

TECHNICAL UNIVERSITY OF CRETE

MASTER THESIS

---

**Development of machine  
learning techniques towards  
spectral dimensionality  
expansion : applications in  
snapshot spectral imaging.**

---

*Author:*

Fragkoulis N.  
LOGOTHETIS

*Comittee:*

Dr. Costas Balas  
Dr. Vassilis Samoladas  
Dr. George Karystinos

*A thesis submitted in fulfillment of the requirements  
for the master degree of Electrical and Computer Engineering*

*in the*

Department of Electrical and Computer Engineering

January 26, 2020

# Declaration of Authorship

I, Fragkoulis N. LOGOTHETIS, declare that this thesis titled, “Development of machine learning techniques towards spectral dimensionality expansion : applications in snapshot spectral imaging.” and the work presented in it are my own.

I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

---

Date:

---

*“Failure is an option here. If things are not failing, you are not innovating enough.”*

Elon Mask

TECHNICAL UNIVERSITY OF CRETE

# *Abstract*

Technical University of Crete

Department of Electrical and Computer Engineering

Electrical and Computer Engineering

**Development of machine learning techniques towards spectral  
dimensionality expansion : applications in snapshot spectral  
imaging.**

by Fragkoulis N. LOGOTHETIS

Hyperspectral imaging (HSI) is an emerging technology that integrates conventional imaging and spectroscopy to attain both spatial and spectral information from an object. The spectral images, collected in the spectral cube, are tens of hundreds and the information we receive from them is crucial for many applications, such as bio-medical technology, remote sensing, microscopy etc. Nevertheless, the current state of the art includes HSI systems which need long acquisition time, something that prevents them from observing any dynamically developing phenomena. Also, they are expensive and sizeable, which makes them inaccessible to many important applications. To address these limitations, a theoretical real-time snapshot spectral imaging system (SNSI) that captures a small number of spectral bands, and by using dimensionality expansion techniques, provides real time HSI, is investigated in this study. A comparative study of the state of the art was conducted under the assumption of various hardware architectures (RGB narrow/wide, three, six, nine, and twelve spectral channels). Furthermore, two new algorithms called K-Fourier and 2Level are proposed and compared in terms of minimizing the estimation error of the uncaptured spectral images. The novelty of the proposed methods stems mainly from the reduction of the dimensionality of the space needed to be reconstructed. Experiments on standard color charts show that the K-Fourier and non-linear kernels outperform the other competing methods. Looking at the same problem from another perspective, the most feature-rich training yields to higher estimation accuracy. On that account, a great amount of band selection techniques, which are based on similarity-measurement, dynamic programming, and evolutionary formulas were analyzed and compared. Genetic Algorithms turned out to be the most promising feature selection technique, since it dramatically improves the space reconstruction error. Moreover, we introduce a nonlinear transformation of reflectance values to ensure that the estimated reflection spectra fulfill physically motivated boundary conditions. Ending up, these findings set the basis for the development of a powerful SNSI system. Medical diagnosis is expected to be a leading application of this novel approach.

## *Acknowledgements*

After one and a half years of truly hard, yet exciting work and before presenting my master thesis, I cannot but express my gratitude towards the people who were by my side and made everything I achieved possible.

First and foremost, I would like to thank my advisor, *Prof. Costas Balas*, for exposing me to research and pointing me towards exciting topics. I am also grateful towards the other two members of the committee. I had numerous intriguing discussions with *Prof. Vassilis Samoladas* and *Prof. George Karistinos* about technical issues. I owe a lot to many, among my other Professors of TUC for their wise advice when I had to make critical decisions.

Nothing would have been possible without my family. I wish to thank my family and my close friends, for their advising and support all these years of my studies.

# Contents

<b>Declaration of Authorship</b>	<b>i</b>
<b>Abstract</b>	<b>iv</b>
<b>Acknowledgements</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Introduction to Hyper Spectral Imaging - Estimation . . . . .	1
1.2 Light as electromagnetic wave and also as particle - Photoelectric Effect . . . . .	3
1.3 Spectroscopy/ spectrometry . . . . .	5
1.4 Imaging spectroscopy . . . . .	5
1.5 Hyperspectral Imaging . . . . .	6
1.6 Single Exposure or Instantaneous Spectral Imagers . . . . .	8
1.7 Color vs. Spectral Imaging . . . . .	9
1.8 Spectral Imaging Applications . . . . .	10
1.9 What is Spectral Estimation . . . . .	10
1.9.1 Spectral Acquisition Chain of a Snapshot Spectral Imager	13
<b>2 Spectral Estimation - Problem Statement and Methods</b>	<b>16</b>
2.1 Introduction . . . . .	16
2.2 Measures of Spectral Similarities . . . . .	18
2.2.1 Goodness of Fit Coefficient (GFC) . . . . .	19
2.2.2 Spectral Angle Mapper (SAM) . . . . .	19
2.2.3 Euclidean Distance (ED) . . . . .	19
2.2.4 Root Mean Square Error (RMSE) . . . . .	19
2.3 Problem Statement . . . . .	20
2.4 Spectral Estimation Algorithms . . . . .	24
2.4.1 Estimation of Spectral Reflectances Using Wiener Method	24
2.4.2 Estimation of Spectral Reflectances Using Projection on PCA . . . . .	24
2.4.3 Sparse Dictionary Learning (K-SVD) . . . . .	26

2.4.4	Estimation of Spectral Reflectances Using Hybrid Method	27
2.4.5	Estimation of Spectral Reflectances Using Adaptive(Local)- Wiener Method . . . . .	29
2.4.6	Estimation of Spectral Reflectances Using Multivariate Linear Regression Method . . . . .	31
2.4.7	Estimation of Spectral Reflectances Using Artificial Neu- ral Networks . . . . .	32
2.4.8	Kernel Regression . . . . .	35
2.4.9	K-Fourier . . . . .	37
2.4.10	Two-Level Spectral Estimation (2Level) . . . . .	40
2.4.11	Spectral Matching Method (Search) . . . . .	41
<b>3</b>	<b>Spectral Estimation - Experiments</b>	<b>43</b>
3.1	Datasets, Acquisition, and Pre-processing . . . . .	43
3.2	Training and Evaluation . . . . .	45
3.3	Tuning Hyper-Parameters . . . . .	47
3.4	Results . . . . .	49
3.5	Discuss . . . . .	50
3.5.1	Respecting physical constraints . . . . .	56
<b>4</b>	<b>Band Selection - Feature Extraction/Selection</b>	<b>62</b>
4.1	Introduction to Band Selection . . . . .	62
4.2	Data Preprocessing . . . . .	63
4.3	Properties of Similarity-Based Band Selection . . . . .	64
4.4	Band Selection Algorithms . . . . .	65
4.4.1	MEV-SFS . . . . .	65
4.4.2	Linear prediction (LP) . . . . .	66
4.4.3	Orthogonal Subspace Projection (OSP) . . . . .	67
	Band Used as the Initial for Band Selection for OSP and LP . . . . .	67
4.4.4	Band selection using Optimal Linear Interpolation (LIBS)	68
4.4.5	Genetic Algorithm (GA) . . . . .	70
4.5	Practical Considerations . . . . .	71
<b>5</b>	<b>Band Selection - Experiments</b>	<b>73</b>
5.1	Information about experiments and Results . . . . .	73
5.2	Discuss . . . . .	77
<b>6</b>	<b>Conclusions and Future Work</b>	<b>87</b>

**Bibliography**

# List of Figures

1.1	A Spectral Cube (source: Quantative BioImaging Laboratory,(QBIL) [1].)	1
1.2	Light as an electromagnetic wave (source: Physicstuff [2].) . . .	3
1.3	Spectroscopy using a simple prism (source: Astronomy, Cosmos [3]). . . . .	4
1.4	Principle of (remote sensing) imaging spectroscopy (source: Markelowitz [4]). . . . .	6
1.5	Electromagnetic spectrum (source: Wikiversity [5]). . . . .	7
1.6	Typical (hyper)spectral imaging approaches. (A) Point scan. (B) Line scan (i.e. “pushbroom”). (C) Wavelength scan. (D) snapshot (source: Wang, Yu and P. Reder [6]). . . . .	8
1.7	RBG vs Spectral Cube (source: [7]). . . . .	9
1.8	Spectral and spatial domain . . . . .	11
1.9	Spectral Acquisition Chain of a Snapshot Spectral Imager . . . . .	13
2.1	Synopsis of the spectral model of the acquisition process in a MS system (source: Alamin Mansouri [33]). . . . .	20
2.2	Hybrid model with different values of w. . . . .	28
2.3	Two typical examples of the training sets for the given candidate samples. Channel number C=6-Adaptive Wiener (source: Hui-Liang Shen [32]). . . . .	30
2.4	Model of a biological neuron (source: Wikipedia). . . . .	32
2.5	The Neural Network for the spectral estimation problem. . . . .	33
2.6	Model of a biological neuron (source: Wikipedia). . . . .	34
2.7	Flow Chart of K-Fourier. . . . .	38
2.8	Flow chart of Two Level Spectral Estimation Algorithm. . . . .	40
3.1	Macbeth and Munsell Color Pallets . . . . .	43
3.2	ZWO ASI178MC Spectral Sensitivity. . . . .	46
3.3	Experiment 1. Average RMSE plotted versus the number of spectral channels. . . . .	49
3.4	Experiment 2. Average RMSE plotted versus the number of spectral channels. . . . .	50

3.5	Experiment 3. Average RMSE plotted versus the number of spectral channels. . . . .	50
3.6	Experiment 3 with “arctanh” transformation. Average RMSE plotted versus the number of spectral channels. . . . .	58
4.1	MVS-SFS (source: [64]) . . . . .	65
4.2	LIBS algorithm. . . . .	68
4.3	Genetic Algorithm Flow Chart . . . . .	70
5.1	[Experiment 3/ BS: GA]. Average RMSE plotted versus the number of spectral channels. . . . .	75
5.2	[Experiment 3/ BS: OLI]. Average RMSE plotted versus the number of spectral channels. . . . .	75
5.3	[Experiment 3/ BS: LP]. Average RMSE plotted versus the number of spectral channels. . . . .	76
5.4	[Experiment 3/ BS: OSP]. Average RMSE plotted versus the number of spectral channels. . . . .	76
5.5	[Experiment 3/ BS: MVS-SFS]. Average RMSE plotted versus the number of spectral channels. Blue arrow indicates the algorithm that has the lowest average RMSE. . . . .	77
5.6	[Experiment 3/ BS: GA with “arctanh” transformation]. Average RMSE plotted versus the number of spectral channels. . . . .	77

# List of Tables

3.1	The channel sequences that are used to conduct the experiments.	46
3.2	[Experiment 1] Training Set : Munsell patches / Test Set: Munsell patches. (10-fold cross validation) . . . . .	51
3.3	[Experiment 2] Training Set : Munsell patches/ Test Set: Macbeth patches. . . . .	52
3.4	[Experiment 3] Training Set : TUC dataset patches / Test Set: Munsell patches. . . . .	53
3.5	[Experiment 3 with “arctanh” transformation] Training Set : TUC dataset patches / Test Set: Munsell patches. . . . .	59
5.1	Three narrow bands that are selected bands from the band selection algorithms. . . . .	74
5.2	Six narrow bands that are selected bands from the band selection algorithms. . . . .	74
5.3	Nine narrow bands that are selected bands from the band selection algorithms. . . . .	74
5.4	Twelve narrow bands that are selected bands from the band selection algorithms. . . . .	74
5.5	Training Set: TUC dataset/ Test Set: Munshell/ Band Selection Method GA. . . . .	78
5.6	Training Set: TUC dataset/ Test Set: Munsell/ Band Selection Method: OLI. . . . .	79
5.7	Training Set: TUC dataset/ Test Set: Munsell/ Band Selection Method: LP. . . . .	80
5.8	Training Set: TUC dataset/ Test Set: Munsell/ Band Selection Method: OSP. . . . .	81
5.9	Training Set: TUC dataset/ Test Set: Munsell/ Band Selection Method: MEV-SFS. . . . .	82
5.10	Training Set: TUC dataset/ Test Set: Munsell/ Band Selection Method: Genetic Algorithm/ “arctanh” transformation. . . . .	83

# List of Abbreviations

<b>CI</b>	<b>C</b> olor <b>I</b> maging
<b>SNR</b>	<b>S</b> ignal to <b>N</b> oise <b>R</b> atio
<b>SNSI</b>	<b>S</b> Napshot <b>S</b> pectral <b>I</b> maging
<b>HSI</b>	<b>H</b> yper <b>S</b> pectral <b>I</b> maging
<b>MS</b>	<b>M</b> ulti <b>S</b> pectral
<b>LR</b>	<b>L</b> inear <b>R</b> egression
<b>ANN</b>	<b>A</b> rtificial <b>N</b> eural <b>N</b> etworks
<b>DFT</b>	<b>D</b> iscrete <b>F</b> ourier <b>T</b> ransform
<b>TUC</b>	<b>T</b> echnical <b>U</b> niversity of <b>C</b> rete
<b>NNLS</b>	<b>N</b> on <b>N</b> egative <b>L</b> east <b>S</b> quares
<b>POCS</b>	<b>P</b> rojections <b>O</b> nto <b>C</b> onvex <b>S</b> ets
<b>BS</b>	<b>B</b> and <b>S</b> election
<b>LP</b>	<b>L</b> inear <b>P</b> rojection
<b>OSP</b>	<b>O</b> rthogonal <b>S</b> ubspace <b>P</b> rojection
<b>LIBS</b>	<b>L</b> inear <b>I</b> nterpolation for <b>B</b> and <b>S</b> election
<b>GA</b>	<b>G</b> enetic <b>A</b> lgorithms

# Physical Constants

Speed of Light  $c = 2.997\,924\,58 \times 10^8 \text{ m s}^{-1}$

Planck constant  $h = 6.626\,070\,04 \times 10^{-12} \text{ m}^2 \text{ kg/s}$

*Dedicated to my family...*

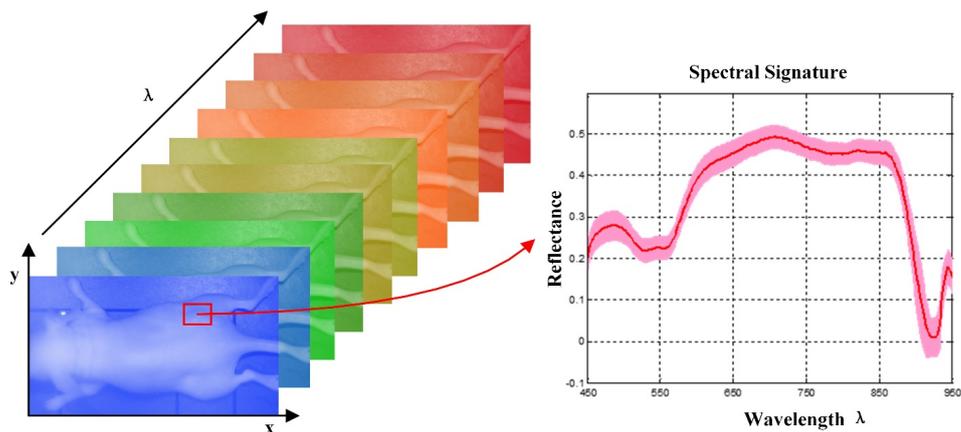


FIGURE 1.1: A Spectral Cube (source: Quantative BioImaging Laboratory,(QBIL) [1].)

## Chapter 1

# Introduction

## 1.1 Introduction to Hyper Spectral Imaging - Estimation

Over the last two decades, the field of optical imaging has developed very rapidly, providing color imaging (CI) systems with very high resolution that nowadays allow 3D imaging, as well as video capture with a very high frame rate. These systems try to emulate human vision in order to reproduce a result (image) that resembles the actual scene as it is perceived by a human eye. Usually these systems produce three-dimensional data (RGB is more commonly used), where each of the three dimensions represents the intensity and chrominance of light.

Although CI systems can provide an accurate representation of the scene at hand, there is a great deal of information that can not be perceived with these systems; information that may be hidden in the visible spectrum or beyond it,

i.e. UV or Infrared regions of the spectrum. In the direction to obtain the hidden information, spectral imaging systems are used.

Spectral Imaging (SI) is the application of reflectance spectroscopy to every pixel in a spatial image. Every spatial image captured represents a different wavelength within the electromagnetic spectrum and each pixel represents the spectral power distribution of the scene at that point. The stack of images created from this system is called *Spectral Cube* (Fig. (1.1)), and the data are represented in multidimensional, spanning spatial and spectral dimensions ( $x$ ,  $y$ ,  $\lambda$ ).

Spectroscopy can be used to detect individual absorption features due to specific chemical bonds in a solid, liquid, or gas. Solids can be either crystalline (i.e. minerals) or amorphous (such as glass). Every material is formed by chemical bonds and has the potential for detection with spectroscopy. Actual detection is dependent on the spectral coverage, spectral and spatial resolution, and signal-to-noise (SNR) of the SI system, the abundance of the material, and the strength of absorption features for that material in the wavelength region measured.

SI systems are widely used nowadays in numerous fields such as medicine, astronomy, industry, military etc. and due to the continuous need to improve the preexisting techniques and methods, many innovations and advances were developed.

The information provided by these systems, which is usually not discernible to the human eye, allows useful facts and phenomena to be revealed. Unfortunately, SI systems are expensive and sizable, which makes them inaccessible for many applications. The acquisition and computational time that is needed is very high, which prevents the system from observing any dynamically developing phenomena.

To address these limitations, real-time Snapshot Spectral Imaging (SNSI) system needs to be developed that allow simultaneous Hyperspectral Imaging (HSI) and has lower cost and size. We have developed and proposed a novel real-time theoretical HSI system that can simultaneously acquires a small number of spectral bands and provide spectral information in any desired wavelength within the visible and infrared spectrum.

## Electromagnetic Wave

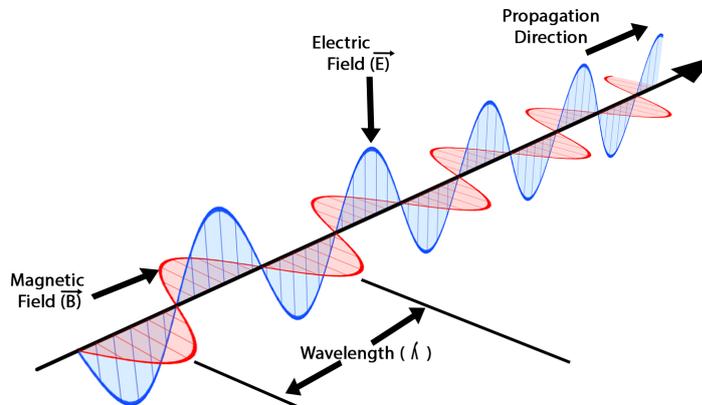


FIGURE 1.2: Light as an electromagnetic wave (source: Physics stuff [2].)

## 1.2 Light as electromagnetic wave and also as particle - Photoelectric Effect

Starting from the beginning of that science, it is important to be clarified that light is one of the fundamental phenomena of electromagnetism that behaves as wave and also as photon particle. It is characterized as wave when is emitted and propagated and as particle when it is absorbed. As an electromagnetic wave, it has both electric and magnetic field components, which oscillate in a fixed relationship to one another, perpendicular to each other and perpendicular to the direction of the energy and wave propagation as it is illustrated in Fig. (1.2). Each electromagnetic wave is characterized by the frequency or wavelength of its wave. The electromagnetic spectrum, in order of increasing frequency and decreasing wavelength, consists of radio waves, microwaves, infrared radiation, visible light, ultraviolet radiation, X-rays and gamma rays. The eyes of various organisms sense a somewhat variable but relatively small range of frequencies called the visible spectrum of the light. Higher frequencies correspond to proportionately more energy carried by each photon; for instance, a single gamma ray photon carries far more energy than a single photon of visible light. This phenomenon was described first time by Einstein, who realized that the energy( $E$ ) of any photon is given by the equation

$$E = h \frac{c}{\lambda}$$

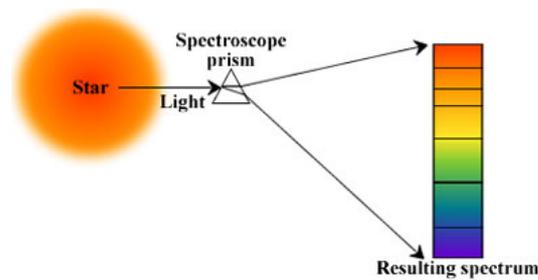


FIGURE 1.3: Spectroscopy using a simple prism (source: Astronomy, Cosmos [3]).

where  $h$  is the Planck constant,  $c$  is the speed of the light in vacuum and  $\lambda$  is the photon's wavelength.

In classical physics, electromagnetic radiation is considered to be produced when charged particles are accelerated by forces acting on them. Electrons are responsible for emission of most electromagnetic radiation because they have low mass, and therefore are easily accelerated by a variety of mechanisms. Rapidly moving electrons are most sharply accelerated when they encounter a region of force, so they are responsible for producing much of the highest frequency electromagnetic radiation observed in nature. Quantum processes can also produce electromagnetic radiation, such as when atomic nuclei undergo gamma decay, and processes such as neutral pion decay.

As it was mentioned before, light behaves as particle when is absorbed. In other words, that is exactly the reason of the photoelectric effect, in which “packets” of energy, called photons increase the maximum kinetic energy of the photoelectrons emitted from a given metal. Einstein got the Nobel Price in 1921 for his paper publication advancing the hypothesis that light energy is carried in discrete quantized packets to explain experimental data and the photoelectric effect. Photoelectric effect led both the scientific community and the industry to implement image sensors, which could read those “packets” of energy and produce a real digital image characterized by high color fidelity.

In this work, focusing in the duality of light is needed. The science of spectroscopy could bring out interesting information about the interaction between radiation and the chemical structure of the objects as a function of the wavelength  $\lambda$ .

### 1.3 Spectroscopy/ spectrometry

Spectroscopy was originally the study of the interaction between radiation and the structure of matter as a function of the wavelength  $\lambda$ . In fact, historically, spectroscopy referred to the use of visible light dispersed according to its wavelength, e.g. by a prism (Fig. (1.3)). Later, the concept was expanded greatly to comprise any measurement of a quantity as function of either wavelength or frequency. Thus, it can also refer to interactions with particle radiation or to a response to an alternating field or varying frequency  $\nu$ . A further extension of the scope of the definition added energy ( $E$ ) as a variable, once the very close relationship

$$E = h\nu$$

for photons was realized as it mentioned in Section 1.2. Spectrometry is the spectroscopic technique used to assess the concentration or amount of a given species. In those cases, the instrument that performs such measurements is a spectrometer or spectrograph.

Spectroscopy/spectrometry is often used in physical and analytical chemistry for the identification of substances through the spectrum emitted from or absorbed by them. Spectroscopy/spectrometry is also heavily used in astronomy and remote sensing. Most large telescopes have spectrometers, which are used either to measure the chemical composition and physical properties of astronomical objects or to measure their velocities from the Doppler shift of their spectral lines.

### 1.4 Imaging spectroscopy

Imaging spectroscopy (also spectral imaging or chemical imaging) is the application of reflectance spectroscopy to every pixel in a spatial image. In remote sensing situations (Fig. (1.4)), the surface materials mapped must be exposed in the optical surface (e.g., to map surface mineralogy it must not be covered with vegetation), and the diagnostic absorption features must be in regions of the spectrum that are reasonably transparent to the atmosphere. The optical surface is the same as what the geologist sees in the field with his or her eyes. Spectroscopy can be used in laboratories on hand samples, in the field with portable field spectrometers (spatial resolution in the millimeter to several meter range), from aircraft, and in the future from satellites. The aircraft systems can image large areas in short time (2 sq. km per second), producing spectra for

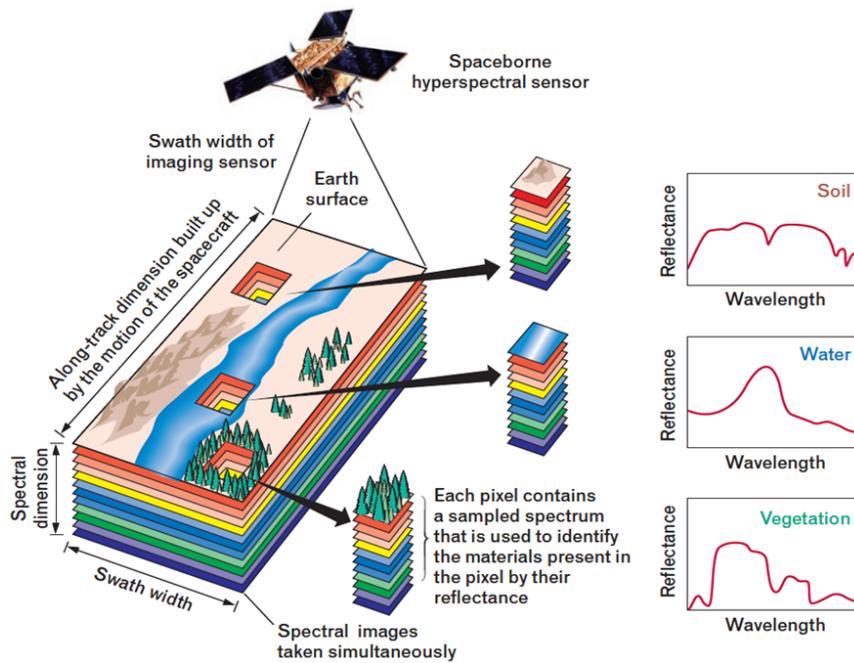


FIGURE 1.4: Principle of (remote sensing) imaging spectroscopy (source: Markelowitz [4]).

each pixel that can be analyzed for specific absorption bands and thus specific materials. These measurements can then be used for the unambiguous direct and indirect identification of surface materials and atmospheric trace gases, the measurement of their relative concentrations, subsequently the assignment of the proportional contribution of mixed pixel signals (e.g., the spectral unmixing problem), the derivation of their spatial distribution (mapping problem), and finally their study over time (multi-temporal analysis).

Imaging spectroscopy can be considered as the equivalent of color photography, but each pixel needs to acquire many bands of light intensity data from the spectrum, instead of just the three bands of the RGB color model. More precisely, it is the simultaneous acquisition of spatially co-registered images in many spectrally contiguous bands. In order to improve the spectroscopy imaging the scientific community turned into the HSI.

## 1.5 Hyperspectral Imaging

Hyperspectral Imaging collects and processes information from across the electromagnetic spectrum. Unlike the human eye, which just sees visible light, HSI is more like the eyes of the “mantis shrimp”, which can see visible light as well as from the ultraviolet to infrared. Hyperspectral capabilities enable the “mantis

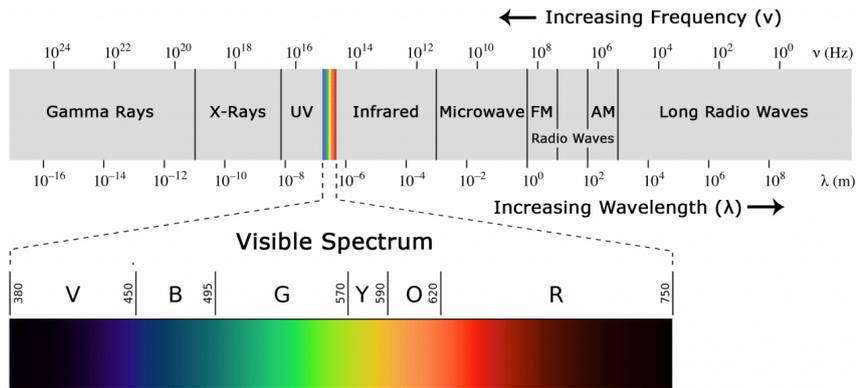


FIGURE 1.5: Electromagnetic spectrum (source: Wikiversity [5]).

shrimp” to recognize different types of coral, prey, or predators, all which may appear as the same color to the human eye.

Humans build sensors and processing systems to provide the same type of capability for application in agriculture, mineralogy, physics, surveillance and other fields of science. Hyperspectral sensors collect information as a set of “images”. Each image represents a range of the electromagnetic spectrum (Fig. (1.5)), and is also known as a spectral band. Hyperspectral sensors look at objects using a vast portion of the electromagnetic spectrum. Certain objects leave unique “fingerprints” across the electromagnetic spectrum. These “fingerprints” are known as *spectral signatures* and enable identification of the materials. For example, having the spectral signature for oil helps mineralogists find new oil fields.

The precision of these sensors is typically measured in spectral resolution, which is the width of each band of the spectrum that is captured. If the scanner picks up a large number of fairly small wavelengths, it is possible to identify objects even if said objects are only captured in a handful of pixels. However, spatial resolution is a factor in addition to spectral resolution. If the pixels are too large, then multiple objects are captured in the same pixel and become difficult to identify. If the pixels are too small, then the energy captured by each sensor-cell is low, and the decreased signal-to-noise ratio reduces the reliability of measured features.

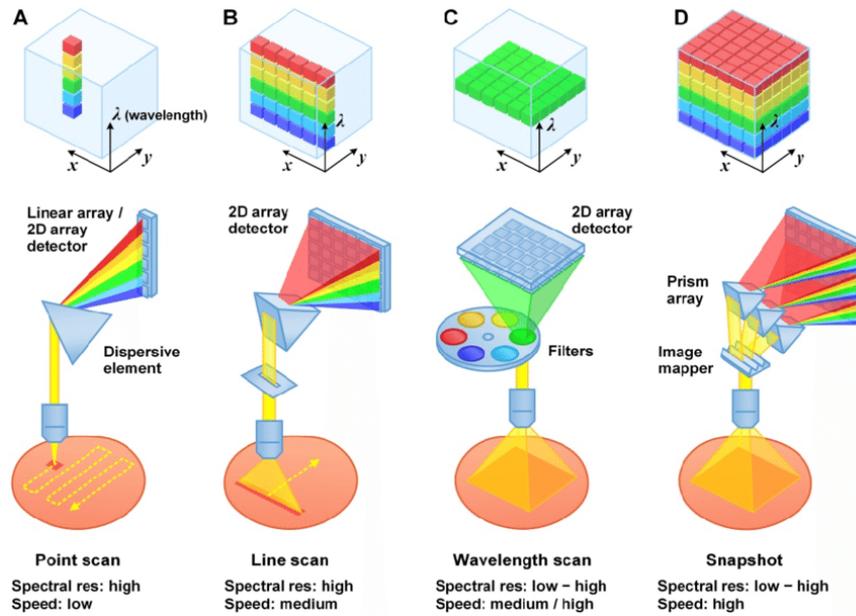


FIGURE 1.6: Typical (hyper)spectral imaging approaches. (A) Point scan. (B) Line scan (i.e. “pushbroom”). (C) Wavelength scan. (D) snapshot (source: Wang, Yu and P. Reeder [6]).

## 1.6 Single Exposure or Instantaneous Spectral Imagers

In order to record spectral images, a possible architecture includes tunable filtered-based systems. Those systems record spectral images in a time-sequentially manner and obtain the spectra from post hoc assembly of the time-sequential data. This reveals a major disadvantage of these systems. When we have phenomena that are changing on a time scale that is shorter than the duration required for recording the spectral cube, the SI systems described cannot perform accurately. Therefore, the scene recorded from the SI system must be static. Otherwise problems will be created in the co-registration of the spectral images, which is needed in order to provide accurate spectra.

There are numerous of applications (i.e biomedical and etc) that require spectral imaging and analysis of transient moving scenes. This is why single shot or instantaneous spectral imagers (Fig. (1.6.D)) need to be developed. Specifically, there are four techniques to develop a spectral scanner. The most simple structure of a HS imager is the point scan (Fig. (1.6.A)), where a line array detector acquires all the spectral information. However, point scan technique is characterized by really low speed. Furthermore, line scan and wavelength scan are often used, but it is more profitable to use the snapshot idea in order to handle quick changing phenomena. SNSI systems have many advantages over

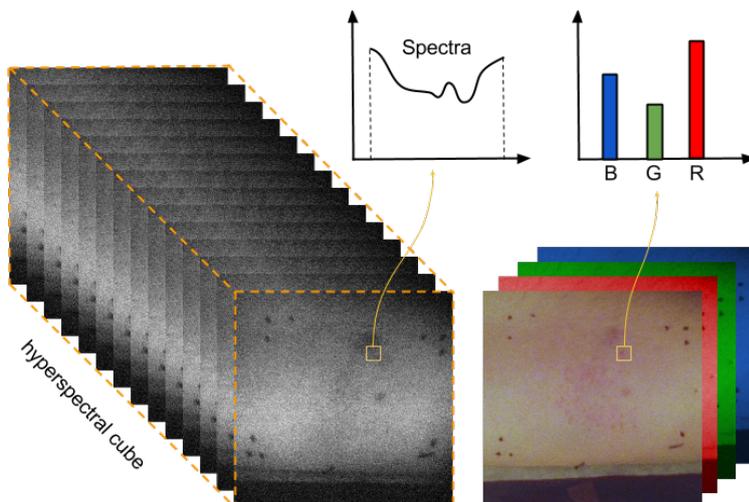


FIGURE 1.7: RGB vs Spectral Cube (source: [7]).

classical SI systems such as fast acquisition of accurately registered images, high device robustness and reliability, low cost etc. But, there are many trade-offs in order to achieve all that. Due to current technological limitations there is a trade-off between spatial and spectral resolution. Specifically, a SNSI system can capture either a small number of spectral images with high resolution or a great number of spectral images with really low spatial resolution.

Ending up, in the case of stationary and invariant scenes the point, line, and wavelength scanners are more preferable due to the fact that spatial and spectral resolution are superior to SNSI systems. On the other hand, SNSI scanners are a unique solution to observe dynamically developed phenomena.

## 1.7 Color vs. Spectral Imaging

Color cameras emulate the human vision for color reproduction and are real-time devices since they record three spectral bands (RGB) simultaneously. But, unfortunately, the human vision has some limitations that are shared by these devices. They allocate the incoming light into the three RGB coordinates, thus missing the important spectral information. So, objects with the same RGB values might have different spectral components. This phenomenon is known as metamerism. Therefore, due to this effect, the RGB imaging systems don't have the ability to distinguish material with the same color appearance but different chemical composition. In the SI systems for each pixel a series of images corresponding to many wavelengths are appeared. With these images, the unique fingerprint of each object could be extracted and the metamerism effect is disappeared (1.7).

## 1.8 Spectral Imaging Applications

As it mentioned before, there are numerous of application that the HSI can be used due to the importance of the spectral information.

HSI is nowadays commonly used in biomedicine. For instance, DYSIS[8] is a new company, which found out a new colposcope with an advanced cervical scan. The DYSIS colposcope is a high resolution digital tool with an adjunctive map. The DYSIS spectral map is generated by a proprietary technology that measures the aceto-whitening reaction and summarises it in the form of an intuitive map. The DYSIS map is overlaid on the live image of the cervix to help with the identification of the most relevant biopsy sites.

HSI is also useful to detect the chemical composition of plants, which can be used to detect the nutrient and water status of wheat in irrigated systems. On a smaller scale, NIR HSI can be used to rapidly monitor the application of pesticides to individual seeds for quality control of the optimum dose and homogeneous coverage.

In the food processing industry, HSI, combined with intelligent software, enables digital sorters (also called optical sorters) to identify and remove defects and foreign material (FM) that are invisible to traditional camera and laser sorters. By improving the accuracy of defect and FM removal, the food processor's objective is to enhance product quality and increase yields.

In astronomy, HSI is used to determine a spatially-resolved spectral image. Since a spectrum is an important diagnostic, having a spectrum for each pixel allows more science cases to be addressed. In astronomy, this technique is commonly referred to as integral field spectroscopy, and examples of this technique include FLAMES<sup>1</sup> and SINFONI<sup>2</sup> on the Very Large Telescope, but also the Advanced CCD Imaging Spectrometer on Chandra X-ray Observatory uses this technique.

From all the above examples, the importance of HSI is undisputed.

## 1.9 What is Spectral Estimation

As we mentioned before the RGB imaging devices miss important spectral information. Additionally, objects with different spectral information may have

---

<sup>1</sup>“FLAMES – Fibre Large Array Multi Element Spectrograph”.

<sup>2</sup>“SINFONI – Spectrograph for integral Field Observations in the Near Infrared”.

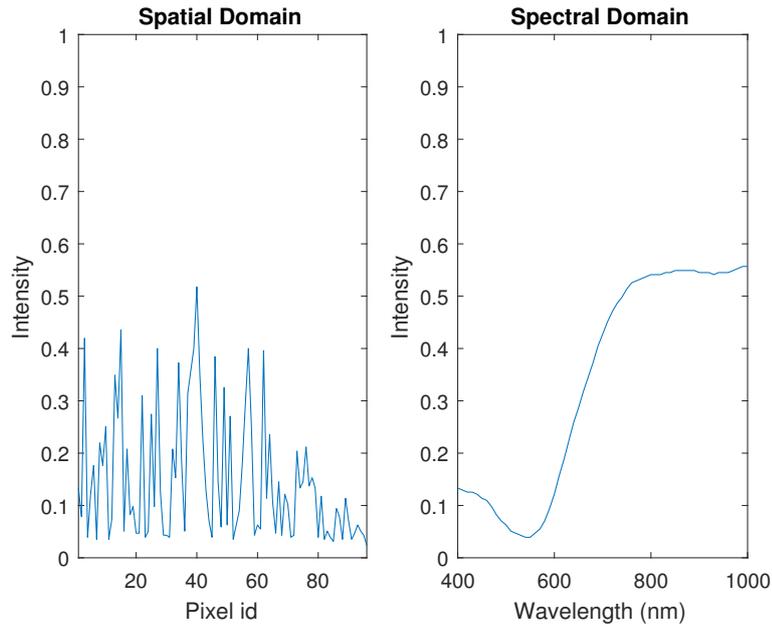


FIGURE 1.8: Spectral and spatial domain

same RGB values, which means that RGB devices cannot detect the chemical difference (metamerism). Furthermore, as we studied in Section 1.6, SNSI is a new trend of technology, combined with a variety of advantages (e.g Speed of acquisition). Looking at the same problem from another perspective, it is more cheaper and quicker to acquire a static number of images in specific wavelengths and estimate all the other bands, using a spectral estimation algorithm. An important question is why is better to estimate signals in spectral domain and not in spatial domain. The spatial domain is characterized from features with high spatial frequency (i.e. edges) that are those that change greatly in intensity over short image distances. These big, non-periodic intensity variation is very hard to be estimated if we also consider the fact that a natural scene is in most cases unique. There is no actual pattern or collection of images that can describe these variations accurately because of the randomness of natural scene characteristics (from objects to lighting conditions). On the other hand, spectral domain signals in the UV-VIS-NIR (400nm to 1000nm) are broad with a few peaks and valleys, as it is illustrated in figure 1.8. The absorption in this spectral range, comes from the electronic transitions (according to the selection rules for each atom) that are happening in the complex molecular structures of matter and the combination of them give these smooth characteristics to the signal. Most approaches for building a real time HSI device have the already mentioned trade-off between spectral and spatial resolution that stems

---

by optoelectronic limitations. As the spectral band number increases the spatial resolution decrease and vice versa. The estimation of missing information (spectral or spatial) is a strategic decision for designing these systems. Taking the above into consideration we choose to preserve the spatial resolution and estimate the missing spectral information as it is quite obvious that the estimation procedure is less demanding. Our goal is find most appropriate algorithms in terms of minimizing the estimation error.

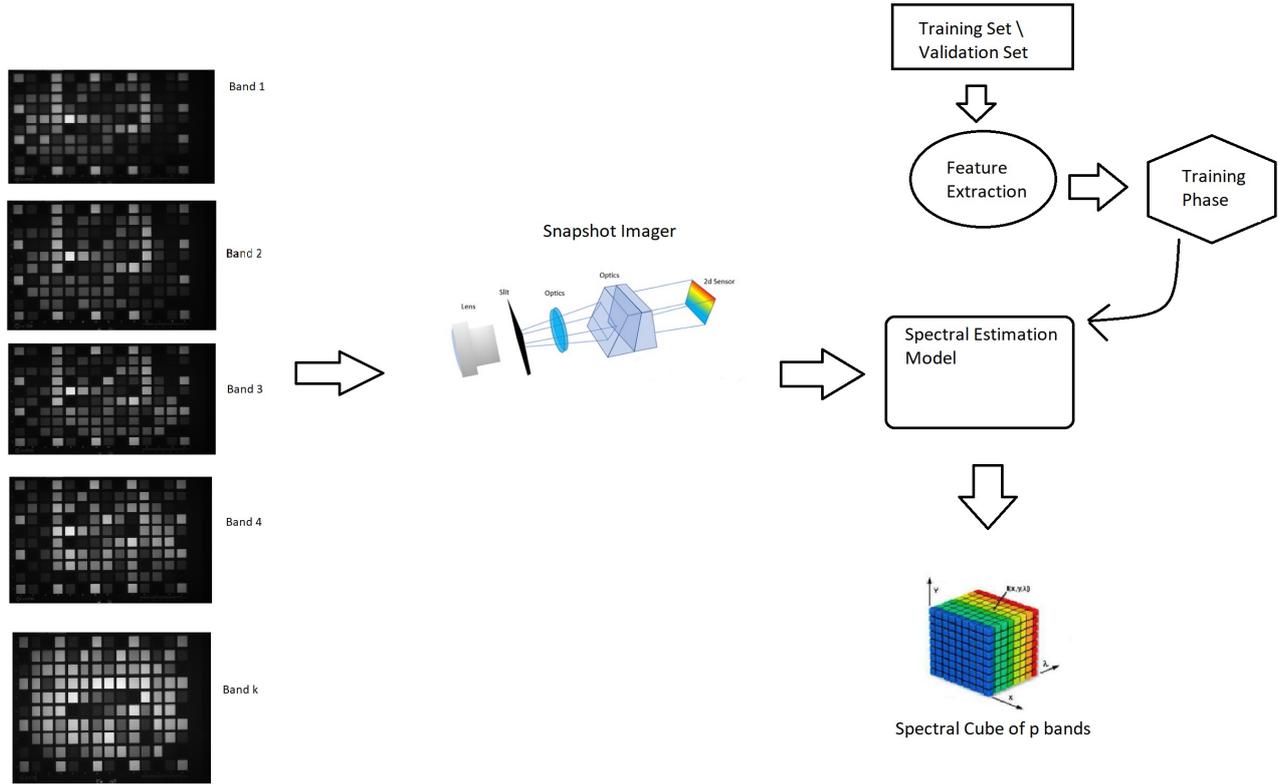


FIGURE 1.9: Spectral Acquisition Chain of a Snapshot Spectral Imager

### 1.9.1 Spectral Acquisition Chain of a Snapshot Spectral Imager

The above figure shows the procedure that take place in order to build a powerful SNSI system.

Firstly, the SNSI system acquires a low number of spectral images, let this number be equal to  $k$ . In this case, the spectral bands that are captured from the SNSI scanner are  $\text{Band}_1, \text{Band}_2, \dots, \text{Band}_k$ . The identity and the number of these bands are investigated in chapter 4.

Secondly, a pre-trained spectral estimation model which takes as input the  $k$  bands and returns as output the bands  $\text{Band}_1, \text{Band}_2, \dots, \text{Band}_p$  will be leveraged. In order to achieve that goal, we should utilize the most informative training set and we should extract the  $k$  more featured bands that force the model to minimize the estimation error. All the above are described analytically at this thesis. In chapter 2 the most well-known spectral estimation methods are reported and in chapter 3 are compared. Conclusions about the effectiveness of each method are made. Also in chapter 3 detailed information

about the experiments are described. In chapter 4 a great amount of feature selection techniques are represented and analyzed in terms of capturing the most informative and distinctive bands.

To make it clear, at this thesis a complete theoretical SNSI system is investigated.



## Chapter 2

# Spectral Estimation - Problem Statement and Methods

### 2.1 Introduction

Hyperspectral imaging is an emerging technology that integrates conventional imaging and spectroscopy to attain both spatial and spectral information from an object. The spectral images collected in the spectral cube are tens of hundreds and the other information we receive from them is crucial for many applications, such as bio-medical technology, remote sensing, microscopy etc. Nevertheless, the current state of the art includes HSI systems which need long acquisition time, something that prevents them from observing any dynamically developing phenomena. Also, are expensive and sizeable, which makes them inaccessible for many important applications. To deal with that, many researchers strive to estimate the spectral domain of the spectral cube by using digital RGB sensors. Using consumer cameras as relatively cheap measurement devices for estimating spectral properties (especially color) has become an interesting alternative to making pointwise spectral measurements. The results obtained with these methods cannot compete with the quality of the traditional devices but they are very attractive since the equipment is relatively cheap and instant measurements are obtained for millions of measurement points. These advantages come, however, at the price of lower-quality measurements, and thus it is of interest to improve the precision of the spectral estimation. Many studies use fixed hardware (RGB cameras), and improvements have to come from better postprocessing algorithms. Wiener and pseudoinverse method [9] are the most traditional ways for spectral reflectance estimation, even in biomedical field [10]–[12]. Murakami et. al. [13] proposed an interesting alternative of dividing into several blocks a three-band image and the spectral estimation is

carried out using the Wiener estimation matrix assigned to each block. Many studies optimized the performance of the above linear methods by defining them locally [14]–[16]. Separate least-squares problems are solved for every target point, where only a pixel or a local neighborhood could be estimated each time. Hence, that methods are computationally expensive and it cannot be leveraged in real-time applications. On the other hand, Mansouri et. al [17] proposed that spectral reflectances can be approximated by a small number of basis functions, that could be extracted from PCA, Fourier, or Wavelets analysis. A local applied version of that method is proposed in [18]. On the other hand, in [19], [20] it was found that linear models could produce adequate results when number of parameters are five to eight. Additionally, the accuracy of reflectance estimation can be degraded by the non-linearity in imaging process, which is due to non-gaussian distribution of the data and the non-linearity of sensors. For that reason, nonlinear kernel based methods considered to be more feasible. Comparative studies of linear and non-linear methods could be found in [21], [22]. Recently, other ways of reconstructing spectral images from RGB images have been explored, including sparse reconstruction and deep learning. Sparse reconstruction methods exploit the sparsity of spectra, recover spectra for every pixels using an learnt overcomplete dictionary [23], [24]. To further improve reconstruction quality, convolutional neural network [25], [26], generative adversarial network [27], and neural network [28] based learning methods which utilize spectral and sometimes spatial information of images have begun to appear.

The drawback of utilizing RGB sensors is the low amount of obtained spectral images, that leads to a three-to-many underdetermined problem. Hence, the accuracy of the aforementioned methods have significant limitations due to the low number of features, therefore, the reflectance reconstruction will always be disputable. Current literature uses also Multi Spectral (MS) imagers, in order to increase the number of the acquired spectral images. In [29], authors used a 16-narrowband wheel-filter MS camera that uniformly covers visible spectrum and the partial least squares method for estimating the missing spectral images. Furthermore, MS imagers was also combined with linear and non-linear methods for reconstructed printed ink's spectrum. Comparative studies could be found in [30], [31]. Moreover, Shen et. al. [32] studied the influence of channels number on performance of local Wiener using liquid crystal tunable filter. The disadvantage of this non-snapshot technology is the disability of observing dynamically developed phenomena, due to the time needed for the acquisition

process. To deal with that, multi-spectral filter arrays (MSFA) or other equivalent snapshot techniques are leveraged. MSFA—based imaging is a compact, convenient, and robust way to acquire and reconstruct spectral images on-shot. The method using a camera with mosaic of the multiple filters on the sensor has become commercially available, e.g., SILIOS technologies developed a manufacturing technique called COLOR SHADES to produce MSFAs. Nevertheless, as the spectral bands increase the spatial resolution decrease and vice versa. Therefore, the optimal number of spectral channels and their spectral sensitivities are strategic decision for designing these systems.

In this study we are focusing on conducting a detail research on estimating the spectral domain using a theoretical one-shot imager equipped with three (RGB) narrow/wide, six, nine, and twelve evenly distributed spectral bands. Various linear and non-linear algorithms are compared in terms of minimizing the root mean square (RMSE) reconstruction error. We propose also two new algorithms for reducing the spectral dimensionality space that need to be estimated and performance comparisons take place. Moreover, respecting the physical constraints of spectra, we suggest a transformation that restricts the output of estimation models in  $[0, 1]$ . Looking at the same problem from another perspective, in order to increase the performance of the models, training procedure should be carried out using the most informative and distinctive bands as features. On that account, a great amount of band selection techniques are analyzed and compared. The reported experimental results are derived from real HS data using standard and measured color charts.

## 2.2 Measures of Spectral Similarities

Before the analyzation of the spectral estimation methods, all mathematical models that measure the similarity of the original reflectance and the estimated reflectance should be taken into consideration. There are plenty of spectral similarity measures but we are going to focus on the most well-known. In the following metrics,  $\mathbf{r}_o$  is a  $p \times 1$  vector and will be the original spectrum. Furthermore,  $\mathbf{r}$  is the estimated spectrum, which has the same size as  $\mathbf{r}_o$ . Note that  $p$  is the number of the reflectance bands, which were estimated.

### 2.2.1 Goodness of Fit Coefficient (GFC)

GFC is defined as the cosine of the angle between the recovered and original signal, thus

$$\text{GFC}(\mathbf{r}_o, \mathbf{r}) = \frac{\mathbf{r}_o^T \mathbf{r}}{\|\mathbf{r}_o\| \|\mathbf{r}\|}. \quad (2.1)$$

### 2.2.2 Spectral Angle Mapper (SAM)

The spectral angle mapper (SAM) determines the spectral similarity between the measured and the reference spectra. The spectra are treated as vectors in a space with dimensionality equal to the number of bands, and the angle that is formed between these vectors is used as a metric of the spectral similarity. Smaller angles represent closer matches to the reference spectrum. SAM has also been used as a feature selection method for selecting an optimal subset of spectral bands. The angle ( $\theta$ ) between pixel vectors as a discrimination measure is given by the following formula:

$$\theta = \text{SAM}(\mathbf{r}_o, \mathbf{r}) = \arccos(\text{GFC}(\mathbf{r}_o, \mathbf{r})) = \arccos\left(\frac{\mathbf{r}_o^T \mathbf{r}}{\|\mathbf{r}_o\| \|\mathbf{r}\|}\right) \quad \text{where } 0 \leq \theta \leq \frac{\pi}{2}. \quad (2.2)$$

SAM is independent of lighting. Also, it is independent to the multiplications of a vector with a natural number since it only increases its length and doesn't change the angle. Hence, SAM is a non-prosthetic distance function.

### 2.2.3 Euclidean Distance (ED)

In mathematics, the Euclidean distance or Euclidean metric is the distance between two points that one would measure with a ruler and is given by the Pythagorean formula. By using this formula as distance, Euclidean space (or even any inner product space) becomes a metric space. The associated norm is called the Euclidean norm. Older literature refers to the metric as Pythagorean metric.

$$\text{ED}(\mathbf{r}_o, \mathbf{r}) = \|\mathbf{r}_o - \mathbf{r}\| = \sqrt{\sum_{i=1}^p (r_{o_i} - r_i)^2}. \quad (2.3)$$

### 2.2.4 Root Mean Square Error (RMSE)

In statistics, the mean square error or MSE of an estimator is one of many ways to quantify the difference between an estimator and the true value of the quantity being estimated. MSE is a risk function, corresponding to the expected value of the squared error loss or quadratic loss. MSE measures the average

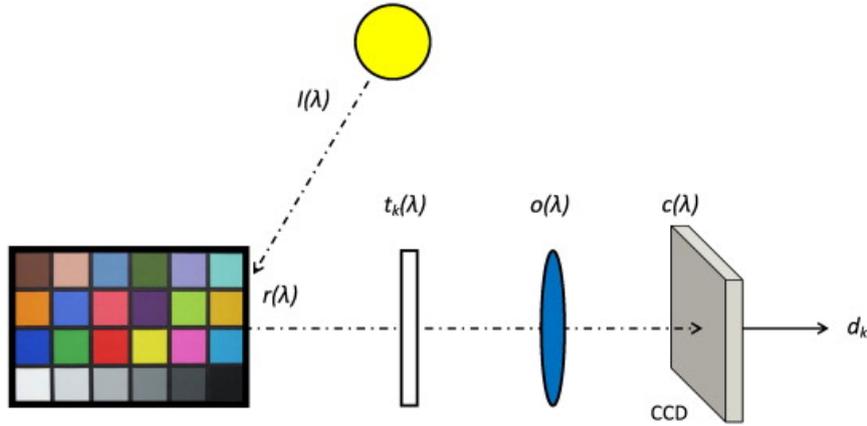


FIGURE 2.1: Synopsis of the spectral model of the acquisition process in a MS system (source: Alamin Mansouri [33]).

of the square of the error. The error is the amount by which the estimator differs from the quantity that is estimated. The difference occurs because of randomness or because the estimator doesn't account for information that could produce a more accurate estimation.

$$\text{RMSE}(\mathbf{r}_o, \mathbf{r}) = \frac{1}{\sqrt{p}} \text{ED}(\mathbf{r}_o, \mathbf{r}) = \sqrt{\frac{\sum_{i=1}^p (r_{o_i} - r_i)^2}{p}}. \quad (2.4)$$

## 2.3 Problem Statement

The generally used spectral model of the acquisition chain in a MS system is illustrated in fig. (2.1), where  $I(\lambda)$  is the spectral radiance of the illuminant,  $r(\lambda)$  is the spectral reflectance of the surface,  $o(\lambda)$  is the spectral transmittance of the optical system,  $t_k(\lambda)$  is the spectral transmittance related to the  $k^{\text{th}}$  filter,  $c(\lambda)$  is the spectral sensitivity of the camera,  $n_k$  represents the spectral noise for the  $k^{\text{th}}$  channel,  $k = [1, 2, \dots, K]$ , and the function  $\mathbf{\Gamma}_k$  collects the nonlinearity of the system. The camera output  $d_k$ , related to the channel  $k$  for a single pixel of the image is given by

$$d_k = \mathbf{\Gamma}_k \left( \int_{\lambda_{\min}}^{\lambda_{\max}} I(\lambda) r(\lambda) o(\lambda) c(\lambda) t_k(\lambda) d\lambda + n_k \right). \quad (2.5)$$

If the noise is assumed to be removed by preprocessing, and assuming a linear opto-electronic transfer function, we can replace  $I(\lambda)$ ,  $c(\lambda)$ ,  $o(\lambda)$ , and  $t_k(\lambda)$  by

the spectral sensitivity  $s_k(\lambda)$  of the  $k_{th}$  channel. Then, Eq. (2.5) becomes

$$d_k = \int_{\lambda_{min}}^{\lambda_{max}} s_k(\lambda)r(\lambda) d\lambda \approx d_k = \mathbf{s}_k^T \mathbf{r} \quad (2.6)$$

where  $\mathbf{s}_k = [s_k(\lambda_1), s_k(\lambda_2), \dots, s_k(\lambda_p)]^T$  is a column vector containing the spectral sensitivity of the acquisition system related to the  $k^{th}$  channel for a single pixel,  $\mathbf{r} = [r(\lambda_1), r(\lambda_2), \dots, r(\lambda_p)]$  is a column vector of the spectral reflectances of the scene for a single pixel, and  $T$  is the transpose vector operator. We use the fact that these spectra represent probabilities (of a photon being reflected from an object point) and that  $r(\lambda_i)$  will therefore lie in the range from 0 to 1. Considering the system with all channels, Eq. (2.6) can be written as

$$\mathbf{d} = \mathbf{S}^T \mathbf{r} \quad (2.7)$$

where  $\mathbf{d}$  is a column vector containing all  $k$  camera outputs  $[d_1, d_2, \dots, d_k]$  for a single pixel, and  $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_k]$  is the matrix containing the channels spectral sensitivities. The final goal is to recover  $\mathbf{r}$  from the camera output according to Eq. (2.7). This is obtained by finding an operator  $\mathbf{Q}$  that solves for the following equation:

$$\mathbf{r} = \mathbf{Q}\mathbf{d}. \quad (2.8)$$

Depending on how the matrix  $\mathbf{S}$  is determined, three paradigms to assess the  $\mathbf{Q}$  matrix exist.

- **Underdetermined Problem:**

If  $\mathbf{S}$  is obtained by a direct physical system characterization (combination of all spectral responses of the acquisition system components, that provided by the manufacturer), then the operator  $\mathbf{Q}$  is the inverse of  $\mathbf{S}$ . However, because  $\mathbf{S}$  is a not usually a square matrix, its inverse does not exist. A choice could be the pseudo-inverse  $\mathbf{Q} = \mathbf{S}^-$ , where  $\mathbf{S}^-$  is the pseudo-inverse matrix of  $\mathbf{S}$ , which is equal to  $(\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T$ .

- **Overdetermined Problem:**

If  $\mathbf{S}$  is obtained indirectly by matching a set of  $J$  color patches for which the theoretical reflectances are already known, we only have to capture an image of these patches with the MS camera. Then we produce a set of corresponding pairs  $(\mathbf{d}_j, \mathbf{r}_{oj})$  for  $j = 1, 2, \dots, J$ , where  $\mathbf{d}_j$  is a vector of dimension  $k$  containing

the camera responses and  $\mathbf{r}_{oj}$  is a vector of dimension  $p$  representing the spectral reflectance of the  $j^{\text{th}}$  patch. The reflectances  $\mathbf{r}_{oj}$  are gathered in the matrix  $\mathbf{R} = [\mathbf{r}_{o1}, \mathbf{r}_{o2}, \dots, \mathbf{r}_{oJ}]$  and the camera outputs for the  $J$  patches are gathered in the matrix  $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_J]$ , in order that  $\mathbf{D} = \mathbf{S}^T \mathbf{R}$ . The operator  $\mathbf{Q}$  is straightforwardly assessed by any optimization method (such methods will be described at the next sections).

• **Interpolation Problem:**

A third paradigm for spectral reflectance estimation consists of direct interpolation of the camera outputs  $d_k$ . In this case,  $\mathbf{S}$  is a selection matrix and specifically a submatrix of the identity matrix. Nevertheless, rigorous conditions about the spectral shape of the filters as well as well-calibrated and normalized data are required for this kind of reconstruction. The reconstruction can be performed by any interpolation operator (linear, spline, etc.).

This study is focusing on the second way (learning from examples). Assuming we have the labeled training set  $(\mathbf{d}_j, \mathbf{r}_{oj})$  for  $j = 1, 2, \dots, J$ , one possibility is to solve the estimation problem using  $L^2$  error on a linear  $p$ -dimensionality space

$$\arg \min_Q \sum_{j=1}^J \|\mathbf{r}_{oj} - \mathbf{Q} \mathbf{d}_j\|^2. \quad (2.9)$$

Overfitting of the model to training data can cause problems. System 2.9 can be interpolated an infinite number of ways and can be interpreted as an ill-posed problem. Additional constraints and regularization terms can be introduced to modify the problem to well-posed. Using Tikhonov regularization, this leads to the variational problem

$$\arg \min_Q \sum_{j=1}^J \left( \|\mathbf{r}_{oj} - \mathbf{Q} \mathbf{d}_j\|^2 + \gamma \|\mathbf{Q}\|_F^2 \right), \quad (2.10)$$

where  $\|\mathbf{Q}\|_F^2$  denotes the Frobenius norm. The selection of the gamma parameter in system 2.10 depends mainly on the noise properties of the measurement device and allows the control of the trade-off between the norm of the solution and the accuracy of the data fitting. Noise in response values can be reduced by averaging over repeated measurements. For flat, spatially homogenous surfaces, averaging can also be done over spatial regions.

In the following section, plenty of methods that are based on second paradigm (overdetermined case) will be analyzed and evaluated.

## 2.4 Spectral Estimation Algorithms

### 2.4.1 Estimation of Spectral Reflectances Using Wiener Method

The purpose of the Wiener estimation (Stigell et al. [9]) is to make estimations from low-dimensional data into high-dimensional data, for example, from six-filter camera responses into reflectance spectra. The Wiener estimation is one of the conventional estimation methods which provides accurate estimations. The purpose of the Wiener model is to assess the matrix  $\mathbf{Q}$  in terms to minimize the square error (MSE) between reference  $\mathbf{r}_o$  reflection signal and the estimated  $\mathbf{r}$ . Let us assume that the data are centered. The solution of the Wiener method is denoted by

$$\mathbf{Q}^{\text{Wiener}} = R_{\mathbf{r}_o\mathbf{d}}(R_{\mathbf{d}\mathbf{d}} + \gamma\mathbb{I})^{-1} = \mathbb{E}[\mathbf{r}_o\mathbf{d}^T](\mathbb{E}[\mathbf{d}\mathbf{d}^T] + \gamma\mathbb{I})^{-1} \approx \frac{1}{J^2}\mathbf{R}\mathbf{D}^T(\mathbf{D}\mathbf{D}^T + \gamma\mathbb{I})^{-1}, \quad (2.11)$$

where  $\mathbb{E}$  is the expected value. In Eq. (2.11),  $R_{\mathbf{r}_o\mathbf{d}}$  is the cross-correlation matrix of vectors  $\mathbf{r}_o$  and  $\mathbf{d}$ . Matrix  $R_{\mathbf{d}\mathbf{d}}$  is the auto-correlation matrix of the vector  $\mathbf{d}$ . Eq. (2.11) could be rewritten as

$$\mathbf{Q}^{\text{Wiener}} \approx \frac{1}{J^2}\mathbf{R}\mathbf{D}^T(\mathbf{D}\mathbf{D}^T + \gamma\mathbb{I})^{-1} = \frac{1}{J^2}(\mathbf{D}^T\mathbf{D} + \gamma\mathbb{I})^{-1}\mathbf{D}^T\mathbf{R}. \quad (2.12)$$

The time complexity of Wiener algorithm is  $O(pk)$ . The training phase costs a constant number of multiplications due to the fact that takes place only one time.

### 2.4.2 Estimation of Spectral Reflectances Using Projection on PCA

Principal component analysis (PCA) was invented in 1901 by Karl Pearson et al. [34] and is a technique extensively used for dimensionality reduction in a data set. It consists of finding an orthogonal basis composed of vectors called principal components. Each component is associated with an energy that indicates the statistical relevance of the vector in the data. Technically speaking, PCA is an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data lies on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. In the field of MS imaging, PCA has been largely used for data compression but also in spectral

reflectance reconstruction. The main difference from Wiener estimation is that we are approaching  $\mathbf{r}$  by using the basis matrix  $\mathbf{B}$  and the weight vector  $\mathbf{a}$  such that each reflectance  $\mathbf{r}$  could be written as

$$\mathbf{r} = \mathbf{B}\mathbf{a}. \quad (2.13)$$

Let us assume that data are centered. Then the  $p \times p$  covariance matrix  $\mathbf{C}$  of  $\mathbf{r}_o$  is given by

$$\mathbf{C} = \mathbb{E}[\mathbf{r}_o \mathbf{r}_o^T] \approx \frac{1}{J} \sum_{i=1}^J \mathbf{r}_{oj} \mathbf{r}_{oj}^T = \frac{1}{J} \mathbf{R}\mathbf{R}^T. \quad (2.14)$$

It is a symmetric matrix and, therefore, it can be diagonalized. Hence,  $\mathbf{C}$  of  $\mathbf{r}_o$  can be rewritten as

$$\mathbf{C} = \mathbf{V}\mathbf{L}\mathbf{V}^T \quad (2.15)$$

where  $\mathbf{V}$  is a matrix of eigenvectors (each column is an eigenvector) and  $\mathbf{L}$  is a diagonal matrix with eigenvalues  $\lambda_i$  in the decreasing order on the diagonal. The eigenvectors are called principal axes or principal directions of the data. Projections of the data on the principal axes are called principal components. Note that the eigenvectors could be also calculated using singular value decomposition. We define the  $p \times m$  matrix  $\mathbf{B}$  as

$$\mathbf{B} = m \text{ first column vectors of matrix } \mathbf{V}. \quad (2.16)$$

The basis functions are themselves functions of wavelength but free of constraints such as being positive or constrained to be limited to the range  $[0 - 1]$ . Eq. (2.7) can be written as

$$\mathbf{d} = \mathbf{S}^T \mathbf{B}\mathbf{a} \quad (2.17)$$

where  $\mathbf{a}$  is a  $m \times 1$  vector that holds the weights that define the particular spectrum which we are trying to reconstruct. Furthermore,  $\mathbf{S}^T$  is a  $k \times p$  matrix which it is assessed using one of the paradigms that are described in Section 2.3. An approximate solution of Eq. (2.17) is

$$\mathbf{a} = (\mathbf{S}^T \mathbf{B})^{-} \mathbf{d}. \quad (2.18)$$

Totally,

$$\mathbf{Q}^{\text{PCA}} = \mathbf{B}(\mathbf{S}^T \mathbf{B})^{-}. \quad (2.19)$$

The time complexity of PCA spectral estimation algorithm is  $O(pm)$ . The training phase costs a constant number of multiplications due to the fact takes

place only one time.

Notice that a mean function is subtracted from the training set before deriving the basis and this mean function should be added when using the basis for the reconstruction. Mansouri et al. [17] studied the performance of PCA spectral estimation, alongside with Fourier and Wavelets analysis.

### 2.4.3 Sparse Dictionary Learning (K-SVD)

Sparse dictionary learning is a representation learning method which aims at finding a sparse representation of the input data (also known as sparse coding) in the form of a linear combination of basic elements as well as those basic elements themselves. These elements are called atoms and they compose a dictionary. Atoms in the dictionary are not required to be orthogonal, and they may be an over-complete spanning set. This problem setup also allows the dimensionality of the signals being represented to be higher than the one of the signals being observed. The above two properties lead to having seemingly redundant atoms that allow multiple representations of the same signal but also provide an improvement in sparsity and flexibility of the representation. One of the key principles of dictionary learning is that the dictionary has to be inferred from the input data. The emergence of sparse dictionary learning methods was stimulated by the fact that in signal processing one typically wants to represent the input data using as few components as possible. Before this approach the general practice was to use predefined dictionaries (such as fourier or wavelet transforms). However, in certain cases a dictionary that is trained to fit the input data can significantly improve the sparsity, which has applications in data decomposition, compression and analysis and has been used in the fields of image denoising and classification, video and audio processing. Recently, other ways of reconstructing spectral images from raw response images (usually RGB images) have been explored, including sparse reconstruction and deep learning. Sparse reconstruction methods exploit the sparsity of spectra, recover spectra for every pixels using an learnt overcomplete dictionary [23], [24], [35]. Given the input dataset  $\mathbf{R} = [\mathbf{r}_{o1}, \mathbf{r}_{o2}, \dots, \mathbf{r}_{oJ}], \mathbf{r}_{oj} \in \mathbb{R}^p$  we wish to find a dictionary  $\mathbf{Z} \in \mathbb{R}^{p \times m} : \mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_m]$  and a representation matrix  $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_J], \mathbf{a}_i \in \mathbb{R}^m$  such that  $\|\mathbf{R} - \mathbf{Z}\mathbf{A}\|_F^2$  is minimized and the representations  $\mathbf{a}_i$  are sparse enough. This can be formulated as the following optimization problem:

$$\arg \min_{\mathbf{Z}, \mathbf{A}} \|\mathbf{R} - \mathbf{Z}\mathbf{A}\|_2^2, \text{ s.t. } \forall i \in [1, m], \|\mathbf{a}_i\|_0 \leq s, \quad (2.20)$$

where  $s$  is the sparsity constraint. The minimization problem above is not convex because of the  $L_0$ -norm and solving this problem is NP-hard. In some cases  $L_1$ -norm is known to ensure sparsity and so the above becomes a convex optimization problem with respect to each of the variables  $\mathbf{Z}$  and  $\mathbf{A}$  when the other one is fixed, but it is not jointly convex in  $\mathbf{Z}, \mathbf{A}$ . The dictionary defined above can be "undercomplete" if  $p < m$  or "overcomplete" in case  $p > m$  with the latter being a typical assumption for a sparse dictionary learning problem. Overcomplete dictionaries do not require the atoms to be orthogonal (they will never be a basis anyway) thus allowing for more flexible dictionaries and richer data representations. As the optimization problem described above can be solved as a convex problem with respect to either dictionary or sparse coding while the other one of the two is fixed, most of the algorithms are based on the idea of iteratively updating one and then the other.

K-SVD is an algorithm that performs SVD at its core to update the atoms of the dictionary one by one and basically is a generalization of K-means. It enforces that each element of the input data is encoded by a linear combination of not more than  $s$  elements. This algorithm's essence is to first fix the dictionary, find the best possible  $\mathbf{A}$  under the above constraint (using Orthogonal Matching Pursuit) and then iteratively update the atoms of dictionary  $\mathbf{Z}$  [36]. The next step is to reform Eq. (2.13) for the dictionary case. Therefore, Eq. (2.13) could be rewritten as

$$\mathbf{r} = \mathbf{Z}\mathbf{a}' \quad (2.21)$$

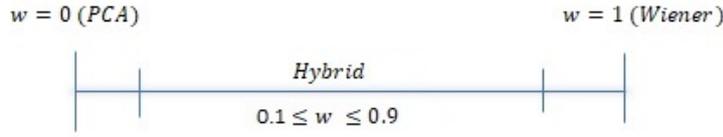
where  $\mathbf{a}' \in \mathbb{R}^m$ . The coefficient vector  $\mathbf{a}'$  can be optimized as

$$\mathbf{a}' = (\mathbf{S}^T \mathbf{Z})^{-1} \mathbf{d}. \quad (2.22)$$

The time complexity of the sparse is  $O(pm)$ . The training phase costs a constant number of multiplications due to the fact takes place only one time.

#### 2.4.4 Estimation of Spectral Reflectances Using Hybrid Method

Hybrid is a finite-dimensional algorithm that is based on combination of models in order to achieve better estimations with the minimum error [37]. This algorithm combines Wiener (Stigell et al. [9]) and PCA (Mansouri et al. [17]).

FIGURE 2.2: Hybrid model with different values of  $w$ .

Specifically, we have already known from Eq. (2.11) that Wiener Estimation is given by

$$\mathbf{r} = R_{\mathbf{r}_o\mathbf{d}} R_{\mathbf{d}\mathbf{d}}^{-1} \mathbf{d} \quad (2.23)$$

additionally, from Eq. (2.13)

$$\mathbf{r} = \mathbf{B}\mathbf{a}. \quad (2.24)$$

As a result, a least squares approximation of  $\mathbf{a}$  is

$$R_{\mathbf{r}_o\mathbf{d}} R_{\mathbf{d}\mathbf{d}}^{-1} \mathbf{d} = \mathbf{B}\mathbf{a}' \implies \mathbf{a}' = \mathbf{B}^{-1} R_{\mathbf{r}_o\mathbf{d}} R_{\mathbf{d}\mathbf{d}}^{-1} \mathbf{d}. \quad (2.25)$$

Note that  $\mathbf{B}$  is a orthogonal matrix, therefore,  $\mathbf{B}^T \mathbf{B} = \mathbb{I}$  where  $\mathbb{I}$  is the identity matrix and  $\mathbf{B}^{-1} = \mathbf{B}^T$ . Ending up, the Eq. (2.24) can be rewritten as

$$\mathbf{r} = \mathbf{B}\mathbf{B}^T R_{\mathbf{r}_o\mathbf{d}} R_{\mathbf{d}\mathbf{d}}^{-1} \mathbf{d}. \quad (2.26)$$

In order to achieve the contribution of the two above methods, we linearly merge Eq. (2.23) and Eq. (2.26), therefore, we can express the Hybrid model as

$$\mathbf{r}'' = w R_{\mathbf{r}_o\mathbf{d}} R_{\mathbf{d}\mathbf{d}}^{-1} \mathbf{d} + (1 - w) \mathbf{B}\mathbf{B}^T R_{\mathbf{r}_o\mathbf{d}} R_{\mathbf{d}\mathbf{d}}^{-1} \mathbf{d} \quad (2.27)$$

where  $w \in [0.1 \ 0.9]$  is scaling factors between Wiener and PCA model that were presented in Eq. (2.26). Hence,  $\mathbf{Q}^{\text{Hybrid}}$  it can be written as

$$\mathbf{Q}^{\text{Hybrid}} = w R_{\mathbf{r}_o\mathbf{d}} R_{\mathbf{d}\mathbf{d}}^{-1} + (1 - w) \mathbf{B}\mathbf{B}^T R_{\mathbf{r}_o\mathbf{d}} R_{\mathbf{d}\mathbf{d}}^{-1} = [w\mathbb{I} + (1 - w) \mathbf{B}\mathbf{B}^T] R_{\mathbf{r}_o\mathbf{d}} R_{\mathbf{d}\mathbf{d}}^{-1}. \quad (2.28)$$

Moreover, the Gradient Multipliers technique will be utilized to find the value of  $w$ . Hence, finding  $w$  could be described as a linear programming problem.

$$\begin{aligned} \min_w \quad & \mathbb{E} \|\mathbf{r}_o - \mathbf{r}''\|^2 \\ \text{subject to} \quad & 0.1 \leq w \leq 0.9. \end{aligned} \quad (2.29)$$

When  $w = 0$  the Hybrid model is equivalent to PCA model and when  $w = 1$  is equivalent to Wiener model as it is illustrated in fig. (2.2). The first constraint ensures that the weightings of the two techniques summarize to 1, while the

second constraint forces positive contribution ( $\geq 0.1$ ) of each method. The objective functions together with the constraints can be solved using linear or quadratic programming. Another important aspect is the selection of the number of  $m$ . Usually, this number is chosen such that the highest amount of signal's energy is covered. In this study, Wiener and PCA models are exploited to build a weighted average combination model. Undeniably, that algorithms could be constituted by any two or more techniques, which are presented at this chapter. The time complexity of any Hybrid algorithm is  $O(pk)$ . The training phase costs a constant number of multiplications due to the fact that takes place only one time.

### 2.4.5 Estimation of Spectral Reflectances Using Adaptive(Local)-Wiener Method

Shen et al. [32] proposed a method to reconstruct spectral reflectance by using a modified Wiener estimation method, without a prior knowledge of the spectral characteristics of the samples being imaged. Shen investigated whether the accuracy of reflectance estimation can be improved if the training samples for calculating the reflectance characteristics are adaptively selected and appropriately weighted. The novelty of the proposed method is mainly in the manner of training sample selection and autocorrelation matrix construction. Suppose that the response of the spectral camera is  $\mathbf{d}$ , and then its corresponding reflectance  $\mathbf{r}$  can be calculated according to the traditional Wiener estimation using Eq. (2.11). The training samples with reflectance  $\mathbf{r}_{oi}$  can then be selected according to their spectral similarity to  $\mathbf{r}$ . In the calculation of spectral similarity, the reflectance is normalized so that its summation is equal to 1. The reason for normalization is that statistical information of the reflectance is mainly decided by its shape, not magnitude. The spectral similarity consists of two terms, i.e., mean spectral distance and maximum spectral distance

$$\text{dist}_i = \alpha \frac{1}{p} \left\| \frac{\mathbf{r}_{oi}}{\|\mathbf{r}_{oi}\|_\infty} - \frac{\mathbf{r}}{\|\mathbf{r}\|_\infty} \right\|_1 + (1 - \alpha) \left\| \frac{\mathbf{r}_{oi}}{\|\mathbf{r}_{oi}\|_\infty} - \frac{\mathbf{r}}{\|\mathbf{r}\|_\infty} \right\|_\infty \quad (2.30)$$

where  $\alpha$  is a scaling factor,  $\frac{1}{p} \|x\|_1$  is the mean value of  $x$ , and the  $\|x\|_\infty$  is the maximum value of  $x$ . For two reflectances with similar shapes both of the mean and maximum distances should not be large. When  $\mathbf{r}$  is very similar to  $\mathbf{r}_{oi}$ ,  $\text{dist}_i$  is close to 0. We select  $L$  training samples according to their spectral similarities and sort the similarity measurements in ascending order, or more specifically,  $\text{dist}_1 \leq \text{dist}_2 \leq \dots \leq \text{dist}_L$ . Among the selected  $L$  training samples

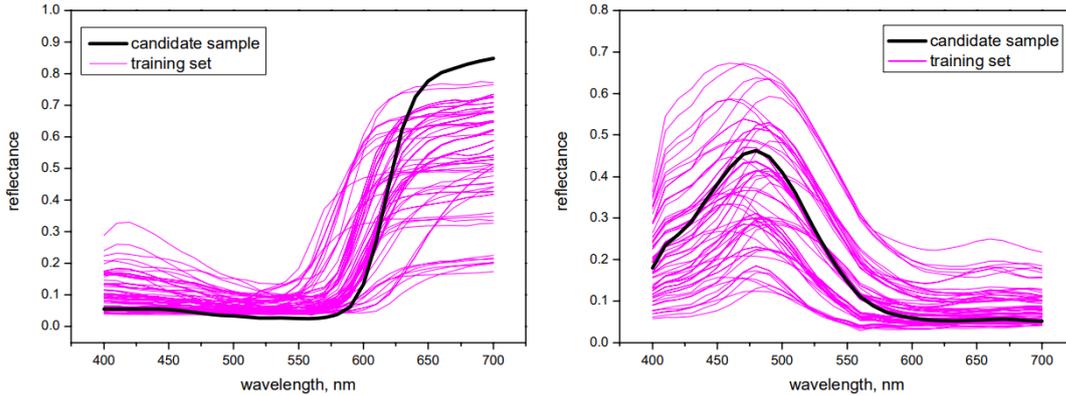


FIGURE 2.3: Two typical examples of the training sets for the given candidate samples. Channel number  $C=6$ -Adaptive Wiener (source: Hui-Liang Shen [32]).

it is reasonable to assume that the spectral characteristic of the more similar  $\mathbf{r}_{oi}$  ( $1 \leq i \leq L$ ) should be more close to that of the candidate sample. Then the recalculation of the Wiener model is needed using the new adapted training set. It is noted that Wiener need to be calculated per pixel in a MS image and thus the proposed method is computationally expensive than the traditional Wiener estimation. This is actually the common shortcoming for almost all adaptive methods.

Figure 2.3 shows the selected training sets for the given candidates. As expected, the shapes of the training samples are similar to those of the candidate samples. The performance of the proposed adaptive Wiener estimation was investigated by Shen for different channel numbers. Experimental results of Shen, indicate that the proposed method is significantly better than the Wiener estimation when the channel number is not large (for example, 6 or 7), while is slightly better than or close to the traditional Wiener estimation when channel number 11 or more.

As it mentioned before, adaptive methods are really computationally expensive. The time complexity of adaptive-Wiener algorithm is  $\Theta(pixels \times (2pk + J \log J))$ , where  $pk$  is a constant number describes the number of multiplications need to be done to calculate two times the vector  $\mathbf{r}$  (one for similarity check and one for the final spectral estimation),  $pixels$  is the number of pixels of the image, and  $J \log J$  is the time complexity of sorting a  $J$ -element similarity matrix (i.e Merge Sort). The last could be improved to  $\Theta(J)$  by finding the  $L^{th}$  smallest element and partitioning all the other elements according to  $L^{th}$  smallest element. Partitioning could be handled using the partition method of quick sort algorithm.

In order to implement adaptive methods we need a large training set to guarantee the adaptation of different surfaces and colors. In other words, adaptive methods can produce decent spectral estimation given that there are similar reflectances in the training set.

### 2.4.6 Estimation of Spectral Reflectances Using Multivariate Linear Regression Method

In spectral estimation, Linear Regression (LR) called also pseudo-inverse method [38]. In statistics, linear regression is a linear approach to modeling the relationship between a scalar response (or dependent variable) and one or more explanatory variables (or independent variables). The case of one explanatory variable is called simple linear regression. For more than one explanatory variable, the process is called multiple linear regression. This term is distinct from multivariate linear regression, where multiple correlated dependent variables are predicted, rather than a single scalar variable. In linear regression, the relationships are modeled using linear predictor functions whose unknown model parameters are estimated from the data. Such models are called linear models. Most commonly, the conditional mean of the response given the values of the explanatory variables (or predictors) is assumed to be an affine function of those values; less commonly, the conditional median or some other quantile is used. Like all forms of regression analysis, linear regression focuses on the conditional probability distribution of the response given the values of the predictors, rather than on the joint probability distribution of all of these variables, which is the domain of multivariate analysis. Linear regression models are often fitted using the least squares approach, but they may also be fitted in other ways, such as by minimizing the “lack of fit” in some other norm (as with least absolute deviations regression), or by minimizing a penalized version of the least squares cost function as in ridge regression (L2-norm penalty) and lasso (L1-norm penalty). Conversely, the least squares approach can be used to fit models that are not linear models. Thus, although the terms “least squares” and “linear model” are closely linked, they are not synonymous. In spectral estimation approach our goal is to find the best linear estimation (regression). So,  $\mathbf{r}$  is given by

$$\mathbf{r} = \mathbf{Q}\mathbf{d} + \mathbf{b} \implies \mathbf{r} = [\mathbf{Q} \ \mathbf{b}] \begin{bmatrix} \mathbf{d} \\ 1 \end{bmatrix} \quad (2.31)$$

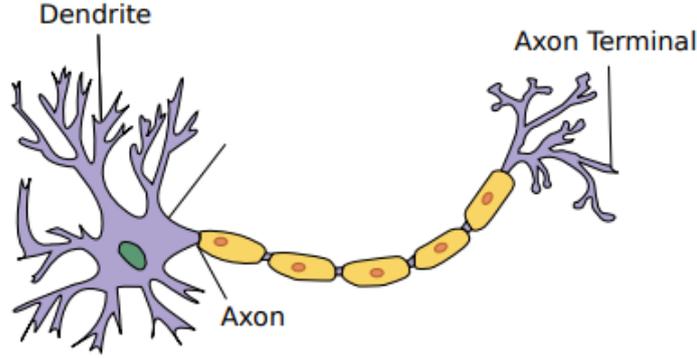


FIGURE 2.4: Model of a biological neuron (source: Wikipedia).

where  $\mathbf{Q}$  is a  $p \times k$  matrix and  $\mathbf{b}$  is a  $p \times 1$  vector. Since, the criterion is to minimize the sum of squared error,  $\mathbf{Q}^{\text{LR}}$  is expressed as

$$\mathbf{Q}^{\text{LR}} = [\mathbf{Q} \mathbf{b}], \quad (2.32)$$

where  $\mathbf{Q}'$  is a  $p \times (k + 1)$  matrix. Additionally, we define  $\mathbf{d}^* = [\mathbf{d} \ 1]^T$  which is a  $(k + 1) \times 1$  vector. Due to this fact, the training set pairs will be  $(\mathbf{r}_{oi}, \mathbf{d}_i^*)$ . A choice to estimate the  $\mathbf{Q}'$  matrix is to minimize the sum of square errors. Therefore,

$$\arg \min_{\mathbf{Q}^{\text{LR}}} \sum_{j=1}^J \left( \|\mathbf{r}_{oj} - \mathbf{Q}^{\text{LR}} \mathbf{d}_j^*\|^2 + \gamma \|\mathbf{Q}^{\text{LR}}\|_F^2 \right) \quad (2.33)$$

The formulation of the problem its the same as Wiener, hence the only difference is that  $\mathbf{d}$  should be replaced by  $\mathbf{d}^*$ . Then, we use Eq. (2.31) to calculate  $\mathbf{r}$ .

### 2.4.7 Estimation of Spectral Reflectances Using Artificial Neural Networks

A. Mansouri et al. [28] proposed Artificial Neural Networks as a way for spectral estimation. Inspired by biological nervous systems, artificial Neural networks (ANNs) aim at reaching their versatility through learning. ANNs are commonly employed in artificial intelligence, machine learning and pattern recognition. There has been substantial research into how the human brains structure achieves such a high level of versatility. This research has provided some important insights, however the conclusions are far from completely explaining the complex functioning of the brain. Even though we have not been able to replicate the brain so far, the field of artificial intelligence offers very effective solutions to many problems by simulating the observations of biological research of various nervous systems. It is estimated, that the average human brain contains 86 billion neurons. Together they form a huge network. Even if we knew

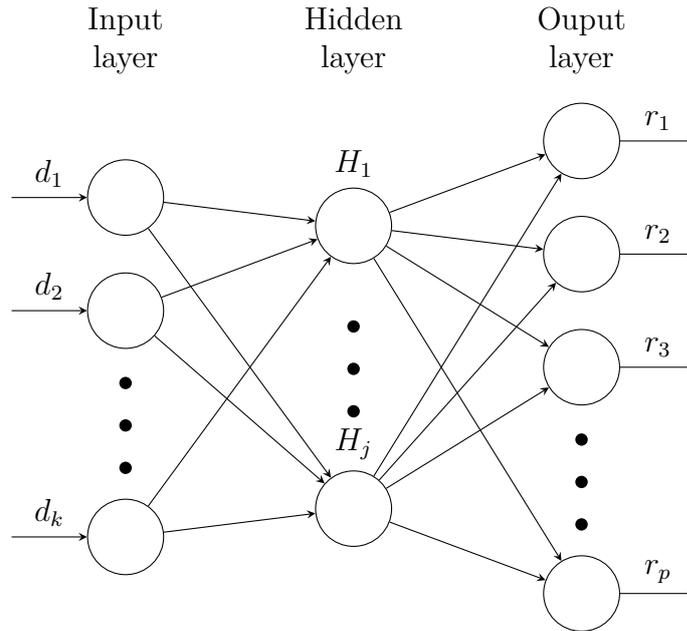


FIGURE 2.5: The Neural Network for the spectral estimation problem.

the detailed inner structure of the human brain, we would still not be able to simulate it with current technology because of its robustness. Our efforts are therefore rather different. We want to build a neural network with a good ratio between its size and its effectiveness.

Generally, ANNs consist of a set of artificial neurons. Formally, an artificial neuron has  $k$  inputs represented as a vector  $\mathbf{d}$ . Inputs in an artificial neuron correspond to the dendrites in a biological neuron, while a single output of an artificial neuron corresponds to the axon in a biological neuron, which is depicted in figure 2.4. Each input  $i$ ,  $1 \leq i \leq k$ , has an assigned weight  $w_1, \dots, w_k$ . Weighted input values are combined and run through an activation function producing some output  $\mathbf{r}$ , as shown in figure 2.6. The network is formed by connecting the neuron output with the input of a different neuron. ANN is therefore effectively described as an oriented graph as shown in figure 2.5, where vertices represent the neurons, and oriented edges represent the output-input connections between them. A set of input neurons consists of the neurons which are the first ones in any complete path in the graph. All input neurons have exactly one input, and all inputs together represent an instance of the problem to be solved by the ANN. A set of output neurons consists of the neurons which are the last ones in any complete path in the graph. All output neurons have exactly one output, and all outputs together represent a possible solution to the problem to be solved by the ANN. A set of hidden neurons consists of the

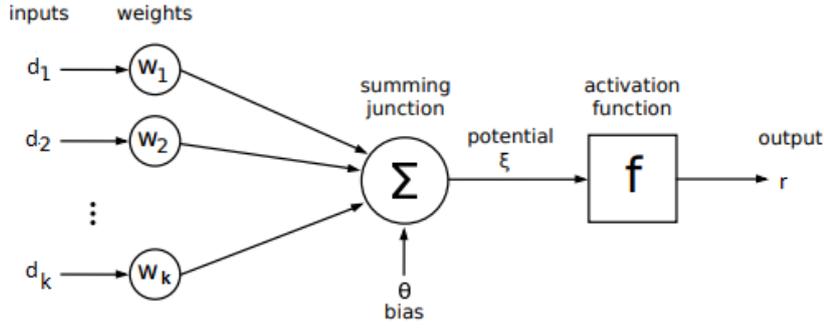


FIGURE 2.6: Model of a biological neuron (source: Wikipedia).

neurons which are not input, nor output neurons. Their number and organization into layers may vary even for the same problem, but is a key feature of the network vastly influencing its performance. An ANN works by feeding the data into the input neurons. The data flows in the direction of oriented edges and ends when the output neurons are hit. The result is interpreted from the values obtained in the output neurons. Formally, an ANN is a 6-tuple  $M = (N, C, k, p, w, t)$ , where

- $N$  is a finite non-empty set of neurons,
- $C \subseteq N \times N$  is a non-empty set of oriented edges between the neurons.
- $k \subset N$  is a non-empty set of neurons in the input layer.
- $p \subset N$  is a non-empty set of neurons in the output layer.
- $w : C \mapsto \mathbb{R}$  is a weighting function.
- $t : N \mapsto \mathbb{R}$  is a function for network bias

Let us consider neuron  $j$  with its input  $\mathbf{d}^{(j)} = (d_1^{(j)}, \dots, d_k^{(j)})$ , weights  $w_1^{(j)}, \dots, w_k^{(j)}$  and bias  $\theta^{(j)}$ . Then the potential of the neuron is computed

$$\xi^{(j)} = \sum_{i=1}^k w_i^{(j)} d_i^{(j)} + \theta^{(j)}. \quad (2.34)$$

Consider the following activation function:

$$f(\xi) = \frac{1}{1 + \exp(-\xi^{(j)})}. \quad (2.35)$$

Then the output  $\mathbf{r}_i$  of the neuron  $j$  is computed:

$$r_i^{(j)} = f(\xi^{(j)}) = f\left(\sum_{i=1}^k w_i^{(j)} d_i^{(j)} + \theta^{(j)}\right). \quad (2.36)$$

Considering ANN containing  $p$  such neurons in the output layer, we obtain the output of the network as  $(r_1, \dots, r_p)$ . During the learning process, ANNs may change their weights, bias, or in some networks even the number of neurons and their setup. In contrast, our definition does not allow such modifications and  $M$  does not change in the process of learning. The Multilayer perceptron (MLP) is a feed-forward neural network consisting of multiple mutually interconnected layers of neurons. The layers are stacked one onto each other. Every neuron in one layer is connected to every neuron in the following layer. The motivation behind designing multilayer networks is to be able to solve more complex tasks. The unit step function is usually not a suitable activation function to be used with the perceptrons. Because of continuity and greater flexibility, the sigmoid function is most commonly used instead. When choosing the most suitable activation function, we want it to be differentiable at every point of its domain, and to be non-linear. Non-linearity is important, because in general we want the output to be non-linearly dependent on the given input. Perceptrons are arranged into  $k \geq 2$  layers. Let us consider a network  $M$  with  $z$  layers. The set of neurons  $C$  is split into mutually distinct subsets called layers  $L_1, \dots, L_z$ . More formally it holds  $\forall i, j : 1 \leq i, j \leq z, (L_i \neq \emptyset \wedge L_i \cap L_j \neq \emptyset) \implies i = j$ . The network layers are stacked one onto each other,  $L_1$  being the input layer,  $L_2, \dots, L_{z-1}$  being the hidden layers and  $L_z$  being the output layer. As shown in figure 2.5, the edges are all oriented in the direction from the input layer  $L_1$  towards the output layer  $L_z$ . Each neuron in layer  $L_i$  is connected to every neuron in layer  $L_{i+1}$ . In other words all neighboring layers form complete bipartite graphs.

The output of the network is computed sequentially, layer by layer. The neural network is trained using the backpropagation technique. The purpose of the back-propagation algorithm is to adjust the synaptic weights of neurons, so that the network produces the desired output. The algorithm describes the process of training (also called learning). The result of this algorithm is a neural network configured to minimize the error when solving given problem. Training must be performed on labeled data ( $\mathbf{d}_i \mapsto \mathbf{r}_{oi}$ ) and therefore is supervised.

### 2.4.8 Kernel Regression

In some cases the linear models seem to be sufficient to provide an accurate estimation (for example, for the measurements of skylight reported in [39]). Nonlinear methods are certainly necessary for nonlinear measuring devices, but even for essentially linear systems they are able to improve the estimation results

(see [40]–[42]). Intuitively, this can be expected, since the spectra of interest usually do not fill the whole linear space. Instead, they are typically restricted to regions of specific geometries. In kernel regression case the minimization problem 2.10 could be reformed as

$$\arg \min_{\mathbf{Q}} \sum_{j=1}^J \left( \|\mathbf{r}_{oj} - \mathbf{Q}\Phi(\mathbf{d}_j)\|^2 + \gamma \|\mathbf{Q}\|_H^2 \right), \quad (2.37)$$

where  $\Phi(\mathbf{d})$  is function of mapping the measurements  $\mathbf{d}$  to a feature space  $H$  and  $\|\mathbf{Q}\|_H^2$  is the norm in a reproduced kernel hilbert space (RKHS) defined by kernel function  $k(\mathbf{x}, \mathbf{z})$ . For instance the polynomial features  $\Phi$  of the vector  $\mathbf{d}$  are computed as

$$\Phi : \mathbf{d} \in \mathbb{R}^k \longrightarrow (1, \mathbf{d}, \mathbf{d}^2, \dots, \mathbf{d}^{W-1}) \in \mathbb{R}^W. \quad (2.38)$$

polynomial feature maps are one example of reproducing kernel Hilbert spaces and were investigated in [43]. Computing actual feature maps is computationally expensive, since the feature vectors can be of a very high dimensionality. In the framework of kernel methods [43], [44], the computation of the mapping  $\Phi : \mathbf{d} \in \mathbb{R}^k \mapsto H$  to feature space  $H$  is avoided by combining it with the subsequent projection step via the positive definite kernel function  $k(\mathbf{x}, \mathbf{z}) = \Phi(\mathbf{x})^T \Phi(\mathbf{z})$ . In that case,  $\mathbf{r}$  is given by

$$\mathbf{r} = \mathbf{k}(\mathbf{d}, \mathbf{d}_j) \mathbf{Q}^{\text{kernel}}, \quad (2.39)$$

where  $\mathbf{d}$  is the actual camera measurement,  $\mathbf{d}_j$  is the  $j^{\text{th}}$  sample of training set,  $\mathbf{k}(\mathbf{d}, \mathbf{d}_j) \in \mathbb{R}^{1 \times J}$ , and  $\mathbf{Q}^{\text{kernel}} \in \mathbb{R}^{J \times p}$ . The above equation is the kernel-based expression of the 2.8. Specifically, the solution of operator matrix  $\mathbf{Q}^{\text{kernel}}$  is given by

$$\mathbf{Q}^{\text{kernel}} = (\mathbf{K} + \gamma \mathbb{I}) \mathbf{R}, \quad (2.40)$$

where  $\mathbb{I}$  is the identity matrix and  $\mathbf{K}$  is the  $J \times J$  Gram matrix of the training data such that

$$\mathbf{K}_{ij} = k(\mathbf{d}_i, \mathbf{d}_j). \quad (2.41)$$

The kernel function  $k(\mathbf{x}, \mathbf{z})$  and the associated RKHS  $H$  play important roles, as  $H$  describes the hypothesis space where one looks for the solution. Though new kernels are being proposed by researchers, the most widely used kernel functions are the classical gaussian kernel

$$k(\mathbf{x}, \mathbf{z}) = \frac{\|\mathbf{x} - \mathbf{z}\|^2}{2\sigma^2}, \quad (2.42)$$

where  $\sigma$  is the bandwidth of the kernel. Small values of  $\sigma$ , correspond to a small effective area of the kernel, and when  $\sigma \rightarrow 0$ , matrix  $\mathbf{K}$  approaches identity matrix  $\mathbb{I}$  and the solution approaches the interpolating model of the training data. For small values of  $\sigma$ , there is a risk of overfitting the training data, leading to poor generalization properties. For large values of  $\sigma$ , higher-order features have a smaller effect, with the risk of oversimplifying the model [43], [44]. Poggio and Girosi have discussed the properties and extensions of estimation with the gaussian kernel in the framework of regularization networks [45]. Gaussian kernel is sometimes called radial basis kernel. The value of a radial basis function kernels depends only on the distance between the argument vector from the training set, rather than their location. Such kernels are also termed stationary. A non-stationary kernel is the polynomial kernel, which assigns different values to pair of points that share the same distance, based on their values. Parameter values must not be negative to ensure that the kernel will be positive definite. The polynomial kernel is formed as

$$k(\mathbf{x}, \mathbf{z}) = (\langle \mathbf{x}, \mathbf{z} \rangle + 1)^W, \quad (2.43)$$

where  $W$  is the degree of the kernel and  $\langle \mathbf{x}, \mathbf{z} \rangle$  the dot product of the vectors. The linear kernel is given by

$$k(\mathbf{x}, \mathbf{z}) = \mathbf{x}^T \mathbf{z}. \quad (2.44)$$

Using the linear kernel is equivalent of using Wiener method ( $\gamma = 0$ ). The computational cost of kernel evaluation depends on the chosen kernel function. It is, therefore, clear that it is computationally more efficient to use standard Wiener estimation and linear pseudoinverse regression instead of linear kernels. It can be assumed that nonlinear features improve the estimation over linear methods also when the sensor system has more than three channels. Large training sets are sometimes necessary to achieve great accuracy.

### 2.4.9 K-Fourier

As it was mentioned, the expansion of a really low dimensional space to a higher yields to solutions that are generally unstable, especially for the RGB hardware case. The MS approach leads to a more efficient spectral estimation model due to the richer feature space. The increment of the spectral dimensionality in filter arrays is not always feasible due to the spectral-spatial trade off. More spectral bands means less spatial resolution, something that is more and more

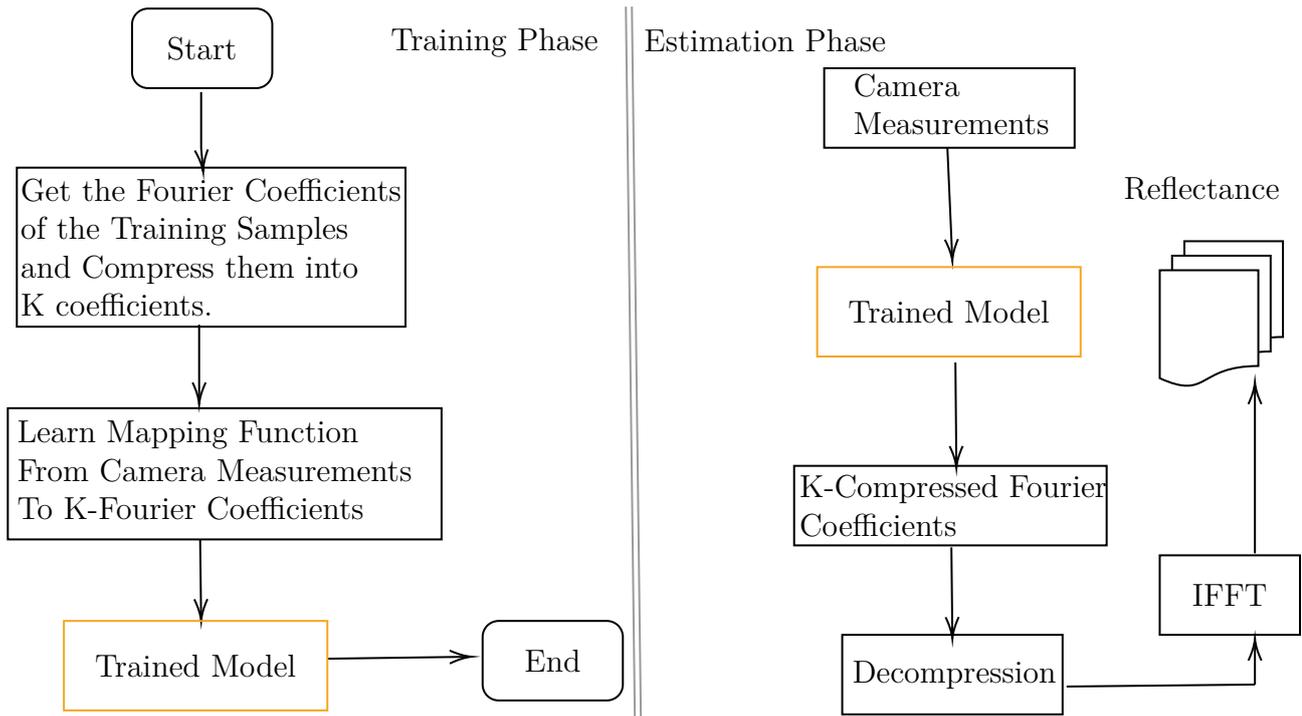


FIGURE 2.7: Flow Chart of K-Fourier.

competitive between mobile and imaging manufacturers. In research domain, there are significant one-shot MS implementations equipped with more than three spectral bands. Since, raising the spectral bands is a demanding approach, new ways to reduce the dimensionality of the estimated space must be engineered. All the above mentioned models target to learn the mapping function from the camera measurements to  $p$ -dimensional reflectance space. Another possible attempt to decline the reconstruction error is to reduce the dimensionality of the reconstructed space. Specifically, the spectral estimation problem is ill-posed since  $\mathbf{Q}$  operator has more rows than columns ( $p \ll k$ ). Models that decline the difference between  $k$  and  $p$  may result in more powerful estimations. At this point, discrete fourier transform (DFT) is leveraged. In mathematics, the DFT converts a finite sequence of equally-spaced samples of a function into a same-length sequence of equally-spaced samples of the discrete-time Fourier transform (DTFT), which is a complex-valued function of frequency. The interval at which the DFT is sampled is the reciprocal of the duration of the input sequence. An inverse DFT is a Fourier series, using the DFT samples as coefficients of complex sinusoids at the corresponding DFT frequencies. It has the same sample-values as the original input sequence. The DFT, therefore, is said to be a frequency domain representation of the original input sequence. Most (if not all) of the signals we deal with in practice are real (such as spectrum). A fourier property that is associated with the real signals is the conjugate symmetric. Specifically,

when

$$\mathbf{x}(\lambda) \in \mathbb{R} \xleftrightarrow{\text{FFT}} \mathbf{X}(-\omega) = \overline{\mathbf{X}(\omega)}, \quad (2.45)$$

where  $\omega$  is a frequency,  $\mathbf{x}(\lambda)$  is the spectrum in reflectance space, and  $\mathbf{X}(\omega)$  is the spectrum in frequency domain. In this case, one half of the spectrum is consisted of positive frequencies, and the other half of negatives. The negative coefficients are conjugate of the positive. This property applies for all the fourier coefficients except of the first one, which its magnitude is the average value of signal  $\mathbf{x}(\lambda)$ . Undeniably, that property yields to a significant improvement. If the second half part of the spectrum could be deduced for the first half, then the problem is transformed to

$$\arg \min_Q \sum_{j=1}^J \left( \| \mathbf{r}_{oj} - \text{IFFT}(\mathbf{X}_K^Q) \|^2 \right), \quad (2.46)$$

where IFFT is the operation of inverse fast fourier transform,  $K$  is the number of half coefficients, and  $\mathbf{X}_K^Q$  is a  $p$ -dimensional vector that contains the entire set of coefficients. In this study, this procedure will be called compression of the real spectrum in frequency domain. Due to the compression, the optimization problem is relied on estimating  $K + 1$  (include the first) coefficients from camera measurement  $\mathbf{d}$ . In that case  $\mathbf{X}_K^Q$  is emerged from

$$\mathbf{X}_K^Q = \psi(\mathbf{Q}^{\text{K-Fourier}} \mathbf{d}), \quad (2.47)$$

where  $\psi$  is the function that takes as input a  $K + 1$  dimensionality coefficients vector and decompresses it into  $p$  dimensional vector.  $\mathbf{Q}^{\text{K-Fourier}}$  could be derived using any non-linear model (gaussian kernel, polynomial kernel, neural networks). Furthermore, the format of the training set must be changed. Training samples must be pairs of  $(\mathbf{d}_j, \text{fft}(\mathbf{r}_{oj})_{1,2,\dots,K+1})$ . The number of  $K$  is crucial for the performance of the algorithm. Given that spectra are broad and smooth signals, the intermediate coefficients will be close to zero. Only the first positive and negative coefficients are expected to have significant impact on reconstruction of spectrum. Hence,  $K$  could be less than the half fourier coefficients. Optimal number of  $K$  could be derived by fitting a model with all possible values of  $K$  (in our case  $K \in [1, \text{floor}(p/2)]$ ) with target to minimize the reconstruction error of the training set. The reconstruction error of K-Fourier is factor of both errors that are produced from DFT operator and the error that stems from the estimation of the fourier coefficients. The former error could be declined when  $K = p/2$  and the last error relies on regression technique. The basic steps of the algorithm are depicted in fig. (2.7).

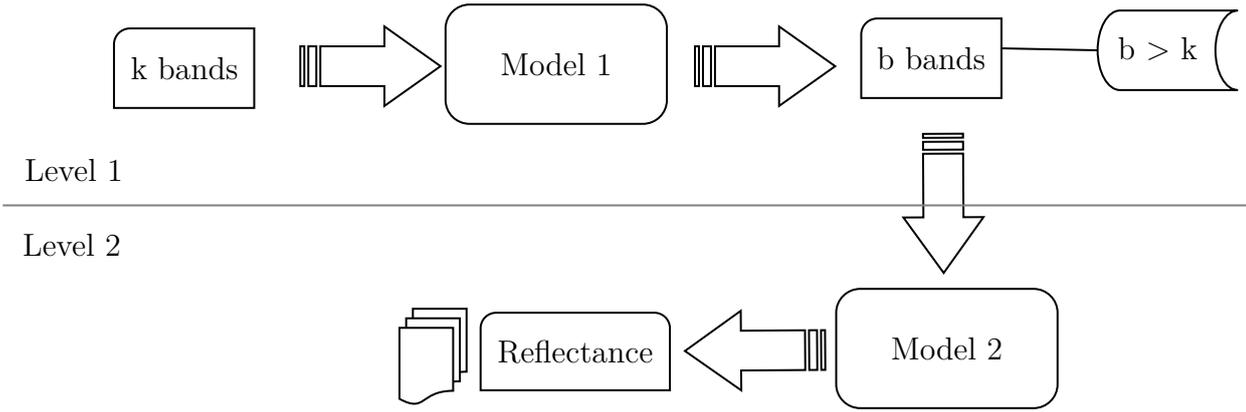


FIGURE 2.8: Flow chart of Two Level Spectral Estimation Algorithm.

### 2.4.10 Two-Level Spectral Estimation (2Level)

The expansion of the spectral dimensionality is more and more important in one-shot spectral imaging. There is a high research effort on reducing the reconstruction error having only a low number of observes. Given a specific number of features, the accuracy of the spectral estimation algorithms can be upper bounded. For instance, Heihhinen et. al [21] created a simulation of a RGB system, in order to find the upper bound of accuracy of gaussian and wiener kernel. There is no doubt, that the upper bound performance could be increased, using more features or reducing the dimensionality of the reproduced Hilbert space. The former is more demanding than the last one. As it is already mentioned, the increment of channels number reduces dramatically the spatial resolution.

Current literature handles spectral estimation as single level procedure. Specifically, the aforementioned algorithms optimize the reconstruction by finding the mapping function from  $k$  bands to  $p$  bands. The bigger  $k$  number, the better the performance. Based on that, we propose to split the estimation procedure into two pieces. The first level is consisted of a model that maps  $k$  dimensionality space to  $b$ , where  $k < b < p$ . After that, the increased  $b$ -number of features will feed the second model in the second level, which projects the increased feature space  $b$ , to the targeted  $p$ -dimensionality space. The whole algorithm is depicted in fig. (2.8). The main goal of that division is to train the second model with more features and reduce the estimated space both at first and second level. The optimal number of  $b$  can be turned out experimentally, such that,

$$\arg \min_{Q,b} \sum_{j=1}^J \left( \| \mathbf{r}_{oj} - \mathbf{Q} \mathbf{b}_j^{\text{Level 2}} \|^2 + \gamma \| \mathbf{Q} \|_F^2 \right), \quad (2.48)$$

where  $\mathbf{b}$  is a  $m \times 1$  that is produced from level 1. This scenario could be adapted from any spectral estimation algorithm.

### 2.4.11 Spectral Matching Method (Search)

In previous subsections, it was mentioned that the estimation accuracy is declined, when the number of features are low. Even though, the reflectance that is expected to be estimated is in the training set, the low amount of features cannot guarantee an accurate estimation. To deal with that, a scenario of substitute of a spectral estimation model with a spectral matching process will be evaluated. Spectral matching method is defined as a matching algorithm, that uses the available features to find the closest spectrum in a dataset. The closest spectrum is used to extract the real reflectance from the observed spectral points. Note that, this scenario cannot outperforms the other spectral estimation models, since there is no dataset with all the available spectra in nature. Furthermore, even though a complete dataset was existed, feature points could have equal distance with more than one spectrum. On the other hand, in some occasions, the matched spectrum may be more accurate than the corresponding estimation from trained spectral models.



## Chapter 3

# Spectral Estimation - Experiments

### 3.1 Datasets, Acquisition, and Pre-processing



(A) Macbeth Color Checker (source: phononet.com [46]).



(B) Munsell Mate Color Spectra Data.

FIGURE 3.1: Macbeth and Munsell Color Pallets

Depending on the available apriori information, we saw that if we have no information about the camera's sensitivities, the most convenient way to learn the mapping function for reconstructing spectral reflectance, is to learn from examples. Both standard color sets and a spectral database from Optical Research Group of Technical University of Crete (TUC) were used as training and test sets in our experiments.

### 1. Macbeth ColorChecker SG. (fig. 3.1a)

Specifically designed to meet the needs of digital photographers, the Digital Macbeth ColorChecker SG target includes the highest quality color reference standards available. Each of the 140 patches was chosen for its location in color space to expand the color gamut, so we can create profiles that capture the full capabilities of any camera and spectral scanner. The ColorChecker Digital SG target includes the colors from the standard ColorChecker target, many of which represent natural objects, such as human skin tone, foliage and blue sky. Additional skin-tone reference colors deliver greater accuracy and consistency over a wide variety of skin tones, and gray scale steps provide accurate control of camera balance to maintain a neutral aspect, regardless of light source. Macbeth ColorChecker SG can also be used to create a white balance with your digital camera to guarantee precise, uniform, neutral white under any lighting condition. From the mentioned color chart we isolated all the patches, except of the spectra with flat reflectance in visible. The remaining spectra are 56.

### 2. Munsell Book of Colors. (fig. 3.1b)

This database obtained from University of Joensuu Color Group [47] and contains in ASCII format the reflectance values of 1269 spectra in one file. Every single entry includes spectral information from 421 channels, between 380 nm and 800 nm with one nanometer step. In order to visualize the color, which each spectrum represents, we designed a artificial pallet of colors like Macbeth ColorChecker.

### 3. TUC spectral library.

This library contains 2654 spectra of various objects that was acquired in the Optoelectronics lab of Technical University of Crete.

The spectra of Macbeth ColorChecker and TUC spectral library was acquired by MUSES-9HS, which was developed from the Optical Research Group at Technical University of Crete (TUC). MuSES-9 HS is a powerful HS imager for

spectroscopy and spectrometry in a wide spectral range from 360 nm (UV) to 1000 nm (Near Infrared). Spectral resolution ranges from 7 nm (for visible or NIR) to 18 nm (for both visible and NIR). More than 110 spectral bands can be acquired in less than 1 min. From the collected images, millions of spectra can be obtained and per-point spectral and color analysis can be performed through a graphical user interface. The main difference we have between the MuSES-9 HS and the simple spectrometer is that MuSES-9 has an acquisition integral of 10nm in contrast with spectrometer, which has lower than 1nm. This affects our data in the following way: if we want to use the data from MuSES-9 we must first make a smoothing process. Also when using spectrometer we need to fit the data in order to remove all kinds of noise the spectrometer might have. One of the main reasons we wanted to use the MuSES-9 HS camera was that we would have all the devices needed in one camera.

The next step is the pre-processing of the data. We have to ensure that the produced spectral cube will be aligned. The spectral cube images registration process is the technique of wrapping one image, so that the features in the other spectral images line up perfectly. The avoidance of that step could be resulted in collection of a spectrum from different adjacent pixels. Furthermore, it should be mentioned that the iteratively acquisition of the same spectral image, may eliminate noise component.

Note that, in this study we focus on visible spectrum [400nm-730nm]. Therefore, we isolated from the aforementioned datasets the spectrum from 400nm to 730nm from all samples.

## 3.2 Training and Evaluation

As we mentioned in the introduction chapter, we want to devise a snapshot imager, which takes  $k$  bands as input and estimates all the  $p - k$  sparse wavelengths. The estimation accuracy of the sparse data relies on both input features and on the trained-model that is used. Shen et. al [32] showed that using more than three bands as input yields to lower reconstruction error of GretagMacBeth ColorChecker. To evaluate the accuracy of the reflectance reconstruction methods in detail, we simulate systems with three narrow or wide bands (RGB case), six, nine and twelve narrow bands. This experiment is conducted in order to investigate the influence of channel number  $k$  on model's accuracy. Since, we are not interested in selecting the most informative features as input bands in this chapter, we select evenly distributed sequences of bands in [400nm, 730nm]

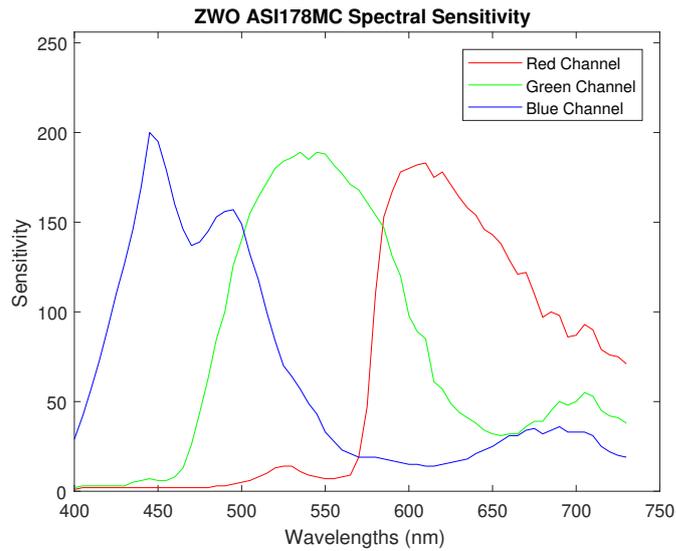


FIGURE 3.2: ZWO ASI178MC Spectral Sensitivity.

for each different channel number, except of the RGB cases, where bands selected according to ZWO ASI178MC. Shen et al. [32] and Heikkinen et. al. [21] employed also evenly distributed bands to conduct their experiments. The following table illustrates the selected bands.

<b>Wide RGB bands</b>	Accodring to ZWO ASI178MC.
<b>Narrow RGB bands</b>	[450nm, 550nm, 600nm]
<b>6 Narrow bands</b>	[400nm, 460nm, 520nm, 580nm, 640nm, 700nm]
<b>9 Narrow bands</b>	[400nm, 440nm, 480nm, 520nm, 560nm, 600nm, 640nm, 680nm, 720nm]
<b>12 Narrow bands</b>	[400nm, 430nm, 460nm, 490nm, 520nm, 550nm, 580nm, 610nm, 640nm, 670nm, 700nm, 730nm]

TABLE 3.1: The channel sequences that are used to conduct the experiments.

Since HS cube contains narrow spectral images, the input parameters of the estimation models have to be also narrow. Nevertheless, a huge effort exists on reconstructing spectral reflectance from digital cameras. The problem in RGB case, therefore, alters from estimating  $p - k$  sparse narrow bands to  $p$  sparse bands, since their is no logical connection between narrow and wide spectral bands. Hence, the model in RGB case learns the mapping from digital wide RGB images to narrow HS images. Both scenarios are evaluated in this study, in order make comparisons of different hardware architectures. In narrow bands case, the training set is created by keeping only the  $k$  values of spectral bands, which are given in table 3.1. Then, the main target is to reconstruct the reflectance from 400nm to 730nm with band step 10 of each pixel of the spectral cube. On the other hand, in RGB wide case, a simulation of digital camera is

developed using the spectral transmittance of ZWO ASI178MC digital camera. From that simulation, the corresponding RGB values of each spectrum was extracted to construct another training set.

The evaluation process is divided into three sub-experiments;

1. Training Set: Munsell patches/ Test Set: Munsell patches (10-fold cross-validation)
2. Training Set: Munsell patches/ Test Set: Macbeth patches
3. Training Set: TUC dataset patches/ Test Set: Munsell patches.

The first experiment was conducted using 10-fold cross validation technique, which used to predict new data that was not used in estimating it, in order to flag problems like overfitting or selection bias and to give an insight on how the model will generalize to an independent dataset. In 10-fold cross-validation, the original sample is randomly partitioned into 10 equal sized subsamples. Of the 10 subsamples, a single subsample is retained as the validation data for testing the model, and the remaining 10-1 subsamples are used as training data. The cross-validation process is then repeated 10 times, with each of the 10 subsamples used exactly once as the validation data. The 10 results be averaged to produce a single estimation error.

Macbeth is a global standard set through photographers. For that reason, the accurate estimation of the 56 isolated patches from Macbeth SG is considered to be the minimum requirement of any spectral estimation model. Experiment 2 will give us an insight about the satisfaction of that demand.

Experiment 3 is an additional experiment to evaluate the spectral estimation models. The best models must produce low RMSE in all experiments.

Tuning hyper-parameters of models is also a demanding process and is crucial for the final performance. The parameters that are selected are depicted in the following section.

### 3.3 Tuning Hyper-Parameters

The hyper-parameters of the models were investigated and are presented in this section. Wiener method and linear regression (pseudoinverse) provides an upper bound performance in ideal conditions (noise free), when  $\gamma = 0$ . However, since we are working in real environments,  $\gamma$  was chosen as  $\gamma = 10^{-5}$ . In some cases, Wiener model was used with  $\gamma = 0$  with adequate results [9]. On the

other hand, PCA's accuracy is based on the number of the  $m$  eigenvectors. In [48] it was shown that for the Munsell set, an six-eight dimensional PCA-basis approximation explained 99.9% of the captured variance. For industrial applications, where high accuracy is needed, it is possible that even the accuracy provided by an eight-dimensional approximation is insufficient. In our experiments we kept the highest six eigenvectors in PCA and Hybrid cases. Adaptive Wiener is a local algorithm, as it is applied on each pixel. In [32] it was investigated the influence of training sample number  $L$  on the accuracy of reflectance reconstruction and it was found that when the obtained spectral channels are six and  $\text{SNR} = 50\text{db}$  the optimal value is  $L = 50$ .

In [21], where gaussian kernel, polynomial, and Wiener kernel were compared, authors were chosen as  $(\sigma, \gamma) = (1.5, 10^{-3})$ . Also, they argued that the parameters chosen are not optimal, but since they are used for every test case, they provide information about the generalization properties of the parameters when the training data changes, noise is random, and the nonlinear transformation is used. In our experiments we adopted that hyperparameters and when polynomial kernel was employed,  $W$  was set to three. Furthermore, for K-fourier,  $K$  value was set to half fourier coefficients ( $p/2$ ) and the model that was used for learning the mapping from camera measurements to fourier coefficient space is the gaussian model. Two level algorithm (2Level) is also a new proposal, that divides reflectance reconstruction problem into two pieces. The first level is consisted of the estimating model of  $b$  bands, where  $b > k$ , and the second level estimates the real reflectance. In this study,  $b$  was set to  $p/2$  evenly distributed bands, and 2Level modification was applied on a third-degree polynomial kernel. Spectral matching method (SEARCH) is a simple searching algorithm that locates the closest available spectrum in a dataset. RMSE was leveraged as distance metric for that method.

The neural network that we designed is not deep. Through experiments, it was turned out that deep networks are not the optimal solution for the spectral estimation problem, due to the fact that spectrum has low amount of features (peaks, valleys, slopes). On contrary, a small neural network with one hidden layer that contains 15 neurons turned out to be a really efficient predictor. Additionally, we used the "momentum" optimizer, "sigmoid" activation function in hidden layer and "ELU" activation function in output layer.

At least, for the sparse representation case the dictionary size was limited to 15 atoms, under a sparsity constraint of 28 non-zero weights per atom. The sparsity

constraint were determined to be ideal via exploration of the parameter space in [24].

### 3.4 Results

Here we report the results of our experiments. Experiment 1, 2, and 3 are given in tables 3.2, 3.3, 3.4, respectively. Each table illustrates analytically the average, max, and min reconstruction error of the corresponding test set, depending on the choice of both estimation algorithm and channels number. Furthermore, figures 3.3, 3.4, 3.5 illustrate how average RMSE is affected by the number of channels.

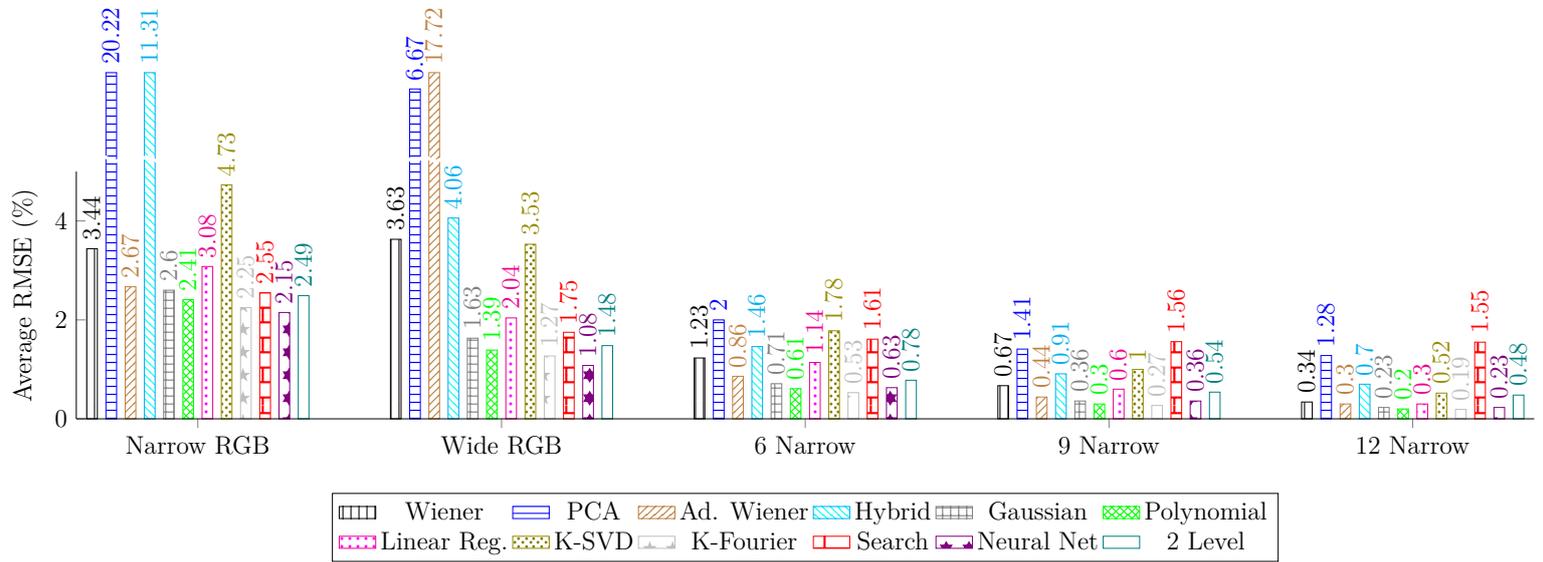


FIGURE 3.3: Experiment 1. Average RMSE plotted versus the number of spectral channels.

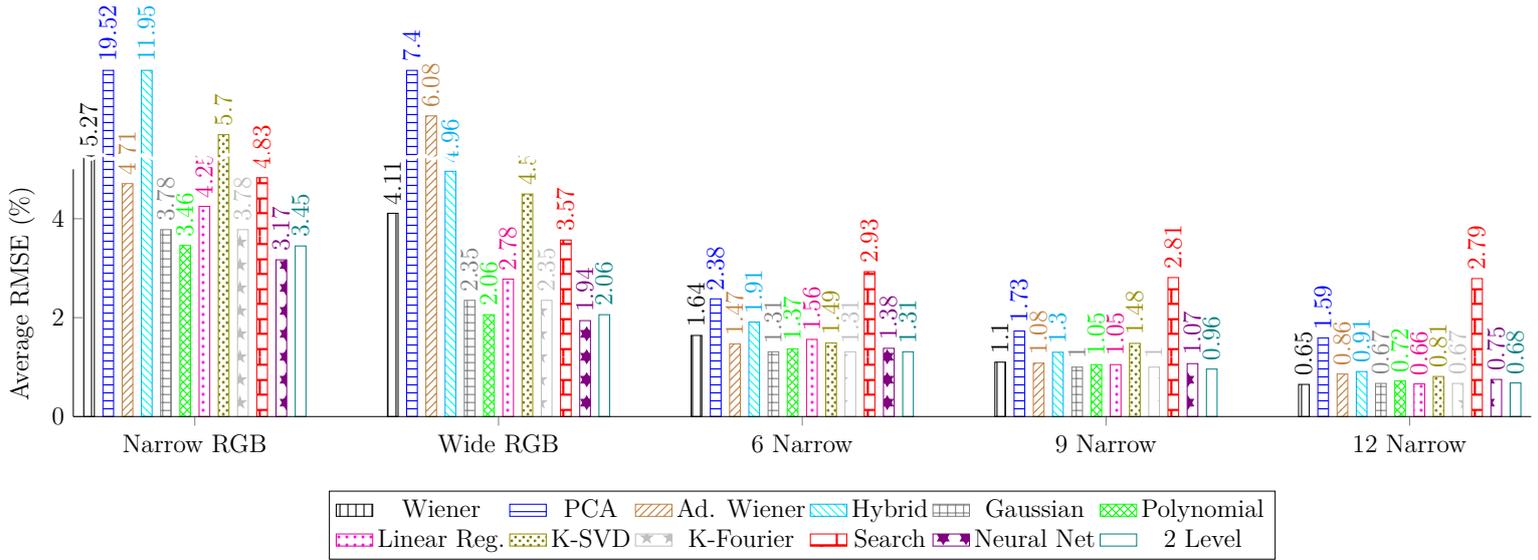


FIGURE 3.4: Experiment 2. Average RMSE plotted versus the number of spectral channels.

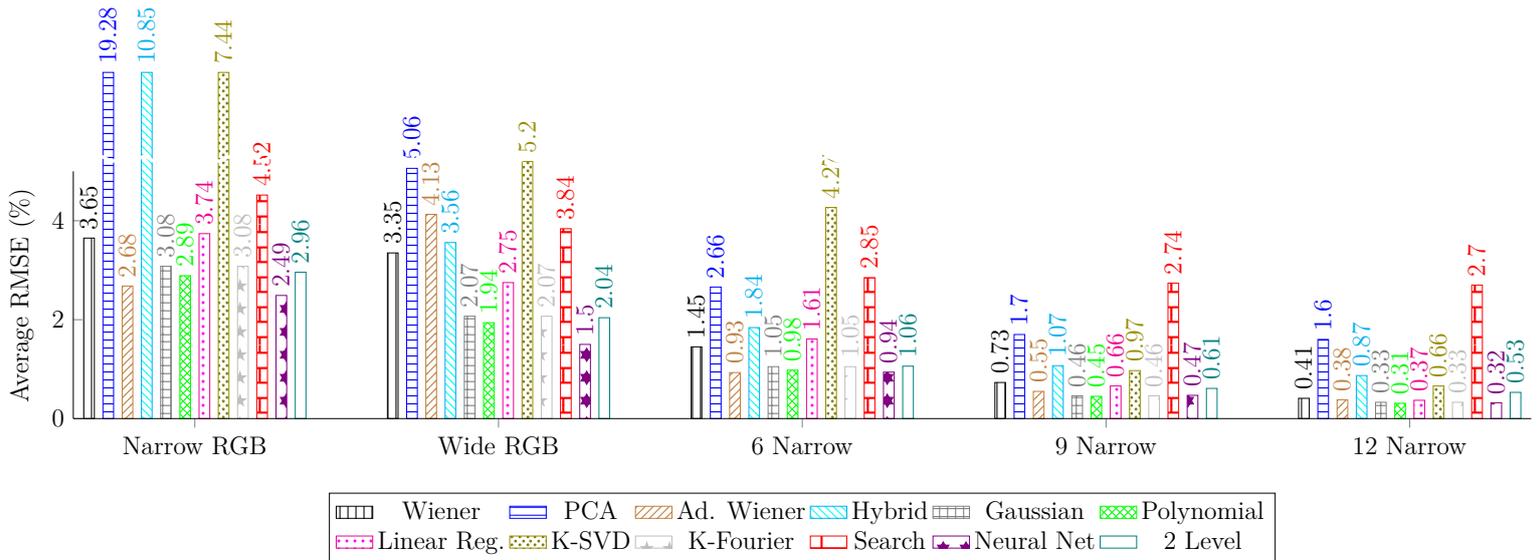


FIGURE 3.5: Experiment 3. Average RMSE plotted versus the number of spectral channels.

### 3.5 Discuss

The execution of the above experiments give us undeniably useful information about the effectiveness of the each method. It can be seen from table (3.2) that the linear regression method gives less average RMSE than Wiener is all cases, except of the cases of 9, 12 narrow bands, where they have approximately equal average RMSE. Through both methods optimize L2-norm, the

		Root Mean Square Error (%)				
		Wide RGB	Narrow RGB	6 Narrow	9 Narrow	12 Narrow
<b>Wiener</b>	<i>Mean</i>	3.63	3.44	1.23	0.67	0.34
	<i>Max</i>	11.70	22.73	4.49	2.32	1.01
	<i>Min</i>	0.54	0.24	0.14	0.10	0.08
<b>Adaptive Wiener</b>	<i>Mean</i>	3.84	2.67	0.86	0.44	0.30
	<i>Max</i>	17.72	21.74	4.16	1.83	0.96
	<i>Min</i>	0.22	0.21	0.14	0.10	0.08
<b>PCA</b>	<i>Mean</i>	6.67	20.22	2.00	1.41	1.28
	<i>Max</i>	19.90	47.95	8.38	4.18	4.08
	<i>Min</i>	0.91	3.25	0.46	0.36	0.32
<b>Hybrid</b>	<i>Mean</i>	4.06	11.31	1.46	0.91	0.70
	<i>Max</i>	14.88	29.85	5.68	2.68	2.07
	<i>Min</i>	0.64	1.76	0.30	0.21	0.18
<b>Gaussian Kernel</b>	<i>Mean</i>	1.63	2.60	0.71	0.36	0.23
	<i>Max</i>	8.74	18.58	2.78	1.40	0.67
	<i>Min</i>	0.18	0.37	0.12	0.09	0.07
<b>Polynomial Kernel</b>	<i>Mean</i>	1.39	2.41	0.61	0.30	0.20
	<i>Max</i>	7.78	17.65	2.67	1.30	0.60
	<i>Min</i>	0.20	0.33	0.11	0.09	0.07
<b>Linear Regression</b>	<i>Mean</i>	2.04	3.08	1.14	0.60	0.30
	<i>Max</i>	10.46	20.71	4.28	2.30	0.94
	<i>Min</i>	0.37	0.65	0.20	0.14	0.08
<b>K-SVD</b>	<i>Mean</i>	3.53	4.73	1.78	1.00	0.52
	<i>Max</i>	12.37	19.57	6.67	3.76	2.00
	<i>Min</i>	0.29	0.38	0.19	0.14	0.10
<b>K-Fourier</b>	<i>Mean</i>	1.27	2.25	0.53	0.27	0.19
	<i>Max</i>	7.23	16.96	2.46	1.15	0.68
	<i>Min</i>	0.20	0.28	0.11	0.09	0.06
<b>Neural Network</b>	<i>Mean</i>	1.08	2.15	0.63	0.36	0.23
	<i>Max</i>	6.10	15.59	2.50	1.49	0.72
	<i>Min</i>	0.23	0.32	0.13	0.10	0.08
<b>2 Level</b>	<i>Mean</i>	1.48	2.49	0.78	0.54	0.48
	<i>Max</i>	7.79	17.68	2.65	1.55	1.36
	<i>Min</i>	0.22	0.36	0.12	0.09	0.09
<b>Search</b>	<i>Mean</i>	1.75	2.55	1.61	1.56	1.55
	<i>Max</i>	8.85	16.78	6.29	6.07	6.00
	<i>Min</i>	0.00	0.00	0.00	0.00	0.00

TABLE 3.2: [Experiment 1] Training Set : Munsell patches /  
Test Set: Munsell patches. (10-fold cross validation)

		Root Mean Square Error (%)				
		Wide RGB	Narrow RGB	6 Narrow	9 Narrow	12 Narrow
<b>Wiener</b>	<i>Mean</i>	4.11	5.27	1.64	1.10	0.65
	<i>Max</i>	9.81	19.22	3.02	2.07	1.21
	<i>Min</i>	1.81	1.40	0.73	0.51	0.26
<b>Adaptive Wiener</b>	<i>Mean</i>	7.40	19.52	2.38	1.73	1.59
	<i>Max</i>	16.57	32.32	7.08	4.18	3.99
	<i>Min</i>	0.91	5.39	1.11	0.82	0.79
<b>PCA</b>	<i>Mean</i>	6.08	4.71	1.47	1.08	0.86
	<i>Max</i>	13.52	18.73	4.62	1.89	2.36
	<i>Min</i>	1.48	0.91	0.64	0.50	0.26
<b>Hybrid</b>	<i>Mean</i>	4.96	11.95	1.91	1.30	0.91
	<i>Max</i>	12.59	24.26	4.09	2.48	2.02
	<i>Min</i>	1.24	3.03	0.95	0.60	0.40
<b>Gaussian Kernel</b>	<i>Mean</i>	2.35	3.78	1.31	1.00	0.67
	<i>Max</i>	7.40	14.73	2.61	1.88	1.20
	<i>Min</i>	0.86	0.86	0.65	0.54	0.24
<b>Polynomial Kernel</b>	<i>Mean</i>	2.06	3.46	1.37	1.05	0.72
	<i>Max</i>	6.16	13.08	2.74	1.93	1.36
	<i>Min</i>	0.72	0.97	0.76	0.48	0.25
<b>Linear Regression</b>	<i>Mean</i>	2.78	4.25	1.56	1.05	0.66
	<i>Max</i>	8.85	17.34	3.09	1.97	1.21
	<i>Min</i>	1.24	1.01	0.48	0.46	0.29
<b>K-SVD</b>	<i>Mean</i>	4.50	5.70	1.49	1.48	0.81
	<i>Max</i>	10.57	15.45	2.82	3.51	1.65
	<i>Min</i>	1.06	1.00	0.70	0.54	0.34
<b>K-Fourier</b>	<i>Mean</i>	2.35	3.78	1.31	1.00	0.67
	<i>Max</i>	7.40	14.73	2.61	1.88	1.20
	<i>Min</i>	0.86	0.86	0.65	0.54	0.24
<b>Neural Network</b>	<i>Mean</i>	1.94	3.17	1.38	1.07	0.75
	<i>Max</i>	6.29	9.13	2.37	1.77	1.82
	<i>Min</i>	0.84	1.05	0.66	0.58	0.28
<b>2 Level</b>	<i>Mean</i>	2.06	3.45	1.32	0.96	0.68
	<i>Max</i>	6.16	13.10	2.41	1.81	1.26
	<i>Min</i>	0.72	1.00	0.69	0.43	0.32
<b>Search</b>	<i>Mean</i>	3.57	4.83	2.93	2.81	2.79
	<i>Max</i>	8.81	15.03	6.16	5.96	5.96
	<i>Min</i>	1.21	1.21	1.21	1.21	1.21

TABLE 3.3: [Experiment 2] Training Set : Munsell patches/  
Test Set: Macbeth patches.

		Root Mean Square Error (%)				
		Wide RGB	Narrow RGB	6 Narrow	9 Narrow	12 Narrow
<b>Wiener</b>	<i>Mean</i>	3.35	3.65	1.45	0.73	0.41
	<i>Max</i>	13.78	28.68	7.40	2.60	1.32
	<i>Min</i>	0.43	0.27	0.13	0.08	0.06
<b>Adaptive Wiener</b>	<i>Mean</i>	4.13	2.68	0.93	0.55	0.38
	<i>Max</i>	25.83	28.92	8.31	3.28	2.07
	<i>Min</i>	0.17	0.19	0.13	0.08	0.07
<b>PCA</b>	<i>Mean</i>	5.06	19.28	2.66	1.70	1.60
	<i>Max</i>	19.29	48.87	10.69	6.12	6.10
	<i>Min</i>	0.70	2.85	0.48	0.34	0.31
<b>Hybrid</b>	<i>Mean</i>	3.56	10.85	1.84	1.07	0.87
	<i>Max</i>	15.18	33.64	8.84	3.29	3.10
	<i>Min</i>	0.54	1.56	0.31	0.18	0.17
<b>Gaussian Kernel</b>	<i>Mean</i>	2.07	3.08	1.05	0.46	0.33
	<i>Max</i>	11.31	24.25	5.85	1.71	1.03
	<i>Min</i>	0.41	0.50	0.12	0.09	0.07
<b>Polynomial Kernel</b>	<i>Mean</i>	1.94	2.89	0.98	0.45	0.31
	<i>Max</i>	11.45	23.96	5.08	1.73	1.03
	<i>Min</i>	0.39	0.37	0.12	0.09	0.07
<b>Linear Regression</b>	<i>Mean</i>	2.75	3.74	1.61	0.66	0.37
	<i>Max</i>	11.61	25.92	6.78	2.58	1.23
	<i>Min</i>	0.92	0.83	0.26	0.12	0.08
<b>K-SVD</b>	<i>Mean</i>	5.20	7.44	4.27	0.97	0.66
	<i>Max</i>	15.69	27.07	23.75	3.73	2.64
	<i>Min</i>	0.35	0.59	0.20	0.11	0.11
<b>K-Fourier</b>	<i>Mean</i>	2.07	3.08	1.05	0.46	0.33
	<i>Max</i>	11.31	24.25	5.85	1.71	1.03
	<i>Min</i>	0.41	0.50	0.12	0.09	0.07
<b>Neural Network</b>	<i>Mean</i>	1.50	2.49	0.94	0.47	0.32
	<i>Max</i>	11.07	24.64	4.03	1.64	1.05
	<i>Min</i>	0.28	0.28	0.16	0.10	0.08
<b>2 Level</b>	<i>Mean</i>	2.04	2.96	1.06	0.61	0.53
	<i>Max</i>	11.41	23.97	5.00	1.89	1.62
	<i>Min</i>	0.47	0.53	0.12	0.08	0.07
<b>Search</b>	<i>Mean</i>	3.84	4.52	2.85	2.74	2.70
	<i>Max</i>	17.40	24.32	8.57	7.55	7.55
	<i>Min</i>	0.29	0.29	0.29	0.29	0.29

TABLE 3.4: [Experiment 3] Training Set : TUC dataset patches  
/ Test Set: Munsell patches.

addition of a constant factor  $\mathbf{b}$  in equation  $\mathbf{r} = \mathbf{Q}\mathbf{d} + \mathbf{b}$  enforces linear regression method to achieve lower reconstruction error. The effectiveness of linear models are analyzed in [19], [20], where it was found that the ensembles of reflectance spectra need to be represented with linear models with at least five to eight parameters. This finding indicates that linear models are impractical for estimating spectral reflectance from RGB sensors, due to the low number of input parameters. The accuracy of the linear models could be dramatically increased by employing more bands (6, 9, 12 narrow bands). Moreover, it is expected that if linear models would be applied locally and not globally, then the reconstruction error could be declined. Fernando et. al. [49] proposed that using smaller subsets based on similar hue values, for example, leads to an improved accuracy of linear models. Separate least-squares problems are solved for every target point, where only the local neighborhood is affecting the solution. In [50] Dicarlo and Wandell proposed the use of global Wiener solution combined with an adjustment term calculated from weighted linear regression with a compact tricube kernel. In [51] an algorithm was proposed that first computes a preliminary global linear estimate and then computes a local linear regression based on the local neighborhood of the preliminary global estimate. A weighted version of this method was proposed in [52]. In our experiments the locally applied wiener (adaptive Wiener) outperforms the global Wiener and linear regression in 3 and 6 narrow bands cases. Using global applied models instead of an adaptive method is more profitable, due to the fact that adaptive Wiener computes different  $\mathbf{Q}$  operator for each pixel, hence, the computational cost is really high and depends on the length of training set. Undeniably, local based methods are considered to be prohibitive for real time applications. The reconstruction accuracy is approximately halved when PCA is employed. Table (3.2) indicates that principal component analysis worsens the estimation as it gives the highest RMSE both in RGB and in six narrow bands case. In [21] authors performed different PCA approximations for a small set and the whole Munsell set of 1269 samples. From these results it can be seen that there is only a marginal improvement in the accuracy of the estimation based on a smaller subset, In that case, we could improve the accuracy of PCA either by applying local PCA or by leveraging an overcomplete sparse dictionary, as it proposed in [53]. The trained dictionary is proved to be more efficient rather than PCA, since mean RMSE is 3.53% versus 6.67% of PCA in wide RGB case and 1% instead of 1.41% of PCA in 6 narrow bands case. Max RMSE is also lower. K-SVD algorithm leads clearly to lower reconstruction error than PCA in all cases but does not surpass linear regression. Reconstruction efficiency of PCA

is improved when it is combined with other methods according to [37]. Hybrid algorithm combines Wiener and PCA methods, which results in slightly lower average RMSE rather than single PCA.

For the table 3.2, it is clear that the nonlinearity provided by the gaussian and polynomial kernel makes them superior to standard linear methods. Specifically, the mapping of the data to the polynomial feature space provides one of the greatest estimations. For instance, in 6 narrow bands case polynomial kernel achieves 0.61% average RMSE, where linear regression has 1.14%. Our experiments show that polynomial kernel outperforms also adaptive Wiener, KSVD, PCA, and Wiener in all cases. Gaussian kernel drops down the max RMSE in RGB cases but polynomial kernel is more powerful when it is employed with more than 6 bands as input. Greater accuracy than polynomial kernel achieved by K-Fourier. According to table 3.2, K-Fourier gives the lowest average, max and min RMSE when channels number is greater than three. This result leads as to comprehend that reducing the dimensionality of the estimated space, could also decline respectively the produced error in this problem. On the other hand, 2 Level algorithm gives higher average and max RMSE compared to the polynomial kernel in all band cases. Neural net has the highest performance in RGB cases, something that confirms that the spectral estimation problem is non-linear in RGB case. The assumption of a linear optoelectronic transfer function is a simplification in order to extract a more comprehensive relation between camera measurements and spectral reflectance. This assumption will be accurate if the measuring devices are ideal. Therefore, non-linear methods outperforms linear due to the fact the former methods are certainly necessary for non-linear measuring devices [21]. Linear models perform better when number of spectral bands are more than nine, but non-linear models produce adequate results in all cases. Neural net, polynomial kernel, and K-Fourier also overpass spectral matching algorithm (search method).

In experiment 2, the majority of RMSE metrics are higher compared to the first experiment. In wide and narrow RGB cases, neural network keeps the first position, as in experiment 1, with 1.94% and 3.17% average RMSE, respectively. When number of bands are six or nine, gaussian kernel, K-Fourier, and 2Level algorithms give the same average RMSE, while 2Level method has the lowest max RMSE. This trait is changed when number of bands are twelve, where Wiener and linear regression gives approximately 0.65% and outperform the other models.

From experiment 3, is turned out that in 9 narrow, and 12 narrow spectral

bands cases, polynomial kernel estimates better the Munsell spectra with 0.45%, 0.31% average RMSE, respectively. Furthermore, in narrow/wide RGB and in 6 narrow bands cases, neural net gives the lowest average RMSE compared to other methods.

Summing up, when the available hardware is an RGB digital imager with wide/narrow filters a trained neural net turns out to be the most profitable choice. If spectral imager contains 6 or 9 evenly distributed bands, gaussian, polynomial kernel, and K-Fourier is more preferable. On the other hand, when spectral bands are more than 12, both linear and kernel models would perform approximately in the same way. As it was mentioned before, increasing the number of spectral channel is embraced with a significant limitation. While the spectral domain is increased, the spatial resolution is declined dramatically. On the other hand, RGB camera is not adequate for spectroscopy purposes, since even using a well-trained neural net, the max RMSE exceeds 6%, when the model is tested on standard Macbeth patches. For that reason, six evenly distributed spectral bands seems to be necessary for generating accurate estimations.

### 3.5.1 Respecting physical constraints

A spectral reflectance of an object or material is defined as the ratio of its luminance to the luminance of a perfect white diffuser material at the same position. The measure of this physical property satisfies then physical constraints. For instance, there is no physical meaning of a negative reflectance. But when performing spectral reconstruction, negative reflectances can appear in the solutions. This is due to the mathematical formulation of the problem. In general, the reconstruction methods seek to minimise a measure of accuracy which sometimes leads to negative values. In our knowledge, the only effort to deal with this constraint is in [54] where the authors apply the Non Negative Least Squares (NNLS) algorithm to build reconstruction operators estimating curves without negative values. As this algorithm is based on the introduction of zeros (instead of negative values) on the constructed operator it is predictable that its accuracy decreases. In any case, the algorithm deals with the problem of non-negative values but not with others constraints. Three constraints are important on a spectral reflectance curve:

1. the curve cannot be negative,

2. it has an upper bound of 1 (we consider here only diffuse materials without specular)
3. the curve must be continuous and smooth.

The reconstruction methods studied on chapter 2 respect the third condition but not the others. In the MS community there exists another problem that is sometimes solved by similar linear methods as the ones presented on chapter 2. This is the problem of filter optimisation. In this context [55] introduced an estimation method for spectral sensitivity curves based on the method of Projections Onto Convex Sets (POCS). The method itself was not new, the first image processing application of POCS was on [56] where the aim was the restoration of images. POCS is a method that allows the use of non-linear maps in simple terms. Every known property of an original signal or image  $f$  can be restricted to lie in a well-defined closed convex set. Thus,  $m$  such properties place  $f$  in the intersection of the corresponding closed convex sets  $C_1, C_2, \dots, C_m$ . Given the projector operators  $P_i$  onto the individual  $C_i$ 's,  $i = 1, \dots, m$ , we find  $f$  by the recursive application of the operators  $P_i$ . The approach is conceptually simple and the major synthesis problem becomes the realization of operators  $P_i$ . [55] used several convex sets for filter transmittance optimization, they are: the non-negative and upper bounded vector set, the noise variance set, the noise outliers set, the passive response set, the smoothness constraint set, the unimodal set and the set of vectors close to the Human Visual Space. Using projection operators onto these sets they solved their problem. As we can see two of these sets are adapted to the constraints of reconstructed spectral curves: the non-negative (and upper bounded) vectors set and the smoothness set. We can think about the application of POCS to spectral reconstruction. However, the nonnegative and upper bounded convex set and the smoothness set are not enough. In fact, they guarantee the reconstruction to be physically feasible but they do not guarantee a good reconstruction accuracy. Consequently, a projection operator minimizing an accuracy criterion must be included. Once done, we could apply POCS for spectral reconstruction. But we must not forget that POCS is an iterative method based on projection operators. Such a method is then not bounded on time and can converge slowly, depending on the case. It is then not adapted for spectral reconstruction where existing methods are bounded and fast. Even if we cannot consider the use of POCS for spectral reconstruction the underlying idea is still interesting. A second solution for this problem proposed by Heikkinen. Heikkinen et. al. [21] introduced the following

transformation for reflectance spectra

$$\mathbf{r}' = \operatorname{arctanh}(2\mathbf{r} - 1), \quad (3.1)$$

where  $\operatorname{arctanh} : [-1, 1] \rightarrow \mathbb{R}$  is used pointwise. The inverse transformation, i.e, the restoration from nonlinear features is written as

$$\mathbf{r} = (1 + \tanh(\mathbf{r}'))/2 \quad (3.2)$$

After this transformation we will apply one of the estimation methods described above to compute an estimate in this space. This solution will afterward be mapped back to the original space of reflection spectra. In this way the values of reflectance estimates are constrained to the  $[0, 1]$  range. We also see that this transformation introduces a nonlinearity into the whole process, even though the estimation in the “arctanh” space might be linear. Third experiment (experiment 3) will be repeated in order to evaluate the performance of the aforementioned spectral estimation algorithms using the above transformation. Results are depicted in table 3.5 and in fig. 3.6.

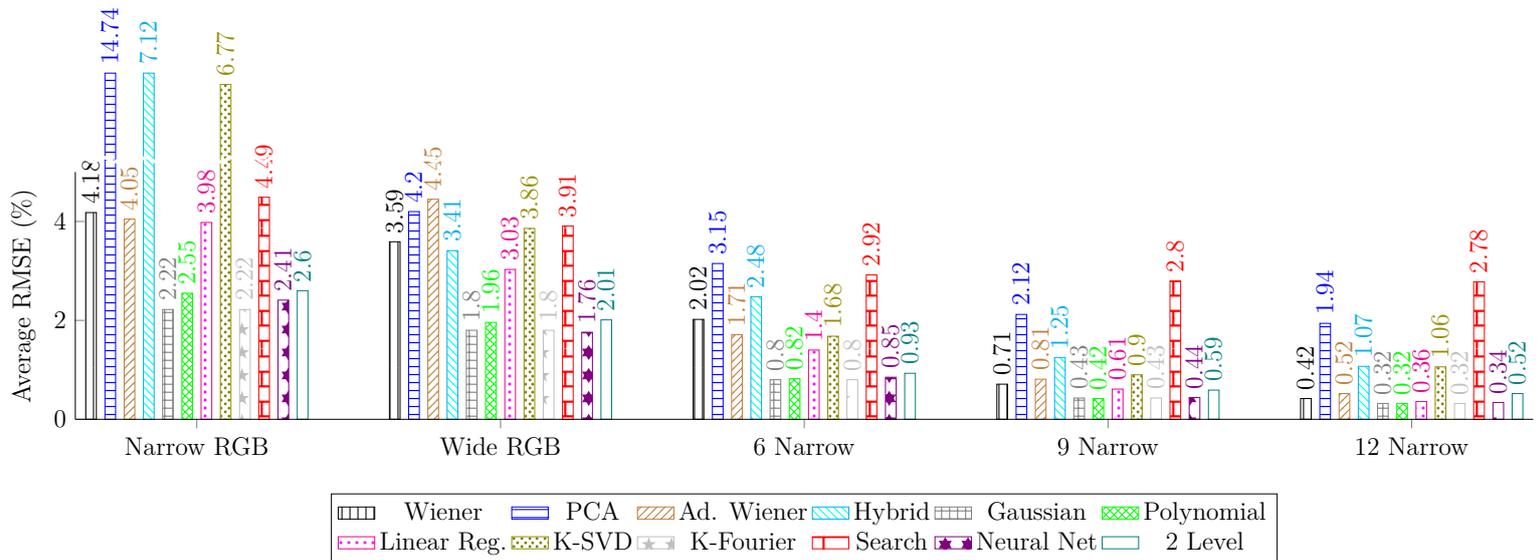


FIGURE 3.6: Experiment 3 with “arctanh” transformation. Average RMSE plotted versus the number of spectral channels.

Comparing table 3.4 with corresponding results of 3.5 it is observed that the average RMSE is raised for Wiener method in all cases. For instance, the average RMSE in 6 narrow bands case is increased from 1.45% to 2.02%. Furthermore, adaptive Wiener follows the same pattern. On the other hand, gaussian and polynomial kernel’s average RMSE is dropped in all band cases. Specifically,

		Root Mean Square Error (%)				
		Wide RGB	Narrow RGB	6 Narrow	9 Narrow	12 Narrow
<b>Wiener</b>	<i>Mean</i>	3.59	4.18	2.02	0.71	0.42
	<i>Max</i>	11.14	22.94	11.03	2.47	1.28
	<i>Min</i>	0.98	0.50	0.41	0.15	0.06
<b>Adaptive Wiener</b>	<i>Mean</i>	4.45	4.05	1.71	0.81	0.52
	<i>Max</i>	15.97	26.49	9.59	3.79	1.90
	<i>Min</i>	0.44	0.35	0.13	0.07	0.07
<b>PCA</b>	<i>Mean</i>	4.20	14.74	3.15	2.12	1.94
	<i>Max</i>	17.22	23.56	9.70	5.59	5.96
	<i>Min</i>	1.76	1.99	0.88	0.61	0.55
<b>Hybrid</b>	<i>Mean</i>	3.41	7.12	2.48	1.25	1.07
	<i>Max</i>	13.03	17.40	10.41	3.54	3.26
	<i>Min</i>	1.33	1.73	0.60	0.35	0.28
<b>Gaussian Kernel</b>	<i>Mean</i>	1.80	2.22	0.80	0.43	0.32
	<i>Max</i>	11.93	23.86	4.36	1.70	1.44
	<i>Min</i>	0.16	0.24	0.09	0.08	0.07
<b>Polynomial Kernel</b>	<i>Mean</i>	1.96	2.55	0.82	0.42	0.32
	<i>Max</i>	12.31	23.94	6.42	1.88	1.45
	<i>Min</i>	0.22	0.28	0.09	0.07	0.06
<b>Linear Regression</b>	<i>Mean</i>	3.03	3.98	1.40	0.61	0.36
	<i>Max</i>	10.84	21.87	8.88	2.74	1.12
	<i>Min</i>	0.93	0.97	0.20	0.09	0.07
<b>K-SVD</b>	<i>Mean</i>	3.86	6.77	1.68	0.90	1.06
	<i>Max</i>	10.02	21.83	5.28	4.12	4.47
	<i>Min</i>	1.35	1.05	0.39	0.17	0.15
<b>K-Fourier</b>	<i>Mean</i>	1.80	2.22	0.80	0.43	0.32
	<i>Max</i>	11.93	23.86	4.36	1.70	1.44
	<i>Min</i>	0.16	0.24	0.09	0.08	0.07
<b>Neural Network</b>	<i>Mean</i>	1.76	2.41	0.85	0.44	0.34
	<i>Max</i>	12.51	22.13	4.00	1.60	1.19
	<i>Min</i>	0.27	0.25	0.12	0.10	0.07
<b>2 Level</b>	<i>Mean</i>	2.01	2.60	0.93	0.59	0.52
	<i>Max</i>	12.28	23.95	6.31	1.81	1.56
	<i>Min</i>	0.21	0.28	0.08	0.07	0.06
<b>Search</b>	<i>Mean</i>	3.91	4.49	2.92	2.80	2.78
	<i>Max</i>	16.17	32.28	9.05	7.99	7.55
	<i>Min</i>	0.29	0.29	0.30	0.29	0.29

TABLE 3.5: [Experiment 3 with “arctanh” transformation]  
Training Set : TUC dataset patches / Test Set: Munsell patches.

in wide RGB case, average RMSE is dropped from 2.07% to 1.80%, in RGB narrow case is dropped from 3.08% to 2.22 and in 6 narrow case is reduced from 1.05% to 0.80%. In 12 narrow narrow band case the performance of that kernels are affected negatively by “arctanh” transformation, and as it depicted in table 3.5 the max RMSE is increased. Generally, except of the case of 12 narrow bands, gaussian and polynomial kernels, linear regression, K-SVD, K-Fourier, and 2 Level models produce more accurate estimations after the transformation. Note that, in this experiment, gaussian and K-Fourier give the lowest reconstruction error. Specifically, in narrow RGB and 6 narrow cases, gaussian and K-Fourier drop the previous lowest average RMSE from 2.49% (neural net) to 2.22% and from 0.94 (neural net) to 0.81%, respectively.

Ending up, the combination of gaussian kernel or K-Fourier with “arctanh” transformation lowers the reconstruction error when the input spectral bands are less than ten.



## Chapter 4

# Band Selection - Feature Extraction/Selection

### 4.1 Introduction to Band Selection

A HS cube contains hundreds of spectral bands with very fine spectral resolution. In particular, the very high data dimensionality presents a challenge to many traditional image analysis algorithms. One approach of reducing the data dimensionality is to transform the data onto a low dimensional space by using certain criteria. For instance, the objective of principal component analysis (PCA) is to maximize the variance of the transformed data (or minimize the reconstruction error). However, these methods usually change the physical meaning of the original data because the channels in the low dimensional space do not correspond to individual original bands but their linear combinations.

Another dimensionality reduction approach is band selection (BS) in order to select a subset of the original bands without losing their physical meaning. That two mechanisms create the main difference between feature extraction and feature selection. Feature extraction and feature selection are two different methods for dimensionality reduction of HS data. Feature extraction methods, such as PCA, project the original HS data into a new low dimensional data by reducing the spectral dimension. Feature extraction methods alter the physical meaning of the HS data during transformation to a new (and lower) dimensional space whereas feature selection methods preserve the original features. Feature selection essentially boils down to carefully selecting a subset of the available wavebands (i.e. waveband selection) that preserves certain traits of the full dataset.

Many band selection methods have been proposed. In terms of object information availability, band selection techniques can be divided into two categories:

supervised and unsupervised. Supervised methods are to preserve the desired object information, which is known a priori, whereas unsupervised methods do not assume any object information. For example, canonical analysis was employed for band selection in [57]; the Jeffries–Matusita distance, divergence, and Bhattacharya distance between classes were used as selection criteria in [58]–[59], respectively. Although these supervised techniques clearly aim at selecting bands that include important object information and the selected bands can provide better detection or classification than those from unsupervised techniques, the required prior knowledge may be unavailable in practice. For instance, the bands that are more informative and they contribute in order to achieve more considerable spectral estimation are unknown and there is no a priori knowledge. Therefore, it is a need to develop reliable unsupervised band selection methods that can generally offer good performance regardless of the types of surfaces they estimate (or classify). Arad and Ben-Shahar [60] first recognized that the quality of HSI recovery from a single RGB image was sensitive to the band selection. To avoid the heavy computational cost of exhaustive search, they proposed an evolutionary optimization based selection strategy.

Because the basic idea of unsupervised band selection methods is to offer the most distinctive and informative bands, a plenty of algorithms have been proposed over the last years. The major difference is that the algorithms are applied in the spatial domain, instead of being applied in the spectral domain for endmember extraction. There are quite few endmember extraction algorithms and a comparative study can be found in [61]. In general, endmember extraction algorithms can be divided into the following two categories: one extracting distinctive pixels based on similarity measurement and the other using the geometry concept, such as simplex. In this chapter, we propose the application of endmember extraction in the spectral domain towards improving spectral reconstruction from low number of observes. Practical considerations such as algorithm initialization, the number of pixels to be involved in the band selection process are investigated.

## 4.2 Data Preprocessing

To select the most distinctive but informative bands, low SNR bands need to be preremoved. This is because they can be very distinctive but not informative. Instead of manual selection, we compute the spectral correlation coefficients between original bands; those bands that have very low correlation coefficients with adjacent bands are considered as bad bands and will be preremoved[62].

The noise component in different bands is varied. If the noise component is larger, a band may look more different from others, although it may not be informatively distinct. Thus, noise whitening is needed, which requires noise estimation. It is known that noise estimation is a difficult task. In [63], it was demonstrated that the net effect of noise whitening and data whitening is similar. Therefore, we apply data whitening to the original bands (after bad band removal), which can be easily achieved by the eigendecomposition of the data covariance matrix. Then, the whitened bands actually participate in the following band selection process. Note that the selected bands are the original ones, not the whitened ones.

### 4.3 Properties of Similarity-Based Band Selection

To select the most distinctive and dissimilar bands, a similarity metric needs to be designated (widely used metrics include distance, correlation, etc). The measurement is taken on each pair of bands. In addition, due to the large number of original bands, the exhaustive search for optimal band combinations is computationally prohibitive. The sequential forward search can save significant computation time [58]. It begins with sub-optimal two bands combination, and then, this band combination is subsequently augmented to three, four, and so on until the desired number of bands are selected. The proposed band selection algorithms using the endmember extraction concept adopt this sequential forward search strategy. The basic steps of that method can be described as follows:

1. Initialize the algorithm by choosing a pair of bands  $\mathbf{b}_1$  and  $\mathbf{b}_2$ . Then, the resulting selected band subset is  $\Phi = \{\mathbf{b}_1 \mathbf{b}_2\}$ .
2. Find a third band  $\mathbf{b}_3$  that is the most dissimilar to all the bands in the current  $\Phi$  by using a certain criterion. Then, the selected band subset is updated as  $\Phi = \Phi \cup \{\mathbf{b}_3\}$ .
3. Continue on Step 2) until the number of bands in  $\Phi$  is large enough.

Another advantage is that, it is less dependent on the number of bands to be selected because those bands that are already being selected do not change with this value; increasing this value simply means to continue the algorithm execution with the bands being selected, whereas decreasing this number simply

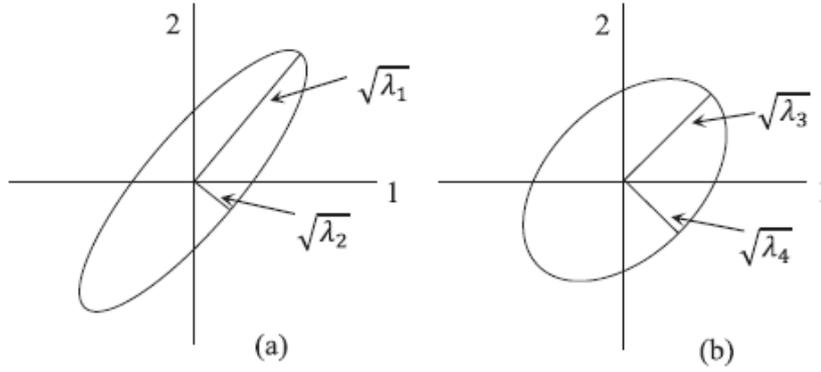


FIGURE 4.1: MVS-SFS (source: [64])

means to keep enough bands from the selected band subset (starting with the first selected band) as the final result.

## 4.4 Band Selection Algorithms

Five band selection algorithms are analyzed and evaluated here. MEV-SFS, Linear Prediction (LP), Orthogonal Subspace Projection (OSP), Band Selection using Linear Interpolation (LIBS), and Genetic Algorithm (GA) are discussed in this section.

### 4.4.1 MEV-SFS

MEV is a BS method originally proposed for MS images. It attempts to find the bands that have a large amount of information and low correlation at the same time. Reference [65] gives an example in 2-D, which is shown in Fig. 4.1. The total variance in Fig. (4.1a) is larger than that in Fig. (4.1b), i.e.,  $\lambda_1 + \lambda_2 > \lambda_3 + \lambda_4$ . However, the correlation of the bands in Fig. (4.1a) is higher than that in Fig. (4.1b); therefore,  $\lambda_1 \times \lambda_2 < \lambda_3 \times \lambda_4$ . The ellipse in Fig. (4.1b) has a larger area than the ellipse in Fig. (4.1a). When extended to a high-dimensional space, MEV attempts to select the ellipsoid, which discourages the selection of band sets with high correlation. It can be proven that the volume of the ellipsoid equals the determinant of the covariance matrix of bands, except for a constant factor [65]. For a band set  $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_p] \in \mathbb{R}^{N \times p}$ , where  $N$  and  $p$  are the numbers of pixels and bands in the set, respectively, assume that the mean value of each band has been removed, then the ellipsoid volume of  $\mathbf{B}$  can be measured by the determinant of the covariance matrix (neglecting the constant factor) [65]. For computational convenience, the square root term is neglected

in the following. Therefore, the criterion of the MEV method is selecting the band set with the maximum determinant of the covariance matrix. The MEV method was originally designed for the Landsat-4 Thematic Mapper, which is a MS sensor and has seven bands [66]. In paper [64], the MVS-SFS is used as a searching strategy. SFS is a simple greedy search algorithm that belongs to the heuristic suboptimal search methods. MEV-SFS starts with calculating the covariance matrix of all bands and selects as initial bands, the band with the maximum variance. The next step is tentatively add the candidate band to the selected band set by turns and determine the band to be selected, i.e.,

$$\Phi = \arg \max_{\Phi_t} \det(\Phi_t \Phi) \quad (4.1)$$

where  $\Phi_t = \Phi \cup \mathbf{b}_t, t = 1, 2, \dots, p$ . MVS-SFS is terminated when desired number of bands are collected. Note that set  $\Phi$  cannot contains duplicate bands.

#### 4.4.2 Linear prediction (LP)

Qian et al. [67] proposed the LP band selection method. The concept in the LP-based band selection was originally used in the UFCLSLU for endmember pixel selection in [68], which means that a pixel with the maximum reconstruction error, using the linear combination of existing endmember pixels, is the most distinctive pixel. The difference here is that, for band selection, there is no constraint imposed on the coefficients of linear combination. Assume that all the pixels in a spectral cube for each band are gathered in  $\mathbf{B}$ , where  $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_p]$  is a  $N \times p$  matrix and there are two bands  $\mathbf{b}_1$  and  $\mathbf{b}_2$  in  $\Phi$  with  $N$  pixels each. To find a band that is the most dissimilar to  $\mathbf{b}_1$  and  $\mathbf{b}_2$ ,  $\mathbf{b}_1$  and  $\mathbf{b}_2$  are used to estimate a third band  $\mathbf{b}$ , i.e.,

$$a_0 + \mathbf{b}_1 a_1 + \mathbf{b}_2 a_2 = \mathbf{b}' \quad (4.2)$$

where  $\mathbf{b}'$  is the estimate or linear prediction of band  $\mathbf{b}$  using  $\mathbf{b}_1$  and  $\mathbf{b}_2$ , and  $a_0, a_1$ , and  $a_2$  are the parameters that can minimize the linear prediction error. Let the parameter vector be  $\mathbf{a} = (a_0 \ a_1 \ a_2)^T$ . It can be determined using a least squares solution

$$\mathbf{a} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (4.3)$$

where  $\mathbf{X}$  is an  $N \times 3$  matrix whose first column is one, second column includes all the  $N$  pixels in  $\mathbf{b}_1$ , and third column includes all the pixels in  $\mathbf{b}_2$ .  $\mathbf{y}$  is an  $N \times 1$  vector with all the pixels in  $\mathbf{b}$ . The band that yields the maximum error  $e_{min}$  (using the optimal parameters) is considered as the most dissimilar band

to  $\mathbf{b}_1$  and  $\mathbf{b}_2$  and will be selected as  $\mathbf{b}_3$  for  $\Phi$ . Obviously, the similar procedure can be easily conducted when the number of bands in  $\Phi$  is larger than two.

### 4.4.3 Orthogonal Subspace Projection (OSP)

Qian et al. [67] proposed the OSP algorithm for band selection approaches. Assume that there are two bands  $\mathbf{b}_1$  and  $\mathbf{b}_2$  in  $\Phi$ . To find a band that is the most dissimilar to  $\mathbf{b}_1$  and  $\mathbf{b}_2$ , an orthogonal subspace of  $\mathbf{b}_1$  and  $\mathbf{b}_2$  is constructed as

$$\mathbf{P} = \mathbb{I} - \mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T \quad (4.4)$$

where  $\mathbb{I}$  is an  $N \times N$  identity matrix, and  $\mathbf{Z}$  is an  $N \times 2$  matrix whose first column includes all the pixels in  $\mathbf{b}_1$  and second column includes all the pixels in  $\mathbf{b}_2$ . Then, the projection  $\mathbf{y}_o = \mathbf{P}^T\mathbf{y}$  is computed, where  $\mathbf{y}$  includes all the pixels in  $\mathbf{b}$  and  $\mathbf{y}_o$  is the component of  $\mathbf{b}$  in the orthogonal subspace of  $\mathbf{b}_1$  and  $\mathbf{b}_2$ . The band that yields the maximum orthogonal component  $\|\mathbf{y}_o\|$  is considered as the most dissimilar band to  $\mathbf{b}_1$  and  $\mathbf{b}_2$  and will be selected as  $\mathbf{b}_3$  for  $\Phi$ . The similar procedure can be easily conducted when the number of bands in  $\Phi$  is larger than two. Compared to the OSP solution, the LP-based approach is computationally more efficient because it involves matrices with relatively smaller size. The remaining problem is to find the initial bands.

#### Band Used as the Initial for Band Selection for OSP and LP

The initial band pair is critical to the performance of the proposed algorithms. Intuitively, we should use the two bands whose dissimilarity is the largest. Instead of the exhaustive search, the following algorithm can be applied to an original  $p$ -band data set.

1. Randomly select a band  $\mathbf{A}_1$ , and project all the other  $p - 1$  bands to its orthogonal subspace  $\langle \mathbf{A}_1 \rangle^\perp$ .
2. Find the band  $\mathbf{A}_2$  with the maximum projection in  $\langle \mathbf{A}_1 \rangle^\perp$ , which is considered as the most dissimilar to  $\mathbf{A}_1$ .
3. Project all the other  $p - 1$  bands to the orthogonal subspace  $\langle \mathbf{A}_2 \rangle^\perp$ , and find the band  $\mathbf{A}_3$  with the maximum projection.
4. If  $\mathbf{A}_3 = \mathbf{A}_1$ ,  $\mathbf{A}_1$  and  $\mathbf{A}_2$  are confirmed to be the pair with the most significant dissimilarity, and the algorithm is terminated; if  $\mathbf{A}_3 \neq \mathbf{A}_1$ , go to the next step.

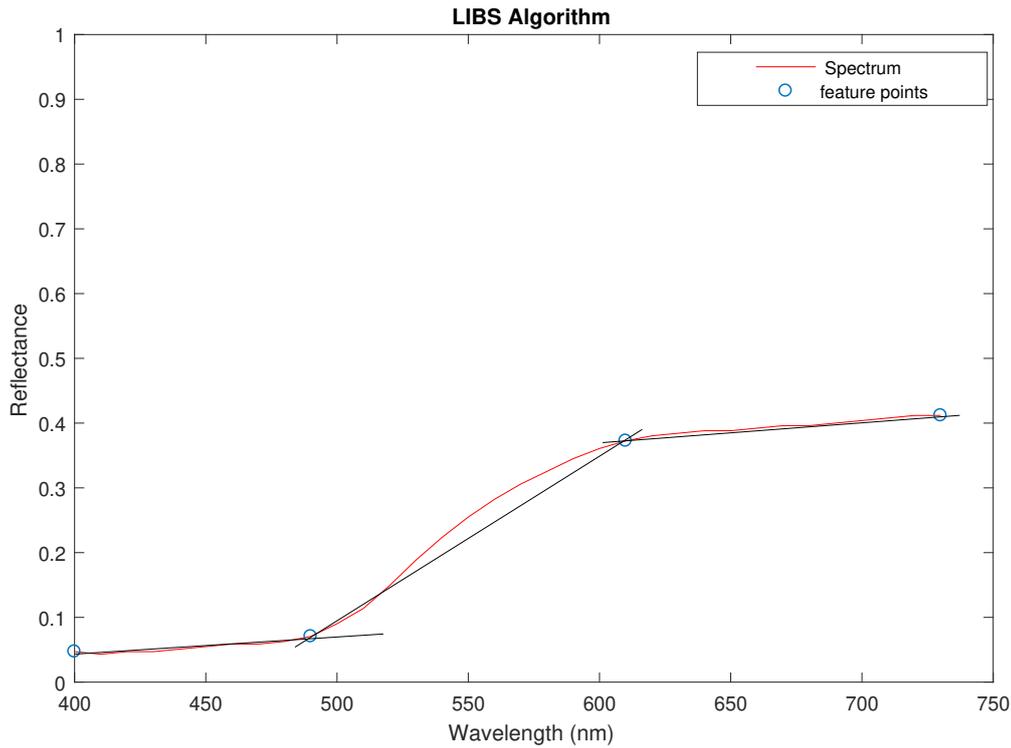


FIGURE 4.2: LIBS algorithm.

5. Continue the algorithm until  $\mathbf{A}_{i+1} = \mathbf{A}_{i-1}$ , then either  $\mathbf{A}_{i-1}$  or  $\mathbf{A}_i$  can be used as the band selection initial  $\mathbf{b}_1$  (or  $\mathbf{A}_{i-1}$  and  $\mathbf{A}_i$  are used as the initial band pair).

Qian et al. [67] found out that this algorithm can always extract the two most distinctive bands, regardless of its initial  $\mathbf{A}_1$ , although it will result in a suboptimal set of bands. In the following experiments we exhaustively searched the two initial-optimal bands, due to the fact that a suboptimal initialization could easily decline the accuracy of the spectral estimation algorithms.

#### 4.4.4 Band selection using Optimal Linear Interpolation (LIBS)

Another idea to fulfill the aim of the band selection is to consider as features the bands of the spectrum that can be united using linear interpolation technique, and that union would yield to the reconstruction of the spectrum with the minimum error. This procedure is known as optimal piecewise linear interpolation. In some cases, there is no single line that could pass accurately through a set of points in a 2D space. Considering the example in fig. (4.2), the only way to cover the red signal, is to find more than one optimal lines towards representing

accurately the line. For the aforementioned example, the first line covers the range  $[400nm, 480nm]$ , the second the range  $[480nm, 600nm]$ , and etc. Finding the optimal lines that reduce the reconstruction error of the red signal, is equivalent of finding the blue points that must be connected through lines in order to reproduce the signal. Finding the aforementioned optimal points is computationally prohibitive, especially in case of p-dimensionality spectrum. The calculation cost could be reduced by leveraging dynamic programming. That problem can be solved optimally by breaking it into sub-problems and then recursively finding the optimal solutions to the sub-problems. A possible sub-problem is to minimize  $E_{ij}$ , where  $1 < i < j < p$ , and  $E_{ij}$  is the average square euclidean error that is produced by linearly interpolating the  $i^{th}$  and the  $j^{th}$  points of all samples of the data set. Then all the sub-problems can be nested recursively inside the larger problem (reconstruct accurately all the spectra of the data set). The algorithm of the optimal piecewise linear interpolation is the following:

---

**Algorithm 1** LIBS Algorithm
 

---

```

1: procedure LIBS(number of bands to keep  $k$ )
2:   /*Compute  $Opt[k, j]$  for  $0 < k < j < p^*$ */
3:   for  $j : 1$  to  $p$  do
4:      $Opt[1, j] = E_{1,j}$ 
5:   end for
6:   for  $l : 2$  to  $k - 1$  do
7:     for  $j : 2$  to  $p$  do
8:        $t := E_{1,j}$ 
9:       for  $i : 1$  to  $j - 1$  do
10:         $t := \min(t, Opt[l - 1, i] + E_{i,j})$ 
11:      end for
12:       $Opt[l, j] = t$ 
13:    end for
14:  end for
15: end procedure

```

---

When  $Opt[l, j]$  is computed, the value of  $i$  that minimized the cost function must be recorded in a auxiliary array. Finally, the auxiliary array is used to retrieve all the blue points. The computational cost is equal to the number of  $Opt$  cells that need to be filled, in order to extract the optimal solution.

#### 4.4.5 Genetic Algorithm (GA)

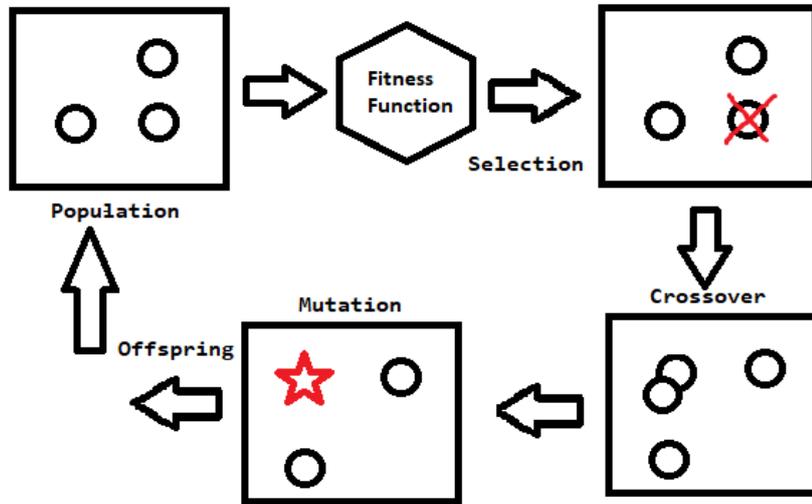


FIGURE 4.3: Genetic Algorithm Flow Chart

As we already mentioned, optical filters are usually used to increase the spectral resolution or sampling in multi/hyper spectral architectures. The accuracy of the spectral data reconstruction strongly relates to the estimation procedure and also to the characteristics of the imaging system including light source, sensor sensitivity and spectral transmission of the filters in the visible range of the spectrum. Although, a naive solution would be the exhaustive search for finding the best answer among all possible filter combinations, it cannot be done in practice. To reduce the computational time, we applied the genetic algorithm to overcome such difficulty. Genetic algorithm is a stochastic general method for solving optimization problems. It uses stimulated mechanisms of biological evaluation systems. In the genetic algorithm, every possible combination of filters as a solution is expressed as chromosome. That chromosomes can be altered or not, until a given “STOP” limit is satisfied or a value of a suitable ‘fitness’ function is reached. The simplest form of the genetic algorithm involves three types of operators: selection operator; which selects chromosomes in the population, crossover operator which takes two individuals and cuts their chromosome at some randomly chosen position for producing two “head” and “tail” segments and bounds them interchangeably, and mutation operator which is a random process where some values (genes) of a chromosomes are randomly changed. A flow chart of the genetic algorithm is given in fig. (4.3). Specifically, for the band selection problem each chromosome is a  $k$  length vector, which is a possible (sub)optimal solution. Each gene of chromosome has a value between  $400nm - 730nm$  and at the beginning of the process the vector is initialized

randomly. A population of the  $z$  chromosomes is created and evaluated from the fitness function at each step of the algorithm. The fitness function determines how fit an individual is (the ability of an individual to compete with other individuals). In our case, the fitness function evaluate a chromosome according to the reconstruction error of a dataset set. Chromosome that makes the best fit contains a set of bands, which gives the lowest reconstruction error among all chromosomes. The spectral estimation model that can be used is any from chapter 2. Ansari et. al. [69] applied genetic algorithm for finding the best set of three to eight filters combinations with specific full width half max. Ansari used Wiener filter estimation for reconstruction of the Munsell spectral data, something that will be adapted in our experiments.

## 4.5 Practical Considerations

1. **Number of Pixels Involved in the Band Selection Process:** Authors of [67] found out that using a small subset of dataset in the band selection process will not change the results in most cases. This is because each band image is spatially highly correlated.
2. **Number of Bands to be Selected:** In practice, it is difficult to know how many bands should be selected. It is more than sure that by increasing the number of the selected bands, the estimation accuracy will be also increased. The number of selected bands is depend on the optical limitations of the hardware and on the application that is studied.



## Chapter 5

# Band Selection - Experiments

### 5.1 Information about experiments and Results

In [60] Arad and Ben-Shahar proved that optimal band selection could boost the quality of HSI recovery from a single RGB image. In other words, they proved that spectral estimation is sensitive to the band selection. This assumption is intuitively confirmed, since the accuracy of any trained model could be enforced by selecting the best features to feed the model. To evaluate the above concept, a real band selection procedure took place. Our goal is to find the set  $\Phi$ , which is consisted of the  $k$  most informative and distinctive bands of Munsell dataset. The next step, is to feed spectral regression algorithms with the bands of set  $\Phi$  to make practical considerations about the improvement of reconstruction (see tables 5.5, 5.6, 5.7, 5.8, 5.9). (Sub)Optimal bands for the cases of narrow three, six, nine, and twelve spectral bands will be extracted from each band selection algorithm and will be employed to repeat experiment 3 of chapter 4 (see tables 5.1, 5.2, 5.3, 5.4). Since, we are purely interested in respecting physical constraints of spectra, we will also evaluate the performance of the “arctanh” transformation on spectral estimation models using an improved band set (see table 5.10). A visualization of average RMSE produced from each BS method and estimation model is given in fig. (5.1, 5.2, 5.3, 5.4, 5.5, 5.6). Since Munsell dataset contains only 1269 spectra, there is no need for reducing the number of samples that are involved in band selection process.

Method	Selected Bands
OSP	[430nm, 540nm, 630nm]
LP	[430nm, 550nm, 670nm]
LIBS	[400nm, 530nm, 730nm]
MVS-SFS	[480nm, 560nm, 730nm]
GA	[460nm, 560nm, 680nm]

TABLE 5.1: Three narrow bands that are selected bands from the band selection algorithms.

Method	Selected Bands
OSP	[430nm, 490nm, 540nm, 590nm, 630nm, 690nm]
LP	[430nm, 500nm, 550nm, 600nm, 670nm, 730nm]
LIBS	[400nm, 480nm, 540nm, 590nm, 640nm, 730nm]
MVS-SFS	[420nm, 480nm, 560nm, 620nm, 720nm, 730nm]
GA	[430nm, 490nm, 540nm, 590nm, 640nm, 700nm]

TABLE 5.2: Six narrow bands that are selected bands from the band selection algorithms.

Method	Selected Bands
OSP	[430nm, 460nm, 490nm, 540nm, 590nm, 630nm, 690nm, 720nm, 730nm]
LP	[400nm, 430nm, 460nm, 500nm, 550nm, 600nm, 670nm, 720nm, 730nm]
LIBS	[400nm, 450nm, 500nm, 540nm, 570nm, 610nm, 650nm, 720nm, 730nm]
MVS-SFS	[420nm, 480nm, 520nm, 560nm, 590nm, 620nm, 660nm, 720nm, 730nm]
GA	[400nm, 450nm, 490nm, 520nm, 550nm, 580nm, 620nm, 660nm, 710nm]

TABLE 5.3: Nine narrow bands that are selected bands from the band selection algorithms.

Method	Selected Bands
OSP	[400nm, 430nm, 460nm, 490nm, 540nm, 560nm, 590nm, 630nm, 660nm, 690nm, 720nm, 730nm]
LP	[400nm, 430nm, 460nm, 500nm, 550nm, 570nm, 600nm, 630nm, 670nm, 710nm, 720nm, 730nm]
LIBS	[400nm, 430nm, 460nm, 500nm, 530nm, 560nm, 590nm, 620nm, 650nm, 710nm, 720nm, 730nm]
MVS-SFS	[400nm, 420nm, 450nm, 480nm, 520nm, 560nm, 590nm, 620nm, 660nm, 700nm, 720nm, 730nm]
GA	[400nm, 440nm, 480nm, 500nm, 530nm, 560nm, 580nm, 600nm, 630nm, 670nm, 700nm, 730nm]

TABLE 5.4: Twelve narrow bands that are selected bands from the band selection algorithms.

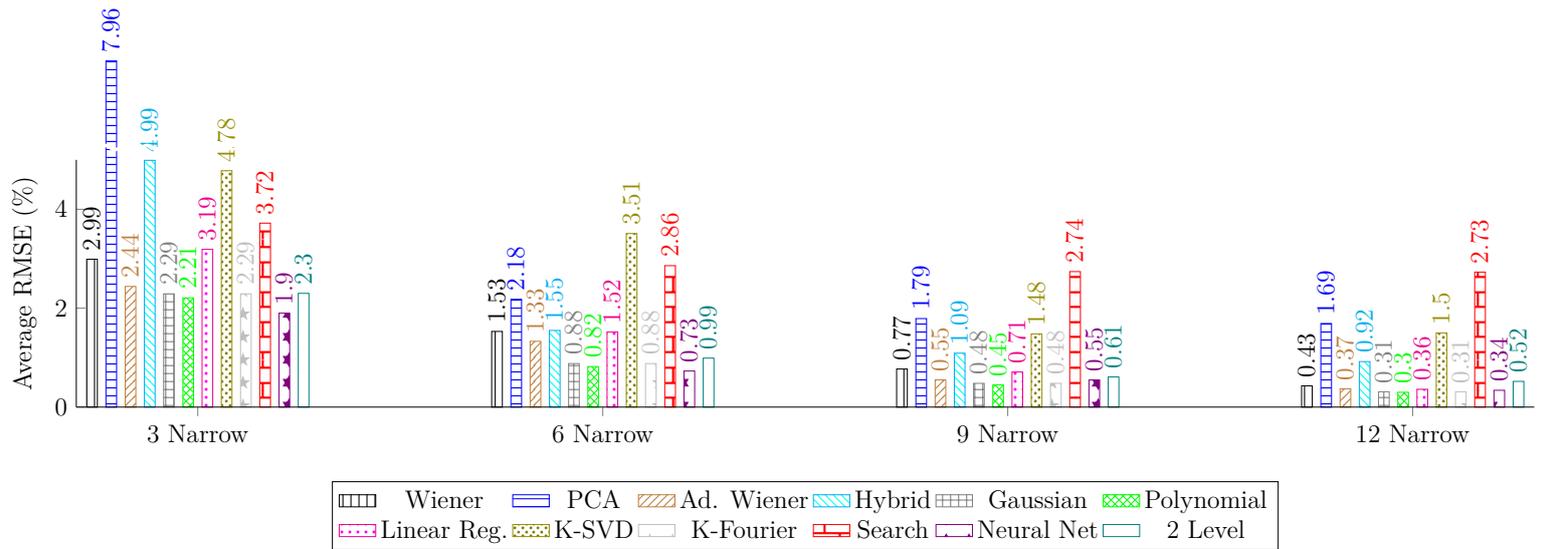


FIGURE 5.1: [Experiment 3/ BS: GA]. Average RMSE plotted versus the number of spectral channels.

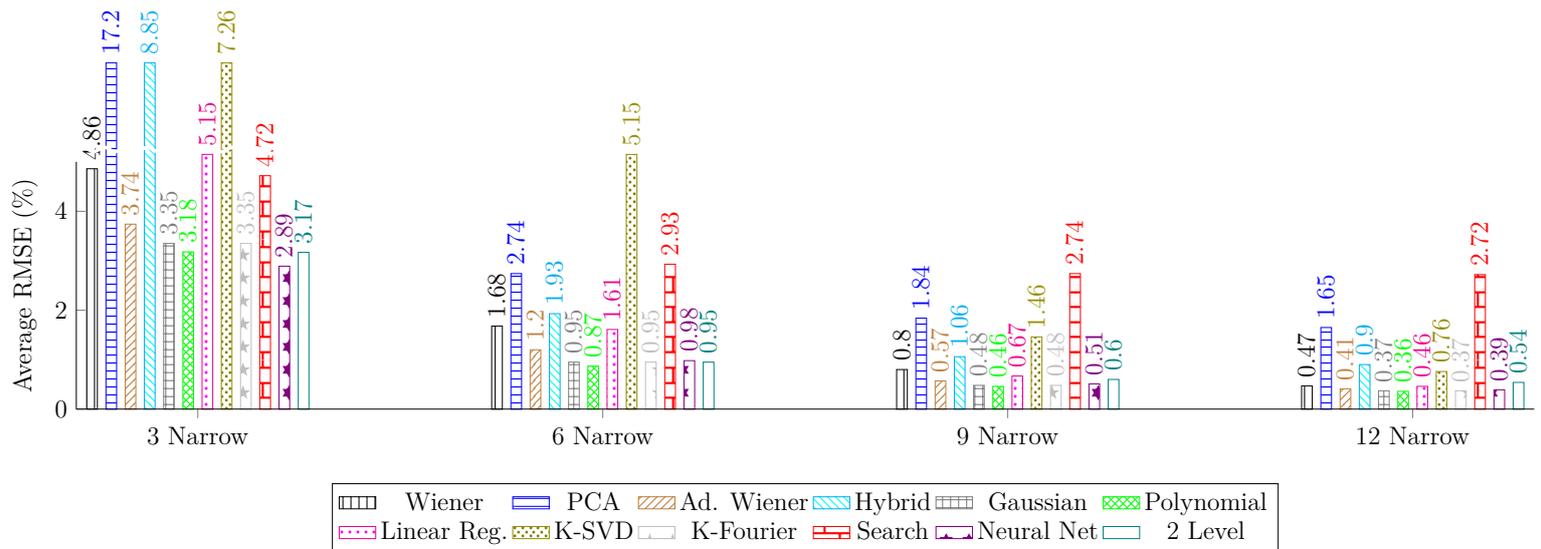


FIGURE 5.2: [Experiment 3/ BS: OLI]. Average RMSE plotted versus the number of spectral channels.

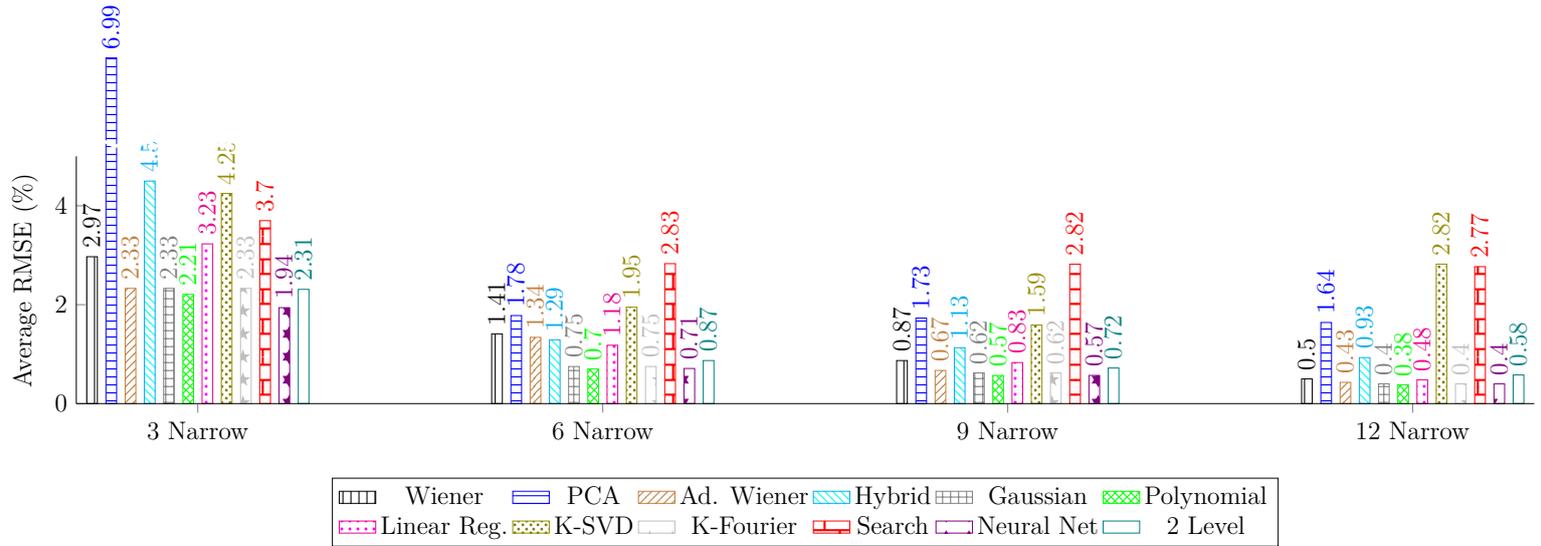


FIGURE 5.3: [Experiment 3/ BS: LP]. Average RMSE plotted versus the number of spectral channels.

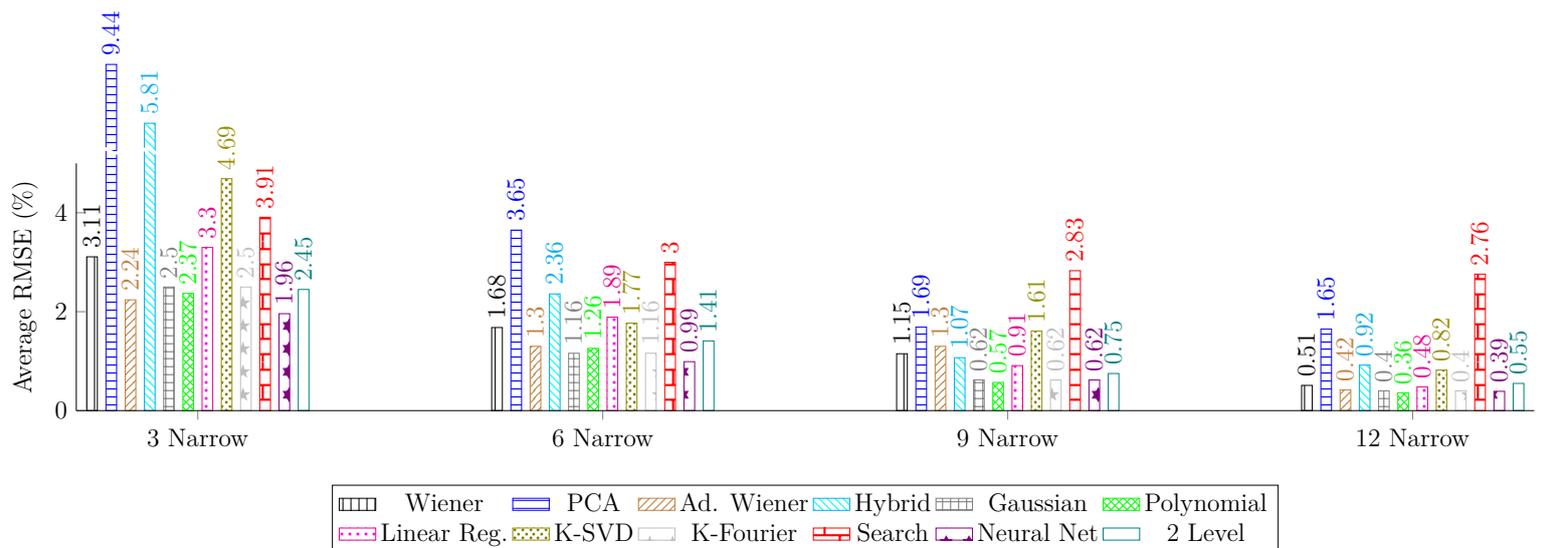


FIGURE 5.4: [Experiment 3/ BS: OSP]. Average RMSE plotted versus the number of spectral channels.

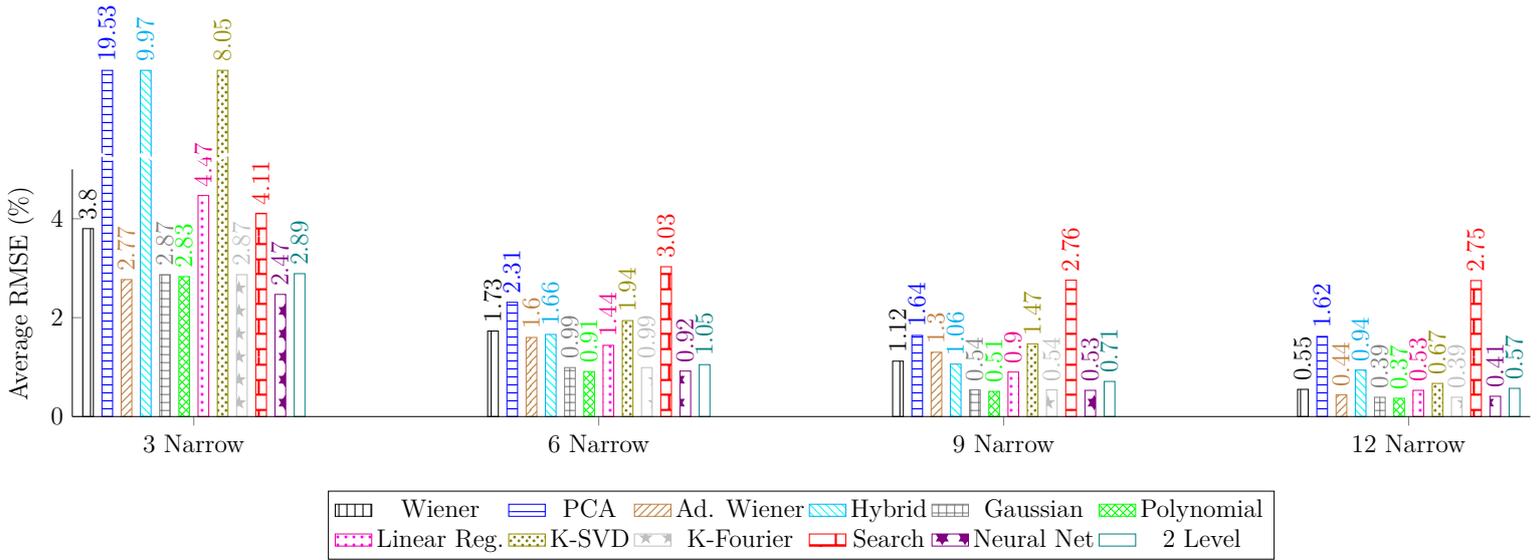


FIGURE 5.5: [Experiment 3/ BS: MVS-SFS]. Average RMSE plotted versus the number of spectral channels. Blue arrow indicates the algorithm that has the lowest average RMSE.

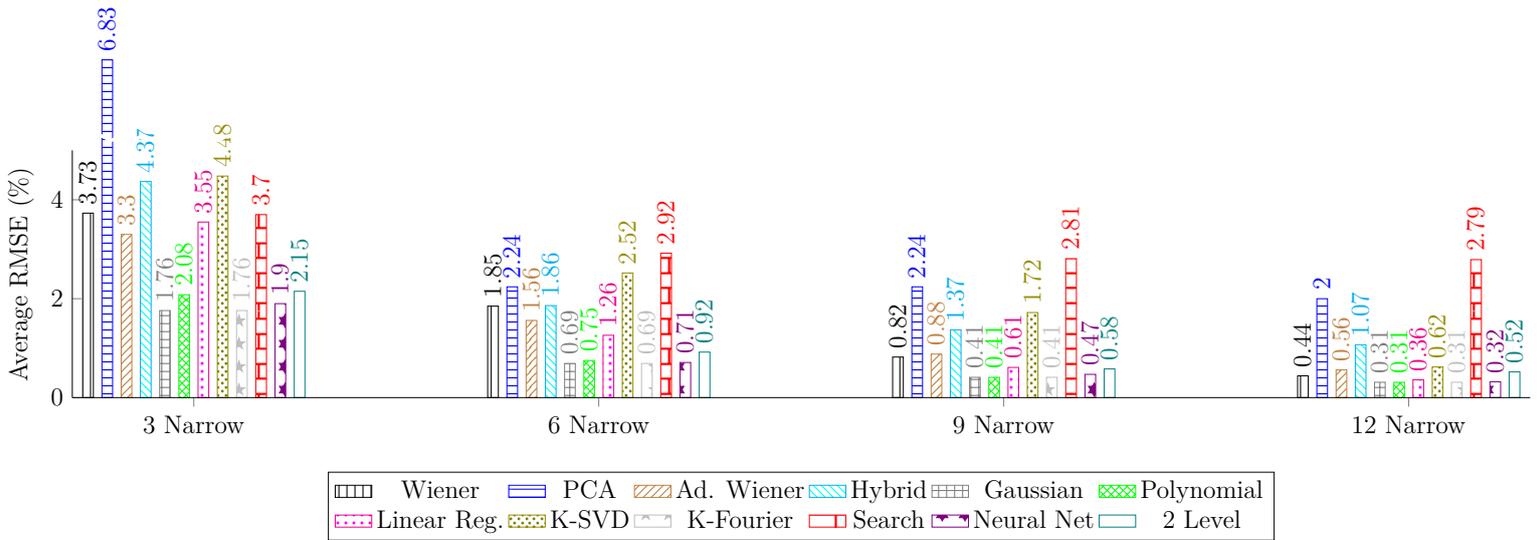


FIGURE 5.6: [Experiment 3/ BS: GA with “arctanh” transformation]. Average RMSE plotted versus the number of spectral channels.

## 5.2 Discuss

This experiment aims at providing an performance analysis of various band selection algorithms towards improving spectral dimensionality expansion. Extracting set  $\Phi$ , therefore, is crucial step for improving the training manner, hence, the most informative and distinctive bands must be identified in the

		Root Mean Square Error (%)			
		3 Narrow	6 Narrow	9 Narrow	12 Narrow
<b>Wiener</b>	<i>Mean</i>	2.99	1.53	0.77	0.43
	<i>Max</i>	13.06	8.37	4.99	1.46
	<i>Min</i>	0.30	0.09	0.09	0.06
<b>Adaptive Wiener</b>	<i>Mean</i>	2.44	1.33	0.55	0.37
	<i>Max</i>	12.79	8.14	2.86	1.50
	<i>Min</i>	0.21	0.12	0.11	0.06
<b>PCA</b>	<i>Mean</i>	7.96	2.18	1.79	1.69
	<i>Max</i>	19.51	11.38	6.45	6.07
	<i>Min</i>	1.03	0.33	0.35	0.34
<b>Hybrid</b>	<i>Mean</i>	4.99	1.55	1.09	0.92
	<i>Max</i>	15.23	9.79	5.52	3.24
	<i>Min</i>	0.64	0.19	0.16	0.18
<b>Gaussian Kernel</b>	<i>Mean</i>	2.29	0.88	0.48	0.31
	<i>Max</i>	11.59	5.76	2.53	0.92
	<i>Min</i>	0.43	0.20	0.08	0.07
<b>Polynomial Kernel</b>	<i>Mean</i>	2.21	0.82	0.45	0.30
	<i>Max</i>	11.46	4.78	1.76	1.16
	<i>Min</i>	0.46	0.17	0.08	0.06
<b>Linear Regression</b>	<i>Mean</i>	3.19	1.52	0.71	0.36
	<i>Max</i>	11.45	7.89	4.78	1.66
	<i>Min</i>	0.95	0.43	0.15	0.05
<b>K-SVD</b>	<i>Mean</i>	4.78	3.51	1.48	1.50
	<i>Max</i>	13.21	22.30	6.65	11.86
	<i>Min</i>	0.43	0.18	0.14	0.09
<b>K-Fourier</b>	<i>Mean</i>	2.29	0.88	0.48	0.31
	<i>Max</i>	11.59	5.76	2.53	0.92
	<i>Min</i>	0.43	0.20	0.08	0.07
<b>Neural Network</b>	<i>Mean</i>	1.90	0.73	0.55	0.34
	<i>Max</i>	9.17	2.66	1.82	1.19
	<i>Min</i>	0.38	0.17	0.13	0.07
<b>2 Level</b>	<i>Mean</i>	2.30	0.99	0.61	0.52
	<i>Max</i>	11.41	4.68	1.73	1.54
	<i>Min</i>	0.56	0.18	0.09	0.06
<b>Search</b>	<i>Mean</i>	3.72	2.86	2.74	2.73
	<i>Max</i>	15.59	7.77	7.77	7.77
	<i>Min</i>	0.29	0.29	0.29	0.29

TABLE 5.5: Training Set: TUC dataset/ Test Set: Munshell/  
Band Selection Method GA.

		Root Mean Square Error (%)			
		3 Narrow	6 Narrow	9 Narrow	12 Narrow
<b>Wiener</b>	<i>Mean</i>	4.86	1.68	0.80	0.47
	<i>Max</i>	17.17	8.40	3.20	1.42
	<i>Min</i>	0.24	0.10	0.07	0.07
<b>Adaptive Wiener</b>	<i>Mean</i>	3.74	1.20	0.57	0.41
	<i>Max</i>	22.92	10.50	2.25	1.85
	<i>Min</i>	0.22	0.12	0.09	0.06
<b>PCA</b>	<i>Mean</i>	17.20	2.74	1.84	1.65
	<i>Max</i>	47.94	9.08	6.98	6.36
	<i>Min</i>	1.95	0.55	0.35	0.31
<b>Hybrid</b>	<i>Mean</i>	8.85	1.93	1.06	0.90
	<i>Max</i>	29.17	8.64	4.86	3.38
	<i>Min</i>	0.65	0.31	0.17	0.17
<b>Gaussian Kernel</b>	<i>Mean</i>	3.35	0.95	0.48	0.37
	<i>Max</i>	14.56	4.97	2.48	1.74
	<i>Min</i>	0.23	0.11	0.08	0.07
<b>Polynomial Kernel</b>	<i>Mean</i>	3.18	0.87	0.46	0.36
	<i>Max</i>	16.37	4.08	2.80	1.53
	<i>Min</i>	0.30	0.13	0.07	0.07
<b>Linear Regression</b>	<i>Mean</i>	5.15	1.61	0.67	0.46
	<i>Max</i>	14.98	8.15	2.76	1.43
	<i>Min</i>	0.81	0.15	0.13	0.08
<b>K-SVD</b>	<i>Mean</i>	7.26	5.15	1.46	0.76
	<i>Max</i>	28.75	11.05	5.46	2.60
	<i>Min</i>	0.66	0.68	0.09	0.11
<b>K-Fourier</b>	<i>Mean</i>	3.35	0.95	0.48	0.37
	<i>Max</i>	14.56	4.97	2.48	1.74
	<i>Min</i>	0.23	0.11	0.08	0.07
<b>Neural Network</b>	<i>Mean</i>	2.89	0.98	0.51	0.39
	<i>Max</i>	22.43	4.17	1.74	1.78
	<i>Min</i>	0.35	0.17	0.10	0.08
<b>2 Level</b>	<i>Mean</i>	3.17	0.95	0.60	0.54
	<i>Max</i>	16.31	3.93	2.70	1.60
	<i>Min</i>	0.29	0.11	0.07	0.07
<b>Search</b>	<i>Mean</i>	4.72	2.93	2.74	2.72
	<i>Max</i>	31.29	9.59	8.32	7.55
	<i>Min</i>	0.29	0.29	0.29	0.29

TABLE 5.6: Training Set: TUC dataset/ Test Set: Munsell/  
Band Selection Method: OLI.

		Root Mean Square Error (%)			
		3 Narrow	6 Narrow	9 Narrow	12 Narrow
<b>Wiener</b>	<i>Mean</i>	2.97	1.41	0.87	0.50
	<i>Max</i>	13.16	5.92	4.96	2.05
	<i>Min</i>	0.31	0.09	0.08	0.05
<b>Adaptive Wiener</b>	<i>Mean</i>	2.33	1.34	0.67	0.43
	<i>Max</i>	12.81	6.55	3.65	2.10
	<i>Min</i>	0.21	0.12	0.09	0.07
<b>PCA</b>	<i>Mean</i>	6.99	1.78	1.73	1.64
	<i>Max</i>	22.63	6.30	6.46	6.15
	<i>Min</i>	0.76	0.32	0.31	0.30
<b>Hybrid</b>	<i>Mean</i>	4.50	1.29	1.13	0.93
	<i>Max</i>	15.23	5.90	5.32	3.15
	<i>Min</i>	0.55	0.19	0.19	0.17
<b>Gaussian Kernel</b>	<i>Mean</i>	2.33	0.75	0.62	0.40
	<i>Max</i>	12.46	2.51	2.57	2.32
	<i>Min</i>	0.46	0.14	0.08	0.07
<b>Polynomial Kernel</b>	<i>Mean</i>	2.21	0.70	0.57	0.38
	<i>Max</i>	12.52	3.14	2.59	2.64
	<i>Min</i>	0.42	0.13	0.09	0.06
<b>Linear Regression</b>	<i>Mean</i>	3.23	1.18	0.83	0.48
	<i>Max</i>	11.07	5.71	4.57	2.13
	<i>Min</i>	0.87	0.16	0.10	0.08
<b>K-SVD</b>	<i>Mean</i>	4.25	1.95	1.59	2.82
	<i>Max</i>	13.74	11.33	6.59	14.50
	<i>Min</i>	0.38	0.23	0.11	0.11
<b>K-Fourier</b>	<i>Mean</i>	2.33	0.75	0.62	0.40
	<i>Max</i>	12.46	2.51	2.57	2.32
	<i>Min</i>	0.46	0.14	0.08	0.07
<b>Neural Network</b>	<i>Mean</i>	1.94	0.71	0.57	0.40
	<i>Max</i>	11.98	2.42	1.94	1.85
	<i>Min</i>	0.26	0.12	0.10	0.06
<b>2 Level</b>	<i>Mean</i>	2.31	0.87	0.72	0.58
	<i>Max</i>	12.50	3.15	2.40	2.32
	<i>Min</i>	0.46	0.12	0.08	0.08
<b>Search</b>	<i>Mean</i>	3.70	2.83	2.82	2.77
	<i>Max</i>	15.59	8.39	8.40	7.69
	<i>Min</i>	0.29	0.29	0.29	0.29

TABLE 5.7: Training Set: TUC dataset/ Test Set: Munsell/  
Band Selection Method: LP.

		Root Mean Square Error (%)			
		3 Narrow	6 Narrow	9 Narrow	12 Narrow
<b>Wiener</b>	<i>Mean</i>	3.11	1.68	1.15	0.51
	<i>Max</i>	18.85	6.72	5.62	2.14
	<i>Min</i>	0.26	0.16	0.07	0.05
<b>Adaptive Wiener</b>	<i>Mean</i>	2.24	1.30	1.30	0.42
	<i>Max</i>	17.13	6.38	6.55	2.90
	<i>Min</i>	0.16	0.12	0.10	0.07
<b>PCA</b>	<i>Mean</i>	9.44	3.65	1.69	1.65
	<i>Max</i>	28.11	19.51	6.30	6.18
	<i>Min</i>	1.22	0.35	0.30	0.31
<b>Hybrid</b>	<i>Mean</i>	5.81	2.36	1.07	0.92
	<i>Max</i>	19.68	10.60	3.65	3.29
	<i>Min</i>	0.80	0.24	0.17	0.17
<b>Gaussian Kernel</b>	<i>Mean</i>	2.50	1.16	0.62	0.40
	<i>Max</i>	16.92	6.36	2.98	2.44
	<i>Min</i>	0.49	0.23	0.07	0.07
<b>Polynomial Kernel</b>	<i>Mean</i>	2.37	1.26	0.57	0.36
	<i>Max</i>	16.95	7.06	2.92	2.29
	<i>Min</i>	0.46	0.24	0.08	0.07
<b>Linear Regression</b>	<i>Mean</i>	3.30	1.89	0.91	0.48
	<i>Max</i>	16.17	3.90	3.18	2.13
	<i>Min</i>	0.82	0.47	0.09	0.07
<b>K-SVD</b>	<i>Mean</i>	4.69	1.77	1.61	0.82
	<i>Max</i>	18.42	7.26	4.98	4.91
	<i>Min</i>	0.40	0.23	0.20	0.09
<b>K-Fourier</b>	<i>Mean</i>	2.50	1.16	0.62	0.40
	<i>Max</i>	16.92	6.36	2.98	2.44
	<i>Min</i>	0.49	0.23	0.07	0.07
<b>Neural Network</b>	<i>Mean</i>	1.96	0.99	0.62	0.39
	<i>Max</i>	16.19	5.94	2.75	2.28
	<i>Min</i>	0.34	0.25	0.09	0.09
<b>2 Level</b>	<i>Mean</i>	2.45	1.41	0.75	0.55
	<i>Max</i>	16.92	7.04	2.57	1.92
	<i>Min</i>	0.53	0.25	0.08	0.07
<b>Search</b>	<i>Mean</i>	3.91	3.00	2.83	2.76
	<i>Max</i>	21.92	14.40	7.80	8.27
	<i>Min</i>	0.29	0.29	0.29	0.29

TABLE 5.8: Training Set: TUC dataset/ Test Set: Munsell/  
Band Selection Method: OSP.

		Root Mean Square Error (%)			
		3 Narrow	6 Narrow	9 Narrow	12 Narrow
<b>Wiener</b>	<i>Mean</i>	3.80	1.73	1.12	0.55
	<i>Max</i>	14.32	6.41	5.29	2.14
	<i>Min</i>	0.20	0.13	0.09	0.07
<b>Adaptive Wiener</b>	<i>Mean</i>	2.77	1.60	1.30	0.44
	<i>Max</i>	21.66	6.21	6.65	3.35
	<i>Min</i>	0.20	0.16	0.08	0.07
<b>PCA</b>	<i>Mean</i>	19.53	2.31	1.64	1.62
	<i>Max</i>	50.95	10.28	6.47	6.16
	<i>Min</i>	2.60	0.35	0.30	0.30
<b>Hybrid</b>	<i>Mean</i>	9.97	1.66	1.06	0.94
	<i>Max</i>	29.12	6.99	4.48	3.14
	<i>Min</i>	1.08	0.21	0.17	0.16
<b>Gaussian Kernel</b>	<i>Mean</i>	2.87	0.99	0.54	0.39
	<i>Max</i>	14.89	5.41	1.83	1.56
	<i>Min</i>	0.31	0.13	0.08	0.07
<b>Polynomial Kernel</b>	<i>Mean</i>	2.83	0.91	0.51	0.37
	<i>Max</i>	15.19	4.93	1.84	1.68
	<i>Min</i>	0.19	0.12	0.10	0.06
<b>Linear Regression</b>	<i>Mean</i>	4.47	1.44	0.90	0.53
	<i>Max</i>	12.55	5.64	2.63	2.17
	<i>Min</i>	0.91	0.14	0.13	0.08
<b>K-SVD</b>	<i>Mean</i>	8.05	1.94	1.47	0.67
	<i>Max</i>	28.86	9.32	5.61	3.26
	<i>Min</i>	0.32	0.21	0.24	0.09
<b>K-Fourier</b>	<i>Mean</i>	2.87	0.99	0.54	0.39
	<i>Max</i>	14.89	5.41	1.83	1.56
	<i>Min</i>	0.31	0.13	0.08	0.07
<b>Neural Network</b>	<i>Mean</i>	2.47	0.92	0.53	0.41
	<i>Max</i>	17.86	5.70	1.79	1.99
	<i>Min</i>	0.32	0.15	0.11	0.08
<b>2 Level</b>	<i>Mean</i>	2.89	1.05	0.71	0.57
	<i>Max</i>	15.12	4.90	1.95	1.76
	<i>Min</i>	0.18	0.12	0.08	0.07
<b>Search</b>	<i>Mean</i>	4.11	3.03	2.76	2.75
	<i>Max</i>	31.29	15.04	7.77	7.55
	<i>Min</i>	0.29	0.29	0.29	0.29

TABLE 5.9: Training Set: TUC dataset/ Test Set: Munsell/  
Band Selection Method: MEV-SFS.

		Root Mean Square Error (%)			
		3 Narrow	6 Narrow	9 Narrow	12 Narrow
<b>Wiener</b>	<i>Mean</i>	3.73	1.85	0.82	0.44
	<i>Max</i>	12.04	11.70	4.47	1.33
	<i>Min</i>	0.47	0.33	0.15	0.05
<b>Adaptive Wiener</b>	<i>Mean</i>	3.30	1.56	0.88	0.56
	<i>Max</i>	11.63	12.84	3.63	2.91
	<i>Min</i>	0.27	0.13	0.10	0.07
<b>PCA</b>	<i>Mean</i>	6.83	2.24	2.24	2.00
	<i>Max</i>	17.37	10.70	6.66	5.79
	<i>Min</i>	1.86	0.70	0.62	0.56
<b>Hybrid</b>	<i>Mean</i>	4.37	1.86	1.37	1.07
	<i>Max</i>	11.09	11.25	5.49	3.23
	<i>Min</i>	1.70	0.43	0.40	0.29
<b>Gaussian Kernel</b>	<i>Mean</i>	1.76	0.69	0.41	0.31
	<i>Max</i>	6.83	3.67	1.74	1.17
	<i>Min</i>	0.22	0.08	0.08	0.06
<b>Polynomial Kernel</b>	<i>Mean</i>	2.08	0.75	0.41	0.31
	<i>Max</i>	11.00	4.25	2.88	1.25
	<i>Min</i>	0.33	0.12	0.08	0.06
<b>Linear Regression</b>	<i>Mean</i>	3.55	1.26	0.61	0.36
	<i>Max</i>	10.78	9.96	3.18	1.60
	<i>Min</i>	0.98	0.18	0.10	0.07
<b>K-SVD</b>	<i>Mean</i>	4.48	2.52	1.72	0.62
	<i>Max</i>	11.07	7.46	6.40	1.99
	<i>Min</i>	1.33	0.74	0.11	0.14
<b>K-Fourier</b>	<i>Mean</i>	1.76	0.69	0.41	0.31
	<i>Max</i>	6.83	3.67	1.74	1.17
	<i>Min</i>	0.22	0.08	0.08	0.06
<b>Neural Network</b>	<i>Mean</i>	1.90	0.71	0.47	0.32
	<i>Max</i>	12.39	2.27	1.90	1.25
	<i>Min</i>	0.27	0.10	0.08	0.06
<b>2 Level</b>	<i>Mean</i>	2.15	0.92	0.58	0.52
	<i>Max</i>	10.89	4.16	2.81	1.51
	<i>Min</i>	0.32	0.10	0.08	0.06
<b>Search</b>	<i>Mean</i>	3.70	2.92	2.81	2.79
	<i>Max</i>	15.59	8.55	8.32	8.32
	<i>Min</i>	0.29	0.29	0.29	0.29

TABLE 5.10: Training Set: TUC dataset/ Test Set: Munsell/  
Band Selection Method: Genetic Algorithm/ “arctanh” trans-  
formation.

full set of bands. In Chapter 3, evenly distributed bands was selected to constitute set  $\Phi$ , since we were more interested in evaluating the performance of model-based spectral estimation methods rather than locating the most feature-rich wavelengths. If set  $\Phi$  is consisted of unique and informative wavelengths, then the average error is expected to be declined, otherwise, the band selection method is not able to locate the proper features. To make comparisons with experiment 3 of chapter 3, neural net and polynomial kernel, gaussian kernel and K-Fourier are employed since was turned out to be the most powerful estimation models in previous chapter.

Overall, it can be seen from tables 5.5, 5.6, 5.7, 5.8, 5.9 that GA produces the most considerable spectral estimations. On the other hand, OLI follows counterbalancing direction.

OLI is a piecewise linear interpolation algorithm that considers as critical bands, the wavelength-points that minimize the reconstruction error of the input set, when they are interpolated. This assumption does not improve the estimation error of Munsell dataset. Technically speaking, when set  $\Phi$  is consisted of evenly distributed bands, the average RMSE of neural net was 2.49% with three narrow bands, while OLI increases that number to 2.89%. Equivalently, average error of six spectral bands climbs from 0.94% to 0.98%, while for nine and twelve bands the reconstruction error seems to be the same. Slightly better results are obtained by employing MVS-SFS. For three narrow bands, max RMSE of neural net drops from 24.64% to 17.86%, but when the amount of captured bands are greater than three, MVS-SFS produces richer estimations. Changing physical meaning is the main reason beyond the defectiveness of MVS-SFS. On the other hand, preserving the original features, lead to more effective estimation models. For instance, OSP algorithm reduces the average RMSE from 2.49% to 1.96% of neural net but worsens the average error when number of spectral bands are six. Furthermore, the same pattern is followed when number of bands are nine or twelve. Average RMSE is declined when LP method is employed. Even though, in the case of three narrow bands is reduced from to 1.95% (close to OSP), in case of six narrow bands error is fallen from 0.94% to 0.71% leveraging neural net. That trend is changed in cases of nine and twelve bands, since max RMSE is raised compare to evenly distributed bands. The drawback of that algorithms, however, is that relied on crucial assumptions. Technically speaking, OSP, LP and MVS-SFS apply a heuristic assumption to identify the first band from the full set, something that affects negatively the final performance of the algorithm when that selection is not optimal. To deal with that, GA is also

employed in our research. The structural manner of GA algorithm is based on evolutionary process, where only the capable solutions are survived. GA is the most profitable band selection method, as it improves the average RMSE of neural net from 2.49% to 1.90% and from 0.94% to 0.73% of three and six narrow spectral bands, respectively. In cases of nine and twelve bands the average estimation error is approximately equal using either evenly distributed bands or GA, when polynomial kernel is leveraged.

From table 5.10 it is revealed that applying “arctanh” transformation using set  $\Phi$  as produced from GA, improves sharply the performance of gaussian and polynomial kernel, K-Fourier, and Neural Net. According to tables 5.5 and 5.10, it is indicated that in case of three bands, average RMSE of gaussian kernel drops from 2.29% to 1.76%, while for polynomial kernel’s same error is reduced from 2.21% to 2.08%. Neural net preserves the same average error, but max RMSE is increased. When the available hardware consisted of six informative narrow bands, max RMSE of gaussian is sharply moderated from 5.76% to 3.67% and from 4.78% to 4.25% when polynomial kernel is employed, respectively. Additionally, the average error of neural net is also declined from 0.55% to 0.47%. Average and max RMSE is also reduced in case of nine narrow bands for the same spectral dimensionality expansion model and seems to converge to reconstruction error without “arctanh” transformation, when twelve spectral bands hardware is available. Note that, gaussian kernel and K-Fourier produce equal results in all occurrences.

Summing up, GA band selection method offers the most informative and distinctive  $k$  bands from a full set of  $p$  bands. The combination of neural net with GA yields to a significant reduction of RMSE rather than considering evenly distributed bands. Moreover, when “arctanh” is employed to enforce the physical constraints of spectrum, the combination of gaussian or K-Fourier with GA yields to the most promising results.



## Chapter 6

# Conclusions and Future Work

Throughout this thesis we studied and compared methods of spectral estimation alongside with band selection methods.

In Chapter 1, a plenty information about spectroscopy and hyper-spectral imaging were provided. Furthermore, we shew that simple exposure imagers reveal a major disadvantage. When we have phenomena that are changing on a time scale that is sorter than the duration required for recording the spectral cube, the HS system cannot perform accurately. On that account, a need for a snapshot spectral imager that acquires in appropriate time the spectral data was revealed. Then, we utilized spectral estimation methods due to the fact that snapshot imagers cannot obtain a huge number of wavelengths without rapid loss of the spatial resolution.

In Chapter 2, all the well-known spectral estimation methods were compared according to the quality of their estimation using RMSE metric. Two new algorithms called K-Fourier and 2level proposed by us, in order to overcome shortcomings of the other methods.

In Chapter 3, detailed experiments were carried out using Munsell, Macbeth ColorChecker color charts, and TUC dataset. When the available hardware was an RGB digital imager with wide/narrow filters a trained neural net turned out to be the most profitable choice. Nevertheless, digital RGB camera is not adequate for spectroscopy purposes, since even using a well-trained neural net, the max RMSE exceeds 6%, when the model was tested on standard Macbeth patches. If spectral imager contains 6 or 9 evenly distributed bands, gaussian, polynomial kernel, and K-Fourier is more preferable. On the other hand, when

spectral bands are more than 12, both linear and kernel models would perform approximately in the same way. Furthermore, the combination of gaussian kernel or K-Fourier with “arctanh’ transformation lowers the reconstruction error when the input spectral bands are less than ten. Consequently, linear methods are sufficient when spectral channels are more than ten. On the other hand non-linear kernels, K-Fourier, and neural net perform accurately in the most of the cases.

Looking at the same problem from another perspective, the most feature-rich training yields to higher estimation accuracy. On that account, a great amount of band selection techniques are represented in chapter 4 and tested in chapter 5, which are based on similarity-measurement, dynamic programming, and evolutionary formulas were analyzed and compared. Genetic Algorithm (GA) turned out to be the most promising feature selection technique, since it dramatically improves the space reconstruction error.

A major advantage of acquiring two-dimensional spectral information from a series of complete images is that both spectral and spatial information can be depicted and displayed comprehensively with the aid of the so-called “thematic maps” which are constructed by annotating different artificial colors to pixel clusters belonging to different classes. Referring particularly to biomedical sciences, the diagnostic value of these thematic maps depends on whether a certain spectral class indicates (or not) the presence of a certain structural condition or a pathology. The diagnostic thematic maps as the end-product of an integrated SI platform would constitute the basis for the development of efficient “optical biopsy” techniques. A major obstacle, however, for the realization and generalization of this concept to a wide range of medical applications is the long time required for the capturing and processing of SI data. With today’s SI technological solutions and concepts, the capturing and post-processing of the spectral cube (composed by tens of images or millions of pixel spectra) is a computational intensive and time consuming procedure. Indicatively, typical post-capturing processing times are in the range of several minutes, which restricts the applicability of the method to time invariant and stationary targets. Moreover, it should be emphasized that the existing tunable filter-based SI systems record spectral images in a time-sequential manner and obtain the spectra from post hoc assembly of the acquired data. Thus, independently of the post-processing times, the today’s scanning SI systems are fundamentally unsuitable for analyzing moving targets or targets whose spectral content is changing with the time. Thus, the clear and unmet need for alternative SI methods, is handed

in this study and solutions are proposed. The ability of displaying a diagnostic map in nearly video rates is expected to substantially advance analytical sciences and change biomedical practices. The availability of the thematic map in real time will, for example, guide medical actions, such as biopsy sampling and/or treatment, during the diagnostic scanning/interrogation of the tissue. That will be our future work.



# Bibliography

- [1] “Quantative bioimaging laboratory,(qbil),” [Online]. Available: [http://feilab.org/Research/Research\\_HSI.htm](http://feilab.org/Research/Research_HSI.htm).
- [2] “Physicstuff,” [Online]. Available: <http://physicstuff.com/tag/3d/>.
- [3] “Astronomy,cosmos,” [Online]. Available: <http://astronomy.swin.edu.au/cosmos/S/Spectroscopy>.
- [4] “Hyperspectral imaging,” [Online]. Available: <http://www.markelowitz.com/Hyperspectral.html>.
- [5] “Wikiversity,” [Online]. Available: [https://en.wikiversity.org/wiki/Electromagnetic\\_wave](https://en.wikiversity.org/wiki/Electromagnetic_wave).
- [6] Y. Wang, N. P. Reder, S. Kang, A. Glaser, and J. Liu, “Multiplexed optical imaging of tumor-directed nanoparticles: A review of imaging systems and approaches,” vol. 1, pp. 369–388, Aug. 2017.
- [7] G. Lua and B. Feia, “Medical hyperspectral imaging : A review,” 2014.
- [8] “Dysis company,” [Online]. Available: <https://dysismedical.com/products/colposcope/>.
- [9] P. Stigell, K. Miyata, and M. Hauta-Kasari, “Wiener estimation method in estimating of spectral reflectance from rgb images,” *Pattern Recognition and Image Analysis*, vol. 17, pp. 233–242, 2007.
- [10] I. Nishidate, T. Ishizuka, K. Yoshida, S. Kawauchi, S. Sato, and M. Sato, “Visualization of hemodynamics and light scattering in exposed brain of rat using multispectral image reconstruction based on Wiener estimation method,” in *Advanced Microscopy Techniques IV; and Neurophotonics II*, E. Beaufrepaire, P. T. C. So, F. Pavone, and E. M. Hillman, Eds., International Society for Optics and Photonics, vol. 9536, SPIE, 2015, pp. 122–132. DOI: [10.1117/12.2183759](https://doi.org/10.1117/12.2183759). [Online]. Available: <https://doi.org/10.1117/12.2183759>.
- [11] I. Nishidate, T. Maeda, K. Niizeki, and Y. Aizu, “Estimation of melanin and hemoglobin using spectral reflectance images reconstructed from a digital rgb image by the wiener estimation method,” *Sensors (Basel, Switzerland)*, vol. 13, pp. 7902–15, Jun. 2013. DOI: [10.3390/s130607902](https://doi.org/10.3390/s130607902).

- [12] S. Chen and Q. Liu, "Modified wiener estimation of diffuse reflectance spectra from rgb values by the synthesis of new colors for tissue measurements," *Journal of biomedical optics*, vol. 17, p. 030 501, Mar. 2012. DOI: [10.1117/1.JBO.17.3.030501](https://doi.org/10.1117/1.JBO.17.3.030501).
- [13] Y. Murakami, M. Yamaguchi, and N. Ohyama, "Piecewise wiener estimation for reconstruction of spectral reflectance image by multipoint spectral measurements," *Applied optics*, vol. 48, pp. 2188–202, May 2009. DOI: [10.1364/AO.48.002188](https://doi.org/10.1364/AO.48.002188).
- [14] H. Li, Z. Wu, L. Zhang, and J. Parkkinen, "Sr-lla: A novel spectral reconstruction method based on locally linear approximation," in *2013 IEEE International Conference on Image Processing*, IEEE, 2013, pp. 2029–2033.
- [15] J. Liang and X. Wan, "Optimized method for spectral reflectance reconstruction from camera responses," *Optics Express*, vol. 25, p. 28 273, Nov. 2017. DOI: [10.1364/OE.25.028273](https://doi.org/10.1364/OE.25.028273).
- [16] B. Cao, N. Liao, Y. Li, and H. Cheng, "Improving reflectance reconstruction from tristimulus values by adaptively combining colorimetric and reflectance similarities," *Optical Engineering*, vol. 56, p. 053 104, May 2017. DOI: [10.1117/1.OE.56.5.053104](https://doi.org/10.1117/1.OE.56.5.053104).
- [17] A. Mansouri, T. Sliwa, J. Y. Hardeberg, and Y. Voisin, "Representation and estimation of spectral reflectances using projection on pca and wavelet bases," *Color Research & Application*, vol. 33, no. 6, pp. 485–493, DOI: [10.1002/col.20442](https://doi.org/10.1002/col.20442). eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/col.20442>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/col.20442>.
- [18] F. Agahian, S. A. Amirshahi, and S. H. Amirshahi, "Reconstruction of reflectance spectra using weighted principal component analysis," 2008.
- [19] L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," *J. Opt. Soc. Am. A*, vol. 3, no. 10, pp. 1673–1683, Oct. 1986. DOI: [10.1364/JOSAA.3.001673](https://doi.org/10.1364/JOSAA.3.001673). [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-3-10-1673>.
- [20] J. P. S. Parkkinen, J. Hallikainen, and T. Jaaskelainen, "Characteristic spectra of munsell colors," *J. Opt. Soc. Am. A*, vol. 6, no. 2, pp. 318–322, Feb. 1989. DOI: [10.1364/JOSAA.6.000318](https://doi.org/10.1364/JOSAA.6.000318). [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-6-2-318>.
- [21] V. Heikkinen, R. Lenz, T. Jetsu, J. Parkkinen, M. Hauta-Kasari, and T. Jääskeläinen, "Evaluation and unification of some methods for estimating

- reflectance spectra from rgb images,” *J. Opt. Soc. Am. A*, vol. 25, no. 10, pp. 2444–2458, Oct. 2008. DOI: [10.1364/JOSAA.25.002444](https://doi.org/10.1364/JOSAA.25.002444). [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-25-10-2444>.
- [22] V. Heikkinen, T. Jetsu, J. Parkkinen, M. Hauta-Kasari, T. Jaaskelainen, and S. D. Lee, “Regularized learning framework in the estimation of reflectance spectra from camera responses,” *J. Opt. Soc. Am. A*, vol. 24, no. 9, pp. 2673–2683, Sep. 2007. DOI: [10.1364/JOSAA.24.002673](https://doi.org/10.1364/JOSAA.24.002673). [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-24-9-2673>.
- [23] A. Robles-Kelly, “Single image spectral reconstruction for multimedia applications,” in *Proceedings of the 23rd ACM International Conference on Multimedia*, ser. MM ’15, Brisbane, Australia: ACM, 2015, pp. 251–260, ISBN: 978-1-4503-3459-4. DOI: [10.1145/2733373.2806223](https://doi.org/10.1145/2733373.2806223). [Online]. Available: <http://doi.acm.org/10.1145/2733373.2806223>.
- [24] B. Arad and O. Shahar, “Sparse recovery of hyperspectral signal from natural rgb images,” vol. 9911, Oct. 2016, pp. 19–34, ISBN: 978-3-319-46477-0. DOI: [10.1007/978-3-319-46478-7\\_2](https://doi.org/10.1007/978-3-319-46478-7_2).
- [25] Z. Shi, C. Chen, Z. Xiong, and D. Liu, “Hscnn+: Advanced cnn-based hyperspectral recovery from rgb images,” Jun. 2018, pp. 1052–10 528. DOI: [10.1109/CVPRW.2018.00139](https://doi.org/10.1109/CVPRW.2018.00139).
- [26] S. Koundinya, H. Sharma, M. Sharma, A. Upadhyay, R. Manekar, R. Mukhopadhyay, A. Karmakar, and S. Chaudhury, “2d-3d cnn based architectures for spectral reconstruction from rgb images,” Jun. 2018, pp. 957–9577. DOI: [10.1109/CVPRW.2018.00129](https://doi.org/10.1109/CVPRW.2018.00129).
- [27] A. Alvarez-Gila, J. Van De Weijer, and E. Garrote, “Adversarial networks for spatial context-aware spectral image reconstruction from rgb,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 480–490.
- [28] A. Mansouri, F. Marzani, and P. Gouton, *Neural networks in two cascade algorithms for spectral reflectance reconstruction*, Oct. 2005.
- [29] H.-L. Shen, H.-J. Wan, and Z.-C. Zhang, “Estimating reflectance from multispectral camera responses based on partial least-squares regression,” *J. Electronic Imaging*, vol. 19, p. 020 501, Apr. 2010. DOI: [10.1117/1.3385782](https://doi.org/10.1117/1.3385782).
- [30] Y. Hu, J. Hernández-Andrés, J. Nieves, E. Valero, J. Romero, M. Schnitzlein, and D. Nowack, “Evaluation and optimization of spectral estimation algorithms for printer inks,” Jun. 2011.

- [31] E. Valero, Y. Hu, J. Ndez-André, T. Eckhard, J. Nieves, J. Romero, M. Schnitzlein, and D. Nowack, “Comparative performance analysis of spectral estimation algorithms and computational optimization of a multispectral imaging system for print inspection,” vol. 39, Feb. 2014. DOI: [10.1002/col.21763](https://doi.org/10.1002/col.21763).
- [32] H.-L. Shen, P.-Q. Cai, S.-J. Shao, and J. H. Xin, “Reflectance reconstruction for multispectral imaging by adaptive wiener estimation,” *Opt. Express*, vol. 15, no. 23, pp. 15 545–15 554, Nov. 2007. DOI: [10.1364/OE.15.015545](https://doi.org/10.1364/OE.15.015545). [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-15-23-15545>.
- [33] C. S. Chane, A. Mansouri, F. S. Marzani, and F. Boochs, “Integration of 3d and multispectral data for cultural heritage applications: Survey and perspectives,” *Image and Vision Computing*, vol. 31, no. 1, pp. 91–102, 2013, ISSN: 0262-8856. DOI: <https://doi.org/10.1016/j.imavis.2012.10.006>. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0262885612001862>.
- [34] K. P. F.R.S., “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*, vol. 2, no. 11, pp. 559–572, 1901. DOI: [10.1080/14786440109462720](https://doi.org/10.1080/14786440109462720). eprint: <https://doi.org/10.1080/14786440109462720>. [Online]. Available: <https://doi.org/10.1080/14786440109462720>.
- [35] Y. Fu, Y. Zheng, L. Zhang, and H. Huang, “Spectral reflectance recovery from a single rgb image,” *IEEE Transactions on Computational Imaging*, vol. 4, pp. 382–394, 2018.
- [36] J. Liu, W. Liu, Q. Li, S. Ma, and G. Chen, “Evaluation of k-svd with different embedded sparse representation algorithms,” in *2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, Aug. 2016, pp. 426–432. DOI: [10.1109/FSKD.2016.7603211](https://doi.org/10.1109/FSKD.2016.7603211).
- [37] H.-L. Shen, J. H. Xin, and S.-J. Shao, “Improved reflectance reconstruction for multispectral imaging by combining different techniques,” *Opt. Express*, vol. 15, no. 9, pp. 5531–5536, Apr. 2007. DOI: [10.1364/OE.15.005531](https://doi.org/10.1364/OE.15.005531). [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-15-9-5531>.
- [38] W.-F. Zhang and D.-Q. Dai, “Spectral reflectance estimation from camera responses by support vector regression and a composite model,” *J. Opt. Soc. Am. A*, vol. 25, no. 9, pp. 2286–2296, Sep. 2008. DOI: [10.1364/JOSAA](https://doi.org/10.1364/JOSAA).

- 25.002286. [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-25-9-2286>.
- [39] S. Peyvandi, S. H. Amirshahi, J. Hernández-Andrés, J. L. Nieves, and J. Romero, “Spectral recovery of outdoor illumination by an extension of the bayesian inverse approach to the gaussian mixture model,” *J. Opt. Soc. Am. A*, vol. 29, no. 10, pp. 2181–2189, Oct. 2012. DOI: [10.1364/JOSAA.29.002181](https://doi.org/10.1364/JOSAA.29.002181). [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-29-10-2181>.
- [40] D. R. Connah and J. Y. Hardeberg, “Spectral recovery using polynomial models,” in *Color Imaging X: Processing, Hardcopy, and Applications*, R. Eschbach and G. G. Marcu, Eds., International Society for Optics and Photonics, vol. 5667, SPIE, 2005, pp. 65–75. DOI: [10.1117/12.586315](https://doi.org/10.1117/12.586315). [Online]. Available: <https://doi.org/10.1117/12.586315>.
- [41] J. Y. Hardeberg, “Acquisition and reproduction of color images - colometric and mulispectral approaches,” PhD thesis, • 2001.
- [42] V. Cheung, S. Westland, C. Li, J. Hardeberg, and D. Connah, “Characterization of trichromatic color cameras by using a new multispectral imaging technique,” *J. Opt. Soc. Am. A*, vol. 22, no. 7, pp. 1231–1240, Jul. 2005. DOI: [10.1364/JOSAA.22.001231](https://doi.org/10.1364/JOSAA.22.001231). [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-22-7-1231>.
- [43] *Statistical Learning Theory*. Wiley, 1998.
- [44] *Learning with Kernels*. MIT, 2002.
- [45] A. J. Smola and B. Schölkopf, “A tutorial on support vector regression,” *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, Aug. 2004, ISSN: 1573-1375. DOI: [10.1023/B:STCO.0000035301.49549.88](https://doi.org/10.1023/B:STCO.0000035301.49549.88). [Online]. Available: <https://doi.org/10.1023/B:STCO.0000035301.49549.88>.
- [46] “Macbeth color checker,” [Online]. Available: <http://photonet.com>.
- [47] “Munsell database,” [Online]. Available: <https://www.uef.fi/web/spectral/munsell-colors-matt-spectrofotometer-measured>.
- [48] J. P. S. Parkkinen, J. Hallikainen, and T. Jaaskelainen, “Characteristic spectra of munsell colors,” *J. Opt. Soc. Am. A*, vol. 6, no. 2, pp. 318–322, Feb. 1989. DOI: [10.1364/JOSAA.6.000318](https://doi.org/10.1364/JOSAA.6.000318). [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-6-2-318>.
- [49] F. Ayala, J. F. Echávarri, P. Renet, and A. I. Negueruela, “Use of three tristimulus values from surface reflectance spectra to calculate the principal components for reconstructing these spectra by using only three eigenvectors,” *J. Opt. Soc. Am. A*, vol. 23, no. 8, pp. 2020–2026, Aug.

2006. DOI: [10.1364/JOSAA.23.002020](https://doi.org/10.1364/JOSAA.23.002020). [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-23-8-2020>.
- [50] J. M. DiCarlo and B. A. Wandell, "Spectral estimation theory: Beyond linear but before bayesian," *J. Opt. Soc. Am. A*, vol. 20, no. 7, pp. 1261–1270, Jul. 2003. DOI: [10.1364/JOSAA.20.001261](https://doi.org/10.1364/JOSAA.20.001261). [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-20-7-1261>.
- [51] H.-L. Shen and J. H. Xin, "Spectral characterization of a color scanner by adaptive estimation," *J. Opt. Soc. Am. A*, vol. 21, no. 7, pp. 1125–1130, Jul. 2004. DOI: [10.1364/JOSAA.21.001125](https://doi.org/10.1364/JOSAA.21.001125). [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-21-7-1125>.
- [52] —, "Spectral characterization of a color scanner based on optimized adaptive estimation," *J. Opt. Soc. Am. A*, vol. 23, no. 7, pp. 1566–1569, Jul. 2006. DOI: [10.1364/JOSAA.23.001566](https://doi.org/10.1364/JOSAA.23.001566). [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-23-7-1566>.
- [53] R. Wu, Y. Li, X. Xijiong, and Z. Lin, "Optimized multi-spectral filter arrays for spectral reconstruction," *Sensors*, vol. 19, p. 2905, Jun. 2019. DOI: [10.3390/s19132905](https://doi.org/10.3390/s19132905).
- [54] T. L. A. Imai F. H. and E. A. Day., "Comparison of the accuracy of various transformations from multi-band image to spectral reflectance.," Rochester, USA., Tech. Rep., 2002.
- [55] *Set Theoretic Estimation in Color Scanner Characterization. Journal of Electronic Imaging*. 1996.
- [56] *Image Restoration by the Method of Convex Projections: Part 1 – Theory. IEEE Transactions on Medical Imaging*. ●, 1982.
- [57] C.-I. Chang, Q. Du, T.-L. Sun, and M. L. G. Althouse, "A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification," *IEEE Trans. Geoscience and Remote Sensing*, vol. 37, pp. 2631–2641, 1999.
- [58] A. Ifarraguerri and M. W. Prairie, "Visual method for spectral band selection," *IEEE Geoscience and Remote Sensing Letters*, vol. 1, no. 2, pp. 101–106, Apr. 2004, ISSN: 1545-598X. DOI: [10.1109/LGRS.2003.822879](https://doi.org/10.1109/LGRS.2003.822879).
- [59] S. D. Backer, P. Kempeneers, W. Debruyne, and P. Scheunders, "A band selection technique for spectral classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 3, pp. 319–323, Jul. 2005, ISSN: 1545-598X. DOI: [10.1109/LGRS.2005.848511](https://doi.org/10.1109/LGRS.2005.848511).
- [60] B. Arad and O. Ben-Shahar, "Filter selection for hyperspectral estimation," *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 3172–3180, 2017.

- [61] P. Bajcsy and P. Groves, "Methodology for hyperspectral band selection," *Photogrammetric Engineering and Remote Sensing*, vol. 70, no. 7, pp. 793–802, 2004, ISSN: 0099-1112. DOI: [doi:10.14358/PERS.70.7.793](https://doi.org/10.14358/PERS.70.7.793). [Online]. Available: <https://www.ingentaconnect.com/content/asprs/pers/2004/00000070/00000007/art00001>.
- [62] S. Cai, Q. Du, and R. J. Moorhead, "Hyperspectral imagery visualization using double layers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3028–3036, Oct. 2007, ISSN: 0196-2892. DOI: [10.1109/TGRS.2007.894922](https://doi.org/10.1109/TGRS.2007.894922).
- [63] Q. Du, H. Ren, and C.-I. Chang, "A comparative study for orthogonal subspace projection and constrained energy minimization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1525–1529, Jun. 2003, ISSN: 0196-2892. DOI: [10.1109/TGRS.2003.813704](https://doi.org/10.1109/TGRS.2003.813704).
- [64] W. Zhang, X. Li, Y. Dou, and L. Zhao, "A geometry-based band selection approach for hyperspectral image analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, pp. 1–16, Mar. 2018. DOI: [10.1109/TGRS.2018.2811046](https://doi.org/10.1109/TGRS.2018.2811046).
- [65] X. Geng, V. Sun, L. Ji, and Y. Zhao, "A fast volume-gradient-based band selection method for hyperspectral image," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, pp. 7111–7119, Nov. 2014. DOI: [10.1109/TGRS.2014.2307880](https://doi.org/10.1109/TGRS.2014.2307880).
- [66] C. Sheffield, "Selecting band combinations from multispectral data," in •.
- [67] Q. Du and H. Yang, "Similarity-based unsupervised band selection for hyperspectral image analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 4, pp. 564–568, Oct. 2008, ISSN: 1545-598X. DOI: [10.1109/LGRS.2008.2000619](https://doi.org/10.1109/LGRS.2008.2000619).
- [68] D. C. Heinz and Chein-I-Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529–545, Mar. 2001, ISSN: 0196-2892. DOI: [10.1109/36.911111](https://doi.org/10.1109/36.911111).
- [69] K. Ansari, J.-B. Thomas, and P. Gouton, "Spectral band selection using a genetic algorithm based wiener filter estimation method for reconstruction of munsell spectral data," *Electronic Imaging*, vol. 2017, pp. 190–193, Jan. 2017. DOI: [10.2352/ISSN.2470-1173.2017.18.COLOR-059](https://doi.org/10.2352/ISSN.2470-1173.2017.18.COLOR-059).