure. This is partly due to the inability to cluster some samples into classes and sub-classes with at least two samples. 12 classes and 33 sub-classes are generated through the material clustering.



(m): The corresponding spectral reflectance R of the five selected samples.

Figure 5.6: Illustration of the spectral screening for the sub-class *Dark reflective ceramic* and the corresponding samples. The three spectral metrics SA, SID and SGA* are calculated for every sample pair in four spectral ranges. The spectral similarity can be observed in (a-l), where (a-d) represents the spectral metric SA, (e-h) SID and (i-l) SGA*. The colors represent the similarity between each pair of material spectra by using the different thresholds (scales located to the far right). Yellow indicates that the two spectra are not similar, bright blue rather similar, and dark blue very similar. (m) illustrates the corresponding spectral reflectance R of the five samples. All five samples are clustered into the same sub-class even if there the metrics are indicating a dissimilarity in the spectral range of VNIR.

5.3.3 Cross-library Spectral Screening

To determine and assign suitable labels to the different material clusters, two publicly available spectral libraries that contain similar materials as KLUM are used: the ASTER library [14] and the LUMA library [124]. Samples that represent urban materials are selected from both spectral libraries by extracting 82% of the samples from the LUMA library and 4% from the ASTER library. The ASTER library contains few urban materials and does not, for some samples, cover the same spectral range as KLUM. Other publicly available spectral libraries, such as the Santa Barbara library [102], the USGS library [120] and the SlyRUM library [265] contain mainly ground or roof materials and are therefore not utilized for this

purpose. In general, the under-representation of facade material samples in publicly available spectral libraries makes this task challenging to determine and compare the labels (as previously seen in Table 3.12).

The three spectral metrics are again used to determine the spectral similarity scores between samples from KLUM and samples from the two chosen spectral libraries. The average spectrum of each KLUM sub-class is calculated to determine the spectral similarity for each sample pair from the libraries ASTER and LUMA. This enables the possibility to sort and rank the spectral similarity of the samples. The 10 samples that resemble the chosen KLUM sub-class the most according to the three metrics are extracted for further analysis. This is exclusively done in the SWIR spectral range to ignore the visible spectrum (i.e. the material color). Furthermore, as few materials are not present in neither the ASTER library nor the LUMA library (e.g. neither contains the material *Plaster*), the field observations and the imagery are used for labeling affected samples.

5.3.4 Intra-class and Intra-sub-class Spectral Evaluation

Finally, to assess and evaluate the quality of the spectral intra-class and intra-sub-class similarities based on the generated clusters, unsupervised classification and clustering in addition to statistical analysis are performed. Firstly, by utilizing unsupervised clustering algorithms, the KLUM's classes and sub-classes can be evaluated and compared. The unsupervised clustering algorithm k-means [148, 152] is chosen since a fixed amount of clusters can be chosen. The squared Euclidean distance is used as the distance metric and the cluster initialization is k-means++ [13]. Thus, the distribution of the classes and sub-classes of KLUM can be compared by setting k to 12 and 33 respectively. To evaluate the k-means clusters, the measure purity is calculated [43, 231]. A cluster is pure if it has a purity score of 1 (range of $[0, 1]$) as it indicates that all samples in that cluster belong to the same class. The average purity score can be calculated across all generated k-means clusters for evaluation with the following equation. Given a set of clusters $\Omega = \omega_1, \omega_2, ..., \omega_k$, a set of classes $C = c_1, c_2, ..., c_k$ and the number of clusters k, purity can be calculated by determining the classes that have the most samples in each cluster as:

$$purity(\Omega, C) = \frac{1}{N} \sum_{k} \underset{j}{max} |\omega_k \cap c_j| \tag{5.18}$$

Additionally, as the spectra are high-dimensional data, PCA is employed to reduce the dimensionality [184]. The first few principal components that cover 99.9% of the data variability are chosen. Secondly, t-SNE is likewise employed as it is suitable for visualizing high-dimensional data [234]. The parameters are tuned to determine the optimal cluster distinction and the distance metric is set to the Euclidean distance, the exaggeration (size of natural clusters in the data) to 12 and the perplexity (effective number of local neighbors of each point) of 30. For these two analyses, the clusters are assessed in different parts of the spectrum. This is done for the entire spectrum, the full SWIR spectral range and the spectral range corresponding to sensor SWIR2. The final assessment consists of a statistical analysis, where the intra-class and intra-sub-class standard deviations σ are determined.

5.4 Results and Discussion

Firstly, the spectral library KLUM is presented with an overview of the available material classes and sub-classes. The used classification scheme is likewise presented which was chosen to enable a more refined categorization. Selected classes and sub-classes are further presented. Secondly, the spectral similarity is assessed for classes from different spectral libraries. This is followed by the intra-class and intra-sub-class spectral assessment of KLUM's samples. Lastly, the measurement setup and the used methodology are discussed.

5.4.1 The Taxonomy of KLUM

181 material samples are at the end clustered into 12 distinct classes and 33 sub-classes, presented in Table 5.1. KLUM contains 97 facade, 46 ground and 38 roof samples. The majority of the classes consist of more than one sub-class. The average effective incidence angle θ for those 181 samples is for horizontal surfaces 22.91° with a standard deviation of 3.85° and for the vertical surfaces 69.81° with a standard deviation of 7.26°.

From the original 326 samples, 56% are included in KLUM. To summarize, 25% of the samples had to be removed after the post-processing either due to noise, incapability to recalculate the reference spectrum E^* or abnormally high spectral reflectance R. Further 19% were excluded after the spectral screening due to the incapability of clustering those into classes or sub-classes with more than two samples.

The three largest material classes are *Ceramic* which consists of 45 samples, followed by *Concrete* with 38 and *Granite* with 16. The sub-classes are named based on the hierarchical material classification scheme, seen in Table 5.2. This scheme is inspired by those used in the spectral libraries LUMA [124], SlyRUM [265] and Santa Barbara [100, 102]. The LUMA and SlyRUM libraries utilize a low-level labeling system with several ranks as both were acquired close-ranged and split up the ranks of the classification scheme into sample identification number, class/sub-class, material, color, status and dimension. The Santa Barbara spectral library splits the scheme into four different ranks where each rank describes the sample in more detail (as seen in Table 3.3). KLUM's scheme provides a description of the usage (facade, ground or roof), the sample color, the surface structure, texture and coating (e.g., reflective or matte) in addition to the status of the material (new or weathered). The classification scheme in Table 5.2 shows all the different labels one sample can be described with. One sample can be described with several of the surface structure/texture/coating labels (e.g. natural and cracked) but with only one label from the remaining ranks. The color hue, saturation and brightness are not described for simplification and each sample is assigned with only one color (the most dominant color). This scheme enables the possibility to split and select a refined sub-class, such as *Painted wood*, and the option to perform material classification for specific cases.

As Chapter 3 is dedicated to the analysis of the taxonomy of spectral libraries, KLUM's taxonomy is firstly further analyzed. Lastly, the two classes *Ceramic* and *Wood* are briefly discussed and presented. Both classes contain sub-classes that highlight the impact of the surface structure/texture/coating as the spectral features can be significantly different depending on the surface. Selected samples from both classes are analyzed. The influence the surface coating has is also presented as one further example.

Table 5.1: The classes and sub-classes of KLUM, here shown with their corresponding ID and final count. The class labels are based on the material compositions while the sub-class labels are describing the dominant features (e.g. the color or the surface coating).

Class	Class ID	Sub-class	Sub-class ID	Count
Asphalt	A	Grey asphalt	A0	4
Brick (clay)	B	Red brick	B0	8
		Beige brick	B1	4
		Painted brick	B2	2
Mortar	C	Grey mortar	C0	3
Ceramic	D	Glazed ceramic (bricks)	D0	4
		Dark reflective ceramic	D1	5
		Glazed ceramic (roof tiles)	D2	12
		Red matte ceramic	D3	8
		Dark matte ceramic	D4	7
		Black glazed ceramic	D5	4
		Grey ceramic	D6	5
Concrete	E	Bright concrete (ground bricks)	E0	7
		Grey concrete (blocks)	E1	5
		Weathered porous concrete	E2	4
		Bright concrete	E3	5
		Grey concrete	E4	9
		Painted concrete	E5	8
Granite	F	Biotite granite	F0	8
		Muscovite granite	F1	8
Limestone	G	White limestone	G0	4
		Colored limestone	G1	5
Metal	H	Paint-sprayed metal	H0	7
		Painted metal	H1	3
Plaster	I	Weathered bright plaster	I0	3
		Colored plaster	I1	4
		Bright plaster	I2	5
Sandstone	J	Red sandstone	J0	4
		Weathered sandstone	J1	6
		White sandstone	J2	2
Conglomerate	K	Conglomerate	K0	11
Wood	L	Varnished wood	L0	2
		Painted wood	L1	5

Table 5.2: The hierarchical classification scheme of KLUM. The additional rank that describes the specific sub-class is not included (rank 1.1, see Table 5.1). A sample is described with only one label from every rank except for rank 4, where several labels can be used.

1. Material	2. Usage	3. Color	4. Surface structure/texture/coating	5. Status
Asphalt	Facade	Beige	Bare	New
Brick (clay)	Ground	Black	Burnt	Weathered
Ceramic	Roof	Blue	Corrugated	
Concrete		Brown	Cracked	
Conglomerate		Green	Fine roughness	
Granite		Grey	Glazed	
Limestone		Pink	Matte	
Metal		Red	Mossy	
Mortar		White	Natural	
Plaster		Yellow	Painted	
Sandstone			Paint-sprayed	
Wood			Popcorn	
			Porous	
			Reflective	
			Smooth	
			Uneven	
			Varnished	

5.4.1.1 Analysis of the Chosen Taxonomy

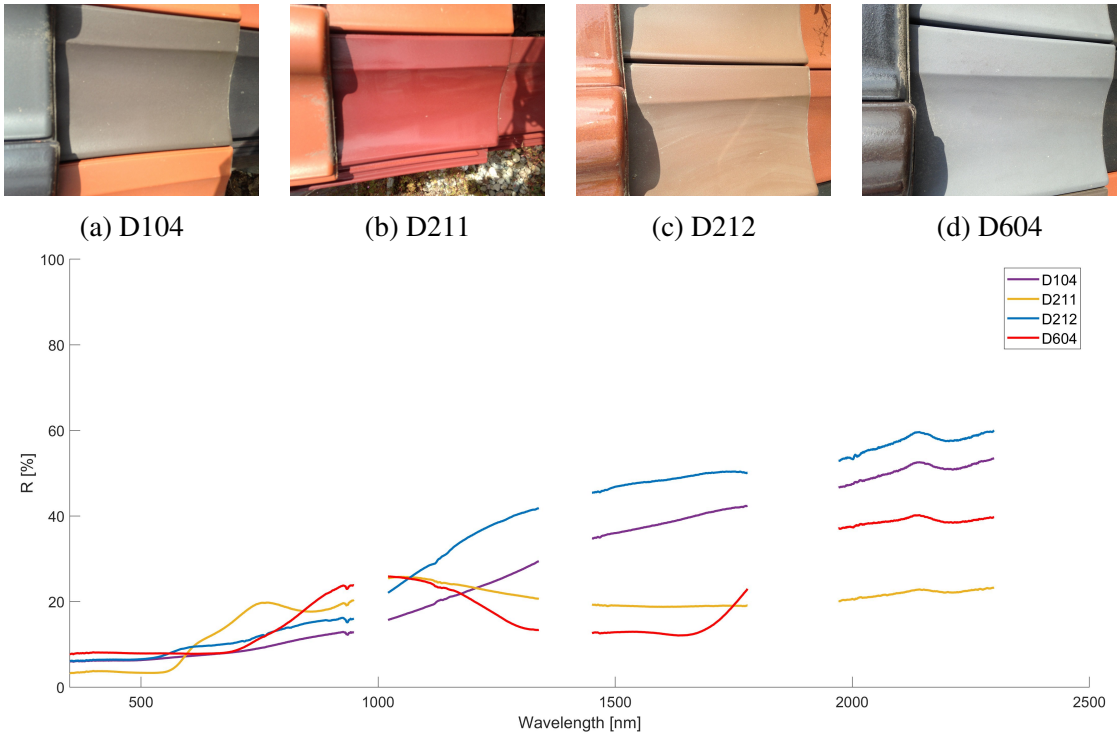
The taxonomy and classification scheme of KLUM is comparable to those of the LUMA and SlyRUM libraries [124, 265]. These libraries are likewise based on close-range acquired samples and can therefore describe the samples on a low level. The highest rank in the classification scheme contains information about the material, unlike those that specialized in land cover and land-use classes. KLUM does, in comparison to LUMA and SlyRUM, provide information about the surface structure/texture/coating in addition to the usage (facade, ground or roof) but KLUM does not describe the color on a low level (e.g. *Light grey*).

A comparison can be made with how KLUM labels the samples from the classes *Asphalt* and *Concrete* with the taxonomy provided in Tables 3.13 and 3.14 respectively. KLUM describes *Asphalt* e.g. as *Asphalt - Grey asphalt - Ground - Grey - Fine roughness - Weathered* and *Concrete* as *Concrete - Bright concrete - Facade - Grey - Bare - Porous - Weathered* or *Concrete - Grey concrete (blocks) - Ground - Grey - Bare - Porous - New*. The taxonomy used for *Asphalt* is to a greater extent comparable with the existing taxonomy. As seen in Table 3.14, the existing taxonomy of *Concrete* differs as each library describes *Concrete* differently. KLUM is likewise contributing to a complex taxonomy of *Concrete*.

The taxonomy and labeling of several classes are not consistent, as some sub-classes are describing the color, others the color and the surface, others only the surface and some the usage or the sample shape (e.g. *Block*). The taxonomy of KLUM could therefore be refined for better consistency.

5.4.1.2 Ceramic

Ceramic is the class in KLUM with the most sub-classes and samples; seven sub-classes and 45 samples. The majority of these samples were acquired at a local building supplier which enabled the possibility to study the impact the color and surface structure/texture/coating have. Figure 5.7 shows this influence, where four samples with different colors from three sub-classes are displayed and visualized; *Dark reflective ceramic* (sample D104), *Glazed ceramic (roof tiles)* (samples D211 and D212) and *Grey ceramic* (sample D604). A spectral variation can be observed in spectral range of VNIR and SWIR. This example highlights the importance of providing information about the surface of a material sample as it appears that the chemical component of the used color coating is influencing the spectral characteristics.



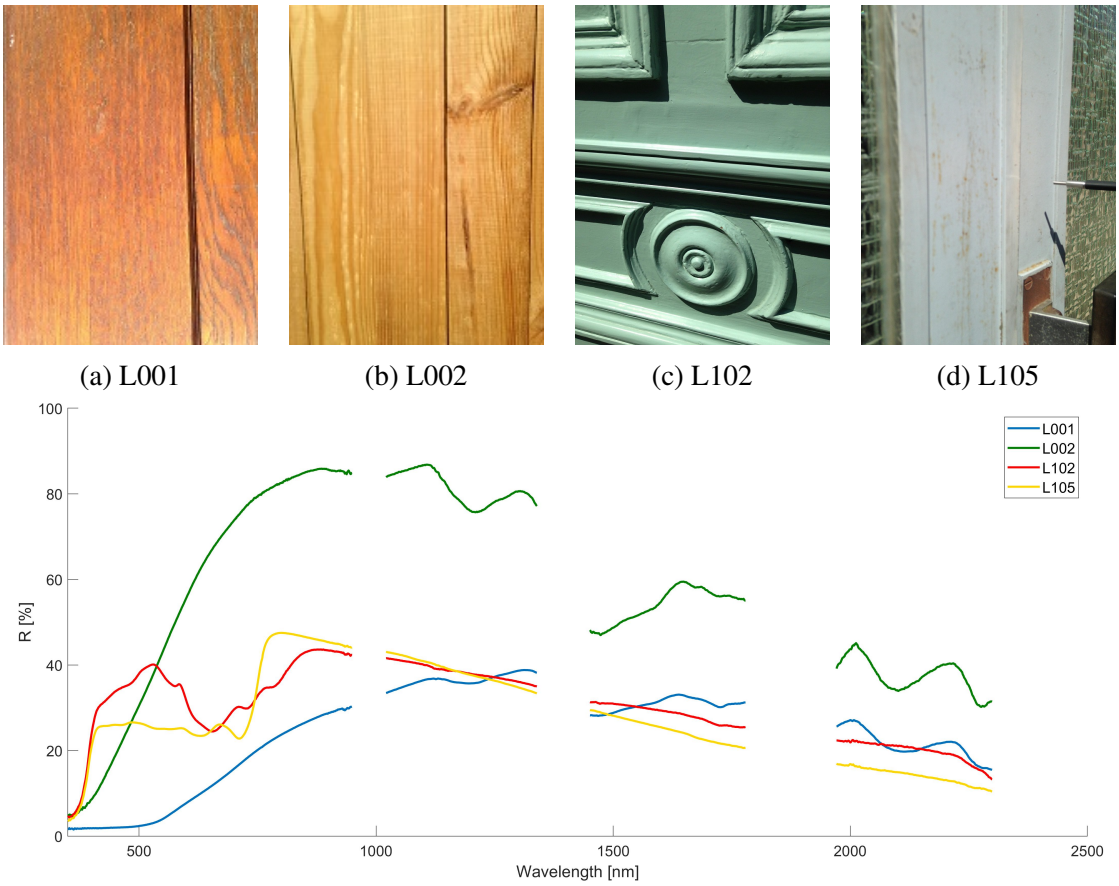
(a) D104 (b) D211 (c) D212 (d) D604

(e): The corresponding spectral reflectance R of the samples (a-d).

Figure 5.7: Visualization of how the material color influences the spectrum. (a-d): Four samples that represent the same material class *Ceramic* from three different sub-classes but have different colors. (e): The corresponding spectra of the samples. A difference is noticeable in the spectral range of VNIR and SWIR, indicating that the color coatings are most likely created with different chemical compositions.

5.4.1.3 Wood

The class *Wood* consists of seven samples and contains two sub-classes which are defined based on the different surface coatings, *Varnished wood* and *Painted wood*. Two samples from each sub-class are chosen for further analysis, samples L001 and L002 from the sub-class *Varnished wood* in addition to L102 and L105 from *Painted wood*. The corresponding imagery and spectral reflectance R are shown in Figure 5.8. Here, it can be observed that the spectral features are distinctly different due to the two kinds of coatings. The influence of the painted color can be observed as a decreasing R throughout the SWIR spectral range. Since these two sub-classes have completely different spectral characteristic features in this observed spectral range, it would be challenging to classify them both as *Wood*. However, they are defined as the same class in KLUM since the underlying material composition is the same.



(a) L001 (b) L002 (c) L102 (d) L105

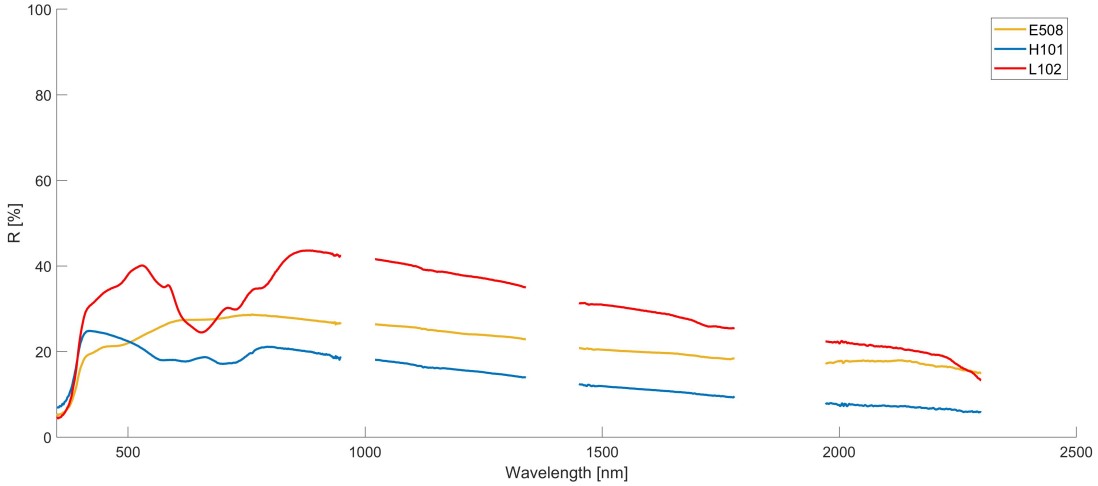
(e) The corresponding spectral reflectance R of the samples (a-d).

Figure 5.8: Visualization of the impact the sample surface can have. All four samples represent the material class *Wood*. (a-b): Imagery of the sub-class *Varnished wood*. (c-d): Imagery of the sub-class *Painted wood*. (e): The corresponding spectral reflectance R of the samples. Due to the different coatings, the spectral reflectance R varies and an intra-class spectral dissimilarity can be observed.

5.4.1.4 Sub-classes with Painted Surface

The influence of the coating has been established in Figure 5.8, where the spectral features are significantly distinct for the two sub-classes of *Wood*. This is likewise supported by three sub-classes that contain samples with painted surfaces. The corresponding imagery and spectrum for the three chosen sub-classes are shown in Figure 5.9: *Painted concrete* (sample E508), *Painted metal* (sample H101) and *Painted wood* (sample L102). Based purely on the spectral characteristic features, the three sub-classes should have been labeled as the same class, e.g. as *Paint*. However, paint per definition is not an urban material. Paint can be used to enhance or alter a surface but can on its own not act as a building or urban material. Each of the three sub-classes is therefore defined as different classes since the underlying material composition differs, even if the spectral features are very similar. This further extends the importance of labeling and accounting for the surface coating, which the majority of spectral libraries does not. This distinct decreasing spectral reflectance R throughout the SWIR spectral range can be observed in other spectral libraries where whose libraries have not specified the coating. This can be observed in e.g the samples *Old Metal Deck* from the UPMSpeclib library [172], *Asphalt shingle roof* from the ASTER library [14] (as seen in Figure 3.1) and the samples V002 and V005 which are labeled as *PVC roofing sheet* from the LUMA library [124]. By not providing information about the surface coating, these would contribute to misclassification.

(a) E508 (b) H101 (c) L102



(d) The corresponding spectral reflectance R of the samples (a-c).

Figure 5.9: Visualization of the impact the surface coating has on the spectral features. (a-c): Imagery of *Painted concrete* (sample E508), *Painted metal* (sample H101) and *Painted wood* (sample L102). (d): The corresponding spectral reflectance R of the samples. The material composition of the three samples differs but due to having the same coating, the spectral reflectance R is similar and an inter-class spectral similarity can be observed.

5.4.2 Cross-library Spectral Similarity Assessment

By retrieving the 10 best-ranked samples that resembled the chosen KLUM sub-class's average spectral reflectance R the most from either the ASTER library or the LUMA library, it was possible to either confirm and validate the labels or receive an indication of the suitability of assigned labels. This proved a challenge due to two reasons. Firstly, few of KLUM's materials are not present in the two external spectral libraries, such as the materials *Plaster* and *Conglomerate*. Therefore, the 10 best-ranked samples were accompanied by lower spectral similarity scores, indicating that they are dissimilar. For those cases, the assigned labels could not be validated and the initial assignments based on the field observations and KLUM's imagery had to be sufficient.

Secondly, the inter-class spectral similarity was unexpectedly high as different materials received high spectral similarity scores which made it challenging to distinguish the materials. The best-ranked samples from the libraries ASTER and LUMA were occasionally several different materials, even if the majority of the material classes were represented by more than five samples from the libraries ASTER and LUMA. A general tendency could however be interpreted. Figure 5.10 displays the five samples from either the ASTER library (dashed line) or the LUMA library (dotted line) that received the highest spectral similarity scores for KLUM's sub-class *Grey asphalt* according to the spectral metric SGA* in the SWIR spectral range. Here, the calculated scores with SGA* were within a small range for four different materials: *Limestone*, *Ceramic*, *Asphalt* and *Cement* (presented here in the ranked order). As these scores were similar, it complicated the comparison and validation of the assigned labels. This is likely due to either incorrect labeling or similar chemical compositions. Furthermore, this indicates that different materials have similar spectral features in this spectral range and that it could be preferable to use parts in the spectral range which are further away, such as LWIR [124]. This would likewise lessen or completely remove the influence the surface coating has (e.g. paint). However, as such equipment is not always available and the majority of spectral libraries, just like KLUM, are providing samples in the VNIR and SWIR spectral range, some awareness and consciousness are needed for material classification and label assignment.

5.4.3 Spectral Similarity Assessment of KLUM

To further evaluate the spectral intra-class and intra-sub-class similarity, the unsupervised classifier k-means and the visualization tool of high-dimensional data provided with t-SNE are firstly employed. Both approaches are assessed with and without a PCA-encoded representation for the different spectral ranges. Lastly, the statistical assessment of spectral intra-class and intra-sub-class similarity is performed by calculating the average standard deviation σ of each class and sub-class.

5.4.3.1 Assessment with k-means

The spectral similarity is firstly assessed with k-means on intra-class and intra-sub-class levels. Figure 5.11 displays the distribution of the 12 k-means clusters for five different spectral ranges: the full spectral range (350 - 2500 nm), the VNIR spectral range (350 - 1000 nm), the spectral range corresponding to the sensor SWIR1 (1001 - 1800 nm), the spectral range corresponding to the sensor SWIR2 (1801 - 2500 nm) and the spectral range that corresponds to the sensors SWIR1 and SWIR2 (1001 - 2500 nm). It can

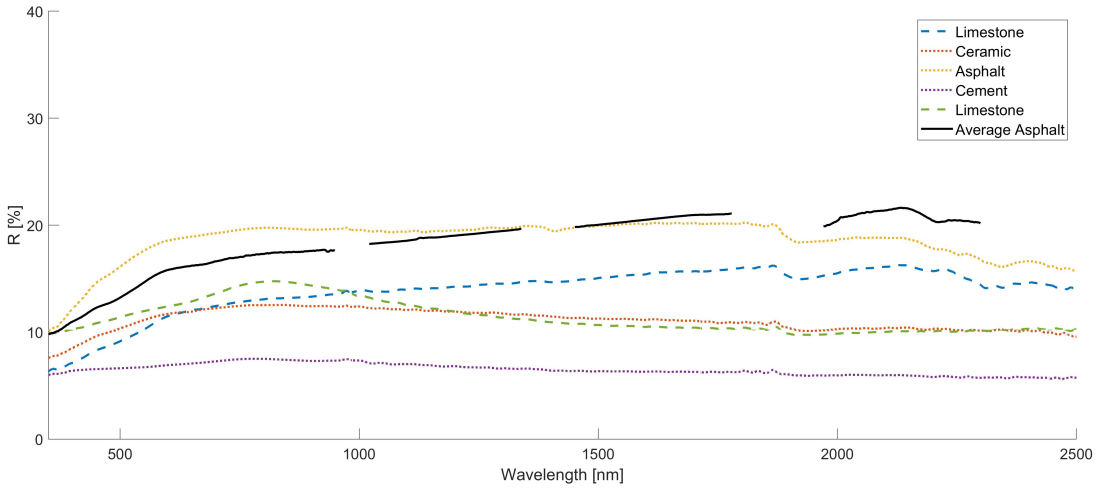
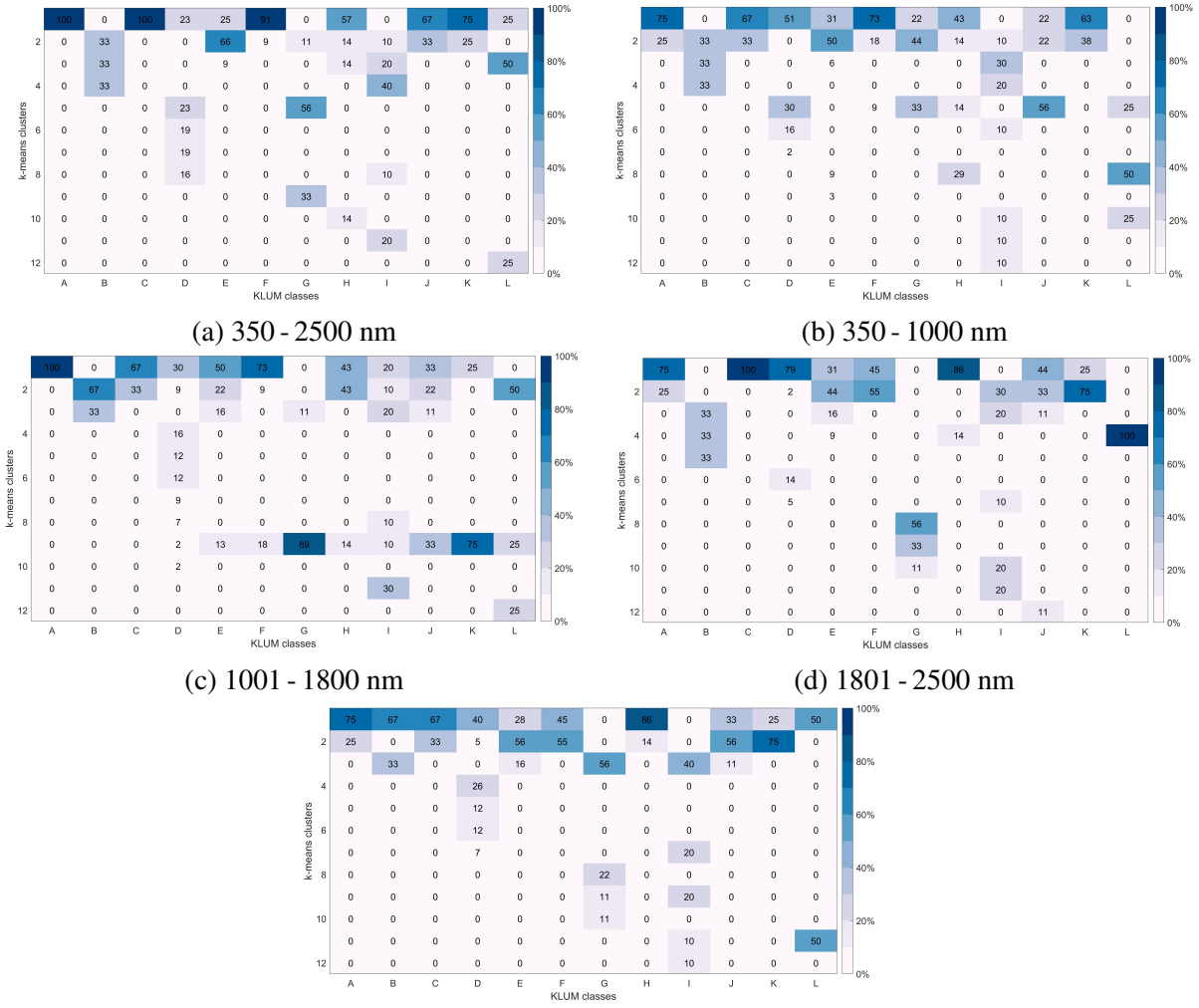


Figure 5.10: The spectral reflectance *R* of the five samples from the libraries ASTER and LUMA that achieved the highest spectral similarity score with the sub-class *Grey asphalt* (black line) according to the metric SGA*. The dashed lines represent samples from the ASTER library and the dotted lines represent samples the LUMA library. According to the scores, the spectral reflectance *R* of *Asphalt* is very similar to the spectral reflectance *R* of samples representing the material classes *Limestone*, *Ceramic* and *Cement*. This indicates that there is a high inter-class similarity.

be observed that two larger clusters have been generated that contain samples from almost every class, regardless of the spectral range. Few distinct clusters can be observed across all five spectral ranges, in particular for the classes *Ceramic* (class D), *Limestone* (class G), *Plaster* (class I) and *Wood* (class L). This indicates that those classes could potentially contain unique sub-classes. No spectral range achieves a majority of only distinct clusters, highlighting that the material distinction is a complex challenge. However, the VNIR spectral range (Figure 5.11b) appears to generate the least amount of distinct clusters compared to the other spectral ranges. This is further supported by observing the calculated purity scores. For the five spectral ranges, the purity score is calculated to 0.75, 0.70, 0.73, 0.78 and 0.79 respectively. The spectral range corresponding to the sensor SWIR2 (1801 - 2500 nm) and the spectral range that corresponds to the sensors SWIR1 and SWIR2 achieve almost the same purity score.

The distribution of the 33 clusters that represent KLUM's sub-classes in the spectral range of SWIR2 (1801 - 2500 nm) with and without PCA encoding is visualized in Figure 5.12. Here it can be observed that more distinct clusters are present as several clusters represent only one sub-class. However, several larger clusters can again be observed. The utilization of a PCA does not significantly improve the distinction between the sub-classes. Contrary, fewer distinct clusters are formed (e.g. 19 compared to 22 distinct clusters). This is likewise supported by the purity score. For the two generated k-means clusters, the purity score is calculated to 0.8 without PCA and 0.77 with PCA. The larger clusters that were generated indicate that the spectral features are very similar for the majority of classes. Thus, this demonstrates again that there is a strong inter-class spectral similarity between the material classes and that spectral features might not be sufficient to distinguish them.



(a) 350 - 2500 nm

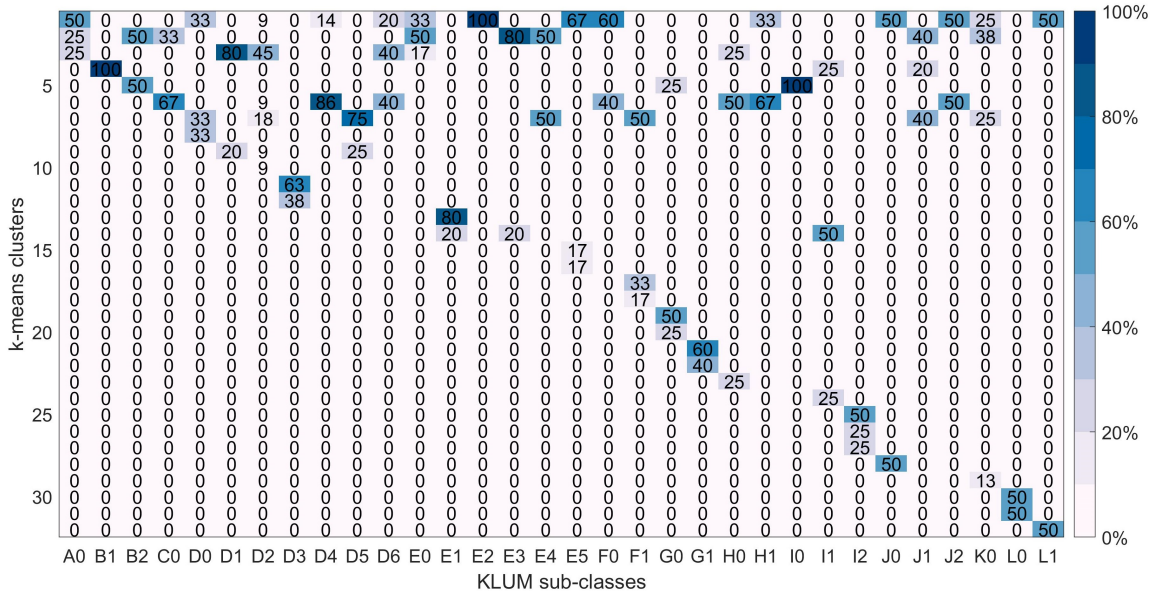
(b) 350 - 1000 nm

(c) 1001 - 1800 nm

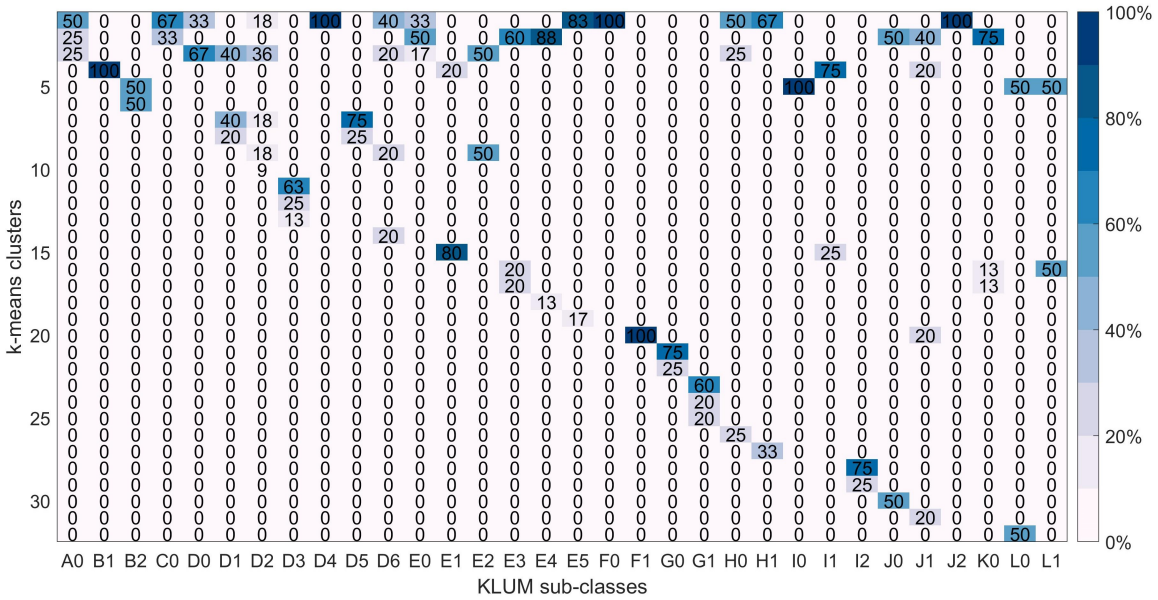
(d) 1801 - 2500 nm

(e) 1001 - 2500 nm

Figure 5.11: Visualization of the k-means clustering and its distribution among the 12 classes. (a-e): The k-means clusters for different spectral ranges. The horizontal axis represents the assigned KLUM labels and the vertical axis the generated k-means clusters. The color displays the distribution of the assigned labels. The ideal class clustering would consist of a distribution of the k-means clusters along the diagonal. The purity score for (a-e) is 0.75, 0.70, 0.73, 0.78 and 0.79 respectively



(a) Without PCA encoding



(b) With PCA encoding

Figure 5.12: Visualization of the k-means clustering and its distribution among the 33 material subclasses in the spectral range 1801 - 2500 nm. (a): Without PCA encoding. (b): With PCA encoding. The horizontal axis represents the assigned KLUM labels and the vertical axis the generated k-means clusters. The color represents the distribution of the assigned material labels. The ideal class clustering would consist of a distribution of the k-means clusters along the diagonal. The purity score for (a-b) is 0.8 and 0.77 respectively.

5.4.3.2 Assessment with t-SNE

The spectral similarity between the classes is likewise supported by the t-SNE distribution as few clusters are given, as seen in Figure 5.13. Here, the t-SNE distribution has been determined using three different spectral ranges: the full spectral range (350 - 2500 nm), the spectral range of SWIR (1001 - 2500 nm) and the spectral range of sensor SWIR2 (1801 - 2500 nm). Few distinct clusters can be observed in the three distributions. The classes *Ceramic* (class D), *Granite* (class G) and *Concrete* (class E) do generate a few distinct clusters with some of their samples. This indicates that those classes could contain sub-classes with distinct spectral features. The different spectral ranges do not significantly improve the distinction, but few clusters are more apparent in the SWIR spectral range, especially for the class *Ceramic*. t-SNE supports alongside the k-means clustering that the spectral features of these material samples are very similar since few clusters are distinct.

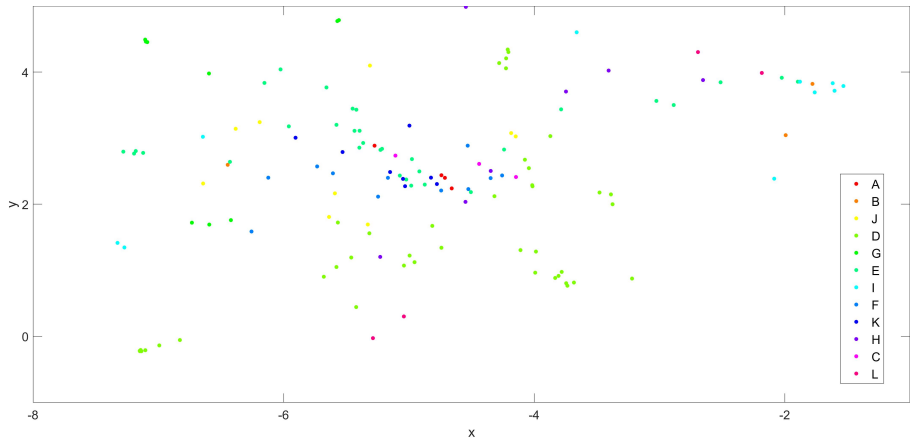
5.4.3.3 Internal Statistical Assessment

The final statistical assessment of the intra-class spectral similarity consisted of determining the average standard deviation σ of every class and sub-class over the full spectral range (350 - 2500 nm), visualized in Figure 5.14. The majority of the classes and sub-classes have similar spectral features since the average σ is less than 5%. The classes that do have a noticeable larger spectral variation are *Plaster* and *Wood* (classes I and L respectively) which likewise can be observed in the corresponding sub-classes (sub-classes I2 and L0). The sub-classes *Bright plaster* (I2) and *Varnished wood* (L0) are the two sub-classes with the largest spectral variation. This is due to two reasons, a large material color variation and few samples, causing a larger spectral variation. Overall, this assessment can conclude that the intra-class and sub-class spectral similarity is similar.

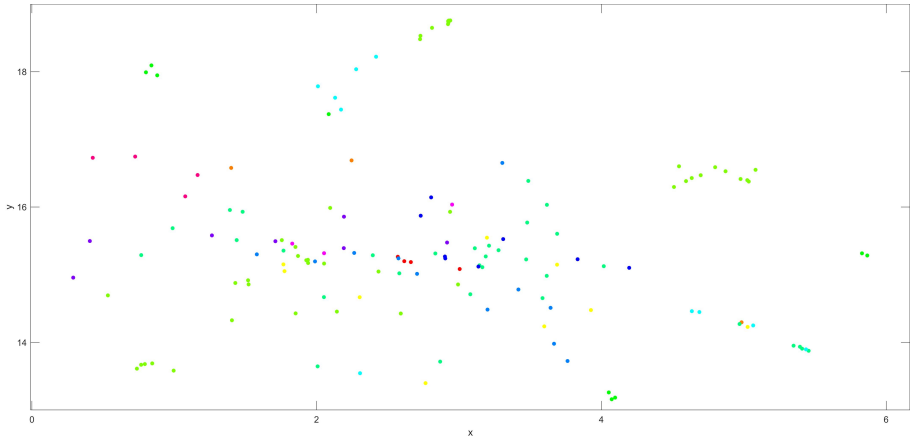
5.4.4 Potential Improvements

To conclude this chapter, the measurement setup and methodology are evaluated to determine challenges and areas of improvement. As KLUM was acquired *in-situ* in an uncontrolled environment, few aspects could neither be measured nor controlled. This includes determining the BRDF, the incoming solar irradiance $S(\lambda)$ which fluctuates throughout the acquisition due to passing clouds and the atmospheric water vapor absorption [81, 228]. To determine the BRDF during the *in-situ* acquisitions, the solar incidence angle θ and viewing angle onto the sample surface have to be recorded and known. The solar incidence angle θ is documented but not the viewing angle since the appropriate equipment needed to determinate the BRDF for samples acquired in an uncontrolled environment was not available.

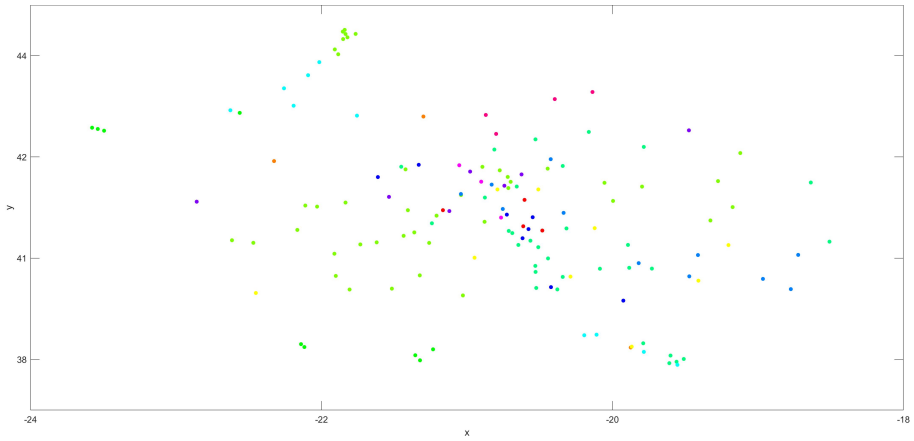
Setting those aspects aside, 44% of the initially acquired samples experienced issues due to either signal clipping, noise, high spectral reflectance R or sensor malfunction. 25% of the 326 acquired samples had to therefore be excluded from further processing which is a significant loss. The measurement setup and the chosen methods to correct said issues will therefore be evaluated and analyzed. Furthermore, as the intent was to have at least two samples in each class and sub-class, 26% of the remaining 245 samples were excluded after the spectral screening. This is likewise a significant loss and will be further discussed. Additionally, the spectral range and its importance is discussed.



(a) Spectral range 350 - 2500 nm.



(b) Spectral range 1001 - 2500 nm.



(c) Spectral range 1801 - 2500 nm.

Figure 5.13: Visualization of the low-dimensional spaced clusters generated with t-SNE with the 12 KLUM material classes in three spectral ranges. (a): The full spectral range. (b): The spectral range of SWIR. (c): The spectral range of sensor SWIR2. The x-axis and y-axis represent the x-coordinate and y-coordinate in the new low-dimensional space. The ideal outcome with t-SNE would be 12 distinct clusters. This would indicate that the original spectra of material classes are distinguishable in the original (high-dimensional) space. However, few prominent clusters are observable and t-SNE indicate therefore that the material classes are not distinct.

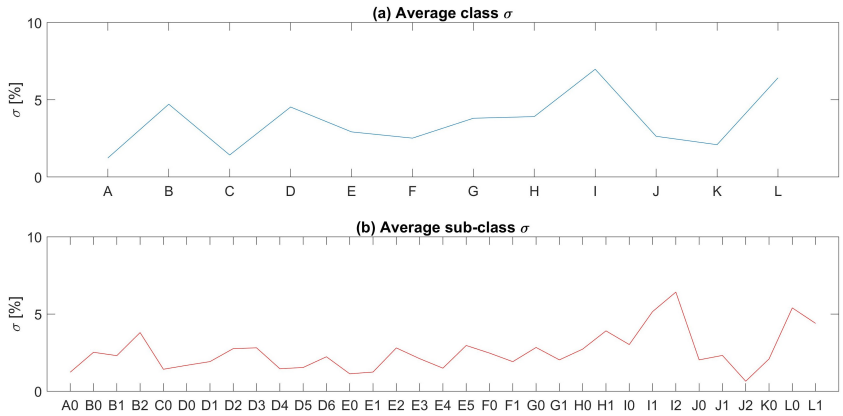


Figure 5.14: Visualization of the average standard deviation σ for the full spectral range (350 - 2500 nm). (a) represents the average standard deviation σ of the material classes and (b) the sub-classes. The material classes *Plaster* and *Wood* (classes I and L respectively) appear to have the largest spectral variation. This can likewise can be observed in the corresponding sub-classes.

Even if the measurement setup and the methodology did experience a few unexpected challenges and could be further improved, the samples that are included in KLUM have passed the quality controls and are therefore satisfactory.

Calibration Procedure

The radiometric calibration with the 95% reflectance plate was always performed before the acquisition of an intra-set L, placed according to Figure 5.1. The incoming solar irradiance $S(\lambda)$ and its incidence angles should be the same for the reflectance plate and the surface of the sample. However, as 8% of the acquired samples experienced abnormal high spectral reflectance R of above or close to 100%, it can be concluded that $S(\lambda)$ was not measured as the same for the two surfaces. This is likely due to different inclination angles. The reflectance plate had most likely another inclination angle than the sample surface during calibration, causing abnormally high spectral reflectance R. This occurs to have been a crucial factor for samples with either a bright surface or surrounding since such samples experienced more generally high spectral reflectance R.

Furthermore, 18% of all samples contained noise in at least one part of the spectrum, of which 14% had to consequently be excluded. This is likewise due to improper calibration. Noise can be reduced by using different signal filters. As the spectral characteristic features are significant for spectral libraries, noise reduction would worsen the distinct spectral characteristic features and no noise reduction was therefore performed.

Solar Irradiance Angle and Acquisition

The incoming solar irradiance $S(\lambda)$ was collected throughout the acquisition with the spectrometer Qmini for adjustment of an intra-set L. This approach works under the condition that the incoming $S(\lambda)$ is acquired under the exact circumstance for an intra-set L and that the upward-looking environment is the same. However, the utilized setup could not always allow such a perfect arrangement. The spectrometer

was not perfectly fixed and was therefore sensible to movement. Its acquisition surface was not always directly facing upwards (towards the sun). In short, the $S(\lambda)$ was not always collected in the same condition for an intra-set L. This was noticeable while adjusting an intra-set L based on $S(\lambda)$ since the adjusted intra-set L^* was occasionally less homogeneous than the original intra-set L.

KLUM was acquired *in-situ* and not in a controlled environment, such as a laboratory. KLUM does therefore not cover how the spectral characteristic features differ when the sample is acquired at different solar incidence angles θ. That is, how the spectral characteristic features alter throughout a day depending on the sun's position and the θ.

To calculate the BRDF [174], the solar incidence angle θ and the viewing angle onto the sample surface need to be measured. The solar incidence angle θ was documented but not the viewing angle. Therefore, KLUM could be extended if the samples are acquired anew and the viewing angle is measured.

Instrumental Malfunction

A user cannot prepare for instrumental malfunction since it is unexpected and cannot be accounted for beforehand. The two sensors spectroradiometer FieldSpec-4 and spectrometer Qmini experienced both malfunctions. The FieldSpec-4 experienced two kinds of malfunctions, most significant was the signal clipping defect in addition to the third sensor SWIR2 acquiring a completely flat signal.

The signal clipping affected 19% of all acquired samples. 3% of all samples had to be completely removed and one part of the spectral range had to excluded for 9% of the samples due to signal clipping. This is a significant loss since samples with a part excluded might not be used for further applications such as training data for classification tasks. The recalculation of the reference spectrum E had a success rate of 39%, i.e. for 39% of the samples, it was not required to remove to exclude any part of the spectrum. This indicates that the approach can be improved. The approach is based on the parameters represented by the incoming solar irradiance $S(\lambda)$ and corresponding incidence angle θ. As the acquisition was done *in-situ* in a city, the parameters could not be kept consistently since they alternate depending on the acquisition location (e.g. which direction the facade surface is facing).

One machine learning approach that could be of interest for recalculating the reference spectrum E is given by GAN [85]. GAN could be exploited by using the reference spectrum E which did not have any signal clipping and alter it into a version that did have it. Hence, the reference spectrum E and its version of the clipped reference spectrum E^s could be used for training a GAN. The GAN could then be used to transform a true clipped reference spectrum E^s into a recalculated reference spectrum E^*.

The malfunction that the spectrometer Qmini experienced was caused by connection issues during the acquisition. The $S(\lambda)$ is completely missing for one field day and is partly missing for two other days. The adjustment of $S(\lambda)$ cannot be executed for those corresponding samples which consequently removes one parameter needed to recalculate the reference spectrum E^* in the spectral range of VNIR.

Spectral Range Importance

Since KLUM covers the spectral range of both VNIR and SWIR, different parts of the spectrum could be further analyzed for spectral screening and the unsupervised clustering. The VNIR spectral range

covers around 30% of the observed spectral range and the color of the sample does consequently affect the evaluation. However, a material does not always have a distinct color even if certain materials do tend to be associated with a specific color, such as *Asphalt* (often grey) and *Brick (clay)* (often red). The visible part of the spectrum was therefore either excluded or ignored during the spectral screening and the unsupervised clustering. The SWIR spectral range was used to further extent as previous studies have shown that the spectral range of SWIR and LWIR is more suitable for determining characteristic spectral features [124, 139**?**]. This is further supported by the experience gathered from assembling KLUM, as the characteristic spectral features are more distinct in the upper part of the range SWIR (>1800 nm).

Sample Pre-Selection

As 26% of the 245 samples that remained after the post-processing were excluded, the samples that were acquired *in-situ* should have been pre-selected. This would have ensured that a sufficient number of samples would have been available for each class and sub-class (i.e. more than two samples of the same material). The *in-situ* acquired samples were not selected beforehand and were chosen directly *in-situ* without considering if said sample would extend or enhance the existing classes and sub-classes. Thus, several acquired samples were later deemed as not suitable for KLUM due to the material of the sample (e.g. gravel). Additionally, as each class or sub-class should contain at least two samples to provide a larger variation, several samples had to be excluded due to the incapability of finding at least one more of the same kind of material.

It was not possible to include and cover a wide range of materials in KLUM due to several reasons. KLUM covers only the commonly used building materials located in southern Germany (central Europe), which is just a small fraction of all existing materials. Furthermore, materials are diverse in terms of both colors and surface conditions. Those included in KLUM were selected based on availability and are therefore not representing all urban and building materials.

5.5 Summary

This chapter presented the spectral library KLUM, the first spectral library with a focus on building facade materials. KLUM is publicly available at `https://github.com/rebeccailehag/KLUM_library` which includes the samples, the corresponding metadata and imagery. KLUM consists of 181 samples and comprises of 12 material classes and 33 sub-classes that were acquired *in-situ* in southwestern Germany. KLUM contains 97 facade, 46 ground and 38 roof material samples in the spectral range of VNIR and SWIR. The developed classification scheme and labeling system are based on the observations from the presented material taxonomy in Chapter 3. KLUM applies a low-level labeling system (similar to the taxonomy of e.g. the library LUMA) by providing information about the surface condition (e.g. color and coating) since such details were possible to observe due to the close-ranged acquisition. The extended imagery is available at `https://github.com/rebeccailehag/KLUM_extended_imagery` which contains 181 images acquired at three scales for 61 selected samples of KLUM.

A post-processing workflow of the acquired samples has been developed. It included the adjustment of intra-sets L using incoming solar irradiance $S(\lambda)$ and recalculation of reference spectrum E of those samples that had a clipped signal. The samples were clustered into classes and sub-classes which contained

at least two samples based on their spectral features using the spectral metrics SA, SID and SGA*. To validate and compare the assigned labels of KLUM's samples, the three metrics were used to calculate a spectral similarity score with samples from the spectral libraries LUMA and ASTER. If the three metrics retrieved a low spectral similarity score between the samples, the field observations and the imagery were to be used instead.

The majority of KLUM's classes exhibited intra-class spectral dissimilarity which could be observed by comparing the spectral reflectance R of samples from different sub-classes. As observed in Figures 5.7, 5.8 and 5.9, the spectral characteristic features for one material sample can differ significantly in this observed spectral range due to the color and surface structure/texture/coating. Through analysis with the unsupervised classifier k-means and the visualization tool t-SNE, it can be observed that the spectral range VNIR is not as suitable for class distinction as the range SWIR. Furthermore, the high inter-class spectral similarity could be observed in Figure 5.10. The spectral characteristic features of different materials are remarkably similar, further increasing the possibility of misclassification. Due to these observed spectral similarities and dissimilarities, it can be challenging to classify material classes purely based on the spectral features. Additional information, such as textural features, could help to address these challenges. A case study presented in Chapter 6 uses selected samples from KLUM for classification with both spectral and textural features.

6 Classification of Urban Materials with Spectral and Textural Features

As established in the previous chapters, two challenges involving spectral features of urban materials have been discovered through analysis and assessment: intra-class spectral dissimilarity and inter-class spectral similarity. An intra-class spectral dissimilarity occurs when the spectral features of one material vary across the respective samples, either due to different surface conditions (such as coating or color differences) or to different chemical compositions. A spectral inter-class similarity can be observed when spectral features of different classes are alike which is caused by similar chemical compositions. Both observations can cause misclassification of urban materials since the spectral features are not consistently class-specific. Classification of urban materials could therefore benefit from additional features to improve the distinction. In addition to spectral characteristic features, this chapter introduces therefore additional features to distinguish urban materials: textural features.

Textural features that describe the visual difference of urban materials could enhance and improve the material class distinction. This chapter analyzes the capability of using either spectral or textural features in addition to both kinds of features for material classification in the scope of two case studies. The first case study uses airborne acquired data to classify roofing materials and the second case study uses close-range acquired data to classify samples from the spectral library KLUM.

How the textural features are described depends on the pixel resolution of the used imagery in addition to the used approach. If the pixel resolution is poor, the textural features might not contribute to an improved material distinction since no textural difference might be detected. Depending on the used approach, the textural features will be expressed in different manners. The standardized and commonly used textural filters, such as the edge detection filter [36], generate imagery represented as the filter response with equal size as the original imagery that contains the textural features while the machine learning approach AE [126] extracts the most important textural features from imagery. Thus, the textural features of the imagery can be extracted and expressed in different manners.

The first case study presented in Chapter 6.1 focuses on the classification of roofing materials in the city Perth, Australia with airborne acquired data. For the classification, different segmentation approaches are used and analyzed. The second case study presented in Chapter 6.2 uses selected samples from the spectral library KLUM where the extended imagery is used to describe the textural features. Both case studies use spectral and textural features but these are described and extracted differently.

6.1 Classification of Roofs with Different Texture Expressions and Additional Data

To improve the readability within the context of this thesis, many paragraphs have been rephrased. Most content of this chapter has been published in the articles:

R. Ilehag, D. Bulatov, P. Helmholz and D. Belton. Classification and Representation of Commonly Used Roofing Material Using Multisensorial Aerial Data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII(1), pp. 217-224, 2018

D. Bulatov, E. Burkard, R. Ilehag, B. Kottler, P. Helmholz. From Multi-sensor Aerial Data to Thermal and Infrared Simulation of Semantic 3D Models: Towards Identification of Urban Heat Islands. *Infrared Physics and Technology* 105, pp. 103233:1 - 15, 2020

The City of Melville is located in Perth, Australia and has undergone several studies closely related to UHI [178]. The City of Melville has a hot-summer Mediterranean climate [224] and is therefore sensitive and effected by UHI. The studies mainly focus on assessing future challenges to ensure that the municipality and authorities can create strategies to counter the effect. This includes determining the correlation between heat and land usage, predicting future temperature changes using time series and analyzing the shading and cooling properties of both native and exotic tree species [54]. The case study presented in this chapter consists of localizing and classifying commonly used roofing materials. This knowledge can help to estimate the locations which in the future may face heating-related challenges due to the roofing material and its color [211, 227, 259]. The City of Melville has collected multi-sensor aerial data, including multispectral imagery, TIR imagery and Light Detection and Ranging (LiDAR) point cloud data since combining data from multiple sensors has proven to contribute with additional information for urban studies [62, 131].

In short, this case study intends to classify commonly used roofing materials in the City of Melville by utilizing three types of airborne acquired datasets. Additionally, different roofing segmentation approaches are evaluated on either pixel-level, superpixel-level or building-level. The classification is performed with textural features which are expressed differently depending on the roofing segmentation approach and additional data through multispectral and TIR imagery. To localize the roofs, roof outlines and the LiDAR point cloud are utilized. Two study areas are analyzed, one smaller area for the analysis of the segmentation approaches and one large area for the assessment of the spectral and textural features. The proposed methodology for the smaller study is presented while only the crucial details needed for the classification in the larger area are presented in this chapter. Further information about the methodology for the classification of the larger area is provided in [33].

6.1.1 Datasets

Two study areas are defined in the City of Melville. The smaller study area (marked in Figure 6.1) used in [106] is located within the larger study area (the entire area in Figure 6.1) which is used in [33].

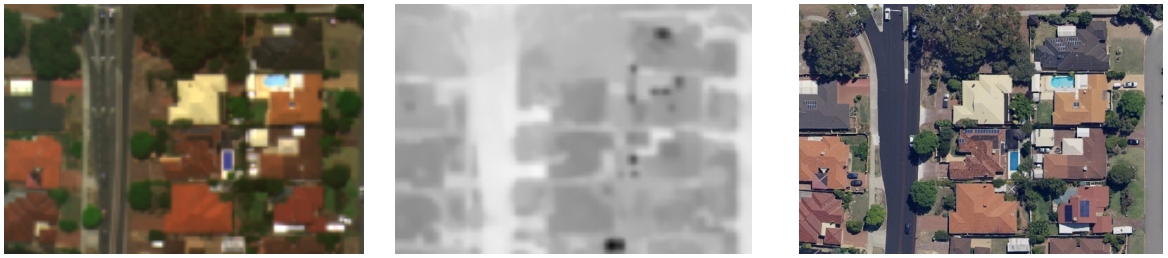
Three kinds of data were acquired during a five-day-long flight campaign in the City of Melville in February 2016: multispectral imagery, TIR imagery and LiDAR point cloud data. The multispectral

Figure 6.1: The two study areas are located in the central-southern parts of the City of Melville. The two study areas are overlapping and the red rectangle indicates where the smaller study area is located.

imagery was acquired with the sensor Micasense Red-Edge[1] from a UAV with a GSD of 0.41 m. The sensor has a global shutter and a field of view of 47.2 °. Furthermore, it has five narrow bands: blue (center at 475 nm and 20 nm bandwidth), green (center at 560 nm and 20 nm bandwidth), red (center at 668 nm and 10 nm bandwidth), red-edge (center at 717 nm and 10 nm bandwidth) and NIR (center at 840 nm and 40 nm bandwidth). The TIR imagery was acquired with the sensor FLIR A615[2] from an aircraft with a GSD of 0.6 m. The sensor has an uncooled microbolometer and captures the thermal radiance within the spectral range of 7.5 -13.5 μm. The TIR imagery was acquired during the night to reduce the thermal radiance impact from the sun during daylight. The LiDAR point cloud data was acquired with the scanner Riegl VZ - 1000[3] from an aircraft. The point cloud was acquired with approximately 2 - 4 points /

[1] https://micasense.com/
[2] www.flir.com
[3] www.riegl.com

m^2 with an accuracy of 0.1 m. The three kinds of data are complemented with the high-resolution RGB aerial imagery acquired in May 2016 by EagleView[4] with a GSD of 0.1 m. The three kinds of imagery are shown in Figure 6.2 where the pixel resolution (the GSD) difference can be observed.



(a) Multispectral imagery (b) TIR imagery (c) High-resolution RGB imagery

Figure 6.2: A selection of buildings in the smaller study area that highlights the pixel resolution (the GSD) difference of the three kinds of imagery. (a): The multispectral imagery. (b): The TIR imagery. (c): The high-resolution RGB imagery.

Roof outlines from 2012 were likewise provided by the City of Melville in addition to a labeled dataset from February 2016. The labeled dataset consists of 10 roofs, where seven roofs have been labeled as *Tiles*, two as *Colorbond*[5] and one as *Zincalume*[6]. These roofing materials are commonly used in Western Australia. *Colorbond* and *Zincalume* are both metallic steel roofs consisting of a mixture of mainly aluminum and zinc, where *Colorbond* is pre-painted steel and *Zincalume* is coated steel. As the provided labeled dataset is small and the class distribution is imbalanced for training and testing a classifier, the dataset is modified. Firstly, *Colorbond* and *Zincalume* are merged into one material class called *Metal*. Secondly, the labeled dataset is extended by manually labeling roofs corresponding to 50 randomly chosen pixels. The dataset is extended to 53 roofs, where 44 roofs are labeled as *Tiles* and nine as *Metal*. These 53 roofs are used for the smaller study area. Three selected areas are chosen in the larger area and those roofs were studied and labeled *in-situ* in October 2018. The larger study area consists of 426 roofs, of which 388 roofs are labeled as *Tiles* and 38 roofs as *Metal*.

6.1.2 Proposed Method

The utilized workflow for the smaller study can be seen in Figure 6.3. Firstly, as the provided roof outlines are dated 2012 and the other datasets from 2016, the outlines have to be updated for consistency. Secondly, once updated, the roofs are segmented into three different levels: pixel-level, superpixel-level and building-level. Thirdly, once segmented, the textural features are calculated and retrieved. Lastly, the roof segments are classified.

For the larger study area, another approach is utilized for updating the roof outlines which is described in [33]. Furthermore, the roofs are only assessed on a building level.

[4] www.eagleview.com
[5] www.steelselect.com.au/materials/colorbond-steel
[6] www.steelselect.com.au/materials/zincalume-steel

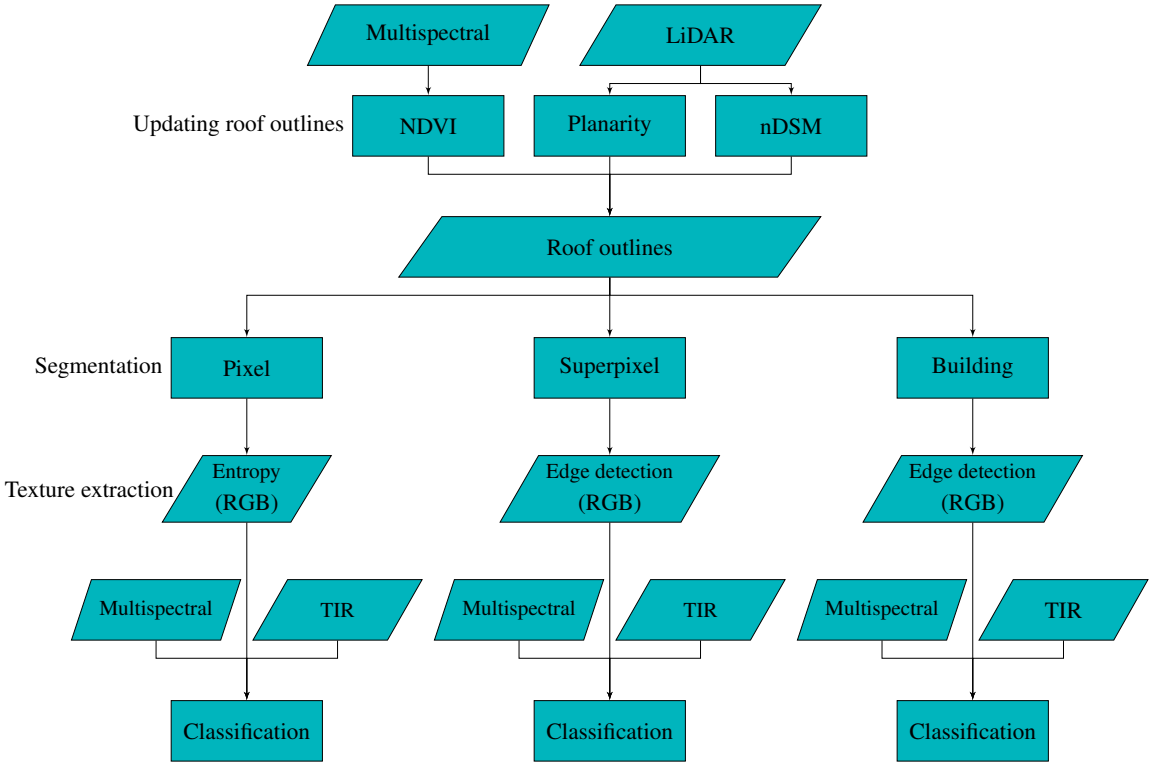


Figure 6.3: The utilized workflow for the analysis of the smaller study area. Firstly, the roof outlines are updated using the multispectral and LiDAR datasets. Once the roof outlines have been updated, the roofs are split into three kinds of segmentation approaches: pixel, superpixel and building. For each kind of segmentation approach, the texture features are extracted. The textural features are, in addition to the multispectral and TIR, used as the input to the RF classifier.

6.1.2.1 Extracting and Updating Roof Outlines

Since the provided roof outlines are from 2012 and the remaining datasets from 2016, it is needed to control if buildings or parts of buildings have been removed or newly built. This is done by utilizing a pixel-based approach. Firstly, to detect new buildings or parts of buildings, a Digital Surface Model (DSM) is sampled with a GSD of 0.5 m using the LiDAR point cloud with a natural neighbor interpolation [22]. By determining the difference between a generated Digital Terrain Model (DTM) and the DSM, a Normalized Digital Surface Model (nDSM) is retrieved [243]. The nDSM is further used to detect elevated objects as it describes the height of objects above the normal (the ground). Here, objects located more than 2.5 m above the normal are deemed as potential newly built buildings or building parts. Normalized Difference Vegetation Index (NDVI) [206], a commonly used ratio, is added to detect non-vegetation. The values of the NDVI are within the range of $[-1,1]$ and the threshold used to separate vegetation and non-vegetation depends on the sensor and the season. In general, a value between $[0.2,0.4]$ is used as a threshold for separation [169]. Here, the NDVI is calculated using the bands NIR and red-edge, as follows:

$$NVDI = \frac{NIR - RedEdge}{NIR + RedEdge} \tag{6.1}$$

Pixels with a NDVI of below 0.2 are deemed as non-vegetation for this case study. As the success rate of determining non-vegetation with the NDVI depends on setting a suitable threshold, an additional measure is calculated with the LiDAR dataset [88, 248], namely planarity. Planarity describes the flatness and levelness of a surface within the range of $[0, 1]$ by using the eigenvalues λ_1, λ_2 and λ_3 of the 3D covariance matrix derived from a local neighborhood. As vegetation and urban materials are not equal in terms of flatness and levelness, planarity can assist with the distinction. Here, pixels with a value above 0.7 are deemed as a non-vegetation. Planarity is calculated using the following equation, where λ_1 is the first, λ_2 the second and λ_3 the third eigenvalue:

$$Planarity = \frac{\lambda_2 - \lambda_3}{\lambda_1} \tag{6.2}$$

According to the regulations defined by the City of Melville, newly built buildings or parts of buildings are those with an area threshold of more than 10 m^2. Thus, new building segments bigger than 10 m^2 are detected by using the thresholds of above 2.5 m according to the nDSM, less than 0.2 according to the NDVI and more than 0.7 according to the planarity measure.

To detect any completely removed buildings or parts of a building, the median nDSM value within the provided roof outlines is determined. If the median value is lower than the set threshold of 2.5 m and the color of the segment differs from the roof's mean color, a building or a part of a building is deemed to have been removed. However, roofs with very low reflection values due to the color are occasionally flagged as removed parts. This is due to a low 3D point density because of poor reflectance properties [95]. Those cases are however manually corrected.

To remove noisy and irregular shapes of the newly added or removed segments, the two morphological operators represented by erosion and dilation [93] are employed, using a kernel of one pixel. Figure 6.4 displays the updated roof outlines in the smaller study area.

Figure 6.4: Visualization of the updated roof outlines in the smaller study area. Unchanged roof outlines are represented in blue, new roof segments in yellow and removed roof segments in orange. No roof was entirely removed in this study area.

6.1.2.2 Roof Segmentation

For the smaller study area, three segmentation approaches are selected to determine the most suitable method. A roof is expressed as either pixels, superpixels or one building. A pixel-level approach requires no further processing but does not consider one roof as one object. Neighboring pixels representing the same roof might not be classified as the same material due to color or shading alternations. Superpixels divide one roof into several sub-parts which are based on differences determined by the utilized superpixel algorithm. As solar panels and solar water heaters can be located on a roof, dividing a roof into several sub-parts is reasonable. A building-level approach assumes that the entire roof is of the same material. Roofs are commonly assigned with one class label since sub-parts, such as solar panels and solar water heaters, are often not labeled and ignored.

For the pixel-level approach, each pixel is assigned with the pixel's values from the multispectral and TIR imagery in addition to the corresponding value from the texture analysis (will be described in the next sub-chapter). The superpixels are generated using the Simple Linear Iterative Clustering (SLIC) algorithm [2]. SLIC groups an image into neighboring regions with similar pixel values by clustering the pixels in a 5D space, defined by the RGB values. Each roof is here split into a maximum of 10 sub-parts and the high-resolution RGB imagery is used for the clustering. Each superpixel is assigned the average value of the corresponding pixels in the multispectral and TIR imagery in addition to the corresponding value from the texture analysis. One roof is defined as one closed object, retrieved from the updated roof outlines. Each roof is assigned the average value of the corresponding pixels in the imagery and the corresponding texture value.

For the larger study area, the building-level approach is utilized. Both the average and the median values are computed for comparison.

6.1.2.3 Extraction of Texture Features

To add supplementary information, the textural features are extracted with texture analysis approaches using the high-resolution RGB imagery. As different segmentation approaches are used, the textural features cannot be expressed similarly for the smaller study area. The textural features for the pixel-level approach are described with entropy [92], a statistical measure of randomness of each pixel. For the superpixel- and building-level approaches, the *Canny* edge detection algorithm is utilized to determine the number of edges in each segment [36]. For each segment, a ratio that expresses the amount of edges is assigned to the corresponding segment.

For the larger study area, the textural features are determined using the MR8 filter bank [237], resulting in eight textural features for each segment.

6.1.2.4 Classification

For both study areas, the classifier RF is used with the implementation of [65]. The parameters are set to a maximum tree depth of five and the minimum number of data points allowed for a split to one. To evaluate the performance of the classifier, an evaluation matrix is generated which consists of the overall accuracy O, the κ-value, the average recall $\overline{R}$, the average precision $\overline{P}$ and the average $\overline{F_1}$-score. The

datasets are split into training and testing datasets to evaluate the performance of the classifier RF. To achieve a balanced class distribution, the number of training samples is set to the same amount for both classes. To ensure consistency, the features are normalized to have a unity-based normalization and unity standard deviation.

In the smaller study area, four random roofs from each class are chosen (8 roofs in total). This represents 44% of the samples of the smaller class *Metal*. 15% of samples in the smaller study area are used as training data. Six spectral features (five multispectral features and one TIR feature) and one textural feature are used to train and test the classifier RF.

In the larger study area, 19 random roofs from each class are chosen (38 roofs in total), representing 50% of the samples from the class *Metal*. 9% of the samples in the larger study area are used for training the classifier. As the roofs in the larger study area have been assigned both average and median feature values, 12 spectral features (10 multispectral features and two TIR features) and 16 textural features are used. Additionally, to assess the importance of the two types of features, the classification is performed using either only spectral or textural features in addition to combining those two. Three scenarios are defined, one which only uses the spectral features (scenario 1), one which only uses the textural features (scenario 2) and one which uses both kinds of features (scenario 3).

6.1.3 Results and Discussion

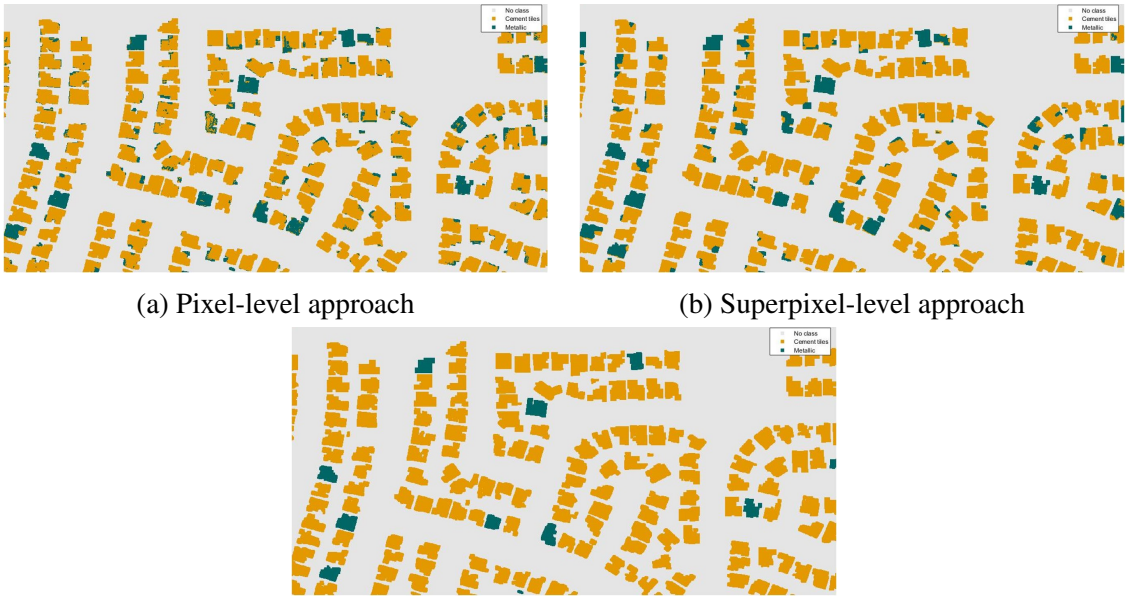
The classification results for the smaller study area are firstly analyzed. The results are visualized in Figure 6.5, where the results achieved with the three segmentation approaches are presented. The evaluation matrix is found in Table 6.1.

Table 6.1: The achieved evaluation scores for the segmentation approaches pixel-, superpixel- and building-level. The best scores are achieved with the building-level approach, shortly followed by the pixel-level approach. Highlighted are the best evaluation scores.

	Pixel	Superpixel	Building
O [%]	94.3	85.2	**97.8**
κ [%]	79.6	33.2	**89.7**
$\overline{R}$ [%]	94.9	77.3	**98.8**
$\overline{P}$ [%]	86.3	62.8	**91.7**
$\overline{F}_1$ [%]	89.8	65.8	**94.8**

As seen in Table 6.1, the best classification results are achieved with the pixel- and building-level approaches. This is supported by all calculated measures stated in the evaluation matrix. The evaluation values indicate however that the class distinction is slightly better with the building-level approach. As the classifier RF is employed, it is possible to assess the band importance. The spectral features represent around 85% of the features that the classifier RF used. The TIR features on the other hand are the least important features, representing less than 1%. This indicates that the classification could have been performed with only the spectral features from the multispectral sensor.

For the larger study area, the overall accuracy O and the κ-value for the different combinations in Table 6.2 indicate that the spectral features are more important than textural features for material classification



(a) Pixel-level approach

(b) Superpixel-level approach

(c) Building-level approach

Figure 6.5: Visualized classification results in the smaller study area where (a) displays the results achieved with the pixel-level approach, (b) the results achieved with the superpixel-level approach and (c) the results achieved with the building-level approach. Orange corresponds to the material class *Tiles* and green *Metal*.

of roofs. Scenario 1, which represents the spectral features, retrieves the best evaluation scores. The scores for scenario 2 which represents the textural features are lower in comparison. The scores for scenario 3 which contains both spectral and textural features achieves a similar overall accuracy O as scenario 1. The κ-values are however lower, indicating that spectral and textural features are incomparable and do not provide an improvement. By observing the band importance acquired by the classifier RF for scenario 3, it can be concluded that the feature from the TIR imagery is again the least important one. The textural features are approximately represented by 28% of the used features, meaning that the spectral features are more important. This is further supported by lower the overall accuracy O and κ-value acquired for scenario 2 where only textural features are used since.

Furthermore, the utilization of median values compared to average values improves the classification results, as seen by both the overall accuracy O and the κ-value. By using the median value instead of the average value, it reduces the influence from other unlabeled objects and the results are improved. Figure 6.6 displays this difference. The roof with *Tiles* in the lower right is misclassified as *Metal* with the average value but correctly classified with the median value. This is due to the solar panels located on the roof as those pixel values influence the roof's mean value.

This case study assumed that the roof could only be two different roofing materials, *Tiles* and *Metal* (*Colorbond* and *Zincalume*) since these material classes were initially given by the City of Melville. However, this two-class task is a simplified model and there are most likely more types of roofing materials present in the study area. The labeled data was created on a building-level with the assumption that one building could only consist of one material. For the pixel-level and superpixel-level approaches, one roof was assigned the same label. However, several buildings have solar panels and solar water heaters on the roofs (e.g. as seen in Figure 6.2). None of those objects had been labeled and they were therefore

(a) Average value (b) RGB imagery (c) Median value

Figure 6.6: A visualization of the difference of using either the average or the median value for classification of the roofing material. The red color corresponds to the material class *Tiles* and green to *Metal*. The roof located in the lower right corner in (a) has been assigned the incorrect label (*Metal*). As the roof is partly covered with solar panels, the average value is screwed.

Table 6.2: Classification results achieved for the two material classes *Tiles* and *Metal* in the larger study area. Scenario 1 that represents the spectral features achieves the best evaluation scores. Scenario 2 represents the textural features and scenario 3 both kinds of features. Scenario 2 achieves the lowest evaluation scores, indicating that textural features are less suitable for distinguishing the two material classes. Highlighted are the best evaluation scores.

		1	2	3
O [%]	Average	**95.1**	83.4	93.2
	Median	**98.4**	84.8	97.3
	Average and median	**97.8**	89.1	96.7
κ [%]	Average	**61.7**	22.8	53.2
	Median	**83.4**	28.3	74.8
	Average and median	**78.9**	35.1	71.1

assigned the same label as the roofing material. This could explain the poorer classification results using the superpixel-level approach since segments might have consisted of only other objects. To provide reliable segments, the additional objects on the roofs should be labeled.

To conclude the case study, it can be observed that spectral features achieve better results than textural features for classification of roofs in the City of Melville. Combining the two does not significantly improve the results, indicating that the spectral features do provide more distinct features. The spectral features were either multispectral or TIR and the spectral correlation that hyperspectral data possesses is therefore not present.

6.2 Close-range Acquired Data for Classification of Urban Materials with Textural and Spectral Features

To improve the readability within the context of this thesis, many paragraphs have been rephrased. Most content of this chapter has been published in the article:

R. Ilehag, J. Leitloff, M. Weinmann, A. Schenk. Urban Material Classification using Spectral and Textural Features Retrieved from Autoencoders, *International Annuals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, V1(2020), pp. 25-32, 2020

As observed in Chapter 5, the material samples of the spectral library KLUM have a high inter-class spectral similarity. This could be observed in the clustering distribution of k-means and t-SNE where few distinct clusters representing one material class existed (e.g. as seen in Figures 5.11 and 5.13). This case study further explores the relationship between spectral and textural features by adding textural features to distinguish material samples from the library KLUM. Selected samples from six material classes are chosen, increasing the classification complexity.

Contrary to the first case study presented in this chapter, this case study uses close-range acquired data. The spectral features are expressed as either the original form, PCA encodings, gradients or compressed spectral representations retrieved from a deep AE [239]. By expressing the spectral features in different manner, it enables the possibility of comparing firstly two dimensionality-reducing approaches (PCA and the compressed representation from the AE) and secondly two approaches that account for the spectral vector difference (the original spectrum and the gradient). The extended imagery of KLUM was acquired from a close distance and the distinct textural features can in detail be described contrary to the previous case study. Furthermore, the textural features from the aforementioned case study were extracted using standardized texture filters where each segment was given a calculated average textural value. Another approach to describe and extract textural features is with an AE [126]. An AE is an unsupervised machine learning method for compressing and decompressing data to retrieve important dimensionality-reduced features. The textural features are retrieved with a stacked denoising CAE [157].

6.2.1 Dataset

From six material classes, 61 samples corresponding to 13 sub-classes are selected for this case study. The material classes *Asphalt*, *Ceramic*, *Conglomerate*, *Limestone*, *Plaster* and *Sandstone* are chosen (see Figure 6.7 for three selected samples). These samples consist of 606 spectra, where each sample represents eight to ten individual measured spectra L_t. The six classes are selected to focus on materials with similar spectral features but with different textural features, such as *Asphalt* and *Conglomerate*. Contaminated samples (e.g. painted surface) are excluded as these features can be observed in this spectral range (as previously discussed and shown in e.g. Figure 5.9). The majority of the samples in the following sub-classes are chosen:

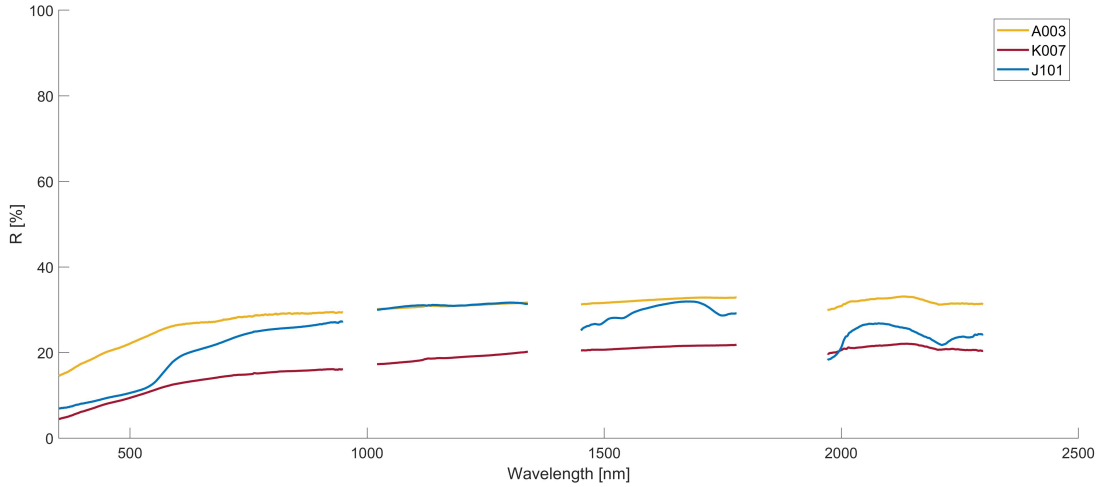
- A0 - *Grey asphalt*
- D1 - *Dark reflective ceramic*
- D2 - *Glazed ceramic (roof tiles)*
- D3 - *Red matte ceramic*
- D5 - *Black glazed ceramic*
- D6 - *Grey ceramic*
- G0 - *White limestone*
- G1 - *Colored limestone*
- I0 - *Weathered bright plaster*
- I1 - *Colored plaster*
- J0 - *Red sandstone*
- J1 - *Weathered sandstone*
- K0 - *Conglomerate*

The extended imagery of KLUM is utilized for this case study as the imagery has a high pixel-resolution of 6144 x 4080 pixels (refer to Chapter 5.1 for further information about the used camera). Three or four images are acquired for each of the 61 samples. Three images of each sample, in total 183 images, are however selected to represent three different scales which are acquired within the distance range of 0.2 - 1.2 m. The texture features can be described and assessed on all three scales or separately on each scale.

(a) A003

(b) K007

(c) J101



(d) The spectral reflectance *R* of the samples (a-c).

Figure 6.7: The extended imagery of KLUM for three chosen samples: (a) A003 (*Asphalt*), (b) K007 (*Conglomerate*) and (c) J101 (*Sandstone*). (d): The corresponding spectral reflectance *R*. The gradients of the spectral characteristic features of *Asphalt* and *Conglomerate* are similar which could make the class distinction with only spectral features challenging. The spectral gaps are due to post-processing which is further explained in Chapter 5.

6.2.2 Feature Expressions

To classify the chosen samples, a framework consisting of expressing and extracting spectral and textural features is utilized. Firstly, the spectral features are either expressed as the original spectra (baseline), as PCA encodings, as gradients or as spectral compressed representations derived from a deep AE. The PCA encodings, the gradients and the spectral compressed representations are derived from the original spectrum. By having the spectral features expressed in four different manners, it is possible to compare the baseline with the alternative expressions and assess the usefulness of the deep AE. Secondly, the textural features are derived from a deep denoising CAE using the extended imagery. Lastly, the two kinds of features are merged to analyze if the material distinction can be improved. The workflow is visualized in Figure 6.8. The workflows for extracting the spectral features are visualized in blue, the workflow for expressing the textural features is visualized in yellow and the workflows relying on the combination of the two features are visualized in green. Nine scenarios are in total defined.

The compressed representations are based either on spectral or textural data and are acquired after encoding (compressing) the said data. To assess the encoder, a decoder and a loss function are needed. The decoder decompresses (decodes) the compressed representation while the loss function measures the lost information (data) between the compressed and decompressed representations. Two AEs are needed and the corresponding architectures differ for the extraction of either spectral and textural features. Nonetheless, few precautions and procedures remain the same for the two AEs.

Firstly, the input data and the compressed representation are normalized with a unity-based normalization to the closed interval of $[0,1]$ to keep the consistency between the AEs. Secondly, 20% of the data is used as testing data, 10% for validation and the remaining 70% for the training of the two AEs. Lastly, the selected samples from KLUM induce a material class imbalance and measures are therefore taken to assure that each material is represented in both the training and testing dataset. The utilized models and the necessary pre-processing procedures are further explained in the following sub-chapters.

Spectral Features

Firstly, as KLUM was acquired *in-situ*, parts of the spectral range had to be removed due to atmospheric influence and the post-processing (refer to Chapter 5.2 for more information). This resulted in spectral gaps in the range as no data is available. The following four spectral ranges are defined for further use:

- 350 - 949 nm (vector size of 600)
- 1021 - 1339 nm (vector size of 319)
- 1451 - 1779 nm (vector size of 329)
- 1971 - 2299 nm (vector size of 329)

Scenario 1 is represented by the original spectra (L_t) which are expressed as continuous ranges with the aforementioned spectral ranges (the spectral gaps are ignored). Scenario 2 uses PCA encodings which are derived from the original spectra. The PCA [184] is employed by using the algorithm available from the Python toolbox *scikit-learn* [185]. The few first principal components that cover 99.99% of the variability of the given data are retrieved. The spectral features for scenario 3 are the first-order gradients

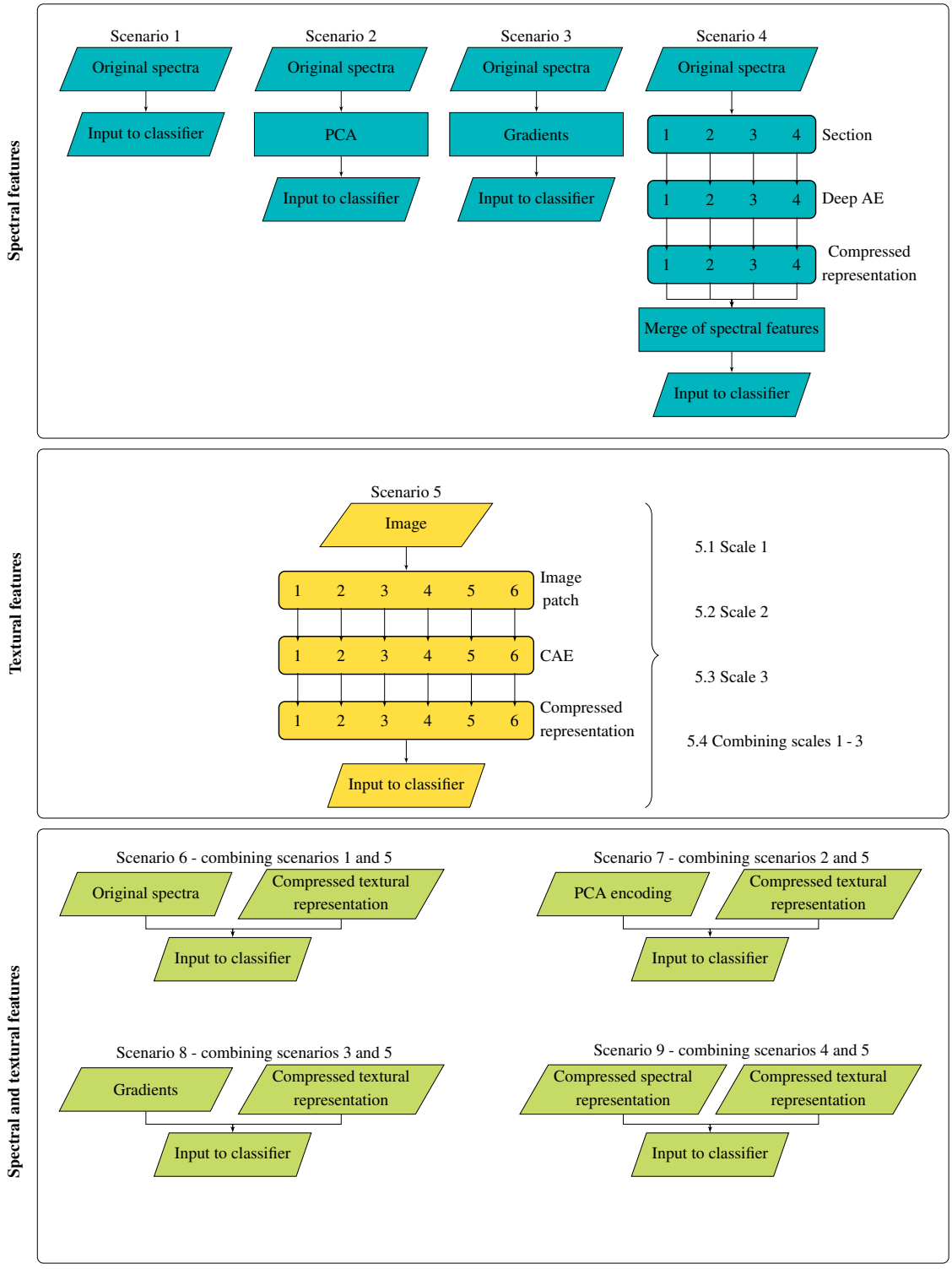


Figure 6.8: The data processing flow for one material sample with the nine different scenarios. Scenarios 1 - 4 utilize only spectral features which are expressed differently. Scenario 5 utilizes only textural features. The four sub-scenarios (scenarios 5.1 - 5.4) are employed to determine the optimal textural representation by using the different scales. The optimal representation is used for scenario 5. Scenarios 6 - 9 utilize both kinds of features by combining the features from scenarios 1 - 4 in addition to the features from scenario 5.

based on the original spectra. The gradients are computed using the finite difference $|f|(x)$ with a point spacing of 1. The central difference is computed as follows, where x is the point of interest:

$$\delta_h|f|(x) = f(x+\frac{1}{2}) - f(x-\frac{1}{2}) \tag{6.3}$$

The forward and backward differences are determined at the endpoints respectively:

$$\Delta_h|f|(x) = f(x) - f(x-1) \tag{6.4}$$

$$\Delta_h|f|(x) = f(x+1) - f(x) \tag{6.5}$$

The spectral features are expressed as compressed spectral representations in scenario 4. These are determined using four individual deep AEs due to the spectral gaps. As the four parts are not equally sized, each deep AE is structured differently. 61 samples and the corresponding 606 individual spectra L_t are available for training and testing of the deep AEs. Since the individual measured spectra L_t for a sample are highly correlated, measures are taken to assure that a sample is present in either the training or the testing dataset. Four deep AEs are constructed by connecting either three or five hidden layers, depending on the size of the input data. Five hidden layers are used to encode and decode the first defined spectral range, 350 - 949 nm, where the deep AE is fitted using a batch size of 20 during 500 epochs (see Figure 6.9 for the used architecture). Three hidden layers are used for the remaining defined spectral ranges as they are smaller and contain less complex spectral features and are more responsive to more compact compression (see Figure 6.10). These are trained during 100 epochs with a batch size of 20. For all deep AEs, the Rectified Linear Unit (ReLU) activation, defined as $y = max(0, x)$, is utilized for each hidden layer except for the last output layer. A linear activation is used for the output layer, defined as $y = x$. Additionally, a sparsity constraint of 10^{-5} is added to prevent that the hidden layers are learning an approximation of the PCA. To assemble the deep AEs, the stochastic optimizer Adam [119] and the mean squared error are chosen to assess the loss. For each of the four deep AEs, a compressed spectral representation expressed as a vector with a size of 8 of its corresponding spectral range is retrieved. The compressed spectral representation over the complete spectral range is represented by a vector with 32 elements (a compression factor of approximately 49).

In short, the spectral features are represented as follows:

- Scenario 1 - original spectrum (vector size of 1577)
- Scenario 2 - PCA encoding (vector size of 10)
- Scenario 3 - gradient (vector size of 1578)
- Scenario 4 - compressed spectral representation (vector size of 32)

Textural Features

To extract and describe the textural features from the imagery of KLUM, a deep denoising CAE is employed. Some pre-processing steps are first done to enhance and improve the ability to encode and decode

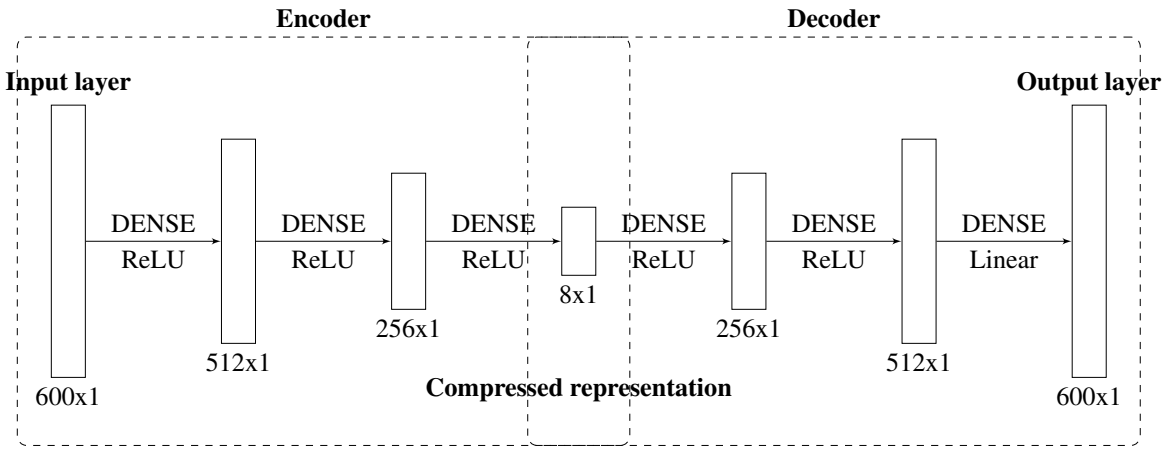


Figure 6.9: The architecture of the deep AE for the spectral range of 350 - 949 nm. Five hidden layers are used to encode and decode the original spectra through dense layers, ReLU activation layers and a final linear activation layer.

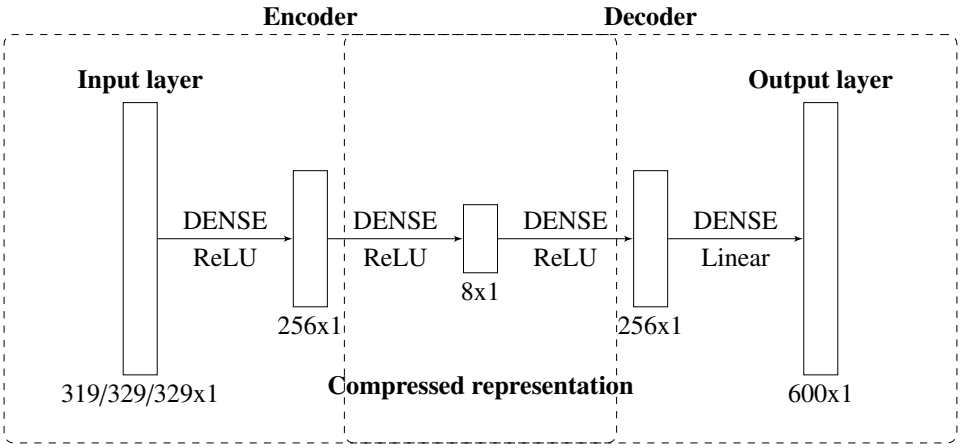


Figure 6.10: The architecture of the deep AEs for the remaining spectral ranges of 1021 - 1339 nm, 1451 - 1779 nm and 1971 - 2299 nm. Three hidden layers are used to encode and decode the original spectra through dense layers, ReLU activation layers and a final linear activation layer.

the imagery. Firstly, the acquired imagery has a pixel-resolution of 6144 x 4080 pixels. The pixel resolution is scaled down with a factor of two into 3072 x 4020 pixels as a high level of detail is neither desired nor necessary for this case study. Secondly, six randomly located patches of 256 x 256 pixels in each image are chosen to capture and extend the textural variation of one image. Each image generates six image patches. Thirdly, the image patches are transformed into a grey-scaling as information from the three color channels is not needed to retrieve textural features. The grey-scale transformation Y' is employed by using the following equation, where R, B and G represent intensity values of the corresponding RGB channels:

$$Y' = 0.2125 \cdot R + 0.7154 \cdot G + 0.0721 \cdot B \tag{6.6}$$

Lastly, normally-distributed noise is added randomly across the image patches to corrupt the patches which increases the robustness. As mentioned in Chapter 4.2, the AE is forced to extract the most important features by adding noise [84, 239].

The encoder consists of eight layers, whereas one layer consists of a 2D convolutional layer with a linear activation, batch normalization and an additional ReLU activation layer. The last encoder layer consists of reshaping (flattening) the layer into a 32-dimensional vector. The decoder consists of the inverse layers to decode the compressed representation into the same shape as the encoded input but with a sigmoid activation, defined as $y = \frac{1}{1+e^{-x}}$, to retrieve a value between $[0, 1]$. The deep denoising CAE is using the stochastic optimizer Nadam [66] and the mean squared error to assess the loss. The architecture of the encoder can be seen in Figure 6.11. The compressed textural representation for one RGB image patch (512 x 512 x 3) is compressed by a factor of 24,576. Figure 6.12 displays the compressed textural representation of five samples.

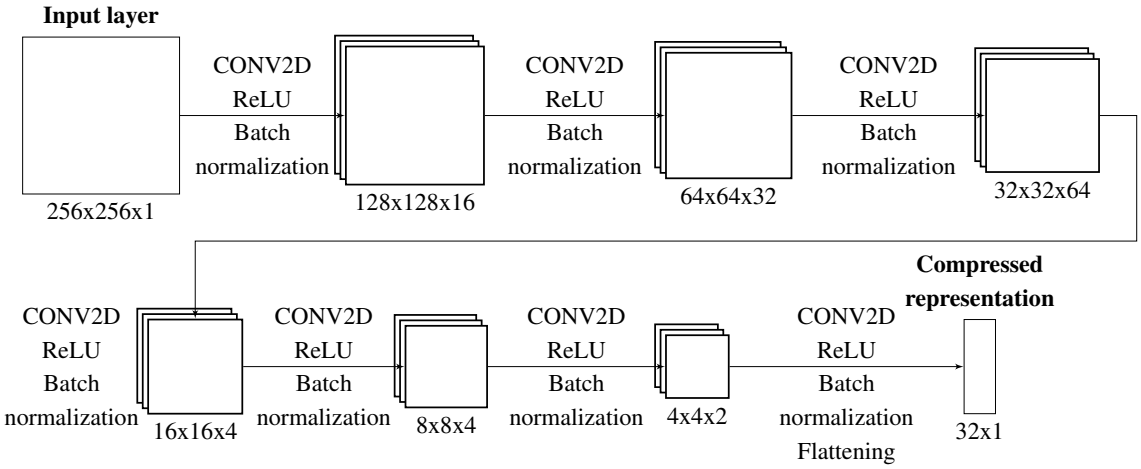


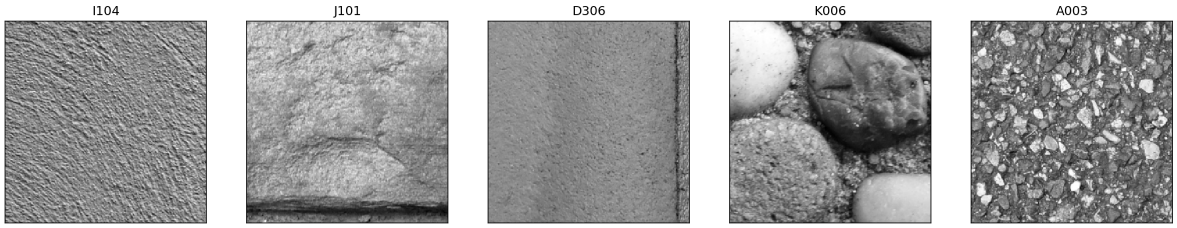
Figure 6.11: The encoder architecture of the CAE for one image patch. The encoder consists of eight layers, whereas one layer consists of a 2D convolutional layer with a linear activation, a batch normalization and an additional ReLU activation layer. The last encoder layer consists of reshaping (flattening) the layer into a 32-dimensional vector.

As the imagery contains images acquired at different detail levels, i.e. acquisition distance of 0.2 - 1.2 m, the extracted features represent distinct information from each scale. Four sub-scenarios (scenarios 5.1 - 5.4) are employed to determine the optimal textural representation of the different scales for urban material classification. These sub-scenarios are extracted from the generated compressed textural representations based on the scale indexing. The four sub-scenarios are defined as follows:

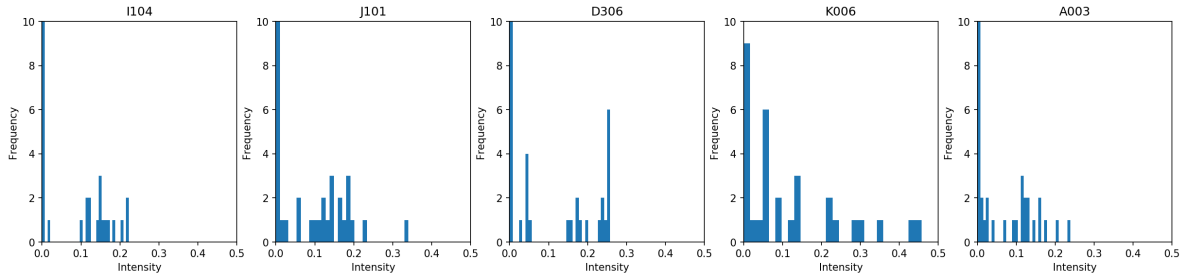
- Scenario 5.1 - scale 1 (acquisition distance of 1.2m)
- Scenario 5.2 - scale 2 (acquisition distance of 0.7m)
- Scenario 5.3 - scale 3 (acquisition distance of 0.2m)
- Scenario 5.4 - all scales

6.2.3 Classification

For the material classification, three supervised classifiers that are available in the Python toolbox *scikit-learn* [185] are employed for the nine scenarios: the classifiers RF [32], HGB [185] and SVM [253]. The parameters of the classifiers are optimally chosen through parameter-tuning. Firstly, the parameters of the classifier RF are set to 100 decision trees, a maximum tree depth of 100 and the number of features to consider for the optimal split to the square root of the number of features. Secondly, the classifier



(a) Original image patches



(b) Textural compressed representation

Figure 6.12: (a): Five image patches at scale 3 representing five samples I204 (*Plaster*), J101 (*Sandstone*), D306 (*Ceramic*), K006 (*Conglomerate*) and A003 (*Asphalt*). (b): The corresponding compressed textural representation expressed as a histogram. It can be observed that the compressed textural representations are expressed differently for the five samples. This indicates that the CAE has achieved distinct expressions.

HGB is set to use a maximum of 100 iterations, a maximum depth of 10, a loss function based on categorical cross-entropy and a learning rate of 0.5. Lastly, the classifier SVM is set to use a radial basis function kernel for the separation of the classes as a simple linear separation would not be enough for high-dimensional data. Furthermore, the regularization parameter is set to 100, the kernel coefficient for the radial basis function to $1/(num(samples) \cdot var(dataset))$ and the regularization parameter for each class as $num(samples)/(num(classes) \cdot bincount(classes))$ to account for the class imbalances.

Some precautions and procedures remain the same for all nine scenarios (as defined in Figure 6.8). Firstly, the spectral and textural features that correspond to the same sample are either used for training or testing. Secondly, due to the class imbalance, a measure is taken to ensure that each material class is present in both the training and testing dataset. A stratified k-fold cross-validation approach is chosen to preserve the material class distribution for training and testing the classifiers. A k-fold cross-validation approach performs the classification several times using different splits which reduces the chance of overfitting the classifier during training. Based on the smallest material class to ensure that each split contains enough samples, $k = 3$ is chosen. The data is split into 80% training and 20% testing data. Lastly, to evaluate the performance of the classifiers, the evaluation matrix and its values are considered. As the k-fold cross-validation approach is chosen, the average evaluation values in addition to the standard deviation σ are calculated and given.

As four scenarios are generated with spectral features, four additional scenarios that merge the spectral and textural features are generated. For each sample, the textural feature is combined with either the corresponding original spectrum, PCA encoding, gradient or retrieved compressed spectral representa-

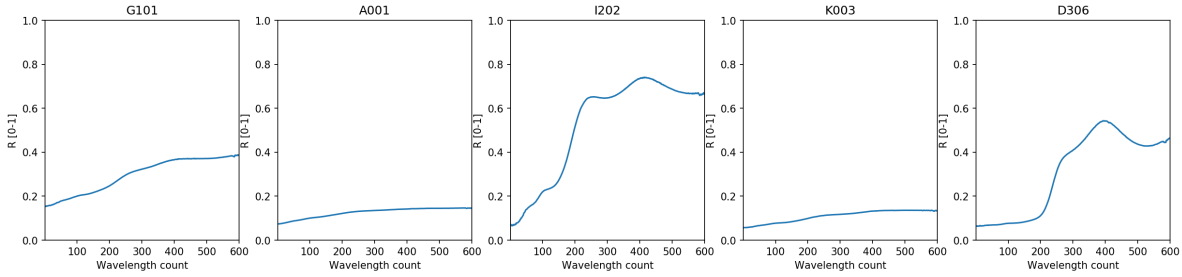
tion (as seen in Figure 6.8). Each textural compressed representation of one image patch is assigned a randomly chosen L_t from the corresponding sample.

6.2.4 Results and Discussion

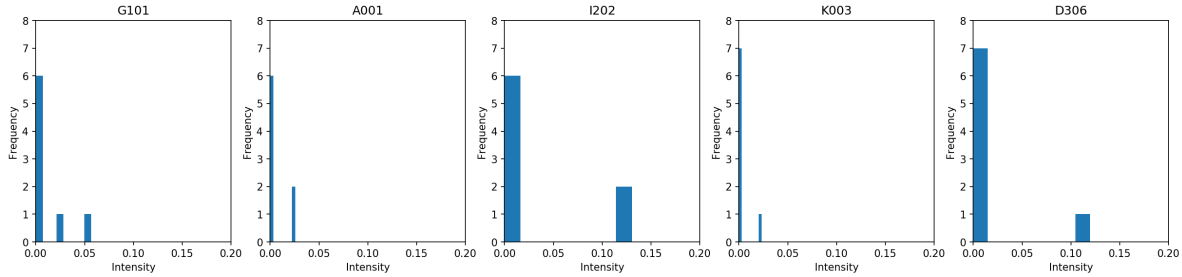
To retrieve a compressed representation, an AE requires both training and testing data which PCA does not. Thus, depending on the size and architecture of the deep AE, the compressed spectral representation requires significantly more time and computing resources than the standard PCA algorithm. Furthermore, the size of the compressed representation has to be chosen to achieve the desired compression ratio. The sizes of the spectral and textural compressed representations were assessed by analyzing and comparing them with respect to two aspects. Firstly, how the size of the compressed representation influenced the AE's ability to decode the compressed representation and secondly, how the size influenced the classification results. Based on these observations, it can be concluded that the size of the compressed representation influences the lossiness of decoded representation more than the classification results. The smaller the compressed representation, the higher the loss and consequently, a lossier decoded representation. However, the difference between using a small and a large compressed representation as input to a classifier does not significantly alter the final results. Smaller compressed representations were opted for to be comparable with the size of the PCA encoding and to achieve a balanced ratio between spectral and textural features.

Figure 6.13 displays the deep AE's ability to encode and decode the original spectra in the spectral range of 350 - 949 nm for five samples. A visual difference can be observed while comparing the original spectra with the corresponding decompressions due to the lossiness of the deep AE. The general spectral characteristic features do remain the same, but smaller deviations can be observed. Furthermore, the compressed spectral representations of the five samples are significantly distinct, here showing the distribution of the intensity, which indicates that they can be utilized for material distinction. However, the spectral features in the first spectral range of 350 - 949 nm have, in comparison to the other spectral sections, a larger spectral variation. Thus, this spectral section required more hidden layers to receive satisfying compressed spectral representations (as likewise observed by the different deep AE architectures).

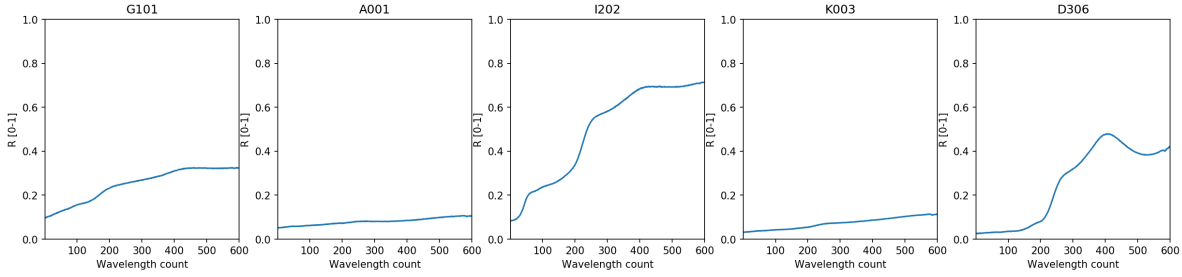
The classification results of the textural features retrieved from the four sub-scenarios 5.1 - 5.4 (which represent the textural features at different scales) are assessed and the results can be seen in Table 6.3. The classification results indicate no significant difference between the four sub-scenarios, as the overall accuracy O is within a small range. The average overall accuracy O (excluding the variation) for the four sub-scenarios is 51% for sub-scenario 5.1, 58% for sub-scenario 5.2, 62% for sub-scenario 5.3 and 56% for sub-scenario 5.4. The overall accuracy O is close to the same for all sub-scenarios. However, the other calculated measures indicate that sub-scenario 5.3 is the optimal representation. The calculated confusion matrices using the results from the classifier SVM for one cross-validation cycle can be seen in Figure 6.14. Here, the cycle with the highest average overall accuracy O across all four sub-scenarios was chosen for comparison of the confusion matrices. Likewise, it can be observed that the sub-scenarios achieve very similar results as the label prediction distributions are similar. Even if sub-scenario 5.3 does achieve better results than the other sub-scenarios, the measures indicate that textural features of one class are frequently mistaken as another class. This can be observed by the high variation for the remaining average evaluation scores recall $\overline{R}$, precision $\overline{P}$ and $\overline{F_1}$-score. This is due to frequent misclassification of the material classes *Asphalt*, *Plaster*, *Sandstone* and *Conglomerate*. This indicates that there is a high



(a) The original spectrum



(b) The spectral compressed representation



(c) The decoded compression

Figure 6.13: (a): The original spectrum in the spectral range of 350 - 949 nm for the five samples G101 (*Limestone*), A001 (*Asphalt*), I202 (*Plaster*), K003 (*Conglomerate*) and D306 (*Ceramic*). (b): The corresponding spectral compressed representation expressed as a histogram. (c): The corresponding decoded compression. Firstly, it can be observed that the compressed spectral representations are expressed differently for the five samples. This indicates that the deep AE has achieved distinct expressions. Secondly, a visual difference can be observed while comparing the original spectra with the corresponding decompressions which is due to the lossiness of the deep AE.

inter-class textural similarity between these material classes. Nonetheless, sub-scenario 5.3 is chosen to represent compressed textural features.

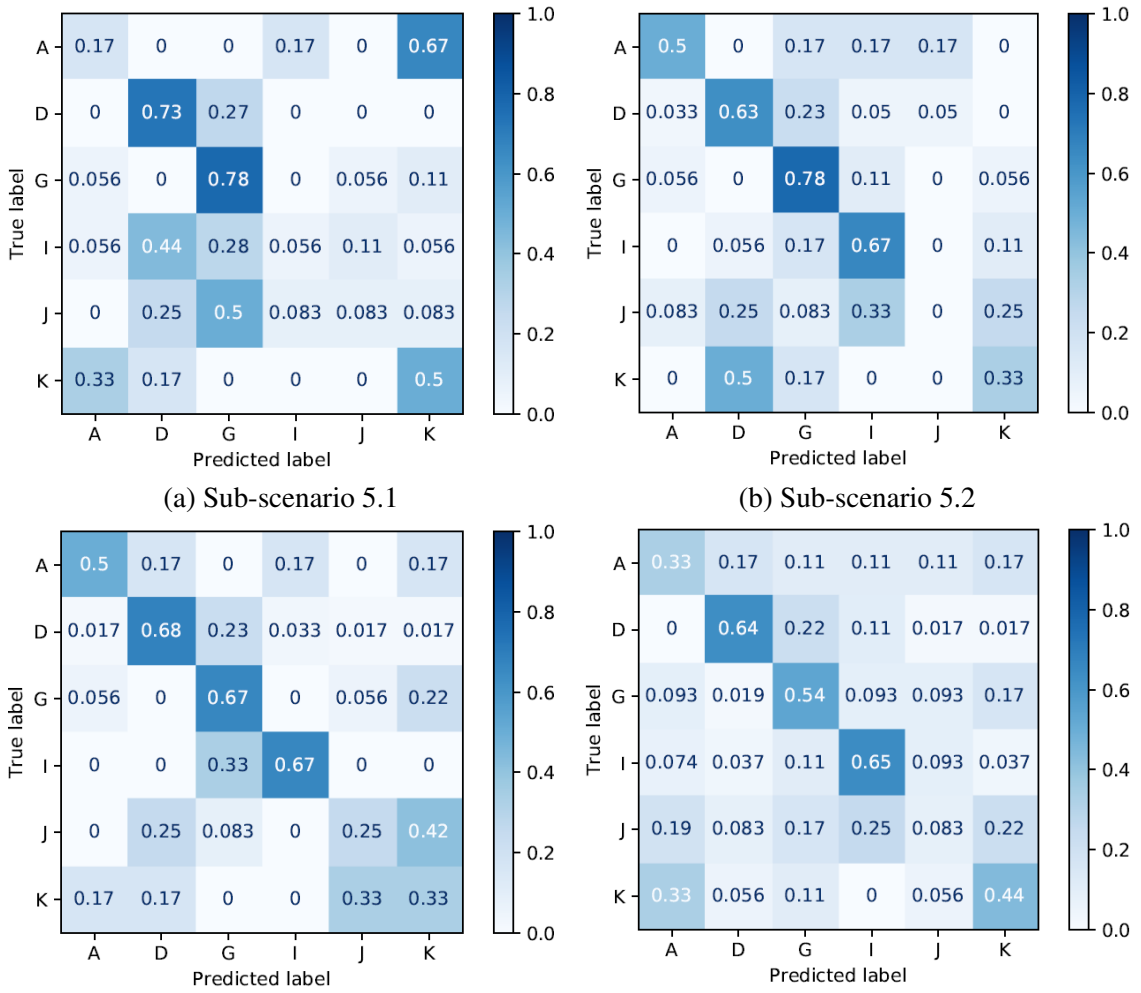
The material classification results with the three classifiers reveal, in general, an overall accuracy O within the range of 70 - 80% as seen in Table 6.4. As spectral characteristic features are similar for urban materials and an inter-class correlation exists, the overall accuracy O around this range is expected. The overall accuracy O differs for the three major scenario types. While analyzing the classification results using either spectral (scenarios 1 - 4) or textural features (scenario 5), the overall accuracy O is noticeable lower when only the textural features are used. This indicates that the spectral features are more important for material classification. Furthermore, the other evaluation scores support this statement as those are

Table 6.3: The achieved classification results for the four sub-scenarios 5.1 - 5.4 with either the classifier RF, HGB or SVM. The evaluation scores display the mean value in addition to the variation expresses as the standard deviation $\pm\sigma$. Highlighted are the best overall accuracy O and the κ-value.

		Textural features			
		5.1	5.2	5.3	5.4
RF	O [%]	53±2	58±3	**63±5**	56±2
	κ [%]	31±4	39±4	**45±6**	37±3
	$\overline{R}$ [%]	33±34	40±30	44±32	38±27
	$\overline{P}$ [%]	35±30	50±28	60±31	43±20
	$\overline{F_1}$ [%]	31±29	38±28	43±29	38±23
HGB	O [%]	50±1	58±1	**60±4**	56±2
	κ [%]	28±26	40±3	**42±5**	37±3
	$\overline{R}$ [%]	30±29	40±28	43±28	40±26
	$\overline{P}$ [%]	30±24	42±23	51±26	40±21
	$\overline{F_1}$ [%]	29±26	39±25	44±25	41±22
SVM	O [%]	50±4	59±4	**64±4**	55±1
	κ [%]	31±2	43±5	**51±4**	40±1
	$\overline{R}$ [%]	33±29	45±26	52±24	42±21
	$\overline{P}$ [%]	35±28	45±22	55±21	41±24
	$\overline{F_1}$ [%]	32±27	43±23	52±22	41±22

lower for scenario 5. Combining spectral and textural features (scenarios 6 - 9) does not improve the overall accuracy O. Contrary, it decreases slightly.

The difference between the four scenarios 1 - 4 is minor, as the evaluation scores are close to the same for all three classifiers across all four scenarios. Scenario 1 (the original spectrum) does achieve slightly lower evaluation scores with a larger variation compared to the other scenarios. The two scenarios with dimensionality-reduced spectral representations, scenarios 2 (PCA encoding) and 4 (compressed spectral representation) achieve similar evaluation scores. The variation and the average scores are within a small comparable range. Scenario 3 (gradients) provide comparable scores and variations as scenarios 2 and 4. By comparing the two approaches that express the spectral differences (scenarios 1 and 3), gradients appear to retrieve better evaluation scores. Secondly, the textural features acquired from scenario 5 represent the compressed textural features from scale 3 and achieve the lowest evaluation scores. They are however more robust than those from scenarios 1 - 4 since the standard deviation σ (the variation) is lower. Lastly, the scenarios 6 - 9 represent the combination of spectral and textural features. The overall accuracy O is slightly lower those for scenarios 1 - 4. This either implies that the spectral and textural features are not comparable or that the assignment of the spectral and textural feature pairs is not suitable. Nonetheless, this indicates that adding the textural features to a spectral dataset does not improve the distinction of classes with similar spectral features, such as *Asphalt* and *Conglomerate*. This is further supported by analyzing the most important features according to the classifier RF for scenarios 6 - 9. For the two scenarios with a higher ratio of spectral features, scenarios 6 and 8, 99% of the features that the classifier RF used were spectral features (98% of the input features are spectral features). Scenario 7 has a larger ratio of textural features but the classifier RF still used around 62% spectral features (76% of the input features are textural features). The ratio between the spectral and textural features are the same for scenario 9 (50%). Nonetheless, the classifier RF used around 69% spectral features. Even when the ratio is the same, the spectral features provide more essential information according to the RF classifier.



(a) Sub-scenario 5.1

(b) Sub-scenario 5.2

(d) Sub-scenario 5.3

(e) Sub-scenario 5.4

Figure 6.14: The confusion matrices for the four sub-scenarios 5.1 - 5.4 with the classifier SVM. The cycle with the highest average overall accuracy O across all four sub-scenarios is here shown. The material class *Sandstone* (J) has no correctly predicated samples in sub-scenario 5.2 which causes a calculation of 0 for the evaluation scores recall R, precision P and the F_1-score.

For scenarios 7 and 9, the variation is in general slightly larger than for the other scenarios which further indicates that the spectral and textural features are not completely comparable.

A noticeable observation is a large variation across all scenarios (larger standard deviation σ). This is due to frequent misclassification of certain material classes. By observing the evaluation scores R, precision P and the F_1-score for scenarios 1 - 4, the material classes *Asphalt*, *Plaster*, *Sandstone* and *Conglomerate* are frequently incorrectly classified. For scenarios 6 - 9, the material classes *Asphalt*, *Sandstone* and *Conglomerate* are frequently incorrectly classified. This indicates that the distinction of the material class *Plaster* is improved by adding textural features. The material classes *Ceramic* and *Limestone* appear to have both distinct spectral and textural features, as those are infrequent misclassified.

The three classifiers achieve evaluation scores within similar ranges. The calculated confusion matrices using the results from the classifier SVM for one cross-validation cycle can be seen in Figure 6.15. The cycle with the highest average overall accuracy O across all scenarios was chosen for comparison of the

confusion matrices. The class distributions across the nine scenarios are for the majority similar except for scenario 5 (as likewise explained by the evaluation scores in Table 6.4).

Table 6.4: The achieved classification results for the nine scenarios 1 - 9 with the classifiers RF, HGB and SVM. The evaluation scores display the average value in addition to the variation expresses as the standard deviation $\pm\sigma$. Highlighted are the best overall accuracy O and the κ-value.

		Spectral features				Textural features	Spectral and textural features			
		1	2	3	4	5	6	7	8	9
RF	O [%]	75±7	**78±8**	75±4	**80±6**	58±12	73±5	75±4	75±3	**76±5**
	κ [%]	66±10	**69±11**	65±5	**72±9**	39±12	63±7	65±5	65±5	**67±7**
	$\overline{R}$ [%]	62±37	66±34	64±35	66±33	40±31	59±32	62±32	62±34	64±32
	$\overline{P}$ [%]	71±36	79±27	68±30	78±31	63±29	72±33	70±32	61±28	71±29
	$\overline{F_1}$ [%]	58±37	65±27	60±30	66±33	40±26	55±36	62±28	59±29	63±30
HGB	O [%]	73±10	**77±5**	71±6	**73±3**	60±14	70±4	**76±8**	70±6	68±2
	κ [%]	62±14	**67±6**	58±11	**64±5**	44±15	57±6	**65±11**	57±10	56±5
	$\overline{R}$ [%]	58±37	64±36	59±32	64±34	48±27	51±36	63±39	53±36	55±37
	$\overline{P}$ [%]	69±33	70±34	62±24	72±35	55±26	45±33	73±31	56±32	72±36
	$\overline{F_1}$ [%]	56±35	70±34	55±26	62±32	48±24	47±33	61±34	50±32	54±35
SVM	O [%]	**79±9**	**79±10**	**78±2**	76±4	62±13	75±7	76±7	76±5	65±10
	κ [%]	**70±13**	**71±14**	**69±4**	67±6	47±15	66±9	67±9	67±8	50±12
	$\overline{R}$ [%]	71±30	72±32	65±30	67±32	48±30	67±26	68±27	66±25	53±27
	$\overline{P}$ [%]	77±29	77±29	68±27	74±31	50±29	69±29	70±28	67±26	58±26
	$\overline{F_1}$ [%]	68±24	68±24	63±25	64±29	47±15	64±23	65±23	64±22	52±25

The decoded compressed spectral representation is lossy since a visual difference between the original spectrum and the decoded compression can be observed in Figure 6.13. The decoded compressed textural representation is however even lossier. As seen in Figure 6.16, the decoded compression is noticeable lossy as the visual characteristics are no longer distinct. As detailed observed and further analyzed, the size of the compressed representation influenced the lossiness of the decoded compression. When the size of the compressed representation was larger (i.e. 1024 instead of 32), the lossiness of the visual characteristics was much smaller and the characteristics were significantly more distinct. In short, due to utilizing a larger compression factor (24,576), the visual lossiness is apparent. Furthermore, the deep denoising CAE was trained using a total of 52,665 parameters. Contrary, the deep AEs used to retrieve spectral features utilized a significantly larger amount of parameters (882,784, 168,263, 173,393 and 173,906 respectively). As the deep denoising CAE is using significantly fewer parameters for the training, it is noticeable lossier and not optimally structured. The architecture of the deep denoising CAE could be improved to extract more parameters. Nonetheless, this case study focuses on the compressed textural representation and not on the decoded compression, the visual difference is therefore acceptable. The compressed textural representations of the samples do differ (as seen in Figure 6.12), indicating that distinct textural features have been retrieved and can be further used for the class assessment.

While the usefulness and advantages of the spectral approaches could be compared (scenarios 1 - 4), only one approach for extracting textural features was used (scenario 5). An evaluation of the approach was therefore not done. Traditionally, texture analysis is done for an entire image (e.g. as done in the case study presented in Chapter 6.1). Depending on the chosen texture analysis, the texture is expressed through different manners. A comparison of textural features retrieved from either traditional approaches or an AE would be of interest to determine if the machine learning approach has an advantage. Further-

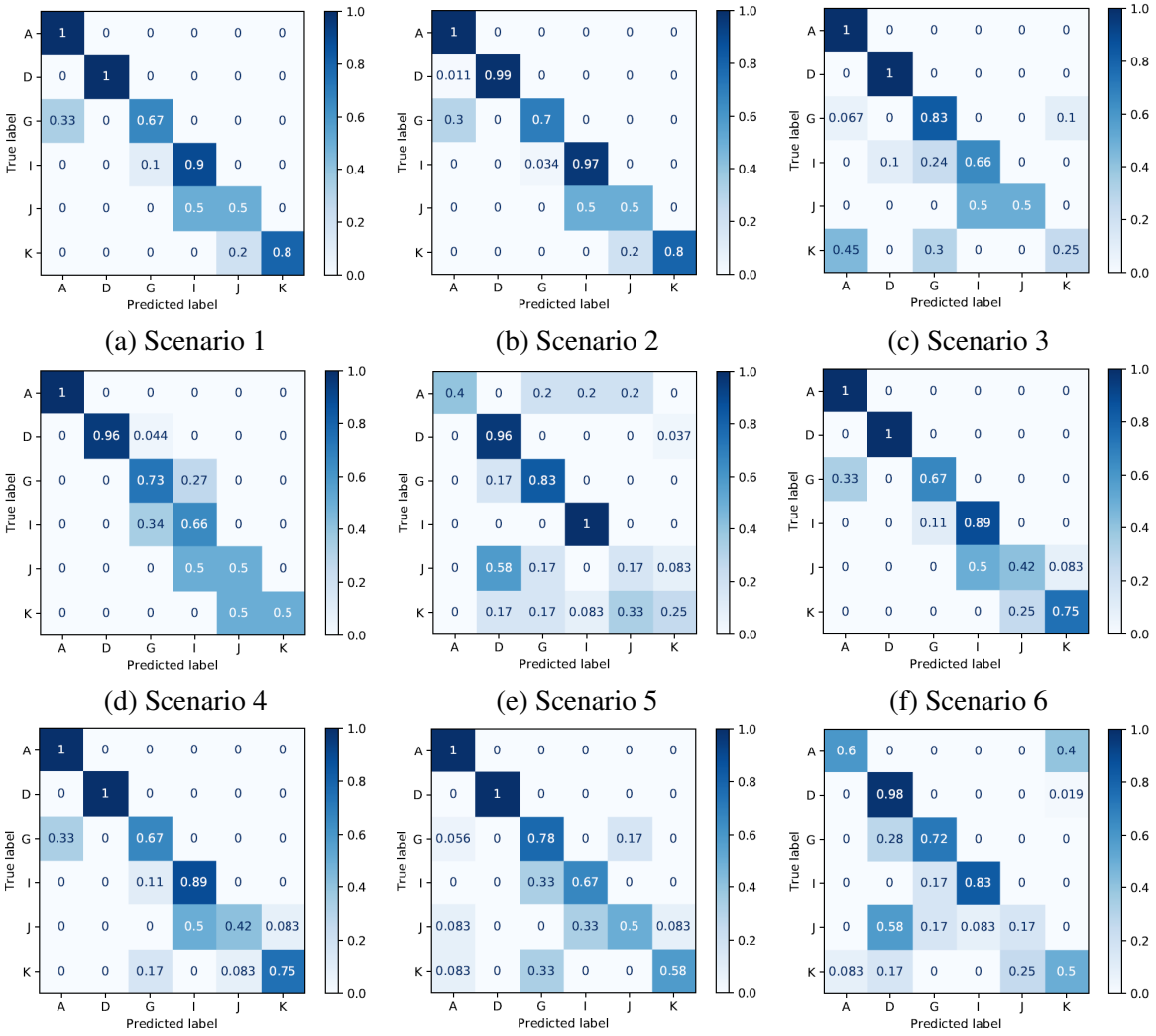
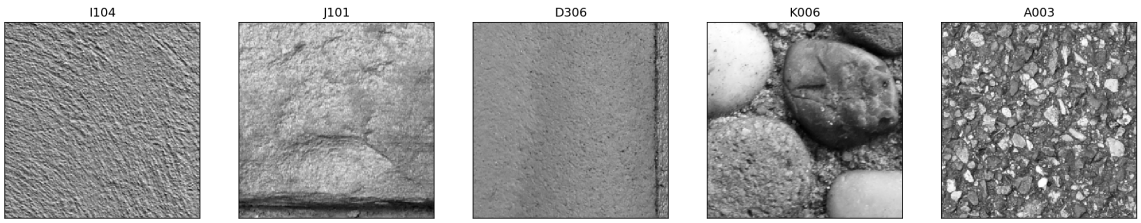


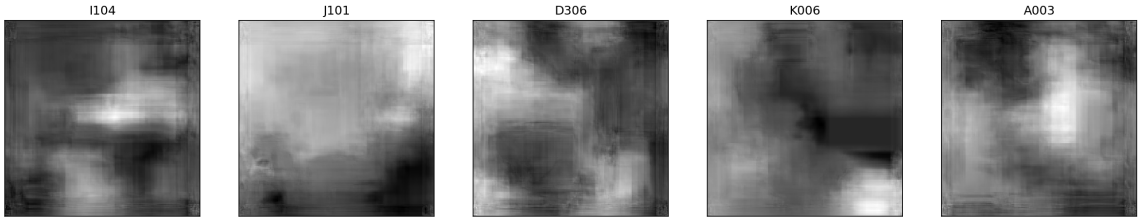
Figure 6.15: The confusion matrices for the nine scenarios using the classifier SVM. The cycle with the highest average overall accuracy O across all nine scenarios is here shown.

more, the acquisition distance (the scale) could be added as an additional layer which would allow the CAE to encode this information and enforce a scale difference.

In general, the occasionally occurring large standard deviation σ, the lower evaluation scores and an overall accuracy O within the range of 70 - 80% across all classification approaches indicate that the size of the dataset is small for this classification task. The dataset does contain a class imbalance, around 46% of the samples belong to the class *Ceramic* while the remaining five classes contain around 4 - 9 samples (7 - 15%). This class imbalance is further prominent during the training of either an AE or a classifier. As the split between training, testing and validation data is randomly chosen, the smaller classes are trained with few samples. If the class itself contains a distinct spectral or textural variation, it cannot be covered through such splits. To counter the class imbalance, the study could have focused on distinguishing material sub-classes instead of material classes since those samples are better distributed in the dataset. This would however consist of an intra-class assessment instead of an inter-class assessment.



(a) Original image patches



(b) Decoded compressed textural representation

Figure 6.16: (a): Five image patches representing five different samples I204 (*Plaster*), J101 (*Sandstone*), D306 (*Ceramic*), K006 (*Conglomerate*) and A003 (*Asphalt*). (b): The corresponding decoded compressed textural representation. It can here be observed that the decoded compressed textural representations are very lossy as they are not visual alike the original image patches due to utilizing a large compression factor of 24,576.

6.3 Summary

This chapter has presented two case studies that have assessed materials using both spectral and textural features from either airborne or close-range acquired data. The first case study assessed roof material classification on three segmentation levels using airborne acquired data. The spectral features were represented as multispectral and TIR data while the textural features were expressed as either statistical and transform-based textures using high-resolution acquired imagery. The second case study consisted of assessing the close-range acquired spectral library KLUM by analyzing different approaches of expressing spectral features in combination with texture features. The spectral features were represented as either the original spectra, PCA encodings, gradients or spectral compressed representations while the textural features were expressed as textural compressed representations.

The two case studies conclude that spectral features are more important than textural features for the assessment of urban materials. This is firstly supported by the evaluation scores which do not significantly increase when the two types of features are combined. Furthermore, by analyzing the evaluation scores retrieved from the classifier RF, the majority of the used features are spectral features. Additionally, by analyzing the evaluation scores and their variation from the second case study, it appears that spectral and textural features are not always comparable. An intra-class textural dissimilarity can likewise be established in addition to the already established intra-class spectral dissimilarity.

Due to the small datasets in both case studies, the class imbalances are prominent during the training of either the classifiers or the AEs. A rigorous balancing of all classes would either further decrease the number of samples or the number of considered classes. The conclusions drawn from this chapter are specific to the two setups. Future work should address the acquisition of larger datasets to include

a larger variation of both spectral and textural features. Nonetheless, the case studies can imply that the combination of spectral and textural features does not significantly improve the distinction of urban materials with similar spectral features.

7 Conclusions and Outlook

This thesis has presented solutions and challenges that concern urban material classification with spectral and textural features by using both close-range and long-range acquired data. This chapter summarizes and discusses the results, contributions and conclusions. Potential future work based on the observations is likewise presented.

7.1 Summary of Contributions

Based on the observations made throughout the thesis, these are the most important and critical conclusions that can be drawn concerning assessing urban materials.

- **Used taxonomy and labeling**
  As observed in Tables 3.13 and 3.14, the taxonomy and labeling of the existing spectral libraries are not consistent. As no standardized taxonomy exists for spectral libraries, the creators of the spectral libraries establish their own system. The comparability is therefore limited since the labeling of samples is mostly based on the α diversity, i.e. the locally available materials. A material sample can be described on a high level which is done for libraries with a focus on land cover classes, where the lowest or second-lowest rank describes the actual urban material. On the other hand, a material sample can be described on a low level which is done for the majority of close-range acquired libraries where the material is described on the highest rank. Hence, it is challenging to assure that the samples that have been labeled as the material from different libraries are truly the same in terms of e.g. chemical composition and surface coating. To establish a standardized system, libraries should label the samples using a β diversity approach to account for regional and global differences.
- **Spectral features**
  Spectral features have been expressed differently to assess the material distinction, as either the original spectra, PCA encodings, CFS encodings, gradients and compressed spectral representations. It can be concluded through the observations that spectral features achieve the highest overall accuracy O when they are expressed as either PCA encodings, gradients and compressed spectral representations (e.g. as seen in Tables 4.3 and 6.3). Nonetheless, it can be observed across all studies that spectral features exhibit intra- and inter-class spectral similarities and dissimilarities respectively. To enhance material distinguishing, further information is needed. Additionally, the spectral range importance has been further studied. The analysis concludes that the upper part of the spectral range of SWIR contains distinct spectral features and contributes therefore to improved material classification. The visible and VNIR spectral ranges contain color information and as the majority of urban and building materials are not defined with one distinct color, limiting the distinction to one of these ranges is not suitable.

- **Textural features**
  To assess textural features for material distinction, they have in this thesis been as described either classical textures retrieved from statistical and transform-based methods or compressed textural representations derived from a machine learning approach. Nonetheless, the achieved overall accuracy O seen in Tables 6.2 and 6.3 demonstrate that textural features are not bound to specific urban materials and thus not distinct, as the overall accuracy O and the additional evaluation measures are low. Based on the presented case studies, they imply that textural features of different material classes are dissimilar since they are incapable of distinguishing the material classes.
- **Spectral and textural features**
  By combining spectral and textural features to assess urban materials, the classification results imply that spectral features are more important. This is firstly supported by the classification results presented in Tables 6.2 and 6.3. The overall accuracy O does not significantly increase when both features are used. Contrary, in comparison with scenarios where only the spectral features are used, the overall accuracy O occasionally decreases. Secondly, this is further supported by analyzing the most important features according to the classifier RF. Across all scenarios that used both kinds of features, the spectral features were deemed as more important. The two scenarios 7 and 9 presented in Chapter 6.2 used a ratio of 76/23% and 50/50% textural/spectral features respectively. Nonetheless, the features the classifier RF used for those two scenarios 62% and 69% spectral features respectively. Even if textural features are well represented, spectral features are considered more important.
- **Inter-class similarity**
  The previously known inter-class spectral similarity could in this thesis be strongly established, which is both unexpected and concerning. Figure 5.10 displays the five samples from ASTER and LUMA most similar to the average spectrum of the sub-class *Grey asphalt* according to the three spectral similarity scores. The spectral similarity can likewise be observed visually, as the samples appear noticeably alike. Even if more than five samples representing *Asphalt* were given by ASTER and LUMA, the five most similar samples corresponded to the material classes *Limestone*, *Ceramic*, *Asphalt* and *Cement* (presented in the ranked order). This is likely due to either incorrect and insufficient labeling or similar chemical compositions and indicates that the spectral characteristic features of different materials are remarkably similar, increasing the probability of misclassification. This is likewise supported by the results retrieved from the assessment of KLUM, as seen in Figures 5.11, 5.12 and 5.13. Additionally, Figure 6.14 further highlights the strong inter-class similarity, both on a spectral and textural level. To conclude, it is challenging to distinguish material classes with both spectral and textural features.
- **Intra-class dissimilarity**
  This study can establish an intra-class dissimilarity by examining spectra from different existing libraries and the developed KLUM. As observed in Figure 3.1, spectra representing the materials *Asphalt* and *Concrete* retrieved from four publicly available spectral libraries vary. This is likewise observed in e.g. Figures 5.8 and 5.7 which display the spectra of different sub-classes that belong to the same material class of KLUM. The intra-class dissimilarity is likely due to either different surface structures, textures and coatings or incorrect labels. This is likewise observed on a spectral and textural level. For studies where both the intra-class dissimilarity and the inter-class similarity are significant, it further increases the chance of misclassification.

- **Datasets**
  Overall, one major weakness is a lack of data as the available and utilized datasets throughout this thesis have all been small. Firstly, the case study presented in Chapter 4.4 acquired 47 samples across 12 material classes where the average class size was approximately 4. However, only 21 samples (210 spectra) were used for the actual classification. The datasets used for the study areas presented in Chapter 6.1 were likewise small. The smaller area contained firstly 10 labeled buildings but was extended to 56 buildings while the larger study area consisted of 426 labeled buildings. The dataset used in the case study in Chapter 4.2 used 61 samples (606 spectra) and 185 images from the spectral library KLUM. Due to these smaller datasets, this thesis cannot draw any strong and rigorous conclusions as the datasets are not large enough. A small dataset limits the possibility of training the classifiers since it does not cover a large variety of spectral and textural features. Furthermore, any possible class imbalance is then more prominent. Depending on how the dataset is split into training and testing data, a major class imbalance can influence the final overall accuracy O as the classifier may not have enough varied samples for the training of the smaller minor class. The majority of the used datasets did contain a class imbalance, thus enhancing the varied overall accuracy O.

7.2 Outlook

The remaining challenges and future work that can be concluded based on the research done in this thesis are here presented. An outlook beyond the scope of this thesis regarding general research within the field of urban materials and remote sensing is also given.

7.2.1 Remaining Challenges and Potential Future Research

The following aspects should be further explored.

- **Extending KLUM**
  The spectral library KLUM and its high-resolution imagery would benefit from an extension of at least e.g. 1000 more samples. This should firstly include acquiring further samples to balance the class distribution since a class imbalance does exist. Secondly, this should include acquiring more samples of each class and sub-class to extend the variation. By increasing the variation, the relationship between spectral and textural features in addition to the intra-class and inter-class similarities could be further studied. Lastly, KLUM should be extended to include more types of urban materials as more than 12 kinds of urban material exist in the city of Karlsruhe.
- **Improving measurement setup for spectra acquisition**
  The measurement setup used in Chapter 4.4 was further extended and improved during the acquisition of KLUM (as described in Chapter 5.1). The measurement setup consisted of acquiring the spectra with the spectroradiometer FieldSpec-4, the incoming solar irradiance with the spectrometer Qmini, the geolocation with a GNSS receiver and the imagery with a standard RGB camera. The setup could be further extended and improved by acquiring information about e.g. the viewing angle upon the sample surface in relationship with the direction of the sun, the acquisition angle and the acquisition distance between the sample surface and sensor. The additional information

could be used to determine further factors, such as the BRDF. Furthermore, the measurement setup could be improved by assuring that the spectrometer is constantly facing the sun directly.

- **Improving classification scheme and labeling with a standardized system**
  The classification scheme of KLUM is not consistent. Few sub-classes are described with the color, others with both the color and the surface structure/texture/coating, others with only the surface structure/texture/coating and some with the usage or shape (e.g. *Block*). The taxonomy could therefore be refined to retain the consistency by basing it on β diversity. Furthermore, the scheme of KLUM differs from those presented in Tables 3.13 and 3.14. The remote sensing community would benefit from a standardized spectral library classification scheme to make the labeling comparable. Future work could consist of further analyzing the used labeling and propose a standardized classification scheme and labeling system, i.e. standardized taxonomy, with a β diversity.
- **Assessing the compressed textural representation**
  The approaches used to describe textural features were not thoroughly compared. The compressed textural representations retrieved from the AE used in the case study presented in Chapter 6.2 were not assessed nor compared with standardized texture analysis approaches such as entropy [92] and *Canny* edge detection algorithm [36] (presented in Chapter 6.1). A further study could include an assessment of the different textural expression features for the assessment of the same dataset to determine the usefulness of the machine learning approach.
- **Further analyzing intra-class and inter-class similarities**
  As both the intra-class dissimilarity and the inter-class similarity were unexpectedly high and prominent, they should be further studied. This could include a study that only focuses on one particular material by exploring the spectral and textural differences, i.e. the intra-class relationship between spectral and textural features. One material is not defined nor bound to specific colors nor surface structures, textures and coatings. On the contrary, one material is usually defined by its chemical components. Nonetheless, the spectral ranges that have been used for this thesis are not sufficient for the distinction as the exterior influences the measured spectrum. The relationship could therefore be explored in the further ranges, e.g. in the range of LWIR [124]. Furthermore, a significantly larger dataset is needed to properly analyze the relationship to cover a larger variety of materials and reduce the class imbalance. To improve the intra-class dissimilarities and the inter-class similarities, further features for material classification could likewise be analyzed. This could include an expanded dataset with both spectral and textural features as those presented in Chapter 6 are limited in size. Additional features could include 3D features, as those have been used to classify materials [87, 264].

7.2.2 Outlook on Automated Remote Sensing of Urban Materials

With the recent advantages of UAV and consequently the small and light-weighted hyperspectral sensors, new kinds of close-ranged scene analyzes are possible. The following aspects could be further explored regarding the automatic classification of urban materials with UAVs within the remote sensing community.

- **Urban material inventory**
- Cities, such as the City of Melville in Perth, may not have an overview of which urban materials that are present in their city. Through the automatic determination of urban materials, the cities can retrieve a material inventory that can be further used as input to 3D city models.
- **Material recycling**
  For a sustainable future, urban materials should be reused. By knowing the existing urban materials in a city, certain materials can be reused instead of newly produced.
- **Heating and cooling demand of buildings**
  The heating demand of buildings is usually determined by using only TIR data to detect thermal bridges but can be combined with material data to calculate the U-value (under the assumption that the materials in the facade are known) [74, 83]. The cooling demand of buildings could consist of detecting facade materials that absorb the incoming radiance.
- **Comfort evaluation in the urban environment**
  To evaluate the comfort of a location, the information about urban materials can contribute to certain parameters. Materials that absorb incoming radiance may be undesirable for cities that have a warm climate while those could be desirable for cities with a cold climate.
- **Analysis of UHI**
  For reliable thermal simulations with 3D city models to study the behavior of an UHI, knowledge about the existing material and the corresponding albedo are needed. With UAVs, the building facade materials can be determined which are usually not detectable with airborne sensors. The material data from UAVs can then be combined with the airborne material data for both local and regional analysis of an UHI.

Bibliography

[1] J. S. Abel and J. Smith. Restoring a clipped signal. In *Proceedings of the ICASSP 91: 1991 International Conference on Acoustics, Speech, and Signal Processing*, pages 1745–1748. IEEE, 1991.

[2] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282, 2012.

[3] S. Aggarwal. Principles of remote sensing. *Satellite Remote Sensing and GIS Applications in Agricultural Meteorology*, pages 23–38, 2004.

[4] R. D. Aines and G. R. Rossman. Water in minerals? A peak in the infrared. *Journal of Geophysical Research: Solid Earth*, 89(B6):4059–4071, 1984.

[5] R. A. W. Albers, P. R. Bosch, B. Blocken, A. A. J. F. Van Den Dobbelsteen, L. W. A. Van Hove, T. J. M. Spit, F. van de Ven, T. Van Hooff, and V. Rovers. Overview of challenges and achievements in the climate adaptation of cities and in the Climate Proof Cities program. *Building and Environment*, 83:1–10, 2015.

[6] G. L. Alexandrino, J. M. Amigo, M. R. Khorasani, J. Rantanen, A. V. Friderichsen, and R. J. Poppi. Unveiling multiple solid-state transitions in pharmaceutical solid dosage forms using multi-series hyperspectral imaging and different curve resolution approaches. *Chemometrics and Intelligent Laboratory Systems*, 161:136–146, 2017.

[7] J. R. Anderson. Land-use classification schemes. *Photogrammetric Engineering*, 37(4):379–387, 1971.

[8] J. R. Anderson, E. Ernst, J. Hardy, J. T. Roach, and R. E. Witmer. *A land use and land cover classification system for use with remote sensor data*, volume 964. U.S. Government Printing Office, 1973.

[9] E. Angelopoulou, S. W. Lee, and R. Bajcsy. Spectral gradient: A material descriptor invariant to geometry and incident illumination. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 2, pages 861–867. IEEE, 1999.

[10] C. D. Anger, S. Mah, and S. K. Babey. Technological enhancements to the compact airborne spectrographic imager (CASI). In *Proceedings of the 1st International Airborne Remote Sensing Conference and Exhibition*, volume 2, pages 205–213, 1994.

[11] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies. Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1(2):205–220, 1992.

[12] L. Armi and S. Fekri-Ershad. Texture image analysis and texture classification methods - A review. *arXiv preprint arXiv:1904.06554*, 2019.

[13] D. Arthur and S. Vassilvitskii. k-means++: The advantages of careful seeding. Technical report, Stanford, 2006.

[14] A. M. Baldridge, S. J. Hook, C. I. Grove, and G. Rivera. The ASTER spectral library version 2.0. *Remote Sensing of Environment*, 113:711–715, 2009.

[15] I. I. Bashter. Calculation of radiation attenuation coefficients for shielding concretes. *Annals of Nuclear Energy*, 24(17):1389–1401, 1997.

[16] S. Bell, P. Upchurch, N. Snavely, and K. Bala. OpenSurfaces: A richly annotated catalog of surface appearance. *ACM Transactions on Graphics*, 32(4):111:1–111:11, 2013.

[17] S. Bell, P. Upchurch, N. Snavely, and K. Bala. Material recognition in the wild with the Materials in Context Database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3479–3487, 2015.

[18] R. E. Bellman. *Adaptive control processes: A guided tour*. Princeton University Press, 2015.

[19] E. Ben-Dor, N. Levin, and H. Saaroni. A spectral based recognition of the urban environment using the visible and near-infrared spectral region (0.4–1.1 μm). A case study over Tel-Aviv, Israel. *International Journal of Remote Sensing*, 22:2193–2218, 2001.

[20] M. H. Bharati, J. J. Liu, and J. F. MacGregor. Image texture analysis: Methods and comparisons. *Chemometrics and Intelligent Laboratory Systems*, 72(1):57–71, 2004.

[21] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and Remote Sensing Magazine*, 1(2):6–36, 2013.

[22] J. D. Boissonnat and F. Cazals. Smooth surface reconstruction via natural neighbour interpolation of distance functions. *Computational Geometry*, 22(1-3):185–203, 2002.

[23] G. Bonifazi, G. Capobianco, and S. Serranti. Asbestos containing materials detection and classification by the use of hyperspectral imaging. *Journal of Hazardous Materials*, 344:981–993, 2018.

[24] M. Borengasser, W. S. Hungate, and R. Watkins. *Hyperspectral remote sensing: Principles and applications*. CRC Press, 2007.

[25] G. Bouchard and B. Triggs. The tradeoff between generative and discriminative classifiers. In *Proceedings of the 16th IASC International Symposium on Computational Statistics*, pages 721–728, 2004.

[26] D. E. Bowker, R. E. Davis, D. L. Myrick, K. Stacy, and W. T. Jones. Spectral reflectances of natural targets for use in remote sensing studies. Technical Report NASA RP- 1139, NASA Langley Research Center, Hampton, VA, United States, 1985.

[27] R. N. Bracewell. *The Fourier transform and its applications*. McGraw-Hill New York, 1986.

[28] R. J. Brachman. What is-a and isn't: An analysis of taxonomic links in semantic networks. *Computer*, 16(10):30–36, 1983.

[29] L. Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996.

[30] L. Breiman. Bias, variance, and arcing classifiers. Technical Report 460, Statistics Department, University of California, Berkeley, CA 94720 US, 1996.

[31] L. Breiman. Some properties of splitting criteria. *Machine Learning*, 24(1):41–47, 1996.

[32] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[33] D. Bulatov, E. Burkard, R. Ilehag, B. Kottler, and P. Helmholz. From multi-sensor aerial data to thermal and infrared simulation of semantic 3D models: Towards identification of urban heat islands. *Infrared Physics and Technology*, 105:103233:1–15, 2020.

[34] R. G. Burns. *Mineralogical applications of crystal field theory*, volume 5. Cambridge University Press, 1993.

[35] R. G. Burns. Origin of electronic spectra of minerals in the visible and near-infrared region. *Remote Geochemical Analysis*, pages 3–29, 1993.

[36] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, 1986.

[37] B. Caputo, E. Hayman, and P. Mallikarjuna. Class-specific material categorisation. In *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV'05)*, volume 1, pages 1597–1604. IEEE, 2005.

[38] M. E. Celebi, H. A. Kingravi, and P. A. Vela. A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems with Applications*, 40(1):200–210, 2013.

[39] U. Cevik, N. Damla, R. Van Grieken, and M. V. Akpınar. Chemical composition of building materials used in Turkey. *Construction and Building Materials*, 25(4):1546–1552, 2011.

[40] C. C. Chang and C. J. Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):27:1–27:27, 2011.

[41] C. I. Chang. *Hyperspectral imaging: Techniques for spectral detection and classification*. Springer Science & Business Media, 2003.

[42] T. Chang and C. C. J. Kuo. Texture analysis and classification with tree-structured wavelet transform. *IEEE Transactions on Image Processing*, 2(4):429–441, 1993.

[43] S. Chawla and A. Gionis. k-means: A unified approach to clustering and outlier detection. In *Proceedings of the 2013 SIAM International Conference on Data Mining*, pages 189–197. SIAM, 2013.

[44] C. H. Chen. *Handbook of pattern recognition and computer vision*. World Scientific, 2015.

[45] X. L. Chen, H. M. Zhao, P. X. Li, and Z. Y. Yin. Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. *Remote Sensing of Environment*, 104(2):133–146, 2006.

[46] R. N. Clark. Chapter 1: Spectroscopy of rocks and minerals, and principles of spectroscopy. In *Manual of Remote Sensing*, volume 3, pages 3–58. John Wiley and Sons, New York, 1999.

[47] R. N. Clark, T. V. V. King, M. Klejwa, G. A. Swayze, and N. Vergo. High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysical Research: Solid Earth*, 95(B8):12653–12680, 1990.

[48] R. N. Clark and T. L. Roush. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research: Solid Earth*, 89(B7):6329–6340, 1984.

[49] R. N. Clark, G. A. Swayze, R. Wise, K. E. Livo, T. Hoefen, R. F. Kokaly, and S. J. Sutley. USGS digital spectral library splib06a. Technical Report 231, U.S. Geological Survey, Reston, Virginia, U.S., 2007.

[50] M. A. Cochrane. Using vegetation reflectance variability for species level classification of hyperspectral data. *International Journal of Remote Sensing*, 21(10):2075–2087, 2000.

[51] T. Cocks, R. Jenssen, A. Stewart, I. Wilson, and T. Shields. The HyMapTM airborne hyperspectral sensor: The system, calibration and performance. In *Proceedings of the 1st EARSeL workshop on Imaging Spectroscopy*, pages 37–42. EARSeL, 1998.

[52] P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3):287–314, 1994.

[53] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[54] Council of the City of Meville. Urban forest strategic plan 2017-2036: Plan A: City-controlled plan. Technical Report 1, City of Melville, Perth, Australia, 2017.

[55] A. Criminisi and J. Shotton. *Decision forests for computer vision and medical image analysis*. Advances in Computer Vision and Pattern Recognition, Springer, London, UK, 2013.

[56] G. R. Cross and A. K. Jain. Markov random field texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-5(1):25–39, 1983.

[57] D. Dai, H. Riemenschneider, and L. Van Gool. The synthesizability of texture examples. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3027–3034, 2014.

[58] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics*, 18(1):1–34, 1999.

[59] R. Das and E. Walia. Partition selection with sparse autoencoders for content based image classification. *Neural Computing and Applications*, 31(3):675–690, 2019.

[60] J. Degerickx, A. Okujeni, M.-D. Iordache, M. Hermy, S. van der Linden, and B. Somers. A novel spectral library pruning technique for spectral unmixing of urban land cover. *Remote Sensing*, 9(6):565:1–24, 2017.

[61] S. S. Deshpande, A. B. Inamdar, and H. M. Vin. Spectral library and discrimination analysis of Indian urban materials. *Journal of the Indian Society of Remote Sensing*, 47(5):867–877, 2019.

[62] A. Dimmeler, H. Schilling, M. Shimoni, D. Bulatov, and W. Middelmann. Combined airborne sensors in urban environment. In *Proceedings of Electro-Optical Remote Sensing, Photonic Technologies, and Applications VII; and Military Applications in Hyperspectral Imaging and High Spatial Resolution Sensing*, volume 8897, pages 88970U:1–9, 2013.

[63] H. Ding, X. Wang, Y. Wang, and H. Luo. Ensemble classification of hyperspectral images by integrating spectral and texture features. *Journal of the Indian Society of Remote Sensing*, 47(1):113–123, 2019.

[64] C. Doersch. Tutorial on variational autoencoders. *arXiv preprint arXiv:1606.05908*, 2016.

[65] P. Dollár. Piotr's computer vision Matlab toolbox (PMT). `https://github.com/pdollar/toolbox`, 2016. [Online, last accessed 29.06.2018].

[66] T. Dozat. Incorporating nesterov momentum into Adam. In *Proceedings of the ICLR Workshop 2016*, volume 1, pages 1–4, 2016.

[67] H. Emami, B. Mojaradi, and A. Safari. A new approach for land surface emissivity estimation using LDCM data in semi-arid areas: Exploitation of the ASTER spectral library data set. *International Journal of Remote Sensing*, 37(21):5060–5085, 2016.

[68] H. Enghoff. What is taxonomy?–An overview with myriapodological examples. *Soil Organisms*, 81(3):441–451, 2009.

[69] K. G. J. Fairbarn. Visible-near infrared (VNIR) and shortwave infrared (SWIR) spectral variability of urban materials. Master's thesis, Naval Postgraduate School, Monterey, CA, US, 2013.

[70] V. C. Farmer and J. D. Russell. The infra-red spectra of layer silicates. *Spectrochimica Acta*, 20(7):1149–1173, 1964.

[71] M. Fauvel, J. Chanussot, and J. Benediktsson. A spatial–spectral kernel-based approach for the classification of remote-sensing images. *Pattern Recognition*, 45(1):381–392, 2012.

[72] X. Feng and S. W. Myint. Exploring the effect of neighboring land cover pattern on land surface temperature of central building objects. *Building and Environment*, 95:346–354, 2016.

[73] M. P. Ferreira, F. H. Wagner, L. E. O. C. Aragão, Y. E. Shimabukuro, and C. R. de Souza Filho. Tree species classification in tropical forests using visible to shortwave infrared WorldView-3 images and texture analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 149:119–131, 2019.

[74] P. A. Fokaides and S. A. Kalogirou. Application of infrared thermography for the determination of the overall heat transfer coefficient (U-value) in building envelopes. *Applied Energy*, 88(12):4358–4365, 2011.

[75] M. Franc. Texture library. `http://textures.forrest.cz/index.php?`, 2010. [Online, last accessed 26.03.2020].

[76] J. Franke, D. A. Roberts, K. Halligan, and G. Menz. Hierarchical multiple endmember spectral mixture analysis (MESMA) of hyperspectral imagery for urban environments. *Remote Sensing of Environment*, 113(8):1712–1723, 2009.

[77] J. H. Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, pages 1189–1232, 2001.

[78] R. G. Garrett, C. Reimann, D. B. Smith, and X. Xie. From geochemical prospecting to international geochemical mapping: A historical overview. *Geochemistry: Exploration, Environment, Analysis*, 8(3-4):205–217, 2008.

[79] P. Geladi and B. R. Kowalski. Partial least-squares regression: A tutorial. *Analytica Xhimica Acta*, 185:1–17, 1986.

[80] J. Geng, J. Fan, H. Wang, X. Ma, B. Li, and F. Chen. High-resolution SAR image classification via deep convolutional autoencoders. *IEEE Geoscience and Remote Sensing Letters*, 12(11):2351–2355, 2015.

[81] P. J. Gibson and C. H. Power. *Introductory remote sensing: Principles and concepts*. Psychology Press, 2000.

[82] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock. Imaging spectrometry for earth remote sensing. *Science*, 228(4704):1147–1153, 1985.

[83] D. González-Aguilera, S. Lagueela, P. Rodríguez-Gonzálvez, and D. Hernández-López. Image-based thermographic modeling for assessing energy efficiency of buildings façades. *Energy and Buildings*, 65:29–36, 2013.

[84] I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. MIT Press, 2016.

[85] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, pages 2672–2680, 2014.

[86] J. S. Griffith. *The theory of transition-metal ions*. Cambridge University Press, 1964.

[87] E. Grilli and F. Remondino. Classification of 3D digital heritage. *Remote Sensing*, 11(7):847:1–23, 2019.

[88] H. Gross and U. Thönnessen. Extraction of lines from laser point clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(3):86–91, 2006.

[89] L. Guanter, H. Kaufmann, K. Segl, S. Foerster, C. Rogass, S. Chabrillat, T. Kuester, A. Hollstein, G. Rossner, C. Chlebek, et al. The EnMAP spaceborne imaging spectroscopy mission for Earth observation. *Remote Sensing*, 7(7):8830–8857, 2015.

[90] D. Guidici and M. L. Clark. One-dimensional convolutional neural network land-cover classification of multi-seasonal hyperspectral imagery in the San Francisco Bay Area, California. *Remote Sensing*, 9(6):629:1–25, 2017.

[91] M. A. Hall. *Correlation-based feature subset selection for machine learning*. PhD thesis, Department of Computer Science, University of Waikato, New Zealand, 1999.

[92] R. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6):610–621, 1973.

[93] R. M. Haralick and L. G. Shapiro. *Computer and robot vision*. Addison-Wesley Reading, 1992.

[94] E. Hayman, B. Caputo, M. Fritz, and J. O. Eklundh. On the significance of real-world conditions for material classification. In *Proceedings of the European Conference on Computer Vision*, pages 253–266, 2004.

[95] M. Hebel. *Änderungsdetektion in urbanen Gebieten durch objektbasierte Analyse und schritthaltenden Vergleich von Multi-Aspekt ALS-Daten*. PhD thesis, Technical University of Munich, 2012.

[96] U. Heiden, S. Roessner, K. Segl, and H. Kaufmann. Analysis of spectral signatures of urban surfaces for their identification using hyperspectral HyMap data. In *Proceedings of the IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas (Cat. No. 01EX482)*, page 173–177, Nov. 2001.

[97] U. Heiden, K. Segl, S. Roessner, and H. Kaufmann. Determination and verification of robust spectral features for an automated classification of sealed urban surfaces. *Imaging Spectroscopy: New Quality in Environmental Studies*, pages 165–174, 2005.

[98] U. Heiden, K. Segl, S. Roessner, and H. Kaufmann. Determination of robust spectral features for identification of urban surface materials in hyperspectral remote sensing data. *Remote Sensing of Environment*, 111(4):537–552, 2007.

[99] S. Hemalatha and S. M. Anouncia. Unsupervised segmentation of remote sensing images using FD based texture analysis model and ISODATA. *International Journal of Ambient Computing and Intelligence (IJACI)*, 8(3):58–75, 2017.

[100] M. Herold, M. E. Gardner, and D. A. Roberts. Spectral resolution requirements for mapping urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9):1907–1919, 2003.

[101] M. Herold and D. A. Roberts. Spectral characteristics of asphalt road aging and deterioration: Implications for remote-sensing applications. *Applied Optics*, 44:4327–4334, 2005.

[102] M. Herold, D. A. Roberts, M. E. Gardner, and P. E. Dennison. Spectrometry for urban area remote sensing–Development and analysis of a spectral library from 350 to 2400 nm. *Remote Sensing of Environment*, 91(3-4):304–319, 2004.

[103] H. Huang, X. Hu, Y. Zhao, M. Makkie, Q. Dong, S. Zhao, L. Guo, and T. Liu. Modeling task fMRI data via deep convolutional autoencoder. *IEEE Transactions on Medical Imaging*, 37(7):1551–1561, 2017.

[104] G. F. Hughes. On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 14(1):55–63, 1968.

[105] G. R. Hunt. Spectral signatures of particulate minerals in the visible and near infrared. *Geophysics*, 42(3):501–513, 1977.

[106] R. Ilehag, D. Bulatov, P. Helmholz, and D. Belton. Classification and representation of commonly used roofing material using multisensorial aerial data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII(1):217–224, 2018.

[107] R. Ilehag, J. Leitloff, A. Schenk, and M. Weinmann. Urban material classification using spectral and textural features retrieved from autoencoders. *International Annuals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, V1(2020):25–32, 2020.

[108] R. Ilehag, A. Schenk, and S. Hinz. Concept for classifying facade elements based on material, geometry and thermal radiation using multimodal UAV remote sensing. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII(2):145–151, 2017.

[109] R. Ilehag, A. Schenk, Y. Huang, and S. Hinz. KLUM: An urban VNIR and SWIR spectral library consisting of building materials. *Remote Sensing*, 11(18):2149:1–28, 2019.

[110] R. Ilehag, M. Weinmann, A. Schenk, S. Keller, and S. Jutzi, B. Hinz. Revisiting existing classification approaches for building materials based on hyperspectral data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII(3):65–71, 2017.

[111] A. K. Jain. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8):651–666, 2010.

[112] S. Jalali, D. M. Parapari, and M. J. Mahdavinejad. Analysis of building facade materials usage pattern in Tehran. *Advanced Engineering Forum*, 31:46–62, 2019.

[113] C. C. Johnson, N. Breward, E. L. Ander, and L. Ault. G-BASE: Baseline geochemical mapping of Great Britain and Northern Ireland. *Geochemistry: Exploration, Environment, Analysis*, 5(4):347–357, 2005.

[114] R. Kaden and T. H. Kolbe. City-wide total energy demand estimation of buildings using semantic 3D city models and statistical data. In *Proceedings of the 8th International 3D GeoInfo Conference*, pages 163–171, 2013.

[115] S. Kalogirou. Thermal performance, economic and environmental life cycle analysis of thermosiphon solar water heaters. *Solar Energy*, 83(1):39–48, 2009.

[116] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T. Y. Liu. LightGBM: A highly efficient gradient boosting decision tree. In *Proceedings of the Advances in Neural Information Processing Systems*, pages 3146–3154, 2017.

[117] J. M. Keller, S. Chen, and R. M. Crownover. Texture description and segmentation through fractal geometry. *Computer Vision, Graphics, and Image Processing*, 45(2):150–166, 1989.

[118] N. Keshava. Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries. *IEEE Transactions on Geoscience and Remote Sensing*, 42(7):1552–1565, 2004.

[119] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[120] R. F. Kokaly, R. N. Clark, G. A. Swayze, K. E. Livo, T. M. Hoefen, N. C. Pearson, R. A. Wise, W. M. Benzel, H. A. Lowers, R. L. Driscoll, and A. J. Klein. USGS Spectral Library Version 7: U.S. Geological Survey Data Series. Technical Report 1035, U.S. Geological Survey, Reston, Virginia, U.S., 2017.

[121] T. H. Kolbe. Representing and exchanging 3D city models with CityGML. In *3D Geo-Information Sciences*, pages 15–31. Springer, 2009.

[122] T. H. Kolbe, G. Gröger, and L. Plümer. CityGML: Interoperable access to 3D city models. In *Geo-Information for Disaster Management*, pages 883–899. Springer, 2005.

[123] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas. Supervised machine learning: A review of classification techniques. *Emerging Artificial Intelligence Applications in Computer Engineering*, 160:3–24, 2007.

[124] S. Kotthaus, T. E. L. Smith, M. J. Wooster, and C. S. B. Grimmond. Derivation of an urban materials spectral library through emittance and reflectance spectroscopy. *ISPRS Journal of Photogrammetry and Remote Sensing*, 94:194–212, 2014.

[125] B. Kottler, E. Burkard, D. Bulatov, and L. Haraké. Physically-based thermal simulation of large scenes for infrared imaging. In *Proceedings of the 14th International Conference on Compute Graphics Theory and Applications*, pages 25–27, 2019.

[126] M. A. Kramer. Nonlinear principal component analysis using autoassociative neural networks. *American Institute of Chemical Engineers Journal*, 37(2):233–243, 1991.

[127] D. R. Krathwohl and L. W. Anderson. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman, 2009.

[128] A. Krizhevsky and G. E. Hinton. Using very deep autoencoders for content-based image retrieval. In *Proceedings of the 19th European Symposium on Artificial Neural Networks*, volume 1, pages 1–7. Citeseer, 2011.

[129] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. 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.

[130] S. Kullback. *Information theory and statistics*. Courier Corporation, 1997.

[131] U. Kumar, C. Milesi, R. Nemani, and S. Basu. Multi-sensor multi-resolution image fusion for improved vegetation and urban area classification. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL(7):51–58, 2015.

[132] T. Kutser, I. Miller, and D. L. B. Jupp. Mapping coral reef benthic substrates using hyperspectral space-borne images and spectral libraries. *Estuarine, Coastal and Shelf Science*, 70(3):449–460, 2006.

[133] R. Kwitt and P. Meerwald. Salzburg texture image database (STex). `http://wavelab.at/sources/STex/`, 2018. [Online, last accessed 26.03.2020].

[134] G. E. Kyriakodis and M. Santamouris. Using reflective pavements to mitigate urban heat island in warm climates–Results from a large scale urban mitigation project. *Urban Climate*, 24:326–339, 2018.

[135] S. Lachérade, C. Miesch, X. Briottet, and H. Le Men. Spectral variability and bidirectional reflectance behaviour of urban materials at a 20 cm spatial resolution in the visible and near-infrared wavelengths. A case study over Toulouse (France). *International Journal of Remote Sensing*, 26(17):3859–3866, 2005.

[136] S. Lachérade, C. Miesch, F. Lemaître, X. Briottet, H. Le Men, D. Boldo, and C. Valorge. Analysis of the spectral variability of urban materials for classification. a case study over Toulouse (France). In *Proceedings of the ISPRS Joint Symposium URBAN05 and URS05*, pages 1–3. Citeseer, 2005.

[137] R. Lan, Z. Li, Z. Liu, T. Gu, and X. Luo. Hyperspectral image classification using k-sparse denoising autoencoder and spectral-restricted spatial characteristics. *Applied Soft Computing*, 74:693–708, 2019.

[138] S. Lazebnik, C. Schmid, and J. Ponce. A sparse texture representation using local affine regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1265–1278, 2005.

[139] A. Le Bris, N. Chehata, X. Briottet, and N. Paparoditis. Spectral band selection for urban material classification using hyperspectral libraries. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, III(7):33–40, 2016.

[140] D. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. In *Advances in Neural Information Processing Systems*, pages 556–562, 2001.

[141] T. Y. Lee and Y. J. Kaufman. Non-Lambertian effects on remote sensing of surface reflectance and vegetation index. *IEEE Transactions on Geoscience and Remote Sensing*, 24:699–708, 1986.

[142] W. Li, H. Fu, L. Yu, P. Gong, D. Feng, C. Li, and N. Clinton. Stacked autoencoder-based deep learning for remote-sensing image classification: A case study of African land-cover mapping. *International Journal of Remote Sensing*, 37(23):5632–5646, 2016.

[143] Y. Li, Q. Wang, C. Wu, S. Zhao, X. Xu, Y. Wang, and C. Huang. Estimation of chlorophyll a concentration using NIR/red bands of MERIS and classification procedure in inland turbid water. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3):988–997, 2011.

[144] A. Liaw and M. Wiener. Classification and regression by randomForest. *R News*, 2/3:18–22, 2002.

[145] Z. Lin, Y. Chen, X. Zhao, and G. Wang. Spectral-spatial classification of hyperspectral image using autoencoders. In *Proceedings of the 2013 9th International Conference on Information, Communications & Signal Processing*, pages 1–5. IEEE, 2013.

[146] C. Linnaeus. *Systema naturae*. Stockholm Laurentii Salvii, 1758.

[147] J. Liu, Z. Wu, Z. Wei, L. Xiao, and L. Sun. Spatial-spectral kernel sparse representation for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(6):2462–2471, 2013.

[148] S. Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982.

[149] J. Ma, D. W. Sun, H. Pu, J. H. Cheng, and Q. Wei. Advanced techniques for hyperspectral imaging in the food industry: Principles and recent applications. *Annual Review of Food Science and Technology*, 10:197–220, 2019.

[150] J. Ma, D. W. Sun, H. Pu, Q. Wei, and X. Wang. Protein content evaluation of processed pork meats based on a novel single shot (snapshot) hyperspectral imaging sensor. *Journal of Food Engineering*, 240:207–213, 2019.

[151] S. A. Macenka and M. P. Chrisp. Airborne visible/infrared imaging spectrometer (AVIRIS) spectrometer design and performance. In *Proceedings of Imaging Spectroscopy II*, volume 834, pages 32–43. International Society for Optics and Photonics, 1987.

[152] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297. Oakland, CA, US, 1967.

[153] W. H. Maes and K. Steppe. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends in Plant Science*, 24(2):152–164, 2019.

[154] A. K. Mahlein. Plant disease detection by imaging sensors–Parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease*, 100(2):241–251, 2016.

[155] J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *International Journal of Computer Vision*, 43(1):7–27, 2001.

[156] J. Mao and A. K. Jain. Texture classification and segmentation using multiresolution simultaneous autoregressive models. *Pattern Recognition*, 25(2):173–188, 1992.

[157] J. Masci, U. Meier, D. Cireşan, and J. Schmidhuber. Stacked convolutional auto-encoders for hierarchical feature extraction. In *Proceedings of the International Conference on Artificial Neural Networks*, pages 52–59. Springer, 2011.

[158] A. Materka and M. Strzelecki. Texture analysis methods–A review. *Technical University of Lodz, Institute of Electronics, COST B11 Report, Brussels*, 10(1.97):1–33, 1998.

[159] B. McKelvey. *Organizational systematics–Taxonomy, evolution, classification*. University of California Press, 1982.

[160] A. D. McNaught and A. Wilkinson. *Compendium of chemical terminology*, volume 1669. Blackwell Science Oxford, 1997.

[161] S. K. Meerdink, S. J. Hook, D. A. Roberts, and E. A. Abbott. The ECOSTRESS spectral library version 1.0. *Remote Sensing of Environment*, 230(111196):1–8, 2019.

[162] M. Mehta, R. Agrawal, and J. Rissanen. SLIQ: A fast scalable classifier for data mining. In *Proceedings of the International Conference on Extending Database Technology*, pages 18–32. Springer, 1996.

[163] P. K. Mehta and P. J. M. Monteiro. *Concrete microstructure, properties and materials*. University of California, Berkeley, US, 2017.

[164] S. Mei, J. Ji, J. Hou, X. Li, and Q. Du. Learning sensor-specific spatial-spectral features of hyperspectral images via convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 55(8):4520–4533, 2017.

[165] G. Meister. *Bidirectional reflectance of urban surfaces*. PhD thesis, University of Hamburg, 2000.

[166] F. Melgani and L. Bruzzone. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8):1778–1790, 2004.

[167] G. V. Middleton. Chemical composition of sandstones. *Geological Society of America Bulletin*, 71(7):1011–1026, 1960.

[168] F. Mirzapour and H. Ghassemian. Improving hyperspectral image classification by combining spectral, texture, and shape features. *International Journal of Remote Sensing*, 36(4):1070–1096, 2015.

[169] L. M. Montandon and E. E. Small. The impact of soil reflectance on the quantification of the green vegetation fraction from NDVI. *Remote Sensing of Environment*, 112(4):1835–1845, 2008.

[170] D. J. Mulla. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, 114(4):358–371, 2013.

[171] C. Müller. Use of cement in concrete according to European standard EN 206-1. *HBRC Journal*, 8(1):1–7, 2012.

[172] N. E. M. Nasarudin and H. Z. M. Shafri. Development and utilization of urban spectral library for remote sensing of urban environment. *Journal of Urban and Environmental Engineering*, 5:44–56, 2011.

[173] A. Ng. Sparse autoencoder – CS294A Lecture notes, 2011.

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

[175] R. Nouvel, J. M. Bahu, R. Kaden, J. Kaempf, P. Cipriano, M. Lauster, K. H. Häfele, E. Munoz, O. Tournaire, and E. Casper. Development of the CityGML application domain extension energy for urban energy simulation. In *Proceedings of the 14th Conference of the International Building Performance Simulation Association*, pages 559–564, 2015.

[176] T. Novack, L. Vorbeck, H. Lorei, and A. Zipf. Towards detecting building facades with graffiti artwork based on street view images. *ISPRS International Journal of Geo-Information*, 9(2):98:1–17, 2020.

[177] T. Ojala, T. Maenpaa, M. Pietikainen, J. Viertola, J. Kyllonen, and S. Huovinen. Outex-new framework for empirical evaluation of texture analysis algorithms. In *Proceedings of the Object Recognition Supported by User Interaction for Service Robots*, volume 1, pages 701–706. IEEE, 2002.

[178] T. R. Oke. The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society*, 108(455):1–24, 1982.

[179] A. Okujeni, S. van der Linden, and P. Hostert. Extending the vegetation–impervious–soil model using simulated EnMAP data and machine learning. *Remote Sensing of Environment*, 158:69–80, 2015.

[180] A. Okujeni, S. van der Linden, and P. Hostert. Berlin-Urban-Gradient dataset 2009–An EnMAP preparatory flight campaign. Technical Report 1.2, Potsdam : GFZ Data Services, 2016.

[181] OpenStreetMap Wiki contributors. Texture Library – OpenStreetMap Wiki. `https://wiki.openstreetmap.org/w/index.php?title=Texture_Library&oldid=1098488`, 2014. [Online, last accessed 26.03.2020].

[182] OpenStreetMap Wiki contributors. OpenStreetMap Wiki. `https://www.openstreetmap.org`, 2017. [Online, last accessed 26.03.2020].

[183] W. Ouerghemmi, S. Gadal, G. Mozgeris, D. Jonikavičius, and C. Weber. Urban objects classification by spectral library: Feasibility and applications. In *Proceedings of the 2017 Joint Urban Remote Sensing Event (JURSE)*, pages 1–4. IEEE, 2017.

[184] K. Pearson. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901.

[185] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[186] F. L. Pedrotti, L. M. Pedrotti, and L. S. Pedrotti. *Introduction to optics*. Cambridge University Press, 2017.

[187] J. C. Petersen. Chemical composition of asphalt as related to asphalt durability. *Developments in Petroleum Science*, 40:363–395, 2000.

[188] R. C. Pettersen. The chemical composition of wood. In *The Chemistry of Solid Wood*, pages 57–126. ACS Publications, 1984.

[189] M. Poletto, A. J. Zattera, and R. M. C. Santana. Structural differences between wood species: evidence from chemical composition, FTIR spectroscopy, and thermogravimetric analysis. *Journal of Applied Polymer Science*, 126(S1):E337–E344, 2012.

[190] B. A. Pour, M. Hashim, and M. Marghany. Using spectral mapping techniques on short wave infrared bands of ASTER remote sensing data for alteration mineral mapping in SE Iran. *International Journal of the Physical Sciences*, 6(4):917–929, 2011.

[191] D. M. Powers. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1):37–63, 2011.

[192] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C*. Cambridge University Press Cambridge, 1988.

[193] J. C. Price. How unique are spectral signatures? *Remote Sensing of Environment*, 49(3):181–186, 1994.

[194] J. C. Price. Examples of high resolution visible to near-infrared reflectance spectra and a standardized collection for remote sensing studies. *Remote Sensing*, 16(6):993–1000, 1995.

[195] Q. Pu, L. Jiang, J. Xu, H. Chu, Y. Xu, and Y. Zhang. Evolution of pH and chemical composition of pore solution in carbonated concrete. *Construction and Building Materials*, 28(1):519–524, 2012.

[196] J. Qin, K. Chao, M. S. Kim, R. Lu, and T. F. Burks. Hyperspectral and multispectral imaging for evaluating food safety and quality. *Journal of Food Engineering*, 118(2):157–171, 2013.

[197] N. R. Rao, P. K. Garg, and S. K. Ghosh. Development of an agricultural crops spectral library and classification of crops at cultivar level using hyperspectral data. *Precision Agriculture*, 8(4-5):173–185, 2007.

[198] P. H. Raven, B. Berlin, and D. E. Breedlove. The origins of taxonomy. *Science*, 174(4015):1210–1213, 1971.

[199] L. J. Rickard, R. W. Basedow, E. F. Zalewski, P. R. Silverglate, and M. Landers. HYDICE: An airborne system for hyperspectral imaging. In *Imaging Spectrometry of the Terrestrial Environment*, volume 1937, pages 173–179. International Society for Optics and Photonics, 1993.

[200] M. K. Ridd. Exploring a VIS (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: Comparative anatomy for cities. *International Journal of Remote Sensing*, 16(12):2165–2185, 1995.

[201] J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, 1978.

[202] D. A. Roberts and M. Herold. Imaging spectrometry of urban materials. *Infrared Spectroscopy in Geochemistry, Exploration and Remote Sensing, Mineral Association of Canada, Short Course Series*, 33:155–181, 2004.

[203] S. Robila. An investigation of spectral metrics in hyperspectral image preprocessing for classification. In *Proceedings of the Geospatial goes global: From your neighborhood to the whole planet. ASPRS Annual Conference*, pages 7–11, 2005.

[204] S. Roessner, K. Segl, U. Heiden, and H. Kaufmann. Automated differentiation of urban surfaces based on airborne hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7):1525–1532, 2001.

[205] L. Rokach. Ensemble-based classifiers. *Artificial Intelligence Review*, 33(1-2):1–39, 2010.

[206] J. W. Rouse, R. H. Haas, J. A. Schell, and D. W. Deering. Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publication*, 351:309–317, 1974.

[207] N. Salamati, C. Fredembach, and S. Süsstrunk. Material classification using color and NIR images. In *Proceedings of the 17th Color and Imaging Conference*, volume 1, pages 216–222. Society for Imaging Science and Technology, 2009.

[208] J. W. Salisbury. Spectral measurements field guide. Technical Report ADA362372, Earth Satellite Corporation, 1998.

[209] J. Sánchez and E. Quirós. Semiautomatic detection and classification of materials in historic buildings with low-cost photogrammetric equipment. *Journal of Cultural Heritage*, 25:21–30, 2017.

[210] S. Sankaran and R. Ehsani. Introduction to the electromagnetic spectrum. In *Imaging with Electromagnetic Spectrum*, pages 1–15. Springer, 2014.

[211] M. Santamouris. Using cool pavements as a mitigation strategy to fight urban heat island—-A review of the actual developments. *Renewable and Sustainable Energy Reviews*, 26:224–240, 2013.

[212] M. Santamouris, A. Synnefa, and T. Karlessi. Using advanced cool materials in the urban built environment to mitigate heat islands and improve thermal comfort conditions. *Solar Energy*, 85(12):3085–3102, 2011.

[213] M. A. E. Santana and E. Y. A. Okino. Chemical composition of 36 Brazilian Amazon forest wood species. *Holzforschung*, 61(5):469–477, 2007.

[214] M. E. Schaepman, M. Jehle, A. Hueni, P. D'Odorico, A. Damm, J. Weyermann, F. D. Schneider, V. Laurent, C. Popp, F. Seidel, and et al. Advanced radiometry measurements and Earth science applications with the Airborne Prism Experiment (APEX). *Remote Sensing of Environment*, 158:207–219, 2015.

[215] F. L. Schwab. Framework mineralogy and chemical composition of continental margin-type sandstone. *Geology*, 3(9):487–490, 1975.

[216] K. L. Seow, M. Doya, X. Briottet, R. Ceolato, N. Riviere, and S. K. Pang. Spectral BRDF of common urban materials. In *Proceedings of the 2012 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–4. IEEE, 2012.

[217] J. Shafer, R. Agrawal, and M. Mehta. SPRINT: A scalable parallel classifier for data mining. In *Proceedings of the 22nd Vldb Conference*, volume 96, pages 544–555. Citeseer, 1996.

[218] L. Sharan, R. Rosenholtz, and E. Adelson. Material perception: What can you see in a brief glance? *Journal of Vision*, 9(8):784, 2009.

[219] K. D. Shepherd and M. G. Walsh. Development of reflectance spectral libraries for characterization of soil properties. *Soil Science Society of America Journal*, 66(3):988–998, 2002.

[220] W. C. Snyder and Z. Wan. BRDF models to predict spectral reflectance and emissivity in the thermal infrared. *IEEE Transactions on Geoscience and Remote Sensing*, 36(1):214–225, 1998.

[221] J. A. Sobrino, R. Oltra-Carrió, J. C. Jiménez-Muñoz, Y. Julien, G. Soria, B. Franch, and C. Mattar. Emissivity mapping over urban areas using a classification-based approach: Application to the dual-use european security IR experiment (DESIREX). *International Journal of Applied Earth Observation and Geoinformation*, 18:141–147, 2012.

[222] J. A. Sobrino, G. Sòria, M. Romaguera, and J. Cuenca. DESIREX 2008: Estudio de la isla de calor en la ciudad de Madrid. *Revista de Teledetección*, 31:80–92, 2009.

[223] J. A. Stratton. *Electromagnetic theory*, volume 33. John Wiley and Sons, 2007.

[224] A. Sturman and N. J. Tapper. *The weather and climate of Australia and New Zealand*. Oxford University Press, US, 1996.

[225] S. Sultana and A. Bhasin. Effect of chemical composition on rheology and mechanical properties of asphalt binder. *Construction and Building Materials*, 72:293–300, 2014.

[226] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang, and X. Chen. A sparse auto-encoder-based deep neural network approach for induction motor faults classification. *Measurement*, 89:171–178, 2016.

[227] T. Susca, S. R. Gaffin, and G. R. Dell'Osso. Positive effects of vegetation: Urban heat island and green roofs. *Environmental Pollution*, 159(8-9):2119–2126, 2011.

[228] K. Tempfli, N. Kerle, G. Huurneman, and L. L. F. Janssen. *Principles of Remote Sensing – An introductory textbook*. The International Institute of Geo-Information Science and Earth Observation (ITC), Enschede, The Netherlands, 2009.

[229] Q. J. Tian, Q. X. Tong, R. L. Pu, X. Guo, and C. Zhao. Spectroscopic determination of wheat water status using 1650-1850 nm spectral absorption features. *International Journal of Remote Sensing*, 22(12):2329–2338, 2001.

[230] M. E. Tipping and C. M. Bishop. Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(3):611–622, 1999.

[231] C.-O. Truica, F. Radulescu, and A. Boicea. Comparing different term weighting schemas for topic modeling. In *Proceedings of the 2016 18th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, pages 307–310. IEEE, 2016.

[232] M. Tuceryan and A. K. Jain. Texture analysis. In *Handbook of Pattern Recognition and Computer Vision*, pages 235–276. World Scientific, 1993.

[233] S. van der Linden, A. Okujeni, F. Canters, J. Degerickx, U. Heiden, P. Hostert, F. Priem, B. Somers, and F. Thiel. Imaging spectroscopy of urban environments. *Surveys in Geophysics*, 40(3):471–488, 2019.

[234] L. van der Maaten and G. Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(Nov):2579–2605, 2008.

[235] F. van der Meer. Analysis of spectral absorption features in hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation*, 5(1):55–68, 2004.

[236] G. Vane, R. O. Green, T. G. Chrien, H. T. Enmark, E. G. Hansen, and W. M. Porter. The airborne visible/infrared imaging spectrometer (AVIRIS). *Remote Sensing of Environment*, 44(2-3):127–143, 1993.

[237] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62(1-2):61–81, 2005.

[238] P. Vincent, H. Larochelle, Y. Bengio, and P. A. Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th International Conference on Machine Learning*, pages 1096–1103, 2008.

[239] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P. A. Manzagol. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research*, 11(Dec):3371–3408, 2010.

[240] J. Wang, R. Xu, and S. Yang. Estimation of plant water content by spectral absorption features centered at 1,450 nm and 1,940 nm regions. *Environmental Monitoring and Assessment*, 157(1-4):459, 2009.

[241] K. Ward, S. Lauf, B. Kleinschmit, and W. Endlicher. Heat waves and urban heat islands in Europe: A review of relevant drivers. *Science of the Total Environment*, 569:527–539, 2016.

[242] S. A. Waugh, C. A. Purdie, L. B. Jordan, S. Vinnicombe, R. A. Lerski, P. Martin, and A. M. Thompson. Magnetic resonance imaging texture analysis classification of primary breast cancer. *European Radiology*, 26(2):322–330, 2016.

[243] U. Weidner and W. Förstner. Towards automatic building extraction from high-resolution digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4):38–49, 1995.

[244] M. Weinmann. *Reconstruction and analysis of 3D scenes – From irregularly distributed 3D points to object classes*. Springer, Cham, Switzerland, 2016.

[245] M. Weinmann, J. Gall, and R. Klein. Material classification based on training data synthesized using a BTF database. In *Proceedings of the European Conference on Computer Vision*, pages 156–171, 2014.

[246] T. P. Weldon, W. E. Higgins, and D. F. Dunn. Efficient Gabor filter design for texture segmentation. *Pattern Recognition*, 29(12):2005–2015, 1996.

[247] Q. Weng, D. Lu, and J. Schubring. Estimation of land surface temperature–Vegetation abundance relationship for urban heat island studies. *Remote Sensing of Environment*, 89(4):467–483, 2004.

[248] K. F. West, B. N. Webb, J. R. Lersch, S. Pothier, J. M. Triscari, and A. E. Iverson. Context-driven automated target detection in 3D data. In *Proceedings of the Automatic Target Recognition XIV*, volume 5426, pages 133–144. International Society for Optics and Photonics, 2004.

[249] R. H. Whittaker. Vegetation of the Siskiyou mountains, Oregon and California. *Ecological Monographs*, 30(3):279–338, 1960.

[250] L. P. Wilding, N. E. Smeck, and G. F. Hall. *Pedogenesis and soil taxonomy: The soil orders*. Elsevier, 1983.

[251] A. F. Wolf. Using WorldView-2 vis-NIR multispectral imagery to support land mapping and feature extraction using normalized difference index ratios. In *Proceedings of the Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*, volume 8390, pages 83900N:1–7. International Society for Optics and Photonics, 2012.

[252] D. H. Wolpert. Stacked generalization. *Neural Networks*, 5(2):241–259, 1992.

[253] T. Wu, C. Lin, and R. C. Weng. Probability estimates for multi-class classification by pairwise coupling. *Journal of Machine Learning Research*, 5(Aug):975–1005, 2004.

[254] X. Xie. A review of recent advances in surface defect detection using texture analysis techniques. *ELCVIA: Electronic Letters on Computer Vision and Image Analysis*, 7(3):1–22, 2008.

[255] F. Xing, H. Yao, Y. Liu, X. Dai, R. Brown, and D. Bhatnagar. Recent developments and applications of hyperspectral imaging for rapid detection of mycotoxins and mycotoxigenic fungi in food products. *Critical Reviews in Food Science and Nutrition*, 59(1):173–180, 2019.

[256] B. Xiong, S. O. Elberink, and G. Vosselman. Building modeling from noisy photogrammetric point clouds. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II(3):197–204, 2014.

[257] J. Xu, L. Xiang, Q. Liu, H. Gilmore, J. Wu, J. Tang, and A. Madabhushi. Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. *IEEE Transactions on Medical Imaging*, 35(1):119–130, 2015.

[258] Y. Yamaguchi, A. Kahle, H. Tsu, T. Kawakami, and M. Pniel. Overview of advanced spaceborne thermal emission and reflection radiometer (ASTER). *IEEE Transactions on Geoscience and Remote Sensing*, 36(4):1062–1071, 1998.

[259] J. Yang, Z. H. Wang, and K. E. Kaloush. Environmental impacts of reflective materials: Is high albedo a 'silver bullet' for mitigating urban heat island? *Renewable and Sustainable Energy Reviews*, 47:830–843, 2015.

[260] J. Yang, M. S. Wong, M. Menenti, and J. Nichol. Study of the geometry effect on land surface temperature retrieval in urban environment. *ISPRS Journal of Photogrammetry and Remote Sensing*, 109:77–87, 2015.

[261] C. M. Yeum, A. Lenjani, S. J. Dyke, and I. Bilionis. Automated detection of pre-disaster building images from Google Street View. *arXiv preprint arXiv:1902.10816*, 2019.

[262] L. Yu and H. Liu. Feature selection for high-dimensional data: A fast correlation-based filter solution. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pages 856–863, 2003.

[263] J. Yuan, D. Wang, and R. Li. Remote sensing image segmentation by combining spectral and texture features. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1):16–24, 2013.

[264] L. Yuan, J. Guo, and Q. Wang. Automatic classification of common building materials from 3D terrestrial laser scan data. *Automation in Construction*, 110:103017:1–30, 2020.

[265] P. Zambrano, A. Josa, J. Rieradevall, S. Gassó, and X. Gabarrell. Spectral library of rooftop urban materials. Technical report, Universitat Autònoma de Barcelona, 2017.

[266] X. Zhang, H. Dou, T. Ju, J. Xu, and S. Zhang. Fusing heterogeneous features from stacked sparse autoencoder for histopathological image analysis. *IEEE Journal of Biomedical and Health Informatics*, 20(5):1377–1383, 2015.

[267] Z. Zhao and H. Liu. Spectral feature selection for supervised and unsupervised learning. In *Proceedings of the 24th International Conference on Machine Learning*, pages 1151–1157, 2007.

[268] Z. Zhao, F. Morstatter, S. Sharma, S. Alelyani, A. Anand, and H. Liu. Advancing feature selection research – ASU feature selection repository. Technical report, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ, US, 2010.

[269] P. Zhou, J. Han, G. Cheng, and B. Zhang. Learning compact and discriminative stacked autoencoder for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7):4823–4833, 2019.

[270] W. Zhou, Z. Shao, C. Diao, and Q. Cheng. High-resolution remote-sensing imagery retrieval using sparse features by auto-encoder. *Remote Sensing Letters*, 6(10):775–783, 2015.

[271] X. J. Zhu. Semi-supervised learning literature survey. Technical report, University of Wisconsin-Madison Department of Computer Sciences, 2005.

[272] R. J. Zomer, A. Trabucco, and S. L. Ustin. Building spectral libraries for wetlands land cover classification and hyperspectral remote sensing. *Journal of Environmental Management*, 90(7):2170–2177, 2009.