

Doctoral theses at NTNU, 2023:423

Federico Grillini

Reflectance imaging spectroscopy: Fusion of VNIR and SWIR for Cultural Heritage analysis

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Information Technology and Electrical
Engineering
Department of Computer Science



NTNU

Norwegian University of
Science and Technology

Federico Grillini

Reflectance imaging spectroscopy: Fusion of VNIR and SWIR for Cultural Heritage analysis

Thesis for the Degree of Philosophiae Doctor

Gjøvik, December 2023

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

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Information Technology and Electrical Engineering
Department of Computer Science

© Federico Grillini

ISBN 978-82-326-7546-3 (printed ver.)
ISBN 978-82-326-7545-6 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2023:423

Printed by NTNU Grafisk senter

Abstract

Reflectance Imaging Spectroscopy, often referred to as hyperspectral imaging, is an imaging technique that enables the simultaneous capture of spatial and spectral information from a scene without physical contact and in a non-invasive manner. These desirable features make it especially well-suited for applications in Cultural Heritage analysis, where the investigation of historical artifacts should avoid causing irreversible damage.

This thesis is about the revisiting of the imaging pipeline from data acquisition to the processing steps that fuse two independent hyperspectral images captured in separate spectral ranges. The need to address this topic comes from the fact that Visible Near-Infrared (VNIR) and Short-Wave Infrared (SWIR) imaging spectroscopy are being consistently deployed in the field of Cultural Heritage to conduct a series of research tasks including but not limited to analyzing the basic components of historical artifacts (pigments, dyes, binding media, mordants, fiber, etc.), long-term artifact monitoring, assessment during conservation treatments, component mapping, and revealing of hidden patterns not discernible to the human eye. However, VNIR and SWIR hyperspectral images of the same scene are often analyzed independently because of the intrinsic differences present at the image sensor level, which makes data fusion a challenging problem.

The first goal of this thesis is to develop an appropriate imaging setup for the simultaneous acquisition of VNIR-SWIR hyperspectral data with the twofold aim of obtaining high-quality data while preserving the integrity of the studied artifact. Secondly, the spatio-spectral alignment of the two hyperspectral images is addressed. Since the problem of spatial image registration has been extensively studied in the literature, we focus on the factors that may influence its performance in this context. For the spectral alignment, we propose a novel *splicing* correction that smoothly connects hyperspectral images with adjacent or overlapping spectral ranges. We then explore the application of image sharpening (e.g. pansharpening) techniques originally developed for remote sensing on proximally-sensed histor-

ical artifacts, proposing a discussion focused on the negative impact that some algorithms have on subsequent analysis processes such as the classification of spectral signals. Finally, from the hypothesis of having to capture complex artifacts such as glossy paintings, we address the integration of polarimetric imaging in the fusion pipeline, developing an acquisition paradigm for the acquisition of VNIR-SWIR spectral Stokes images that allows the study of spectro-polarimetric quantities such as the correlation between the reflectance and the linear degree of polarization.

In the initial hypothesis, the joint analysis of VNIR and SWIR Reflectance Imaging Spectroscopy data can be thought of as more powerful than the individual analyses conducted separately. However, this hypothesis could not be fully verified within this thesis, and some open questions are left for future explorations regarding its validity.

Sammendrag

Bildespektroskopi av spektral refleksjon, ofte kalt hyperspektral fotografering, er en fototeknikk som gjør det mulig å fange opp romlig (-spatial-) og spektral informasjon fra en scene uten fysiske inngrep på objektet. Denne egenskapen gjør teknikken spesielt godt egnet til bruk innen kulturarvanalyse, der det er viktig å unngå at historiske gjenstander blir påført irreversible skader.

Denne doktorgradsavhandlingen handler om en gjennomgang av arbeidsflyten til fototeknikken fra datainnsamling til prosesseringstrinnene som smelter sammen to uavhengige hyperspektrale bilder tatt i separate spektralområder. Behovet for å ta opp dette temaet kommer av det faktum at ‘Visible Near-Infrared’ (VNIR) og ‘Short-Wave Infrared’ (SWIR) bildespektroskopi stadig brukes innen kulturminnefeltet for å utføre en rekke forskningsoppgaver, inkludert, men ikke begrenset til, analyse av de grunnleggende komponentene i historiske gjenstander (pigmenter, fargestoffer, bindemidler, beisemidler, fiber osv.), langsiktig overvåking av gjenstander, vurdering under konserveringsbehandlinger, komponentkartlegging og avsløring av skjulte mønstre som ikke er synlige for det menneskelige øyet. VNIR- og SWIR-hyperspektrale bilder av samme scene analyseres imidlertid ofte uavhengig av hverandre på grunn av de iboende forskjellene på sensornivå i kameraene, noe som gjør datafusjonen mellom de to bildene til en utfordring.

Det første målet med denne avhandlingen er å utvikle et egnet bildebehandlingssoppsett for parallell innsamling av VNIR-SWIR hyperspektrale data med tilleggs mål om å oppnå data av høy kvalitet mens man samtidig bevarer integriteten gjenstanden man studerer. Med det andre målet tar vi for oss den romlig-spektrale justeringen av de to hyperspektrale bildene. Siden problemet med romlig bilderegistrering har blitt grundig studert i litteraturen, fokuserer vi på de faktorene som kan påvirke ytelsen til registreringen i denne sammenhengen. For spektraljustering foreslår vi en ny *-splicing-* korreksjon som forbinder hyperspektrale bilder med grensende eller overlappende spektralområder på en jevn måte. Deretter utforsker vi anvendelsen av bildeskarphetsteknikker (f.eks. ‘pansharpening’) som

opprinnelig ble utviklet for fjernmåling på historiske gjenstander som er tatt på nært hold, og presenterer en diskusjon med fokus på den negative innvirkningen noen algoritmer har på påfølgende analyseprosesser, for eksempel klassifisering av spektralsignaler. Til slutt, med utgangspunkt i hypotesen om å måtte fange opp komplekse gjenstander som f.eks. blanke malerier, tar vi for oss integreringen av polarimetrisk bildebehandling i fusjonsprosessen, og utvikler et datainnsamlingsparadigme for innsamling av VNIR-SWIR spektrale Stokes-bilder som gjør det mulig å studere spektro-polarimetrisk verdier som korrelasjonen mellom refleksjonen og den lineære polarisasjonsgraden.

Den opprinnelige hypotesen er at en felles analyse av VNIR og SWIR spektroskopidata kan antas å være mer effektiv enn de individuelle analysene som utføres hver for seg. Denne hypotesen kunne imidlertid ikke verifiseres fullt ut i denne avhandlingen, og det gjenstår noen åpne spørsmål for fremtidige undersøkelser for å besvare den.

Acknowledgments

The reaching of a milestone is always a perfect occasion to look back at where we were at the starting point and take a broader look at the path we undertook.

My first and foremost thanks go to Jean-Baptiste Thomas and Sony George, my academic supervisors. When hiking on this path, the guidance you offered me was special. You pushed me forward when I thought I was stuck in the mud, helped me around an obstacle by tracing my steps back, and slowed me down when I was fast-pacing toward a cliff. Maybe enough with the hiking metaphors, all in all, I think we also had our great share of fun.

The NTNU Colourlab will always have a special place in my heart. So I would like to thank the Head of the department Marius Pedersen for creating such a scientifically and humanly wonderful environment. I want to specifically thank Hilda, Jon, Giorgio, and Peter for all the science and non-science-related interactions. To my fellow PhD students and post-docs: Agnese, Mathieu, Irina, Siamak, Ali, Fereshteh, Tanzima, Greg, Olga, Anuja, Markus, Raed, and Simon. You really made the time spent in the office fly by. Thanks for the ping-pong and the breaks after the lunch breaks.

I also want to thank the people with whom I collaborated on projects that started as *side-quests* and eventually turned out to pleasantly take most of my time during my last year. To Lavinia de Ferri, Alexis Pantos, Hartmut Kutzke, Pierre-Jean Lapray, Lyes Aksas, Alban Foulonneau, and Laurent Bigué, thank you for sharing your knowledge with me. I will cherish thinking about our collaborations and how much I could learn from each of you during my trips to Oslo and Mulhouse.

To the PhD students, now Doctors, of the CHANGE-ITN, and to the project itself. Thank you for sort of adopting me and embracing me. Joining the training schools represented for me a great chance for personal growth and I had some of the best experiences in my life.

To my friends Armelle, Chetana, Clara, Didrik, Gabriela, Gloria, Nakul, Thomas,

and Vegard. Thank you for being the great bunch you are and for letting me unplug and chillax.

To my friends back home in Italy: Tommo, Meo, Merco, Deciu, Corso, Fra. Just thank you, hearing from you is just the best.

All the most beautiful words I know cannot describe how thankful I am for my family. You have always been there I and love you.

To Annie, for navigating me through wonderful and just less wonderful times, for designing my posters and slides, for balancing my mood when the light would go off and on, for encouraging me to get a sticker every day. Love you, Beeboop.

Contents

Abstract	iii
Sammendrag	v
Acknowledgments	vii
I Introduction	1
1 Introduction	3
1.1 Motivation and broader context	3
1.2 Research aims and questions	6
1.3 Dissertation structure	8
1.4 List of contributing articles	8
2 Background	11
2.1 Introduction to the chapter	11
2.2 Analysis of Cultural Heritage Artefacts	13
2.2.1 Conservation Science	13
2.2.2 Paintings and Textiles	17

2.3	Light-Matter Interaction	20
2.3.1	Light description and the spectrum	20
2.3.2	Maxwell Equations and refractive index	22
2.3.3	Interference and Diffraction	26
2.3.4	Types of interaction	27
2.3.5	Radiometry	34
2.3.6	Illumination sources	36
2.3.7	Reflection Models	38
2.4	Detection of Visible and Infrared Radiation	42
2.4.1	The Photodetector	42
2.4.2	Materials for Detection	45
2.5	Toward Reflectance Imaging Spectroscopy	49
2.5.1	Imaging Principles	49
2.5.2	The camera as a simple optical system	51
2.5.3	Quantities related to image sharpness	52
2.5.4	Color and multispectral imaging	54
2.5.5	Spectroscopy Principles	60
2.5.6	Collection of spectra	61
2.5.7	Spectral Metrics	63
2.5.8	Hyperspectral Imaging a.k.a. Reflectance Imaging Spec- troscopy	66
2.5.9	Methods for capturing hyperspectral images	67
2.5.10	The spatio-spectral trade-off	68
2.5.11	Quality of RIS systems	69
2.5.12	Acquisition pipeline	70
2.5.13	Calibration pipeline	71
2.5.14	Multivariate analysis for hyperspectral images	74

2.5.15	RIS for Cultural Heritage Analysis	76
2.6	Fusion of VNIR-SWIR data	82
2.6.1	Image Registration	82
2.6.2	Spectral Splicing	85
2.6.3	Hypersharpening	86
2.7	Spectral and Polarization Imaging	92
3	An overview of needs and limitations	97
4	Summary of Articles	107
4.1	Articles grouping	107
4.2	Imaging setup	108
4.3	A1: Hyperspectral VNIR-SWIR image registration: Do not throw away those overlapping low SNR bands	110
4.4	A2: Logistic splicing correction for VNIR-SWIR reflectance ima- ging spectroscopy	113
4.5	A3: Reflectance imaging spectroscopy for the study of archaeolo- gical pre-Columbian textiles	115
4.6	A4: Full VNIR-SWIR hyperspectral imaging workflow for the monitoring of archaeological textiles	116
4.7	A5: Relationship between reflectance and degree of polarization in the VNIR-SWIR: A case study on art paintings with polarimetric reflectance imaging spectroscopy	118
5	Discussion	119
5.1	Revisiting the Research Questions	119
5.2	On the order of \mathcal{R} , \mathcal{S} , and \mathcal{H}	127
5.3	On the usage of rolling metrics	129
5.4	Visualization opportunities	129

6 Conclusion	133
6.1 Breadth of the Research	133
6.2 Contributions	133
6.3 Impact of the Research	134
6.4 Future directions and perspective	136
Bibliography	138
 II Original Articles	 165
Article A1_c: Hyperspectral VNIR-SWIR image registration: Do not throw away those overlapping low SNR bands	167
Article A2_j: Logistic splicing correction for VNIR–SWIR reflectance imaging spectroscopy	173
Article A3_j: Reflectance imaging spectroscopy for the study of archaeological pre-Columbian textiles	179
Article A4_c: Full VNIR-SWIR hyperspectral imaging workflow for the monitoring of archaeological textiles	213
Article A5_j: Relationship between reflectance and degree of polarization in the VNIR-SWIR: A case study on art paintings with polarimetric reflectance imaging spectroscopy	221

List of Acronyms

- BIDR** Bi-Illuminant Dichromatic Reflectance.
- BRDF** Bi-directional Reflectance Distribution Function.
- CCD** Charge-Coupled Device.
- CH** Cultural Heritage.
- CMF** Color Matching Functions.
- CMOS** Complementary Metal-Oxide Semiconductor.
- DSLR** Digital-Single Lens Reflex.
- EEA** Endmember Extraction Algorithm.
- FORS** Fiber Optics Reflectance Spectroscopy.
- FTIR** Fourier-Transform Infrared spectroscopy.
- FWHM** Full Width at Half Maximum.
- GS** Gram-Schmidt.
- GSD** Ground Sampling Distance.
- HSI** Hyperspectral Imaging.
- ICA** Independent Component Analysis.

KLPD Kullback–Leibler Pseudo Divergence.

LCTF Liquid Crystal Tunable Filter.

LED Light Emitting Diode.

LSF Line Spread Function.

MI Mutual Information.

MSI Multispectral Imaging.

MTF Modulation Transfer Function.

NMF Non-negative Matrix Factorization.

PCA Principal Component Analysis.

PFA Polarization Filter Array.

PSA Polarization State Analyzer.

PSF Point Spread Function.

RIS Reflectance Imaging Spectroscopy.

RMSE Root Mean Square Error.

ROI Region Of Interest.

SA Spectral Angle.

SC Spectral Correlation.

SFA Spectral Filter Array.

SIFT Scale-Invariant Feature Transform.

SNR Signal-to-Noise Ratio.

SPD Spectral Power Distribution.

SVD Singular Value Decomposition.

SWIR Short-Wave Infrared.

UV Ultraviolet.

VNIR Visible Near-Infrared.

XRF X-Ray Fluorescence.

Part I

Introduction

Chapter 1

Introduction

1.1 Motivation and broader context

This dissertation is the product of three years of work in a highly multi-disciplinary environment encompassing the fields of heritage science, optics, computer science, and signal processing, in no particular order. Broadly speaking, the work is intended to bring improvements to the development of a specific imaging technique - Reflectance Imaging Spectroscopy (RIS) (also known as Hyperspectral Imaging, HSI) - for its applications in the field of heritage science. With this premise, we can say that the work is done at the service of Cultural Heritage (CH).

The formal definitions of CH are provided by large organizations that regulate preservation practices around the world such as UNESCO and ICOMOS. Although they slightly differ, the two definitions seem to speak to us, people of the present, as a reminder that CH is not about us, but about future generations to come, and our task is to be its stewards.

the legacy of physical artifacts and intangible attributes of a group or society that are inherited from past generations, maintained in the present, and bestowed for the benefit of future generations

- UNESCO [1]

an expression of ways of living, developed by a community and passed on from generation to generation, including customs, practices, places, objects, artistic expressions, and values.

- ICOMOS [2]

Table 1.1: Definitions of Cultural Heritage provided by UNESCO and ICOMOS.

Cultural Heritage is categorized in its definition into *intangible* and *tangible*. The

former includes all forms of oral traditions and legacies that cannot be physically accessed, such as music, languages, poems, and culinary recipes [3]. The latter is the CH in the form that we are more accustomed to admiring in museums and exhibitions and can be split into *movable* (paintings, manuscripts, textiles, statues, etc.) and *immovable* (archaeological excavation sites, buildings, etc.) CH. Given the quality requirements related to the acquisition of images and the limitations of the technology at hand, this dissertation focuses on the application of RIS on small movable artifacts such as paintings and textiles.

The disciplines of heritage science and conservation science analyze CH artifacts with the goal of obtaining a full artifact characterization in the least possible invasive way. Characterizing an artifact can be understood as a way to explain why things appear as such as a result of production processes, but at the same time, reveal information that is not immediately visible to the human eye. Currently, the analytical techniques that are deployed to extract information from the artifacts are also those that bring the largest amounts of irreversible alterations. Finding an analytical technique that achieves the same levels of accuracy, while at the same time protecting the artworks, is then a paramount goal of heritage science. This is where this thesis tries to enter the conversation, by studying how to efficiently implement the fusion of RIS techniques to collect and analyze data of historical artifacts, developing every step with a goal in mind: whatever is implemented must be compliant with heritage safety.

Reflectance Imaging Spectroscopy is indisputably a marvel of innovation. Simultaneously with the development of early imaging technologies, scientists and researchers poured effort into improving their capabilities thanks to the knowledge gained in spectral optics. Nowadays, a variety of optical designs can be deployed to obtain hyperspectral images in which a single pixel or a group of neighboring pixels, if carefully calibrated, can carry information regarding the optical, physical, and chemical properties of materials.

Considering a purely technical viewpoint, the fusion of RIS techniques is something that recently saw a surge in interest in the field of remote sensing, with many research efforts that have been poured into developing dedicated algorithms. At the *proximal sensing* level, another wording of considering acquisitions carried out in a laboratory or controlled environment, two techniques: RIS in the Visible Near-Infrared (VNIR) and in the Short-Wave Infrared (SWIR) are typically analyzed independently. This dissertation tackles the problem of fusing RIS data coming from two different sources, but related to the same scene, in all its aspects, from how to collect the data more efficiently, to the processing steps involved, and to the final visualization. Figure 1.1 illustrates a very high-level ideal workflow that this thesis aims to accomplish.

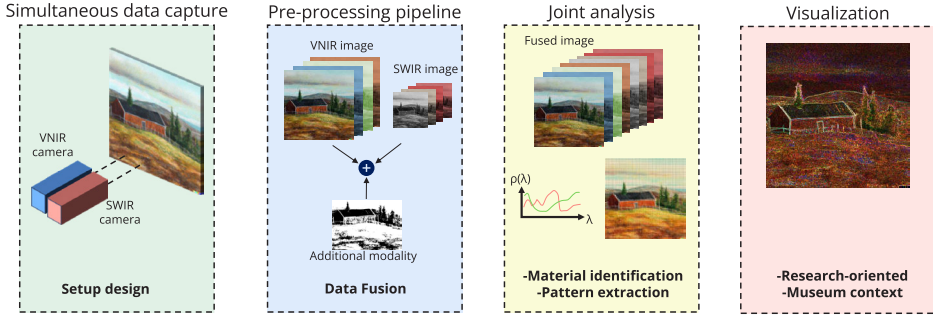


Figure 1.1: High-level workflow of VNIR-SWIR RIS for historical artifacts. The design of the imaging setup is performed aiming to obtain high-quality data while preserving the integrity of the artifacts. In the pre-processing, data fusion is accomplished by spatio-spectrally aligning the two individual hyperspectral images and accounting for any additional modality (in this thesis, polarimetric imaging). The data is then analyzed jointly to produce results that can be visualized in a research or museum context, depending on the necessities.

The need to fuse VNIR and SWIR information in the context of Heritage Science stems from the fact that typical historical artifacts such as paintings and textiles showcase a quite marked differentiation of the information in the two spectral ranges. As we will see later, the paint layer of a painting responds differently throughout the spectrum, being typically reflective in the visible range, and thus allowing us to perceive colors, while letting infrared radiation through. An infrared sensor would therefore be able to unveil structures underneath the paint layer such as underdrawings, *pentimenti*, or typical characteristics of deeper painting layers. For what concerns textiles, an infrared sensor could help us spot more easily different types of fibers that would look highly similar to the unaided eye. Nonetheless, not all elements that can be found in a textile show a segmentation of information, as in the case of typical materials used in conservation interventions such as fillings and markers.

Fusing datasets has the first goal of smoothly bringing different spectral responses to the same place and allowing an easier exploration. Although the fusion of datasets representative of different characteristics can sound redundant, as the interactions that are generated do not necessarily help in improving tasks strictly related to those characteristics, such interactions can generate new features (interpretable with the aid of machine learning) that can open up new paths to characterize historical artifacts.

1.2 Research aims and questions

The ultimate goal of this thesis is to allow conservation science practitioners to deploy VNIR-SWIR RIS in the most advantageous way. The definition of this goal, in relation to the state-of-the-art knowledge at our disposition, leads us to further define five sub-goals related to different aspects:

1. **Data acquisition:** Designing an imaging setup that allows the simultaneous acquisition of VNIR-SWIR data of high quality while respecting the guidelines of conservation and preservation of artworks.
2. **Fusion:** Devising a pipeline that provides VNIR-SWIR RIS data fusion at the pixel level.
3. **Analysis:** Application of algorithms for the spectral analysis of artworks in the extended VNIR-SWIR spectral range, taking advantage of data cross-talk.
4. **Visualization:** Development of visualizations otherwise inaccessible if VNIR and SWIR data are considered separately. Given the scope of this thesis, this topic is not treated as in-depth as the others.
5. **Multimodality:** Addition of modalities on the same imaging setup such as polarization imaging for artifacts with challenging appearance due to a complex surface topography.

To achieve the mentioned goals, we define two main research questions:

- **RQ1: How to efficiently build a fused hyperspectral image starting from two hyperspectral images in the VNIR and SWIR?**

When two images are acquired with different detectors working in the VNIR and SWIR ranges, both systems capture the scene radiance, but the way the information is presented cannot be readily fused. With this question, we want to investigate the best procedures that can be adopted to harmonize the data. It is important to point out that in this case, the word -efficiently- is used broadly as it encompasses concepts of efficiency related to the safe imaging of artifacts, while at the same time obtaining high-quality data with which a high-quality analysis is enabled.

- **RQ2: Can the performances of the typical tasks conducted in RIS analysis for CH be improved by considering the full extension of the spectral data in the VNIR-SWIR ranges?**

Pigment/dye mapping and unmixing are typical tasks performed to analyze artifacts using RIS. They are often conducted considering independent sets of spectral data, thus it is important to explore if processing them jointly can be advantageous.

Given the amount of contributions related to **RQ1**, it is perhaps beneficial to split it into three sub-questions that treat different topics, but work synergically to achieve the same result.

- **RQ1a: What are the factors and decisions that influence the performance of spatio-spectrally aligning two hyperspectral images coming from two different sources and electromagnetic ranges?**

The spatial alignment of hyperspectral images is a process that can be best summarized by a decision tree, with each path along the branches that leads to different performance results. Moreover, spectral alignment is a necessary step since, as it will be explained later in detail (Section 2.6.2), two different sensors capturing the same information, e.g. spectral radiance, will rarely have a matching output. This is translated into spectral anomalies that make the spectra look like they present unnatural discontinuities.

- **RQ1b: What are the implications of sharpening techniques developed in remote sensing when translated to proximal sensing applications?**

The field of RIS for CH significantly draws innovative analytical techniques from the field of remote sensing, which is a huge engine in research for RIS. With this question, we want to explore if the techniques developed for sharpening an image (e.g. increase its resolution) are suitable for a completely different imaging domain, and thus assess if a dedicated approach is needed.

- **RQ1c: Is it possible to develop a paradigm for the joint analysis of VNIR-SWIR RIS and polarimetric imaging?**

Many artworks such as paintings with complex surface textures can be challenging to image due to the high amount of generated specular reflections. Polarization imaging in combination with RIS can offer a gateway to investigating the material properties in those corrupted areas, but it has never been implemented in the full VNIR-SWIR spectral range.

Taken more broadly, this question wants to evaluate the possibilities of including a layer of multimodality to the various spectral acquisitions that can be performed.

1.3 Dissertation structure

This thesis is organized with a two-part design in which the second part collects the published results of the work conducted during the last three years. The list of contributing articles is found at the end of this section. Out of the five contributing articles, three have been submitted to peer-reviewed journals and two have been presented and published in the proceedings of peer-reviewed conferences. An additional supporting conference article is listed as well. This will not be included in the final collection of articles although it represented a meaningful step towards the designing of one of the contributing articles.

The first part of this thesis wants to shift the attention to all the aspects that are involved when RIS is deployed in the study of artworks in a heritage science context. Chapter 1, which is almost over, was a broad introduction to the context in which this work is placed, highlighting the current needs of heritage science to find a sustainable imaging technique and the need for unifying spectral datasets that are often treated independently. Here, we defined the main goals and research questions tackled in this dissertation. Chapter 2 addresses the fundamental background knowledge that is needed to fully understand how RIS works and is thus an information-dense chapter that a knowledgeable reader can browse more quickly. In Chapter 3 we reflect upon the research question with a different angle provided by an increased knowledge gained in the previous chapter. Here, we shall address what are the specific limits and needs of the state of the art, and thus the rationale for which the contributing articles were written. Chapter 4 proposes the common imaging setup deployed in all contributing articles and then gathers the summaries of the articles and their links. The discussion brought in Chapter 5 focuses on the analysis of the contributions of this thesis in relation to the proposed research questions of Chapter 1, while Chapter 6 provides a closing overview and future perspectives in the field of RIS for the analysis of artworks.

1.4 List of contributing articles

The following articles constitute the core contributions of this thesis. The subscript -j- indicates an article published in a journal, whereas -c- refers to conference publications. The research works of Articles **A3_j** and **A5_j** were conducted in multidisciplinary teams with specialized researchers in the field of Cultural Heritage and polarimetric imaging, respectively, whereas the research works of Articles **A1_c**, **A2_j**, and **A5_j** were conducted with the academic advisors.

A1_c Grillini, F., Thomas, J-B., George, S. (2022). *Hyperspectral VNIR-SWIR image registration: Do not throw away those overlapping low SNR bands*. 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote

Sensing (IEEE WHISPERS). DOI: <https://doi.org/10.1109/WHISPERS56178.2022.9955080>.

A2_j Grillini, F., Thomas, J-B., George, S. (2023). *Logistic splicing correction for VNIR–SWIR reflectance imaging spectroscopy*. Optics Letters, vol. 48, pg. 403–406 (Optica Publishing Group). DOI: <https://doi.org/10.1364/OL.478691>

A3_j Grillini, F., De Ferri, L., Pantos, G.A., George, S., Veseth, M. (2023). *Reflectance imaging spectroscopy for the study of archaeological pre-Columbian textiles*. (Under review)

A4_c Grillini, F., Thomas, J-B., George, S. (2023). *Full VNIR-SWIR hyperspectral imaging workflow for the monitoring of archaeological textiles*. Archiving Conference, vol. 20, pg 192–197 (Society for Imaging Science and Technology). DOI: [10.2352/issn.2169-2629.2021.29.276](https://doi.org/10.2352/issn.2169-2629.2021.29.276)

A5_j Grillini, F., Aksas, L., Lapray, P-J., Foulonneau, A., Thomas, J-B., George, S., Bigué, L. (2023). *Relationship between reflectance and degree of polarization in the VNIR-SWIR: A case study on art paintings with polarimetric reflectance imaging spectroscopy*. (Under review)

Supporting article

The following article represented a previous attempt to solve the problem addressed in article **A2_j**. As such, it is decided to not include it in the list of core articles, but it is important to highlight the pedagogical role that this work had within this thesis.

S1_c Grillini, F., Thomas, J-B., George, S. (2021). *Radiometric spectral fusion of VNIR and SWIR hyperspectral cameras*. Color and Imaging Conference, vol. 29, pg. 276–281 (Society for Imaging Science and Technology). DOI: <https://doi.org/10.2352/issn.2169-2629.2021.29.276>

Articles not included

The first period of the doctorate saw the refinement of the work conducted in my Master’s Thesis with the same supervision team. This led to the publication of two journal articles on the topic of spectral unmixing. Although related to the work conducted in this thesis (applications of RIS for CH), it was decided to not include these publications in the list of core articles. The reason for this is to avoid ambiguity with the main topic of this thesis, which focuses on the fusion of two techniques of RIS, while my Master’s thesis focused on the role of imaging models in the unmixing process for pigment mapping in the VNIR range.

Grillini, F., Thomas, J-B., George, S. (2021). *Comparison of imaging models for spectral unmixing in oil painting*. Sensors, vol. 21 (7), pg. 2471–2486 (MDPI). DOI: <https://doi.org/10.3390/s21072471>

Grillini, F., Thomas, J-B., George, S. (2021). *VisNIR pigment mapping and re-rendering of an experimental painting*. Journal of the International Colour Association, vol. 26, pg. 3–10 (AIC). https://aic-color.org/resources/Documents/jaic_v26_01.pdf

Chapter 2

Background

2.1 Introduction to the chapter

The title of this dissertation is quite a mouthful, and if I think about it, up to a few years ago I did not know what half of the words included meant. For this reason, I decided that the chapter that introduces the fundamental background should also tell a story of how I got to know all the different things that are hiding behind this title.

It might sound like a long shot, but I felt that I owed this to my path of studies, which is coming to an end with this dissertation -yes, we never stop learning-. I also soon realized that a book would be more appropriate, but maybe that would be for another time. In this chapter, I will try to go through all the concepts that I have had to stumble upon at least once and try to explain them in a way that I find suitable for a student who starts reading about this broad topic -that student is ideally me a few years ago-.

Taking a closer look at the title, I decided that a good way of starting the chapter would be by splitting it into five (plus one) parts, as I display in Figure 2.1 with different colors.

The first part is represented by the words *Cultural Heritage*. This dissertation, being at the service of Cultural Heritage, could have not started differently. In this part, I will address the needs of Conservation Science in terms of analysis and digitization of artifacts, focusing on two macro-groups of artwork that are recurrent in the dissertation and in the connected research articles: paintings and textiles.

Reflectance Imaging Spectroscopy is a term that indicates a specific imaging tech-

② REFLECTANCE ④ IMAGING SPECTROSCOPY:
⑤ FUSION OF ③ VNIR AND SWIR FOR
① CULTURAL HERITAGE ANALYSIS

Figure 2.1: Title breakdown according to the subdivision proposed to introduce the fundamental theoretical background.

nique. However, I found it easier to separate the *Reflectance* part to introduce the reader to the realm of light-matter interaction. In this physics-dense part, I will try to build on subsequent bricks in order to arrive at a destination point in which it is possible to understand the phenomena that take place when light (including infrared radiation) interacts with paintings and textiles. In writing this part, the bibliography support, whenever something sounds like it should need a literature reference, came from the following sources: *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* by Max Born and Emil Wolf [4], *Building electro-optical systems: making it all work* by Phil Hobbs [5], *Detection of optical and infrared radiation* by Robert Kingston [6], *Introduction to Radiometry and Photometry* by William Ross McCluney [7].

A portion of the electromagnetic spectrum can be termed *VNIR-SWIR*, but let us not get ahead of ourselves, this will be better introduced in the *Reflectance* part. Here, I will take the opportunity to take a closer look at the conversion of visible and infrared radiation into digital numbers stored in a file, which is a marvel of engineering and applications of the laws of physics, chemistry, and electronics. Images and signals are processed by computers as a series of zeros and ones, but before obtaining these long binary strings, several steps in the analog world had to take place. I will try to describe those steps starting from the main component of the process, the photodetector. The main sources in writing this part were *Wide bandgap semiconductors* by Kiyoshi Takahashi, Akihiko Yoshikawa, and Adarsh Sandhu [8] and *Introduction to infrared and electro-optical systems* by Ronald Driggers, Melving Friedman, and Jonatan Nichols [9].

The part related to *Imaging Spectroscopy* will describe concepts related to both imaging and spectroscopy, starting from the fundamentals of how to separately acquire images and spectra, and then fusing the two.

The fifth part, from the word *fusion*, concerns the coming together of two tech-

niques of imaging spectroscopy in different spectral ranges and their data, which potentially can generate a new dataset that is more informative than the sum of its parts. Thus, concepts of image fusion, image registration, and sharpening are discussed.

One last part will then discuss the implementation of Spectral Polarization Imaging as an additional modality that can be used to explore more challenging artifacts.

2.2 Analysis of Cultural Heritage Artefacts

2.2.1 Conservation Science

As we walk through the corridors, spacious halls, and modernly designed rooms of an exhibition, it can become so easy for us, contemporary individuals, to connect with Cultures and people who once inhabited remote lands and times. The exhibited artifacts, together with the information retrieved by archaeologists, humanists, and historians, carry a special power that can instantly transport us back in time and to the other side of the world. What is sometimes even more powerful, is the thought that what we are able to appreciate in an exhibition is only a tiny part of the evidence of the passage of entire Cultures on our planet, with the vast majority of objects and traditions that ended up being destroyed or lost. On the other hand, the artifacts that eventually reach the frames of the exhibition, make it because of the often understated efforts of conservators and preservation scientists.

Conservation Science and Heritage Science [10] are fields of research dating back to the 1800s that aim to study historical artifacts with scientific equipment, in order to formulate hypotheses regarding their origins, manufacturing, and physical and/or chemical composition. The main goal of conservation scientists is to fully characterize historical objects so as to propose the best practices to preserve the artifacts in the best possible conditions or to carry out targeted treatments following the principles of minimal intervention and invasivity [11]. The word *invasivity*, which will appear several times during this dissertation, is in this instance to be intended with the twofold meaning of extension, i.e. the proposed treatments are applied only on the areas that strictly require it, and alteration, i.e. the proposed methods should only alter specific properties of the treated areas.

A full artifact characterization is certainly useful to produce an individual narrative that can help us to place small objects in broader historical contexts, but the other way around is also true. Indeed, gaining all these insights about a specific object can also assist us in indirectly inferring assumptions about how the disappeared societies would live and the knowledge they owned. Take for example the case of an excavation site in northern Italy where archaeologists discovered an ancient

forge. By studying the remnants of slags it is possible to trace back what temperatures were reached to fuse the metals, which is a crucial piece of information when describing the advancement of forging technologies in the Early Middle Ages [12].

As the number of characterized artifacts increases, so does our knowledge regarding which materials and technologies were owned by each Culture. It becomes thus easier to assign the correct *labels* to newly discovered artifacts, although the risk of committing errors is always present, especially in the case of art forgeries. Forgers have been profiting from the sale of realistic fakes that went undetected because of their absolute resemblance with the style of Old Masters. The detection of fakes is a task that can become easier as the checklist of items to be controlled is extended. Of course, the dilemma is clear, should the checklist be public for everyone (including the art forgers) to see?

Conservation Science, a discipline that originated from visual observations and annotations, has swiftly shifted toward a data-based approach in which the nature of the collected data is not unique. If the technological advances achieved in physics and chemistry were pushing the research forward in earlier times, now an additional player, represented by computer science and artificial intelligence is heavily contributing [13]. This makes Conservation Science an inherently multidisciplinary field of research that can be explored from multiple directions [14].

The collection of data from historical artifacts is a topic regulated by international codes of ethics [15] that differentiate for the type of material that is examined (paper, wood, stone, etc.), the goal of the object acquisition (scientific research, digitization, etc.), and the type of collection (invasivity, destructivity, contact, point measurement, scanning). However, given the shape and extension variety of artifacts, the documentation cannot presently cover all the variants in which the acquisition techniques can be adopted. According to the guidelines for the acquisition of Cultural Heritage artifacts, the ideal technique should be completely non-invasive. Invasivity can be defined as the amount of alteration that is provoked by the application of an acquisition technique, and can therefore be ideally quantified by measuring specific properties before and after the acquisition campaign. However, the provoked alteration can be of different natures and caused by a series of sources. Some analytical techniques can only be deployed in specific conditions, and therefore require the physical extraction of samples from the artifacts. In this case, the technique is defined as destructive, as it causes an irreversible mechanical change. Other mechanical variations on the surface of artifacts can be caused by techniques that require to be in contact with the material. On the other hand, there exist techniques that irreversibly alter the chemical properties of the materials because of the radiation that they deploy. As we will see in the following sections, light itself is radiation, and its deployment in an acquisition campaign

can sometimes lead to an alteration of physical properties, hence regulations must be enforced also in cases in which an apparently harmless setup composed of a light source and a camera is adopted.

At the present time, sampling is still a necessary procedure when studying historical artifacts, since it opens the possibility of deploying analytical techniques such as High-Performance Liquid Chromatography and Surface-Enhanced Raman Spectroscopy to accurately retrieve the relative presence of chemical elements within a volume. The high degree of accuracy of such techniques makes them indispensable in studies that involve the usage of complementary and less invasive techniques, as they provide a *true* response that is often deployed as *ground truth* in classification tasks. A plethora of studies are currently ongoing to reduce the gap in the invasivity-accuracy trade-off by improving the analytical performances of non-invasive techniques so that they could operate independently in the future. Realistically, this scenario could be hard to achieve, but a desirable outcome would foresee the usage of non-invasive analytical techniques to pinpoint relevant areas for sampling, thus avoiding unnecessary extractions.

Another well-known trade-off is the one regarding the amount of spatial information that can be collected at the expense of the physical quantity examined. Punctual techniques, which operate integrating on a single spot (typically of a few millimeters in diameter) offer results that are less affected by measurement fluctuations, i.e. *noise*, when compared to their imaging counterparts. This happens, as we will see in the following sections because when the number of deployed sensors (pixels) is increased, more factors concur to the creation of noise, which is statistically more present. Imaging techniques, which usually deploy scanning or single-shots, offer however a certain user-friendliness in interpreting the data and are significantly faster and easier to perform. Spatial information has additionally the twofold goal of supporting individual punctual measurements and extracting more information exploiting the generation of structural patterns typical of images.

Imaging is nowadays the cornerstone of digitization in museums and archives all over the world. Since the first times in which analog cameras were available, pictures have been taken to document the archeological sites and the contexts in which artifacts were found. Then, more recently, and with the advent of digital cameras, access to imaging technologies increased exponentially, facilitating the usage of imaging for an array of documentation and monitoring tasks.

Today, the standards FADGI [16] in the United States and ISO internationally [17, 18] regulate digitization in archival contexts, setting the benchmark for image quality in official archives. Visualization is also a core value for museums, both in the research and in dissemination to the public side. Photogrammetry is a

well-diffused imaging technique that allows the virtual reconstruction of a physical artifact from the acquisition of hundreds of images (depending on the size of the object) that compose a *cloud* [19]. With Gigapixel imaging [20] it is possible to discover details that would remain hidden from the human eye during a regular exhibition visit because of the limited allowed observation distance. Visualization also includes rendering, and Reflectance Transformation Imaging is a technique that allows the surface reconstruction of artifacts, thus enabling the consequent relighting of an object with a given light source from a certain direction, so as to conveniently explore textural features [21].

Over the years, imaging has also transitioned from a role purely centered around visualization to a more prominent role in the landscape of analytical techniques. What has allowed imaging to reach this hybrid status is the introduction of calibration procedures that allow to perform metrology, both from a standpoint of metric measurements (distances, depths, etc.) and physical measurements (amount of reflected light, colorimetry). Different modalities of imaging are used in daily practices in Conservation Science, from the more common RGB and Ultraviolet (UV) [22] to X-Ray Fluorescence (XRF) [23], Fourier-Transform Infrared spectroscopy (FTIR) [24], and RIS [25, 26]. These techniques, sometimes used in combination, allow performing a variety of tasks, including but not limited to monitoring of artifacts in environmental conditions [27], studying the effectiveness of treatments during conservation interventions [28], monitoring of aging [29], detection and mapping of components [30]. Although most of the imaging techniques listed here are deemed as non-invasive, they still require to use of radiation (typically in the form of light) that is substantially different than the one the artifacts are normally exposed to, thus triggering physical-chemical reactions that could potentially induce permanent changes [31].

The field of research of Conservation Science is at the service of the people to reconstruct the past using the equipment of the present, but not looking into the future could be the most dangerous practice. The imminent impact that climate change will have on the lives of billions of people, with the due prioritization, will also resonate with how their Heritage will be preserved. Research efforts will be needed to study the relationship between deterioration and climate change, for both movable and immovable, indoor and outdoor artifacts. Although this is not the direct goal of this thesis, it is hoped that working within a framework that easily combines information from different spectral ranges will increase the predictive power of models that describe how heritage materials behave when exposed to abnormal conditions, thus enabling preventive conservation actions to minimize risk and damage.

Another important aspect that needs to be addressed is access to technology for

everyone, by reducing costs, while at the same time making the technology more accessible and sustainable [32], by investing in the development of portable instrumentation.

2.2.2 Paintings and Textiles

In the following paragraphs, we will take a closer look at the structures that form two commonly found types of artifacts: oil paintings and textiles. Knowing more in-depth about the physical characteristics of these objects will be helpful in tackling the contents of the next sections and the research articles that follow.

In introducing the two types of artifacts, it is perhaps useful to place them in the context of a practical example that will be a recurring theme in this dissertation and that is appropriate given the topics that are introduced at a later stage.

Let us suppose that during the closing day of the museum, the technical imaging department carries out photography activities on the largest painting and textile piece of the collection. To do so, the ordinary illumination of the museum is switched off, and specifically designed light sources are deployed. A camera is placed on a tripod standing on the floor, the standard reference targets are placed in the scene, and the imaging campaign can commence.

Typical structure of an oil painting

When we visit a painting exhibition, everything we can see of an exposed artwork is the outer surface, the frame that contains it, and sometimes, a protective glass. However, if we had to theoretically cut through the middle of the painting and extract a cross-section, we would observe a very well-defined stratigraphy [33, 34]. More likely, if we had to extract cross-sections from different paintings, we would never find a common agreement on the number of deployed layers, as different painters from different eras received slightly distinct training and the painting techniques changed over the years. However, some elements are crucial to the very definition of an oil painting and will be found in all artworks [35].

The typical, and non-binding, layered structure of an oil painting is schematically reported in Figure 2.2. The first layer that receives attention in the process of composing a painting is the support. Typical supports for oil painting include canvas (lined in different materials such as linen, wool, and silk), wood, paper, metal, and glass [36]. A layer of glue can be applied to facilitate the adhesion of the next layer, which can be sometimes an additional support layer (in the case of paintings that have been consolidated in a conservation treatment [37]), or the ground. The ground, or preparatory layer, is usually composed of *gesso*, a thick chalk-based fluid that forms a protective barrier for the support, preventing the diffusion of

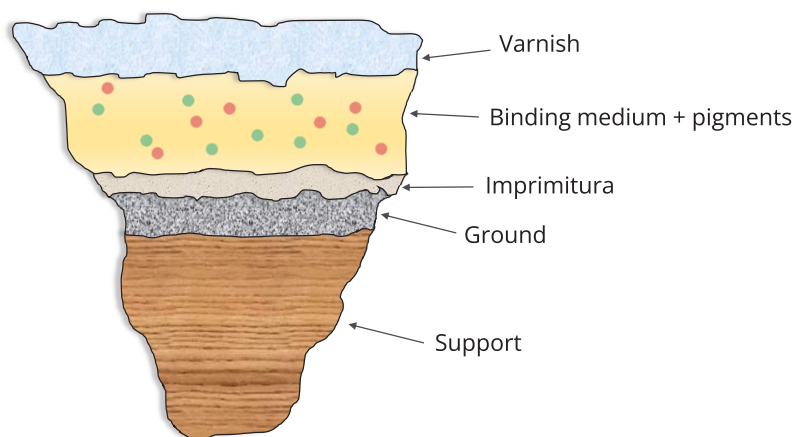


Figure 2.2: Schematic representation of a typical cross-section of an oil painting. The colors and proportions of the layer thicknesses are deliberately unrealistic for visualization purposes.

oil [38]. Sketches and underdrawings [39], usually performed with carbon-based traces, are drawn on the ground layer and covered by a priming glaze (*imprimatura* [40]) which facilitates the adhesion of the paint. In the paint layer, the pigment [41] particles float within the volume of the binding medium [42] (typically linseed oil) and are responsible for the color appearance that is perceived by the human eye. Finally, a transparent layer of resin-based varnish [43] can be applied with the twofold goal of protecting the paint layer from environmental agents such as dust and enhancing the colors, providing sometimes a glossy appearance. Varnishes lose their transparency with time due to the exposure to light and Ultraviolet (UV) radiation, and a yellowing phenomenon can be observed if the varnish layer is not replaced.

As mentioned, the structure and the styles with which the different layers are applied in each and every case are affected by a great variance that represents the development of artistic currents and individual creativity. Most of the variation, which eventually affects the way a painting appears, is found in the manner the paint layer is applied, i.e. the pictorial technique, and in the way pigments are mixed. Different pictorial techniques work sometimes as distinct signatures of artists [44]. Just think about the thick application of *impasto* by Vincent Van Gogh, George Seurat's *pointilism*, or Claude Monet's *Water Lilies* painted *en plein air*. The first deeper layers that compose the painting structure can be approximated to 2D surfaces, whereas the paint layer has often a strong 3D component that affects the visual perception with glossy reflections [19] and shadow areas. When we

consider pigment mixing, it can take place directly on the canvas or on the palette. Particular visual effects of translucency can be created when artists apply different paint layers on top of each other [45].

Typical structure of historical textiles

The first evidence of textile artifacts dates back around the year 8000 BC and is attributed to early civilizations that inhabited the territories of what are known today as the Peruvian Andes [46]. Fibers, the fundamental building blocks of textiles, have been foraged since prehistoric times according to the local availability of flora and fauna. Fibers are long, strong, flexible filaments typically made of cotton, silk, wool, and flax [47, 48]. Their mechanical properties are ideal for the production of yarns, which are the results of spinning and twisting fibers together. Evidence of yarns dated prior to the 12th century shows that a single type of fiber was normally used to produce the yarns, whereas the production of multifilament yarns, obtained by blending different types of fibers, is attributed to populations of Punjab (India) [49].

Yarns can be interlaced in the case of weaving or interlooped in the case of knitting to produce fabrics, which are a specific type of textile recognizable for their planar shape and flexibility. Modern non-woven fabrics can be obtained by industrial processes that directly interlock the fibers together by exercising high mechanical, thermal, or chemical forces. For this dissertation, the studied textiles are produced by weaving. In weaving, two sets of yarns - the warp, in the lengthwise direction, and the weft, in the crosswise direction - are interlaced at a 90° angle to produce more or less detailed patterns. The final product of weaving can be a tapestry, when intricate designs and pictorial representations are directly woven into the fabric, usually by means of a loom. When a pattern is stitched on top of a pre-existing fabric, usually with a needle, then the final textile is called an embroidery. A schematic representation of a plain weave fabric with examples of tapestry and embroidery is reported in Figure 2.3.

Natural fibers per se exhibit limited color variation, predominantly showcasing shades of yellow, brown, and various achromatic tones of white. Since ancient times, more vibrant colors have been obtained by natural dyeing processes. Plants offer multiple sources of dyes from different parts such as roots, leaves, stems, and flowers. Animal-based dyes were extracted from cochineal insects and sea snails, whereas a series of minerals offered colorants in the form of iron oxides and salts [50, 51]. However, not all fibers have a natural disposition to be colored, and require the action of a *mordant*, a substance that aids in the fixation of dye onto the yarn [52]. This process can potentially modify the resulting color, thereby expanding the gamut of achievable shades. A typical dyeing process entails boiling the

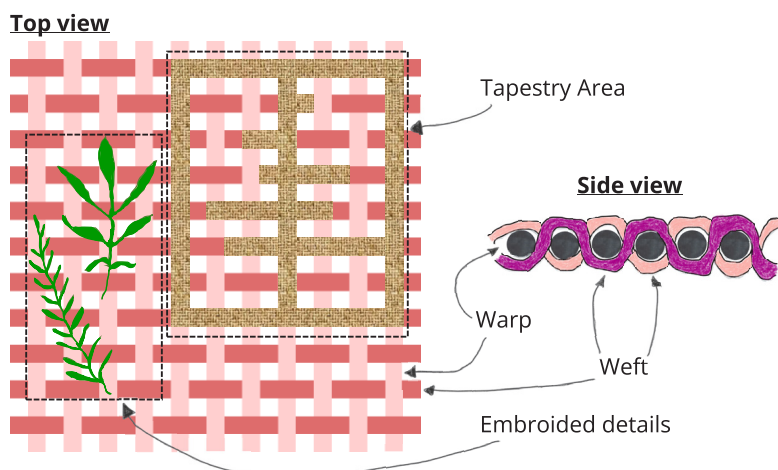


Figure 2.3: Schematic representation of a plain weave fabric from a top view (left) and side view (right). In the tapestry area, yarns of different colors are intrinsically woven in the warp-weft pattern, whereas in the embroidery example, new yarns are sewn onto the base layer of the fabric.

dyeing agent and the mordants, immersing or simmering the yarn in the solution, and subsequently proceeding with rinsing and drying steps.

2.3 Light-Matter Interaction

2.3.1 Light description and the spectrum

Generations of scientists have been puzzled in the attempt to provide a definition to the concept of light, and in present times it is common to alternate between three fundamental descriptions, depending on the phenomena that we are trying to explain. This light ambiguity, after centuries of speculation and research, boils down to the description of light as a swarm of particles, as a wave, or as a ray.

The idea that light consists of microscopic corpuscles had been around for centuries, but it was only in 1900 that the *photons*, the light quanti, were first theorized by the German physicist Max Planck. A photon is a mass-less particle that moves in vacuum at a speed $c \approx 3 \cdot 10^8 \text{ m} \cdot \text{s}^{-1}$ while carrying energy stored as an electromagnetic field oscillating at a frequency ν . The relation between the carried energy E and the frequency is modeled according to Planck's law defined in Equation 2.1,

$$E = h \nu \quad (2.1)$$

in which h is the Planck's constant ($\approx 6.26 \cdot 10^{-34}$ J·s). The quantum description of light as a particle can explain a variety of excitation phenomena such as the Compton scattering [53] effect and the Photo-electric effect [54], but it fails to explain a category of phenomena related to light propagation. Let us briefly describe the Photo-electric effect, which takes place when a highly energetic photon collides with a material such as a metal plate. In *rest* conditions, the electrons of the metal atoms are bound to the material through a binding energy called the *work function* ϕ of the metal. When the photon i strikes, all of its energy $E_i = h\nu_i$ is absorbed by the material. If $E_i > \phi$, then the electron is ejected from the material, becoming a *photo-electron*. The excess energy not used to break the bond is transferred to the photo-electron in the form of kinetic energy $E_k = E_i - \phi$ and assumes a range of values typical of the metal. Albert Einstein was awarded the Nobel Prize in Physics in 1905 for his formulation of the Photo-electric effect.

On the other hand, when light is described as a wave, it is characterized as the periodic oscillation of an electromagnetic field that travels at a speed c along a direction of propagation. An electromagnetic wave is the result of the coupling of an electric field and a magnetic field perpendicular to each other and in turn perpendicular to the direction of propagation. In the simplest case, the wave is said to be monochromatic, which means that the electromagnetic field oscillates at a single frequency ν . The frequency is related to the wavelength λ , i.e. the distance between two consecutive peaks, as described in Equation 2.2.

$$\lambda = \frac{c}{\nu} \quad (2.2)$$

An effect that can be explained by considering the wave formulation of light is the Doppler effect, which manifests itself in astronomy in the cases of red-shift and blue-shift. The observation of moving galaxies and stars is affected by the relative positions of the observer and the target object so that a red-shift (color moving towards longer wavelengths, i.e. lower frequencies) is experienced when the distance between the parts increases, while a blue-shift is observed when the parts get closer to each other.

Wavelength and frequency are convenient parameters for characterizing light, or more in general, radiation. In fact, we define *light* only that very small portion of the radiation that the human visual system is able to detect. In this dissertation, we characterize light by its wavelength using the *nanometer* as the preferred unit. The selection of the unit is made for convenience and depends on having an intuitive identification of the considered radiation. While the International System unit for frequency is the Hertz (Hz), within the field of CH it is not uncommon to characterize the radiation along the spectrum differently according to the deployed ana-

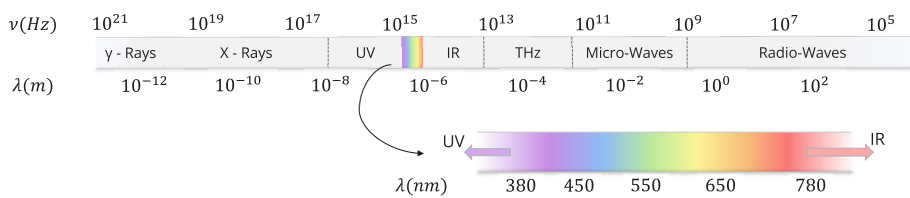


Figure 2.4: Visible radiation occupies a narrow interval in which a single harmonic (doubling of frequency) in the frequency domain is included, whereas most of the other defined types of radiation span several frequency magnitude orders.

lytical techniques and observed quantities. For instance, X-Ray Fluorescence describes re-emission peaks in terms of energy (electron-Volt eV), whereas Fourier-Transform Infrared spectroscopy displays curves in wave numbers (cm^{-1}).

Different types of radiation can be defined following the spectrum (Figure 2.4): Gamma-rays and X-rays are the most energetic known forms of radiation. The Ultraviolet, divided into UV-A, UV-B, and UV-C, spans the portion of the electromagnetic spectrum from 100 nm to 380 nm. Visible light occupies a narrow spectral window from 380 nm to 780 nm, and as we already briefly mentioned in the example of red-shift/blue-shift, the progression from low to high wavelengths follows the color of the primary rainbow as observed from bottom to top. Moving to wavelengths in the order of micrometers, we progressively encounter the three regions of the infrared domain IR-A, IR-B, and IR-C, more known under their secondary nomenclature as Near-Infrared (NIR, 780 nm - 1500 nm), Short-Wave Infrared (SWIR, 1500 nm - 3000 nm), and Far Infrared (FIR, 3000 nm - 10000 nm) [7]. The latter infrared sector is usually selected to represent the domain of thermal radiation. The longest observed wavelengths include Tera-Hertz radiation, microwaves, and radiowaves, with the latter being used in long-distance communications. It is worth mentioning that the extreme theorized wavelengths have not been observed, but are the size of the universe on one end, and the Planck length on the other.

The third description of light as a ray traveling through space is the most intuitive one, but its usage is circumscribed to the description of simpler phenomena related to geometrical optics.

2.3.2 Maxwell Equations and refractive index

The Maxwell equations are a set of four fundamental equations that represent the basics of electromagnetic theory. We have already defined light as an oscillating electromagnetic field that travels at a speed c , but with this set of equations, it

is possible to observe how the electric field \vec{E} and the magnetic field \vec{H} propagate through time and space. Let us briefly introduce the four equations and their physical implications in electromagnetism theory in the following.

The first of the Maxwell equations is also known as flux theorem or Gauss' law. In the differential form reported in Equation 2.3, the theorem states that the divergence ($\nabla \cdot$, the measure of how much a vector field *spreads out* or *diverges* from a given point, computed as the sum of the partial derivatives of the vector field with respect to the spatial dimensions) of the electric field is proportional to the local density of the electric charge ρ .

$$\nabla \cdot \vec{E} = \frac{\rho}{\mathcal{E}_0 \mathcal{E}_R} \quad (2.3)$$

where \mathcal{E}_R is the *permittivity* of the material, i.e. the tendency of the material's positive and negative charges to separate when subjected to an electric field. The permittivity expressed with the subscript ₀ represents the permittivity of *free-space*, a perfectly ideal vacuum where there are no electric charges nor currents ($\mathcal{E}_0 \approx 8.854 \cdot 10^{-12} F \cdot m^{-1}$). In other words, Gauss' law describes the well-known observation that electric charges of the same sign repel each other.

In the same way with which the first Maxwell equation concerns solely the electric field, the second equation (Equation 2.4) focuses on the magnetic field. Also similarly, the second equation was formulated by Gauss as a consequence of the first. In this case, the divergence of the magnetic field is zero.

$$\nabla \cdot \vec{H} = 0 \quad (2.4)$$

If the divergence of a field in a given point is 0, it means that the vectors neither converge nor diverge, or if put in another way, the vector lines do not start or end at any point in space. When we observe a magnetic field, we observe indeed closed lines. The main implication of this law is that the existence of a magnetic monopole (an individual magnetic charge) is forbidden, and it is only possible to observe dipoles.

The third and fourth Maxwell laws concern the interactions of the concatenated electric and magnetic fields. The equation known as Faraday's law of induction serves as the basis for the third equation in Maxwell's set. The Faraday-Maxwell equation in its differential form (Equation 2.5) states that the *curl* ($\nabla \times$, the measure of how much a vector field *revolves* around a given point in space, computed as a linear combination of the partial derivatives with respect to the spatial dimen-

sions) of the electric field is proportional to the partial derivative of the magnetic field with respect to time.

$$\nabla \times \vec{E} = -\mu_0\mu_R \frac{\partial \vec{H}}{\partial t} \quad (2.5)$$

where μ_R is the permeability, i.e. the tendency of a material to be magnetized by the application of a magnetic field, and μ_0 is the permeability of free-space ($\mu_0 \approx 1.257 \cdot 10^{-6} \text{ N} \cdot \text{A}^{-2}$). In practical terms, the application of a magnetic field *induces* the charges of a dielectric material to separate. The charge separation generates an electric field which in turn can generate a current.

The last Maxwell equation symmetrically concerns the curl of the magnetic field expressed in terms of a temporal derivative of the electric field. This law, also known as Ampère's circuital law, can be formulated as illustrated in Equation 2.6.

$$\nabla \times \vec{H} = \varepsilon_0\varepsilon_R \frac{\partial \vec{E}}{\partial t} + \vec{J} \quad (2.6)$$

in which \vec{J} is the electric current. This law describes how the application of an electric field and/or of a current can generate a magnetic field.

The concept of *free-space*, or vacuum, represents an unnatural condition on our planet, but it has served as a precious mental gym for the formulation of laws that govern the behavior of electromagnetic radiation in our daily lives. As already mentioned, no electric charges nor currents are allowed to exist in vacuum. With this simple but powerful statement, it is possible to simplify greatly Maxwell's first and fourth equations (Equation 2.3 and Equation 2.6), as the charge density ρ and the current \vec{J} can be set to 0. In this configuration, it is possible to combine the Maxwell equations to obtain a wave equation for the electric field as Equation 2.7 describes.

$$\frac{c^2}{\varepsilon_R\mu_R} \nabla^2 \vec{E} = \frac{\partial^2 \vec{E}}{\partial t^2} \quad (2.7)$$

The Laplacian operator ∇^2 represents the divergence of the divergence, so it describes the rate of change of a field vector at a given point in space.

The formulation for the magnetic field is similar, but in practice seldomly considered, as the strength of the magnetic field for optical radiation is a few orders of magnitude weaker than the strength of the electric field.

This brings us to the first characterization of the wave equation, as the direction in which the electric field oscillates becomes of paramount importance. Indeed we define *polarization* as the direction of the electric field. In the simplest case, the electric field oscillates on a plane along the direction of propagation, and thus it is said to have *linear* polarization. In some cases, the direction of the electric field can rotate on a plane. If the magnitude is constant for every angle of rotation, the polarization state is said to be *circular*, whereas if it shows perpendicular maxima and minima points it is said to be *elliptical*. There exist other states of polarization, but most of the light that enters our eyes is *unpolarized*, which is an unachievable condition for a single wave, but not for the superimposition of many waves. Unpolarized light, in fact, is light that contains components with a random distribution of polarization states.

The second observation that we can make about Equation 2.7 is that the speed of propagation does not have a dependence on the considered frequency or wavelength (or photon energy), but rather it will always be c in vacuum. The Maxwell equations predict the speed of light in vacuum to be equivalent to a universal constant described in Equation 2.8.

$$c = \frac{1}{\sqrt{\mathcal{E}_0 \mu_0}} \quad (2.8)$$

In a medium, the material properties $\mathcal{E}_R \mu_R$ act as a speed scaling factor and give rise to a fundamental quantity in optics, the refractive index n (Equation 2.9).

$$n = \frac{c}{v} = \sqrt{\mathcal{E}_R \mu_R} \quad (2.9)$$

The refractive index describes the change in the speed of light within a material, but another important consequence, as we will briefly see later, is that it also quantifies the change in the direction of propagation of said radiation. Exploiting the relationship between the refractive index and speed of a wave, Equation 2.10 incorporates the relationship between the refractive index, the frequency, and the wavelength of a wave, thus displaying that the refractive index of a material depends on the wavelength of the incoming radiation.

$$\lambda = \frac{c}{n\nu} \quad (2.10)$$

2.3.3 Interference and Diffraction

Continuing with the description of light as a wave, a time-varying electric field propagating along the spatial direction x can be written as in Equation 2.11.

$$E(x, t) = E_0 \cos \left(2\pi\nu t - \frac{2\pi}{\lambda}x + \phi_0 \right) \quad (2.11)$$

in which E_0 represents the amplitude of the wave, $2\pi\nu$ is the angular frequency, ϕ_0 is the phase angle, and the wave vector is encapsulated by the quantity $\frac{2\pi}{\lambda}$. The angular phase ϕ_0 represents the starting point of the oscillation in the cosine function, but more intuitively, the phase indicates the position of a point within the wave period. The phase of radiation is a fundamental property that, if observed carefully, enables the measurement of material properties and the description of phenomena such as interference and diffraction.

Solar light and light emitted by commercially available illumination sources contain waves oscillating at different frequencies (polychromatic) and displaced at different phases. Such radiation is said to be *incoherent*. *Coherent* radiation, on the other hand, contains monochromatic waves tuned at the same wavelength *in phase* with respect to each other. The generation of coherent light (with lasers for instance) marked an important breakthrough in the field of quantum physics and communications.

When radiation propagates, we can define its *wavefront* as the locus of points that share the same phase. From the shape of the wavefront is then possible to retrieve the shape of the wave and possibly the type of source that emitted it. When two or more waves meet, the resulting intensity depends on the way the two wavefronts combine. Such interaction is called *interference* and can be of two extreme types with a continuum of solutions in between. Constructive interference is observed when interacting points belonging to two wavefronts share the same phase, and the resulting intensity is typically larger than the sum of its components. Destructive interference takes place when the points of the wavefront are *out of phase*, meaning that the difference between their phases is π . In this instance, the intensity decreases to zero. The classical two-slits experiment first performed by Young is schematically reported in Figure 2.5a.

Ideally, the fringe pattern observed in Young's double slit experiment should display bright points of constant intensity. However, it is evident that after a few orders the brightness of the maxima points quickly decreases. This is due to the effect that a single slit has on the plane wavefront. According to the Huygens-Fresnel principle, each point of the wavefront is on its own a new source of spher-

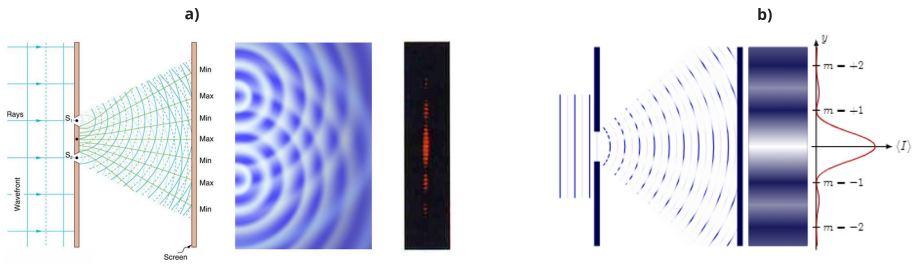


Figure 2.5: **a)** Two-slit experiment displaying the phenomenon of light interference. Image source: University of Central Florida [55] **b)** A diffraction pattern can be obtained by making a light wave interact with a slit with a width comparable to the wavelength of the radiation. Image source: Concepts of Physics [56].

ical waves. When a plane wave encounters a slit with a width comparable to the wavelength of the monochromatic radiation, each point of the slit acts as a secondary source of spherical waves. During *diffraction*, the newly generated spherical waves interact creating the classical diffraction pattern constituted of constructive and destructive interference (Figure 2.5b).

Diffraction represents an important limit in the world of optics and imaging systems. A *diffraction-limited* device such as a lens will produce an image free of all distortions, except diffraction effects that cannot be canceled. In this ideal scenario, two objects can be *resolved* by a lens with a strictly circular diameter D if the viewing angle that separates them is higher than $\sin \theta = \frac{1.22\lambda}{D}$.

2.3.4 Types of interaction

Through the definition of the Maxwell equations and the fundamental wave equation, we have formally defined the refractive index, a fundamental optical property of materials. In the following paragraphs, we will then take a closer look at the possible types of interaction between light and matter. In doing so, it is perhaps convenient to recall the practical example introduced in Section 2.2.2, and focus on the illumination source and the painting that has been selected to conduct an image acquisition campaign.

Refraction

The light that travels from the illumination source is likely to contain many different wavelength components and appear white or yellowish, but for the sake of simplicity, we will assume that it is monochromatic. The refractive index of air is normally $n_0 > 1$ but just slightly and depends on a variety of factors such as the air temperature and altitude (oxygen concentration). A commonly adopted

assumption is to approximate n_0 to 1.

The last assumption follows Fermat's principle of propagation and is at the basis of the most simple description that envisions light as a ray. According to Fermat's principle, light travels from point A to point B following the fastest path, which is a straight line if points A and B are found in the same medium without discontinuities presented by other media. This last statement is quite powerful and it is necessary to stop and delve into its meaning and implications. By *discontinuity*, or *interface* it is intended the junction point between two media with different refractive indexes. In the instance that we have depicted, an interface can be the passage from air to the varnish layer or from the paint layer to the ground. Thus, in the case in which points A and B are found at opposite sides of the interface, Fermat's principle tells us that light travels from A to B following the fastest path in terms of time spent and not distance covered. The change in refractive index that takes place at the interface then *bends* the ray of incoming light.

This phenomenon, called *refraction* is described by Snell-Descartes law (Equation 2.12). The configuration of Snell-Descartes law involves an interface between two materials with different refractive indexes n_1 and n_2 . The surface normal \vec{n} is used as a reference to describe the angle of incidence θ_i and the outgoing angle θ_t .

$$n_1 \sin \theta_i = n_2 \sin \theta_t \quad (2.12)$$

In instances in which $n_2 > n_1$, the outgoing ray *tightens* the normal, while it *opens* the angle otherwise. At normal incidence, i.e. when $\theta_i = 0$, we have that $\theta_t = \theta_i$. The effects of Snell-Descartes law are observable every day when we look outside of thick window glasses, through our spectacles, and even when we swim, but in the context of imaging of a painting we find refractions when the camera objective bends the light rays or when the light coming from the illumination source crosses the various layers of an oil painting, as reported in Figure 2.6.

Reflection

For simplicity, let us assume that the varnish layer is flat, smooth, homogenous, and isotropic (its refractive index does not depend on the observed direction). While an incident ray crosses the interface, a part of it is reflected in a direction that is specular to the incident angle according to Equation 2.13.

$$\theta_i = -\theta_r \quad (2.13)$$

This is known as the reflection law and its demonstration finds a few explanations

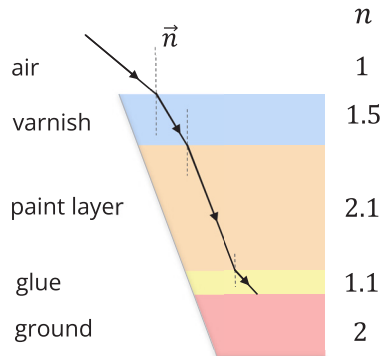


Figure 2.6: Simplified visualization of refraction phenomena taking place within the volume of a painting. In this practical illustration, the layers are assumed to be transparent just to showcase the behavior of light rays when passing through media with different refractive indexes.

related to the interaction of the light wavefront with the electrons of the surface, the development of Maxwell equations, and more recently, the numerical computation of Feynman's integral path which shows how the symmetric direction is the more likely among the infinite possible path directions. A particular case of reflection is enabled when $n_1 > n_2$. By increasing the incidence angle θ_i it is possible to arithmetically find the *critical angle* θ_c for which there is no refracted ray, but only a reflected ray, since $n_1 \sin \theta_c = \pi/2$. This phenomenon is also known as total internal reflection and finds a typical application in the manufacturing of fiber optics.

The division of an incoming ray of light into the paths of reflection and transmission implies that the energy or power of the incoming ray, following the energy conservation principle, is split into two components. The Fresnel coefficients (Equation 2.14 [4]) describe the transmission and reflection of an electromagnetic wave incident on the interface between two media with different refractive indexes n_1 and n_2 . For each component, the coefficients also describe the ratios of the amplitude in two perpendicular polarization components. The s polarization is normal to the plane of incidence, whereas the p polarization is in the plane of incidence. The appropriate combination of s and p components can describe any polarization state, for instance, unpolarized light has an equal amount of them.

$$\begin{aligned}
r_s &= \frac{n_1 \cos \theta_i - n_2 \cos \theta_t}{n_1 \cos \theta_i + n_2 \cos \theta_t} & t_s &= \frac{2n_1 \cos \theta_i}{n_1 \cos \theta_i + n_2 \cos \theta_t} \\
r_p &= \frac{n_1 \cos \theta_t - n_2 \cos \theta_i}{n_1 \cos \theta_t + n_2 \cos \theta_i} & t_p &= \frac{2n_1 \cos \theta_i}{n_2 \cos \theta_i + n_1 \cos \theta_t}
\end{aligned} \tag{2.14}$$

When incident light does not contain a dominant polarization component and is hence unpolarized, the effective reflectance coefficient can be computed as the average between r_s and r_p . A particular phenomenon is observed for partially transparent dielectric materials when θ_r and θ_t are found perpendicular to each other, which in the plane of incidence is translated to $\theta_r + \theta_t = \pi/2$. In this instance, the r_p component of the reflectance goes to 0, thus ascribing all the polarization properties to the r_s component. This is an important result, it tells us that unpolarized light, if incident at a specific angle θ_b can generate a reflection that is completely polarized. The angle θ_b is termed the Brewster angle and can be found for a pair of interfacing media as described in Equation 2.15.

$$\theta_b = \tan^{-1} \left(\frac{n_2}{n_1} \right) \tag{2.15}$$

Figure 2.7 summarizes the reflectance configurations explained in the hypothetical case of an interface between the paint layer and glue in the painting volume (again assuming transparency). In detail, Figure 2.7a displays three instances in which glue is the receiving medium ($n_1 > n_2$): a common reflectance at an angle θ_i , a reflection at the Brewster angle θ_b , and a total reflection at an incidence angle θ_c . Figure 2.7b reports the ratio between the Fresnel coefficients r_p and r_s as a function of the incident angle for a pair of materials with refractive index 1.1 and 2.1 respectively. The Brewster angles are identified at the intersection of the functions with 0, while the critical angle can be identified in the case in which the higher refractive index is established as n_1 .

The most evident products of the interaction of light with an interface at the Brewster angle are specular highlights. Specular highlights represent important visual cues in our everyday life, as we subconsciously use them to determine the direction of illumination and to trigger perceptual reactions such as color constancy. In a painting, their presence is more or less welcomed, depending on the purpose of the imaging campaign. An image of a painting that does not display specular highlights and is thus completely *diffused* might look a bit unnatural and perhaps is not the best for a rendered visualization in a computer graphics context. On the other hand, specular highlights tend to be very bright and therefore can mask

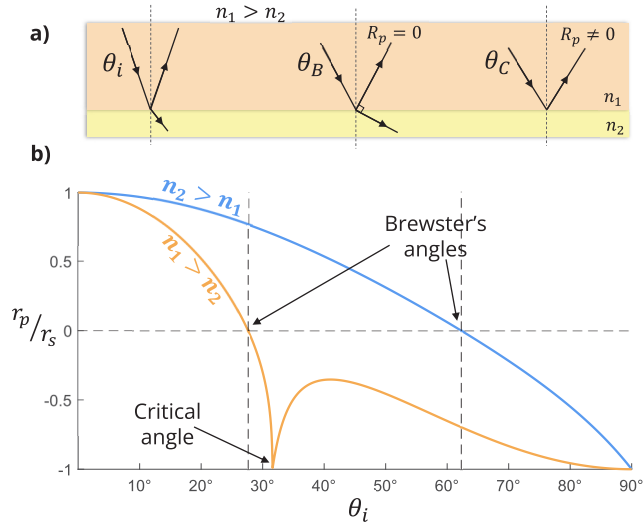


Figure 2.7: (a) Configurations of three characteristic reflections, from left to right: common reflection, interaction at Brewster angle, and total reflection. (b) Ratio r_p/r_s as a function of the incidence angle. When $n_1 > n_2$, it is possible to identify both the Brewster and total reflection angles. Notice that only the real parts of the ratio are reported in this plot.

part of the information carried by the analyzed surface, thus preventing material characterization or an enhanced visualization.

Hardware solutions to limit the presence of specular highlights include the usage of polarization filters in front of the camera objective. In photography, this is sometimes referred to as *cross-polarization*, a practice that boils down to placing a linear polarization filter rotated at an angle that is perpendicular to the direction of polarization of the incoming light. This technique follows the same principles of polarized sunglasses, designed with the first objective of reducing the amount of specular reflection coming from the ocean surface and reaching the eye. Since the polarization of light reflected off water tends to be horizontal, by including a vertically polarized coating (hence the *cross* in the name) the specular highlights do not make it past the spectacles.

Absorption, Transmission, and Fluorescence

To easily display the various instances encountered in the phenomenon of refraction we assumed that the layers of the painting were transparent, but as we easily observe with our eyes, this is not true. Transparent objects like varnish let light through them because of their regular molecular arrangements, whereas opaque

objects possess random arrangements that do not let light pass. The fundamental property that determines if light passes through a medium is absorption. When light propagates within a medium, it transfers its energy to the atoms of the material. Consequently, if light emerges from the medium, its power will be lower than when it entered, following the exponential relationship provided by the Beer-Lambert law (Equation 2.16).

$$\phi = \phi_0 \exp(-\alpha x) \quad (2.16)$$

in which ϕ is the emerging energy flux, ϕ_0 represents the incoming flux, x is the traveled distance within the material or the material thickness, and α is the absorption coefficient. Alternatively, absorption within a material can be described by the absorption length L_α (Equation 2.17), e.g. the value of x for which ϕ becomes 0.

$$L_\alpha = \frac{1}{\alpha} \quad (2.17)$$

In cases in which the thickness of the material is greater than L_α , there will not be any emerging ray on the other side of the medium. The absorption coefficient strongly depends on the material properties and on the wavelength, with the general rule that highly energetic radiation such as X-rays penetrate deeply into materials, as displayed in Figure 2.8. However, when we consider the instance of a paint layer and compare the impact of visible radiation and infrared radiation, it is the latter that penetrates more, despite the fact of possessing lower energy. This is due to the molecular arrangements of pigments that float in the binding medium (typically oil) and the way they interact (or better, do not interact) with the longer wavelengths of infrared radiation, letting it pass more easily. For this reason, infrared radiation is often deployed to analyze paintings in the hope of discovering hidden details such as underdrawing and *pentimenti* that go undetected by the naked eye check.

Some particular materials showcase re-emission phenomena once photons are absorbed. This is due to the electronic structure of the atoms or molecules within the material. When the absorbed energy excites electrons to higher energy levels, they eventually return to their original rest state by emitting a photon with a lower energy than the initially absorbed photon. Fluorescence is thus observed for example when presumed reflected radiation is measured in the visible range (typically in the region of the blue) when UV light is irradiated on a material.

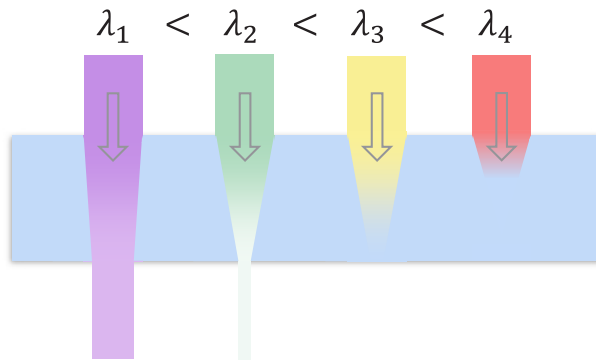


Figure 2.8: Wavelength-dependence of a hypothetical transparent medium that lets through high-energy radiation (the color of the radiation follows the visible light spectrum). The shape of the refracted and transmitted rays has a purely representative purpose to convey the reduction in power of the propagating radiation.

Scattering

The last unrealistic assumption that needs to be removed concerns the topography of the surfaces. So far, we have considered perfectly smooth interfaces, an ideal situation that is found in the real world only in approximated conditions when mirrors and carefully designed optical experimental setups are deployed. All the surfaces that we regularly observe are characterized by various degrees of roughness. To define a surface as smooth, the optical benchmarks state that the size of the rough reliefs must be smaller or comparable to the wavelength of the interacting radiation.

When light interacts with a rough surface, multiple micro-reflections take place simultaneously, thus resulting in the production of infinite reflected rays that are scattered in all directions in the hemisphere of the plane of incidence. Just underneath the surface, light experiences sub-scattering and absorption depending on the material properties. It is this phenomenon that enables the perception of surface colors.

Volume scattering occurs within the medium when light interacts with the particles of the gas or impurities and gets scattered in ways that are quite challenging to predict since the interactions are highly direction-dependent. It is possible to observe this phenomenon in conditions of fog or haze. A schematic representation that highlights the difference between surface and volume scattering is reported in Figure 2.9. Depending on the size of the particles that interact with the radiation, different types of scattering can be observed. When the particle size is comparable

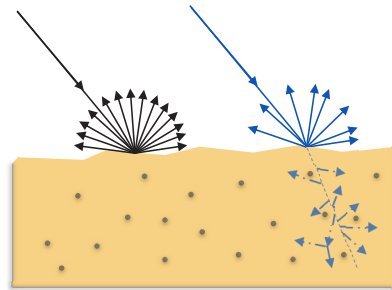


Figure 2.9: Schematization of surface sub-scattering (left) and volume scattering (right).

to the wavelength of the incident radiation, the Mie approximation is typically deployed. Mie scattering does not significantly depend on the wavelength and occurs in broad spectral ranges. In a homogenous medium in which the particles are much smaller than the wavelength of the incident radiation, Rayleigh scattering is observed. In unpolluted air found in the upper layers of the atmosphere, the sky dome appears blue because Rayleigh scattering has a strong inverse dependence on the wavelength (λ^{-4}).

2.3.5 Radiometry

After a short trip into the theoretical aspects of radiation, we have described light as a ray, or as either a swarm of photons or a simple monochromatic wave. In practical instances, however, light is a much more complex entity, with infinite wavefront shapes and combinations of concurring frequencies. Radiometry is the branch of optics that defines the measurement of light-related quantities [7].

To define the fundamental quantities adopted in radiometry, let us set realistic measures for the recurring example of this dissertation, the imaging campaign of two large artifacts during the closing day of a museum. For this section, we can focus on the tapestry. The large textile has a size of 2.25×4.50 m and its bottom edge is hanging at a height of 0.90 m from the floor. A camera is placed on a tripod standing on the floor at a height of 1.50 m and at a distance of 2.85 m from the wall so that the whole fabric is captured in its field of view. This is illustrated in Figure 2.10.

If we are to draw a sphere of radius r centered at the camera location, we will observe that the painting covers a portion of this sphere. From the point of view of the camera objective, the painting is said to *subtend* a *solid angle* Ω that is equal to the ratio between its area and the observation distance. Since the viewing direction is not parallel to the surface normal of the painting (as it can be observed in Figure

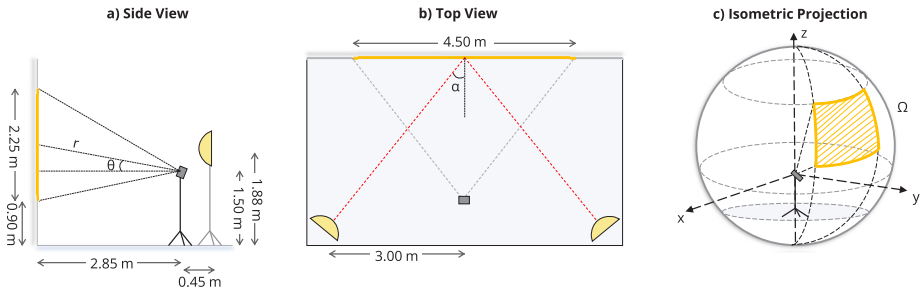


Figure 2.10: Depiction of the practical example in side view (a), top view (b), and isometric projection (c).

2.10a), in the formulation of the solid angle we need to account for the discrepancy angle θ (Equation 2.18).

$$\Omega = \frac{A}{r^2} \cos \theta \quad (2.18)$$

Differently from an ordinary angle, which measures the distance on the unitary circle, a solid angle measures the area covered on the unitary sphere. Similar to the ordinary angle, a solid angle is dimensionless from a mathematical standpoint but is traditionally quantified in steradians [sr]. A sphere will then subtend a solid angle of 4π since its area is $4\pi r^2$ sr, while a hemisphere will logically subtend a solid angle of 2π sr. In our example, the resulting solid angle is $\Omega = 1.26$ sr.

The light that travels from the surface of the painting toward the camera in a given instant has a spectral radiant energy $Q_e(\lambda)$ which is the derivative of the total radiant energy Q_e over the infinitesimal element of the spectrum $d\lambda$. Since the camera shutter remains open to collect photons for a finite integration time dt , we can define the spectral radiant flux (which is the equivalent of power and can be measured in watts [W]) as in Equation 2.19.

$$\Phi_e(\lambda) = \frac{d^2 Q_e(\lambda)}{dt d\lambda} \quad (2.19)$$

The camera shutter, however, can logically capture only those rays that pass through its aperture, which has a finite dimension da so that at precisely the location of the camera shutter we can define a new quantity called spectral irradiance $E(\lambda)$ (Equation 2.20). Spectral exitance can be described with the same formulation, with the only difference being that it is defined from a surface element da located on the surface from which the radiation emerges.

$$E(\lambda) = \frac{d^2\Phi(\lambda)}{da d\lambda} \quad (2.20)$$

Finally, we have to consider that the camera is observing the whole extent of the painting, and therefore the quantity of spectral radiance $L(\lambda)$ (Equation 2.21) allows us to measure the contribution of light over a defined solid angle.

$$L(\lambda) = \frac{d^3\Phi(\lambda)}{da d\lambda \cos \theta d\Omega} \quad (2.21)$$

Spectral radiance is measured with units of $\text{W} \cdot \text{m}^{-2} \cdot \text{nm}^{-1} \cdot \text{sr}^{-1}$, and as we will see in Section 2.5.13, is a fundamental quantity in the realm of imaging spectroscopy that is used as the arrival venue for the conversion from the digital numbers captured by the cameras.

2.3.6 Illumination sources

Temperature is a quantity that is closely related to the agitation of the atoms that constitute matter. As the temperature increases, the particles experience a further acceleration surge, which in turn will produce radiation. Therefore, all bodies with a temperature greater than 0 K (the absolute zero), are radiation emitters. It is not evident for humans, but we can experience our own radiation emission if we hold our hand at about 5 cm from our cheek and we compare the sensation with the other cheek which does not have a held-up hand close by.

The Stefan-Boltzmann law (Equation 2.22) relates the emitted radiance (exitance) M_{bb} of a body and its temperature T through the Stefan-Boltzmann constant σ ($\approx 5.67 \cdot 10^{-8} \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-4}$).

$$M_{bb} = \sigma T^4 \quad (2.22)$$

However, the Stefan-Boltzmann law as reported here applies to a specific class of bodies that emit the maximum possible amount of radiation allowed by their temperature and are therefore fully emitters. In fact, the subscript $_{bb}$ stands for *blackbody*. A blackbody absorbs all incoming radiation, and its color would appear black if exposed to the human eye, although real black bodies are quite a rarity in nature, as only a few materials can approximate this status. A fundamental assumption for a blackbody is that the radiation that it emits does not depend on the angular direction and is therefore constant and homogeneous around a spherical surface. This leads to relating the spectral exitance and the spectral radiance of a blackbody as described in Equation 2.23.

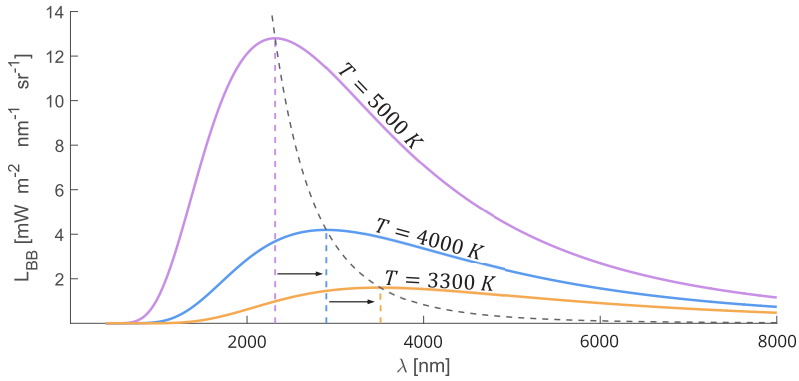


Figure 2.11: Typical spectral radiance curves of blackbodies with varying emitting temperature. The black dashed line that connects the maxima of each curve is obtained by applying Wien's displacement law.

$$M_{bb}(\lambda) = \pi L_{bb}(\lambda) \quad (2.23)$$

The assumption that the radiant energy of blackbodies is quantized allowed the derivation of Planck radiation law (Equation 2.24). in 1901.

$$L_{bb}(\lambda) = \frac{2hc^2}{\lambda^5 \left[\exp\left(\frac{hc}{\lambda kT}\right) - 1 \right]} \quad (2.24)$$

in which k is the Boltzmann's constant ($k \approx 1.38 \cdot 10^{-23} \text{ J} \cdot \text{K}^{-1}$). The derivation of Planck's law allows the visualization of the famous radiance spectra for blackbodies at different temperatures (Figure 2.11). As it is possible to observe, the peak of the spectral curves found at wavelength λ_m progressively shifts towards longer wavelengths as the temperature decreases. The behavior of the peak shift for blackbody emission is characterized by the displacement law formulated by Wien, which is obtained by setting the derivative of the Planck law to zero $\frac{\partial M(\lambda)}{\partial \lambda} = 0$, and is also approximated by the following relationship reported in Equation 2.25.

$$\lambda_m = \frac{2897.8}{T} \mu\text{m} \cdot \text{K} \cdot \text{K}^{-1} \quad (2.25)$$

According to the shapes of the curves displayed in Figure 2.11, it is possible to compare an ideal blackbody spectrum to some available spectra that can be encountered when skimming through a catalog of illumination sources. Figure 2.12

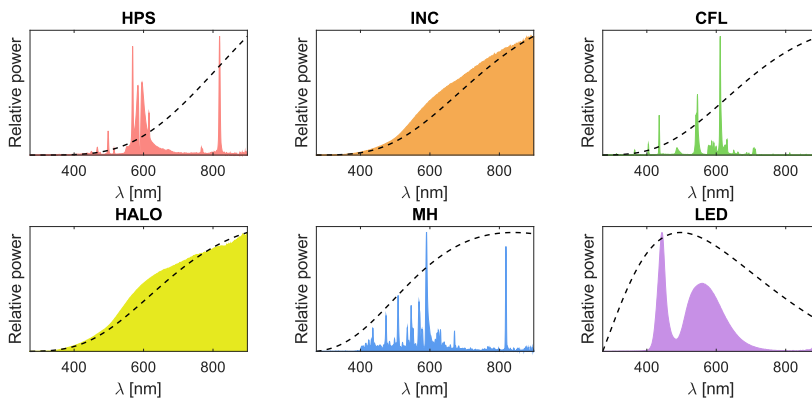


Figure 2.12: Relative spectral power distribution of six commercially available illumination sources. HPS - High-Pressure Sodium (Philips Helios Streetlight). INC - Incandescence (Rona A19 Domestic). CFL - Compact Fluorescence (Globe Twister Domestic). HALO - Halogen (Standard PAR38 Accubeam Domestic). MH - Metal Halide (Philips -Industrial). LED - Light Emitting Diode (Ledtech PAR20 Domestic). The dashed black lines represent the blackbody curves corresponding to the CCT of the source. Data source: Lamp Spectral Power Distribution Database [57].

reports the relative Spectral Power Distribution (SPD) of six commercially available illumination sources.

In a museum context, Light Emitting Diodes (LEDs) are predominantly deployed as main illumination sources because of their high efficiency and low impact on the artworks. On the other hand, their limited spectral range does not enable their use in the context of the acquisition of Infrared information, which can be performed by deploying halogen sources for example. A slow phasing out of halogen lights is ongoing worldwide to contrast inefficient light sources. This has translated into a research effort to develop more sustainable illumination solutions for the capturing of infrared data.

2.3.7 Reflection Models

Computer vision is a branch of computer science that aims to enable computers to interpret images and videos by approximating the capabilities of the human visual system. The main tasks researched and carried out are detection, recognition, segmentation, and scene understanding. Until recently, scene understanding was regarded to be an extremely challenging task, and therefore the ability of humans to resolve these types of problems was considered to be exceptional. Now, with the aid of deep learning, tasks such as image-to-text are carried out with regularity. In the following paragraphs, we will focus on the physical models that have

been developed in an attempt to tackle the recognition and rendering of objects in challenging scenarios. For the sake of keeping the introduction of concepts on track, only relevant models for the appearance and rendering of historical artifacts in the context of technical imaging are considered.

When we observe a scene, whether it is indoor or outdoor, the appearance of the examined objects depends on the physical properties of the materials such as color, translucency, and glossiness, on the topography and texture of the surfaces, on the properties of the illumination (SPD, distance, direction), and on the location of the observer [58, 59]. All these factors concur simultaneously to create an intricate environment in which ambiguous situations are not that rare. The Bi-directional Reflectance Distribution Function (BRDF, Equation 2.26) [60] describes the proportion of intensities between the reflected radiation I_r off a matte surface observed from an angle ω_r compared to the incident radiation I_i coming from an angle ω_i .

$$f(\omega_i, \omega_r, \lambda) = \frac{dL(\omega_r, \lambda)}{dE(\omega_i, \lambda)} = \frac{1}{I_i(\omega_i, \lambda) \cos(\omega_i)} \frac{dI_r(\omega_r, \lambda)}{d\omega_i} \quad (2.26)$$

The BRDF is notably a highly complex model, and cannot be accurately estimated without extensive measurements and expensive setups [61]. However, assuming that a surface is Lambertian can simplify the model by discarding the angular-dependent terms. A Lambertian surface is defined as flat, matte, and diffusive. These three attributes intrinsically contribute to making a Lambertian material isotropic and free from fluorescence phenomena. When the angular terms are discarded, the BRDF coincides with the reflectance (Equation 2.27) of the material under examination.

$$Ref(\lambda) = \frac{I_r(\lambda)}{I_i(\lambda)} \quad (2.27)$$

A classical problem in computer vision is the detection and correction of specular highlights [62–64]. Specular highlights, as we have observed in Section 2.3.4, are generated by the interaction of light with smooth and or glossy materials and usually carry a significant amount of polarized information. A material that presents specular highlights, in the *eyes* of a machine, can be interpreted as two different materials, given the significant difference in recorded signals. Humans, on the other hand, are considered quite good solvers able to use specular reflections at their advantage [65].

The dichromatic reflectance model (Equation 2.28) [66] tries to overcome the lim-

itations imposed by the Lambertian assumption by characterizing the reflected light as the result of the superimposition of a diffuse component from the body of the material and a specular component caused by direct surface reflection. The total intensity I_r after a surface reflection can be modeled by the sum of two intensity components [67]. Each component can also be split into a geometric term $m(\omega_r)$ and a spectral term $c(\lambda)$.

$$I_r(\omega_r, \lambda) = m_b(\omega_r)c_b(\lambda) + m_s(\omega_r)c_s(\lambda) \quad (2.28)$$

The dichromatic model has been initially used to detect specular highlights [66], compute photometric-invariant features [68], and estimate the illuminant of a scene [69] in an attempt of solving color constancy problems. However, the model is based on strong assumptions and is affected by important limitations [70]. A prior segmentation in macro areas of an image is usually needed, the observed media must be opaque, non-homogenous (metals are therefore excluded), and fluorescence-free. Moreover, only a single illumination source with a constant SPD is allowed to exist in the scene. Such a model can be sufficient to characterize flat materials with specular highlights, but it cannot accurately describe the radiance coming from objects that present a complex topography that can lead to the generation of shadows due to the complete or partial occlusion of the illumination source. In such instances, shadowed areas generate ambiguous interpretations, as a naturally dark object cannot be distinguished from a lighter one that receives fewer photons.

The extension and strength of shadows in a scene depend on the relative position between the object and the light source and the physical dimensions of the light source [71]. However, as we will later see when we discuss the different techniques for image acquisition, the distribution of shadows also depends on the way the images are captured (snapshot or scanning). The first differentiation between shadows can be performed between self and cast shadows. Self-shadowing takes place on the surface of the objects that occlude the light source; we can typically observe this on the side of a hill that is not facing the sun. Cast shadows, on the other hand, are observed on objects that do not occlude the illumination. Typically, a shadow is formed by two parts, depending on the degree of occlusion of the light source. When the light source is completely occluded by an obstacle, the observed shadow is an *umbra*, whereas if the occlusion is partial, a *penumbra* is observed. Only umbra can be observed in those instances in which the light source is approximated to a point source, but most real applications deal with extended light sources. The proportion of umbra and penumbra depends on the distance of the shadow from the light source and the height of the relief causing the oc-

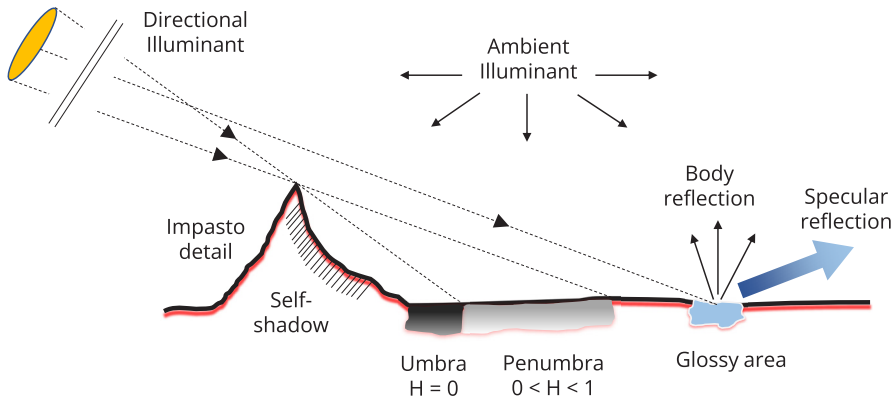


Figure 2.13: Visualization of the shadow formation process and specular reflection in a painting containing impasto and glossy areas. For the sake of representation, the extension and distance of the light source are reported not to scale.

clusion, thus, a varying proportion can only be experienced in cases in which the dimensions of the targets are comparable with said distance. In outdoor scenarios, the light source is found at infinite, and the proportion of umbra and penumbra remains constant regardless of the size of the obstacle.

The Bi-Illuminant Dichromatic Reflectance (BIDR) model (Equation 2.29) [72] ascribes the responsibility of measured signals coming from shadowed areas to the presence of a secondary ambient illumination. The rationale behind the presence of a second illuminant resides in the fact that it is still possible to discern details within the umbra area (provided that the system has a high enough dynamic range, Section 2.5.1). Figure 2.13 schematizes the BIDR and displays the shadow formation process.

While the dominant illuminant is considered to be directional, the ambient illuminant is diffuse and isotropic. The BIDR is based on an extended version of the dichromatic model that includes a constant ambient term, taking inspiration from an earlier shadow detection model [73]. However, in the presence of shadows of varying intensities, this term cannot remain unchanged. The BIDR includes also an occlusion factor $H \in [0, 1]$ for which a value of 0 represents umbra, a value of 1 is considered to be occlusion-free, whereas anything in between is penumbra. Lastly, the two terms $M_{ab}(\lambda)$ and $M_{as}(\lambda)$ represent the body and specular contributions due to the ambient illumination.

$$I(\omega_r, \lambda) = H l_d(\lambda) [m_b(\omega_r)c_b(\lambda) + m_s(\omega_r)c_s(\lambda)] + c_b(\lambda)M_{ab}(\lambda) + c_s(\lambda)M_{as}(\lambda) \quad (2.29)$$

Shadows and specular highlights are visual cues that need to be faithfully reproduced when rendering objects for visualization purposes such as rendering and implementation into virtual/augmented reality environments [74]. On the other hand, when the goal is to physically characterize materials, they can represent challenging obstacles to overcome. Research efforts are still needed to correct areas affected by shadows and specularities towards physical characterization, as for the moment, most of the advances are proposed by deep learning-based techniques, but for visualization purposes [75].

2.4 Detection of Visible and Infrared Radiation

2.4.1 The Photodetector

The pipeline of detection of radiation can be summarized with three fundamental blocks: reception of said radiation, conversion (transduction) of photon energy into an electric current, and analog to digital conversion of the current. The first block of reception, if taken individually, can be carried out by any material that is able to absorb radiation at a given wavelength. However, most materials react passively to the absorption of radiation, and the existence of additional steps in a pipeline is precluded. Semiconductors are a special class of elements that are able to change their conductivity when exposed to illumination. If we take a look at the Periodic table, semiconductors are found spanning from Group III to VI and include elements such as Silicon (Si), Gallium (Ga), Aluminum (Al), and Germanium (Ge).

When observed in their pure state, for example in a lattice of pure Silicon, the valence electrons are tightly bound to the atoms and thus are found in the valence band. If excited, the electrons of the valence band can move to the conduction band, where they are free to move and generate a current. However, the excitation energy required to make this happen must be higher than the band-gap energy. Random thermal excitations can make electrons jump to a higher energy level, leaving a *hole* (a positive charge) behind. However, the subsequent recombination, which generates energy in the form of heat or light, does not allow the sustained flow of current [8]. What we have just described is an *intrinsic* semiconductor: the conduction of current is very limited and the insulation properties are good but not as thorough as those of pure insulators.

What makes semiconductors special is their affinity to the practice of *doping*. Dop-

ing consists of the controlled introduction of elements with a different number of electrons in the valence band to create a concentration of free charges in the lattice. If we stick to the example of Si atoms, we can obtain *n-doping* when an element such as Arsenic (Group V) brings an excess of negative charges (electrons), whereas *p-doping* is obtained when an excess of positive charges (holes) is introduced with elements such as Gallium (Group III). In an *extrinsic* semiconductor, doping allows the movement of free charge carriers, which increases the conductivity of the material. When a voltage is applied, a sustained current flow can be observed.

A particular configuration that allows the flow of current in a unique direction can be obtained with the P-N junction. In the P-N junction, two differently doped semiconductors are brought together. During the recombination phase, the excess electrons that are found on the n-side start flowing toward the p-side, but not all electrons find a hole, or not all holes are filled by an electron. Eventually, the remaining free charges occupy the so-called *depletion region*, an intermediate zone of the P-N junction. The width of the depletion region depends on the concentration of free electrons and holes introduced upon doping, so it is a controllable parameter. What is interesting is that such width can also be controlled at a later stage by varying the voltage at the ends of the P-N junction. Forward bias makes the region more narrow, allowing the current to flow across, while reverse bias widens the region, preventing the flow of current.

When a photon is absorbed by the lattice of the P-N junction, an excited electron moves from the valence to the conduction band and a pair formed by an electron and a hole is created. Electron-hole pairs contribute to the production of a current flow the closer they are found to the depletion region. A photodiode formed by a P-N junction is usually operated in reverse bias (n-side connected to positive pole and p-side connected to negative pole) so that the depletion region is wide and it is statistically more likely that an electron-hole pair is formed near the depletion region.

Main features of a photodetector

Since the main event that triggers the generation of a photocurrent is the absorption of photons, the semiconductor chosen to build a sensor must collect photons efficiently. However, as we have observed in Section 2.3.4, light can be also reflected or transmitted through a material. The quantum efficiency η is probably the most important feature of a detector as it statistically describes the percentage of incoming photons that are effectively converted into free electrons contributing to the photocurrent (constituted of photoelectrons of charge e^-). Since the quantum efficiency is related to light-matter interactions, it is wavelength-dependent, and thus

its examination can help in the evaluation of the suitability of a semiconductive material in a given spectral range.

The generated photocurrent I_{ph} for an incoming photon flux Φ_q can be written as reported in Equation 2.30.

$$I_{ph} = \Phi_q e^- \eta(\lambda) \quad (2.30)$$

This law assumes that the response in current is linear with respect to the incoming flux, which is true to some extent when we consider most sensors deployed nowadays. Luckily for us, the human eye is a nonlinear sensor that has evolved to protect the brain from the generation of currents that would be too strong to sustain.

The photocurrent for a photon flux spanning a wide spectral interval $\Phi_\lambda(\lambda)$ can be expressed as in Equation 2.31.

$$I_{ph} = \int_{\lambda_1}^{\lambda_2} R(\lambda) \Phi_\lambda(\lambda) d\lambda \quad (2.31)$$

In a classical example, the detector is envisioned as a bucket in a constant rain of photon drops. The bucket collects photons during an integration time t_i , and the measurement is repeated N times. Given the random nature of the collection process, it is likely that each reading will be different from each other. Indeed, the measurement fluctuation, denominated photon noise, follows a Poisson distribution in which the average count of photons in a time interval is equal to its variance. Equation 2.32 describes the variance of the photocurrent i_s^2 expressed as a power.

$$i_s^2 = 2e^- I_{ph} \Delta f \quad (2.32)$$

In which Δf is the bandwidth indicating the range of temporal frequencies that can be captured following the Shannon theorem of sampling, thus relating it to the integration time: $\Delta f = \frac{1}{2t_i}$.

The photocurrent can be estimated a priori but not measured directly, as it cannot be distinguished from other currents that concur in the generation of the reading measurement. The dark current I_{dc} is an ever-present current contribution in a sensor that depends on the thermal agitation of the atoms in the lattice. It can be considered as a baseline error and in fact, its measurement is usually carried out in

conditions of total darkness (hence the name). Thermal agitation and the random crossing of the depletion region by free charges follow a Poisson distribution as well. Thus, the shot noise associated with the dark current $i_{s,dc}^2$ can be expressed as in Equation 2.33.

$$i_{s,dc}^2 = 2e^- I_{dc} \Delta f \quad (2.33)$$

Another type of noise generated by thermal energy is the so-called Johnson noise i_j^2 (Equation 2.34). In this instance, the Brownian-like motion of electrons present in all types of conductive elements generates a current that is proportional to the temperature and to the approximated resistance of the detector R_d .

$$i_j^2 = \frac{4kT\Delta f}{R_d} \quad (2.34)$$

Background noise, telegraph noise, reset noise, readout noise, and 1/f noise all contribute to the formation of the total noise in a sensor. However, shot noise and Johnson noise are usually the ones that are more indicated to describe the behavior of a detector, as their typical contributions dominate the others [6]. The Signal-to-Noise Ratio (SNR) plays a crucial role in understanding the performance of a detector by quantifying the relative strengths of the useful signal and the unwanted noise. In many cases, the SNR serves as a primary health check for characterizing detectors, as it provides a clear and concise measure of sensitivity, signal quality, and overall reliability. A high SNR indicates that the detector can reliably detect weak signals and provide accurate measurements, while a low SNR implies that the noise dominates the output. Detectors can also be characterized by their Noise Equivalent Power, which is the power of the smallest detectable signal, assumed to provide a SNR equal to 1.

2.4.2 Materials for Detection

The most used semiconductor for the detection of radiation between 380 nm and 780 nm is by far Silicon. Silicon defined and continues to define the Contemporary Age in which the world is transitioning to be digitally dominated. Photodetectors are only one of the several devices that can be manufactured with Silicon, as most electronic components deploy this precious element whose extraction is regulated by international laws for sustainable reasons and as an economic safety net [76].

Pure Silicon is rarely found on Earth and is characterized by a shiny appearance when exposed to standard conditions of temperature and pressure. Its macro-crystalline structure provides mechanical resistance and a high melting point of

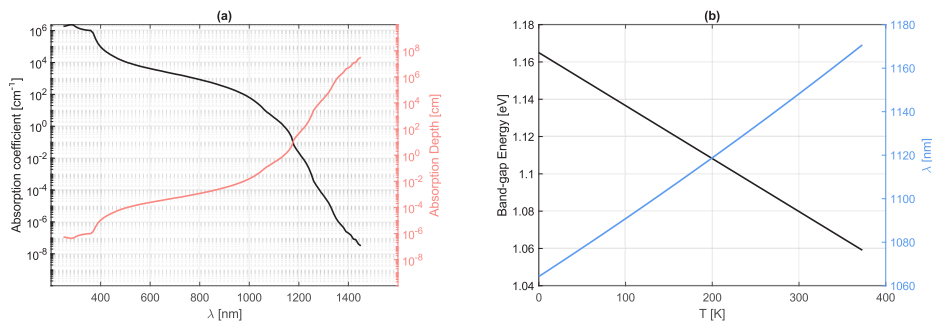


Figure 2.14: (a): Absorption coefficient and absorption depth of Silicon at a temperature $T = 300$ K. Data source: PV Education [77]. (b): Energy band-gap and corresponding wavelength as a function of temperature. The formula used for computing the band-gap is $W_g = 1.165 - 2.84T \cdot 10^{-4}$ [78].

more than 1100°C . It must be pointed out that in its pure state, the properties of Silicon are not ideal for the absorption of visible radiation, as it is possible to observe in Figure 2.14a. Moreover, Silicon works at room temperature as an insulator, since its electrical resistance increases with temperature. Being a good insulator and not a good absorber are two features that so far do not make Silicon an appealing material for the detection of light. Figure 2.14b reports the band-gap energy (expressed in electron volts) and the corresponding cutoff wavelengths as a function of temperature. We notice that the spanned spectral range is found at wavelengths in the NIR, meaning that the more energetic visible radiation can easily excite electrons to the conduction band, while at the same time, not a lot of cooling is required to operate in the most efficient temperature zones.

Doping is a particularly effective practice for Silicon thanks to its electronic arrangement in the valence band. Elements of Group IV and V are indeed suitable for the generation of P-N junctions, as they can draw dopant elements belonging to Groups III-V and IV-V. Another advantage of doping is the resulting decrease in resistivity (and hence an additional increase in conductivity).

As we have observed, the band-gap energy defines the wavelength cutoff for which an incoming photon stops exciting a valence electron. As we observe in Figure 2.14b, the usage of Silicon for the detection of IR radiation (for example at 1500 nm) is out of question since the operating temperature would have to be so high that would eventually damage the electronics. Higher cutoff wavelengths (and lower band-gap energies) can be achieved by using a compound of semiconductive materials. However, not all combinations can generate an efficient new detector. When structures of different materials are merged, their intrinsic geometries need

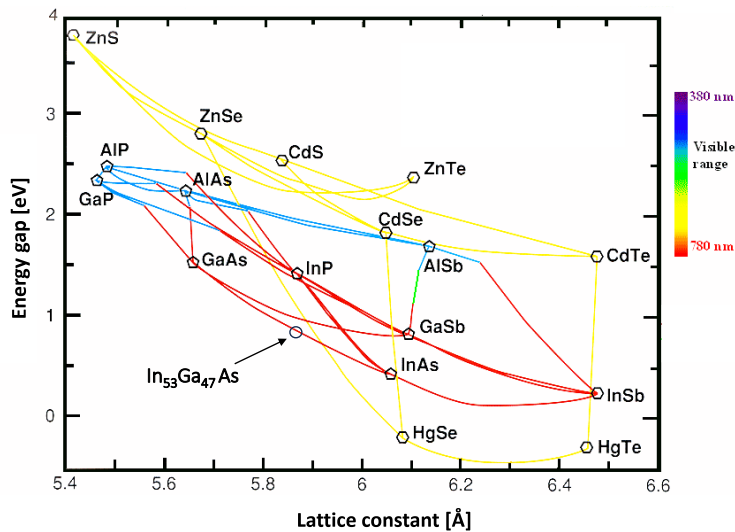


Figure 2.15: Lattice constant (expressed in Armstrong [\AA]) and energy bandgap of commonly used semiconductor compounds. Image source: Prof. Helmut Föll, University of Kiel; Faculty of Engineering [79].

to match. In particular, the parameter that requires attention is the lattice constant, a measure of the intra-distance of atoms in the crystalline structure. Figure 2.15 displays the lattice constant and band gap of commonly used semiconductors, highlighting the binary semiconductors as those compounds that are the result of connecting two materials with an almost vertical line. If the lattice constants differ, the repercussions on the resulting crystalline structure can take the form of strong mechanical strains and so-called dangling bonds. If a chemical bond remains hanging, the likelihood of Shockley-Read generation processes increases, with electrons that are excited to the conduction band and consequently release their energy as heat.

Binary semiconductors formed by the combination of two semiconductive materials, typically feature Group IV symmetry, as the constituting members usually belong to Groups III-V or II-VI. This ensures that the resulting material will have the same number of valence electrons as an atom of Group IV, which is a desirable property when it comes to doping.

Buffer layers of semiconductive materials can be used as a workaround to manufacture photodetectors such as Indium Gallium Arsenide (InGaAs). The lattice constants of InAs and GaAs do not allow the growth of a homogeneous crystal structure, but the introduction of an intermediate layer of Indium Phosphate (InP)

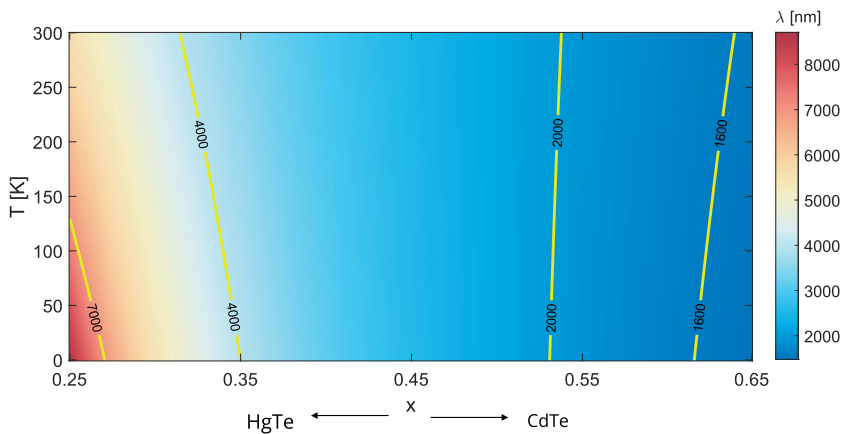


Figure 2.16: Cutoff wavelength as a function of operating temperature and concentration of Cadmium Telluride (CdTe) in a ternary compound of Mercury Cadmium Telluride expressed as $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$.

can facilitate this process. However, the relative concentrations of InAs and GaAs need to be carefully tuned to obtain a lattice matching with InP. This is usually achieved using a $\text{In}_{53}\text{Ga}_{47}\text{As}$ compound. Changing the concentrations of the components also affects the bandgap energy and the cutoff wavelength, as it is possible to observe from the lines that connect the various materials in Figure 2.15.

The operating temperature plays a major role in the detection of IR radiation since it significantly affects the bandgap energy of semiconductor materials and their quantum efficiency. Let us consider the case of the IR sensor Mercury Cadmium Telluride, in which the concentrations of Mercury Telluride and Cadmium Telluride are tuned to generate the compound $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$. Figure 2.16 illustrates how the cutoff wavelength is affected by the operating temperature and by the concentration of Cadmium Telluride [80]. We can observe that in order to sense radiation at wavelengths $\lambda_c < 7000$ nm, a cooling system that brings the operating temperature $T < 150$ K (equivalent to -123 °C) must be deployed, while at the same time, there exists a constraint on the relative abundances of HgTe and CdTe. This is a common feature for detectors of infrared radiation, and one of the main reasons for generally more expensive costs on the market.

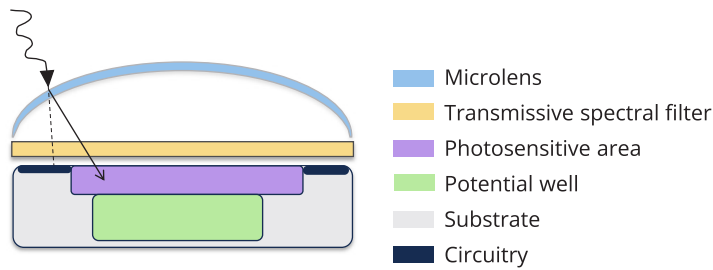


Figure 2.17: Simplified cross-section of a generic pixel highlighting the main elements.

2.5 Toward Reflectance Imaging Spectroscopy

2.5.1 Imaging Principles

Photodetectors exploit the characteristics of semiconductor materials to absorb radiation and convert it into electrical signals. By finding a way of counting the generated electrons, it is possible to effectively measure the amount of incoming radiation, knowing that the two quantities are directly proportional. This measurement, or at least part of it, is carried out in the building block of the image sensor, the pixel.

A pixel (from picture element) can be defined as the assembling of a photosensitive material and the accessories that facilitate the collection of radiation, its selection, and the electronic circuitry that helps in the reading of the measurement [81]. Figure 2.17 illustrates the cross-section of a generic pixel in which the fundamental elements are highlighted.

The first defining parameter for pixels is the fill factor, the proportion of photosensitive area to the total surface of the pixel (pixel size) [82]. Fill factors are generally lower than 100% because of the presence of circuitual elements like wires that do not contribute to the absorption of photons. The efficiency of the light collection can then be improved by adding a microlens that focuses the incoming radiation on the central sensitive area of the pixel [83].

As photons are absorbed in the bulk of the sensitive material during the integration time, the generated electrons flowing across the depletion region are collected in a capacitor (potential well), in which an electric field is able to momentarily *trap* them. The capacity of the well, also called charge capacity, is a parameter that defines the saturation level, or the upper limit for the number of electrons that can be stored. There is then an intrinsic relationship between the saturation level, the brightness of the imaged object, and the integration time. As we have already

discussed in Section 2.4.1, even in conditions of total darkness there will be a flux of electrons across the depletion region. The dark current sets indeed a lower limit for the number of collectable electrons. We can then define the dynamic range of a pixel as the ratio between the saturation level and the dark current expressed as numbers of electrons [84].

Sometimes, a *hot* pixel can act as a saturated pixel even in conditions of low light, because the level of dark current that affects it is anomalously high [85]. *Dead* pixels display a constant response regardless of the incoming radiation level, whereas *clipping* pixels tend to show a saturation behavior without actually reaching the saturation level dictated by the charge capacity.

Pixels can be connected to form stripes and matrices to effectively become image sensors. An image can be indeed defined as an assembling of pixels in which the local variation in signal output defines certain patterns or a distinguishable scene. When pixels are connected, the *pixel pitch* is defined as the center-to-center distance [86]. The pixel pitch is usually larger than the pixel size as gaps typically exist between pixels. The pixel proximity is a good solution to have compact image sensors, but it comes at the expense of possible pixel cross-talk [87] and blooming [88], an effect in which the charges of a saturated pixel overflow to the neighboring ones (horizontal blooming and smearing in the vertical direction). The deployment of the microlens can however mitigate pixel cross-talk.

For the capturing of images, two main technologies are deployed: Charge-Coupled Device (Charge-Coupled Device) [89] and Complementary Metal-Oxide Semiconductor (CMOS) [90]. The main difference between the two configurations resides in how the collected electrons are read out from the potential wells [91]. The pixels of a row in a CCD are connected to perform a sequential reading in which small electric fields are progressively applied so that each potential well discharges its electrons onto the next empty well until all elements in the row are emptied and the measurement is completed. This methodology requires the assistance of a synchronizing clock and reading errors can take place [92]. The reading of a CMOS is less prone to errors because each pixel features a dedicated system for the individual reading of information. Despite the fact that the additional circuitry tends to decrease the fill factor, CMOS sensors are more widespread when imaging technologies are deployed to perform accurate metrology.

Regardless of what technology is adopted to measure the incoming radiation, when the analog current is converted to a binary digital signal a quantization error will be introduced [93]. Quantization is only one of many sources of electronic noise [94], which can be limited by reducing vibrations and implementing better insulation and shielding solutions. At times, the electronic noise sources generate currents

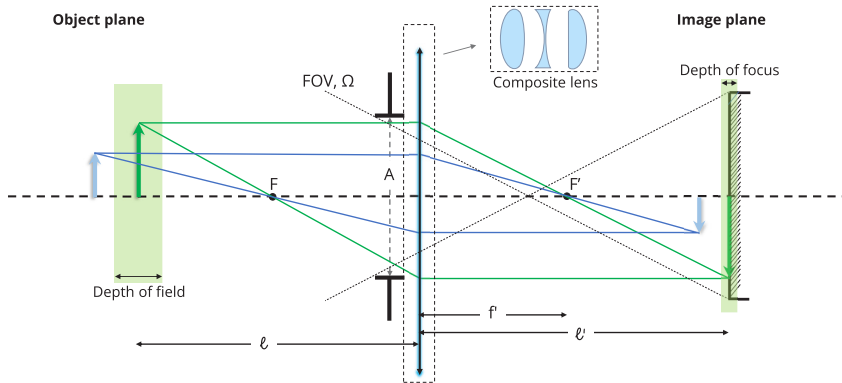


Figure 2.18: Simplification of the optical system within a DSLR camera reduced to a single thin positive lens. The object plane is found to the left of the lens, whereas the image plane is found to its right. If the image sensor is placed in correspondence with the plane where the image of the green object is formed, then all objects found at a distance l (with a tolerance given by the depth of field) from the lens will be images sharply. In this instance, the blue object will appear out of focus on the image detector because found outside of the depth of field.

that are in the same order as the photocurrent, thus resulting in a drastic reduction of the SNR. To avoid this problem, the photocurrent is typically pre-amplified before being converted into a digital signal [95].

2.5.2 The camera as a simple optical system

The microlenses that focus the incoming radiation on the individual pixels are only the last (if we consider the direction of the light path) optical components that we find in a digital camera. If we exclude the optical components designed to display the live view of what is currently captured, the essential optics of a digital camera (Digital-Single Lens Reflex - DSLR) is formed by a composite lens, typically constituted by two or more lenses with different goals of collimation, focusing and zoom [96]. For the sake of simplicity, we can introduce the fundamental parameters of an optical system by considering a composite lens system as a unique resulting lens. A schematization of the fundamental optical distances and terms of a simplified DSLR camera is reported in Figure 2.18.

In a digital camera, the incoming photons are restricted to interact with the lens only in areas close to the main optical axes. The main reason for deploying a pupil (or aperture) in front of the lens is to avoid the refraction of light in areas away from the main optical axes, which are more affected by distortions and aberrations [97]. The second reason is to limit the amount of photons that arrive on the sensor,

thus avoiding saturation. The aperture of a camera can be controlled by tuning the f-number ($f\#$).

To simplify even further, we can consider the composite lens of a camera as a *thin* lens [4]. In this way, we can have an ideal situation in which the *principal planes* coincide with the lens surface. We can then define the *second* (because it is found after the lens) *focal point* as the point on the main optical axes crossed by refracted rays coming from an infinite distance (parallel to the optical axes). Similarly, we define the *first focal point* as the point on the main optical axes that produces a refracted ray parallel to the optical axes. Both points associate focal planes perpendicular to the axes. The focal length is an important parameter of lenses, as it can give information regarding the *power* of a lens expressed in diopters, its magnification, and more importantly, it tells us where to place the image sensor in order to have sharp images. It might seem that sharp images of objects can be obtained only if said objects are found at a defined distance. However, a certain tolerance is inherently present. Objects found within the depth of field, an area around the optimal acquisition distance, are still imaged sharply [98]. The depth of field of a system is inversely proportional to the square of the focal length and the diameter of the aperture pupil.

The angular field of view of a camera is then defined by taking into account the size of the image sensor, the focal distance, and the aperture of the system. A related quantity to the angular field of view is the *étendue* [99], which is proportional to the product of the diameter of the pupil and the solid angle intended as the spread of light that can be collected.

2.5.3 Quantities related to image sharpness

Having already introduced the term *image sharpness*, let us delve more into the theory of it while introducing more advanced concepts related to imaging systems.

When we are asked to judge whether an image is sharp or not, we tend to shift our attention to some particular features and neglect others [100, 101]. Areas of an image that depict uniform scenes such as the clear sky or other artificial smooth materials usually are not of interest, because they contain low spatial frequencies. On the other hand, complex scenes containing high spatial frequency can help us in taking the decision regarding image sharpness. Neglecting for a moment the human sensitivity to various frequency contents, we can define a sharp image as that image that faithfully represents simple elements such as lines and points.

The Point Spread Function (PSF) is a measure adopted to quantify the sharpness of an image system. We can say that during the image formation process, a point in the real-world scene is transferred onto the sensor by a convolution with the PSF

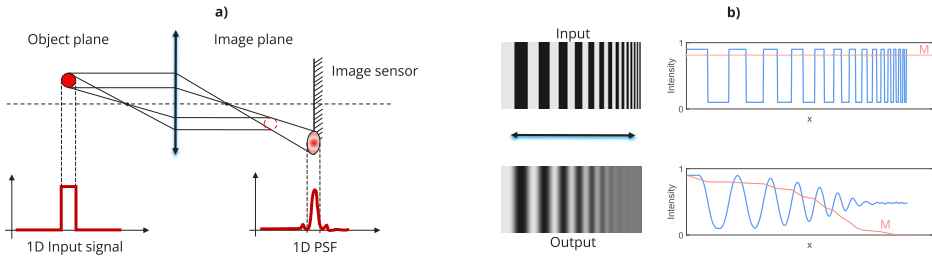


Figure 2.19: **a)** Representation of the PSF produced by a lens. **b)** Typical camera response to stimuli with fixed contrast and varying spatial frequency. The effect of the contrast loss (in red in the plots) is magnified for visualization purposes.

[102]. A punctual light source (usually a laser) is imaged, and the intensity pattern on the image sensor is recorded. The ideal recorded image should represent a circle of uniform intensity with a radius proportional to the magnification of the optical system. Real PSFs, however, are diffraction-limited, meaning that the uniform intensity in the circle cannot be achieved. Instead, a decreasing intensity is usually observed as we move away from the center of the circle. Optical aberrations and distortions also tend to deform the shape of the circle [97]. Astigmatism, a second-order aberration, is caused by a different curvature on specific axes of the lens, thus leading to PSFs with elliptical shapes. Coma, another second-order aberration, makes the PSFs look like a comet (hence the name), exhibiting a trace-like shadow. A schematic representation of PSF is reported in Figure 2.19a.

A fundamental quantity of an imaging system is the spatial resolution, usually intended as the highest spatial frequency that a system is able to resolve. The measurement of this quantity is usually related to an extension of the PSF, the Line Spread Function (LSF). The typical inputs for the measurement are reported in Figure 2.19b, and the resolution is defined as the number of line pairs (black and white) per unit of distance (usually mm). If the LSF produces blurry lines, there would be a spatial frequency for which a pair of black-white lines will be indistinguishable, and therefore the resulting signal will be gray (a mixture of white and black). A factor that can influence at which spatial resolution we encounter the breaking point is the contrast between a pair of lines, a measure related to the ratio of the luminances. The Michelson contrast (Equation 2.35) [103] is usually adopted.

$$M = \frac{I_{MAX} - I_{min}}{I_{MAX} + I_{min}} \quad (2.35)$$

The Modulation Transfer Function (MTF) encapsulates how a system reacts to

varying contrasts and resolutions and is thus a good figure of merit to compare systems for specific applications [104].

Another measure of spatial resolution that sometimes is exchanged, probably erroneously, with the resolution of spatial frequency, is the Ground Sampling Distance (GSD). The GSD represents the size of the minimum spatial detail that can be resolved by the system, as it is a translation of the pixel size into real-world distances. The factors that influence the GSD are the acquisition distance, the angular field of view, and the number of pixels on the sensor [105].

2.5.4 Color and multispectral imaging

Color, along with shape, is probably the main feature we use to communicate the appearance of objects [106]. Yet, unlike shape, it is arguable that color does not exist [107], as it is the product of our perception. Without getting into thorny philosophical discussions, it is possible to state that shape features do depend on perception as well, but they can be unequivocally measured, whereas it is possible to measure physical correlates of color perception [108], but the interpretation of this measurement might be subject to population variability, and its communication is made challenging by factors such as cultural and linguistic background [109, 110].

Human visual perception starts with light entering the pupil, being refracted by the cornea and the lens, traveling through the vitreous humor in the vitreous chamber, and landing on the retina [111] (Figure 2.20a). The retina is the image sensor of the eye and its pixels, the photoreceptors, cover it heterogeneously [112]. No photoreceptors are found in the area of the blind spot, where the optical nerve starts its path toward higher levels of perception that are not treated in this dissertation. While within a classical image sensor all pixels are devoted to the reception of the same type of radiation, the photoreceptors found on the human retina have a precise division of tasks. Rods are wide-band photoreceptors that are sensitive to low levels of light, and thus experience saturation in conditions of daylight. When only rods *fire* signals, no color can be perceived and visual perception is said to be in the scotopic domain. Cones, on the other hand, are photoreceptors with a narrower spectral sensitivity than rods and do not fire in conditions of scotopic vision. Three types of cones, named after the range of wavelength they respond to, are found on the retina. S-cones (short) are sensitive to wavelengths in the range 380-510 nm, M-cones (medium) fire when exposed to wavelength from 420 to 650 nm, whereas L-cones (long) are sensitive to the interval 470-780 nm. Photopic vision is enabled when the intensity of the incoming light is sufficiently high and all rods are completely saturated. Mesopic vision can be defined as that transition area in which mixed activities of rods and cones coexist.

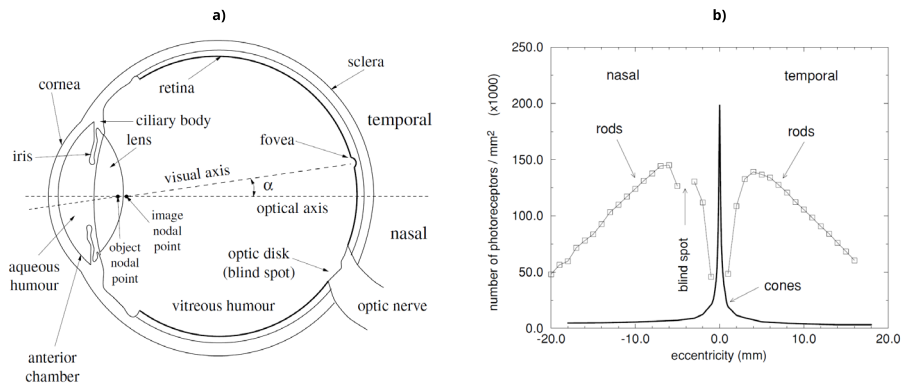


Figure 2.20: **a)** Horizontal section of the human eye with main structures highlighted. **b)** Rods and cones average distribution expressed as a function of angular distance from the fovea. Images source: [112].

As already mentioned, the distribution of photoreceptors is not homogenous on the retina (Figure 2.20b). The proportion of rods to cones is approximately 20:1, while within the cones, L-type and M-type dominate with ranging proportions among a population, and S-type represents only 2% of the cones. The fovea, which subtends an angle of 2° and is found displaced by an angle of 4° with respect to the main optical axes of the eye in the temporal direction, is the area in which humans have the highest concentration of cones and the lowest concentration of rods, which rapidly decreases to zero in the central part, denominated foveola. The fovea is therefore the most exploited area of the retina, as proper color vision is not enabled outside of it. Moreover, the best spatial and temporal resolutions are also achieved in this little region [113]. The human visual system is a peculiar imaging system that simultaneously shows inefficient and efficient properties. It is inefficient because only a very small part of the retina is deployed, but on the other hand, it is extremely efficient because it is able to provide a satisfactory vision despite its poor spatial resolution and thanks to the abilities of its *processor* to use temporal integration and spatiotemporal memory.

The perception of color, however, is not limited to the firing responses of the different types of cones, as more processes occur in the higher levels of vision in the brain [114, 115]. To simplify the problem, the human visual system is considered as a whole, and its spectral sensitivity functions were experimentally measured for the first time in 1931 [116]. In the color-matching experiment, visual observers were asked to match the appearance of monochromatic stimuli by tuning three dials containing three reference primary monochromatic lights (e.g. blue at 435.8 nm, green at 546.1 nm, and red at 700 nm). The resulting Color Matching

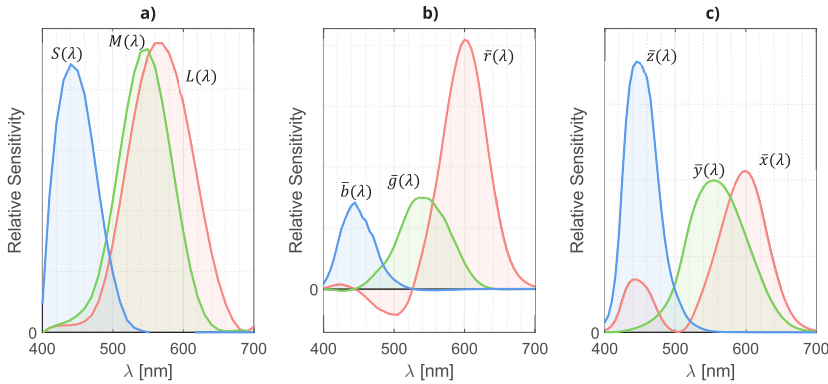


Figure 2.21: **a)** Normalized cones responses. **b)** Color-matching functions obtained from the original 1931 color-matching experiment. **c)** XYZ color matching functions in which $\bar{z}(\lambda)$ mimics the response of $S(\lambda)$ in **a)** and $\bar{y}(\lambda)$ matches $V(\lambda)$.

Functions (CMF) are reported in Figure 2.21b. The main feature of these functions concerns the negativity interval of the matching function related to the red primary $\bar{r}(\lambda)$. Indeed, the monochromatic stimuli in this spectral range could not be matched unless a certain amount of red monochromatic light was added to the input, hence the negative values. The RGB color space is then defined by the deployed amounts of the three primaries. Equation 2.36 shows how to compute the tristimulus RGB values, also called chromaticities coordinates.

$$[r, g, b] = \frac{[R, G, B]}{R + G + B} \quad (2.36)$$

in which R , G , and B are in this case the amount of each component deployed in the color-matching experiment and $r + g + b = 1$.

However, the RGB color space is not a convenient environment when it comes to the direct relationship with the human visual system, and the presence of negative values in $\bar{r}(\lambda)$ represented an additional complexity in the early computations of the 1930s [117]. The color-matching functions were then modified by converting the primaries, and the XYZ color space was defined. The 2° Colorimetric Observer functions reported in Figure 2.21c represent still today the standard proposed by the CIE (Commission Internationale de l'Éclairage). Here, the indication of 2° is deployed to describe the viewing conditions of the color-matching experiment, in which observers could only use the fovea area, hence the 2°.

The conversion of the $[\bar{r}(\lambda), \bar{g}(\lambda), \bar{b}(\lambda)]$ functions was performed so that $\bar{z}(\lambda)$ matches the fundamental S-cones response (confront with Figure 2.21a), while

$\bar{y}(\lambda)$ matches the spectral luminous efficiency function $V(\lambda)$, which represents the average spectral sensitivity to light stimuli of varying wavelengths. The matching with $V(\lambda)$ enables the XYZ color space to be directly related to a measurable physical quantity such as luminance.

The tristimulus values of XYZ (Equation 2.37) can be obtained by integrating the product of the observed object with the SPD of the illuminant and the CMF (constituted by $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$) over the range of wavelengths in the visible range. The value of the constant k is selected so that a white object with constant $Ref(\lambda) = 1$ yields a luminance value $Y = 100$.

$$[\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)] = k \int_{400\text{nm}}^{700\text{nm}} Ref(\lambda) \cdot SPD(\lambda) \cdot CMF(\lambda) d\lambda \quad (2.37)$$

To keep this introduction compact, we cannot explore the various implications of the XYZ color space and its relationship to other color spaces that deal with different aspects of color imaging. Within the scope of this thesis, to describe how a camera *sees* and records color, the XYZ and RGB color spaces are more than sufficient. The XYZ color space represents a good connection environment between physical measurements of visible radiation and human perception, and it is therefore used as a *landing* space in many devices like cameras that capture colors. Readers interested in a better understanding of color science could refer to [118, 119].

In order to *see* colors like a human, a camera must be equipped with a set of sensitivity functions that mimic those of the human visual system. This is obtained by applying spectral filters (typically wide band-pass filters) in front of the photosensitive area of the pixels (confront with Figure 2.17). However, only one spectral filter can be applied on a pixel, thus resulting in a single response value. To obtain a 2D color representation of a scene, the most popular solution is to apply the spectral filters in a prefixed pattern. The most famous pattern, depicted in Figure 2.22a, is the Bayer pattern [120]. Here, the green filters cover the image sensor twice more densely than the red and blue filters because of the finer spatial frequencies that can be sampled with this spectral filter and also in an attempt to emulate the response of the human visual system to wavelengths that would be categorized as *green*, in which both M-type and L-type cones fire at high rates. From the spectral filter array, three image planes with missing values can be extracted and filled by interpolation in a process called demosaicing [121] (Figure 2.22b-c).

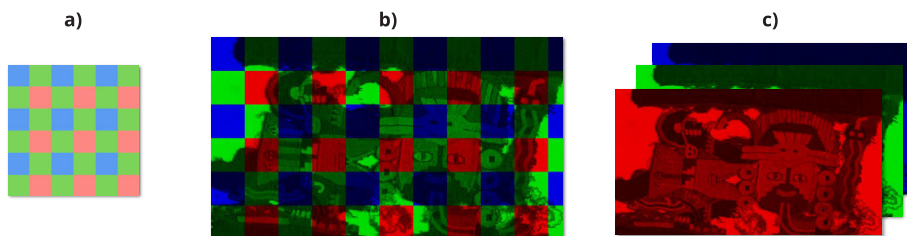


Figure 2.22: a) Typical Bayer spectral filter array b) Laying a Bayer pattern (simplified version reported for visualization) over an image creates three images with gaps. c) Demosaicing creates three distinct image planes with no gaps by interpolation

Multispectral imaging

When a camera is calibrated, the raw values are directly related to physical quantities, and thus it is possible to say that the effect of the camera sensitivity functions is discarded [122]. After demosaicing, a pixel is represented by three values that ideally lay on the irradiance spectrum produced by the interaction of the SPD of the illumination and the reflectance of the observed object. Although a series of spectral characterization methods [123] and spectral reconstruction from RGB [124] exist, a way to retrieve more accurate spectra, is to add more color channels.

The term Multispectral Imaging (MSI) has been used to indicate those imaging systems that produce an image with more than the three classic R, G, and B planes [125]. In the following paragraphs, we will briefly look at various ways in which such images can be obtained.

The three typical methodologies (illustrated in Figure 2.23) to increase the number of channels in a monochrome or standard RGB camera concern the sequential application of color filters, adopting different illumination sources, and the design of more intricate color filter array patterns [126]. Combining the different methodologies is also an option, but it will not be treated for the sake of brevity.

A monochrome camera typically only includes an IR filter to block unnecessary radiation and does not deploy a Bayer color filter array. Not deploying micro-filters provides advantages in terms of quantum efficiency [129], as the removal of optical layers decreases the likelihood of unwanted reflections. However, monochrome responses are not sought after in the field of color imaging. A filter wheel [130] is a technology that deploys a certain number (typically from six to twelve) of wide-band spectral filters that rotate in front of the camera to select the portion of the spectrum to analyze. This methodology is clearly limited to static objects and shows clear limitations on the number of filters that can be mounted on the

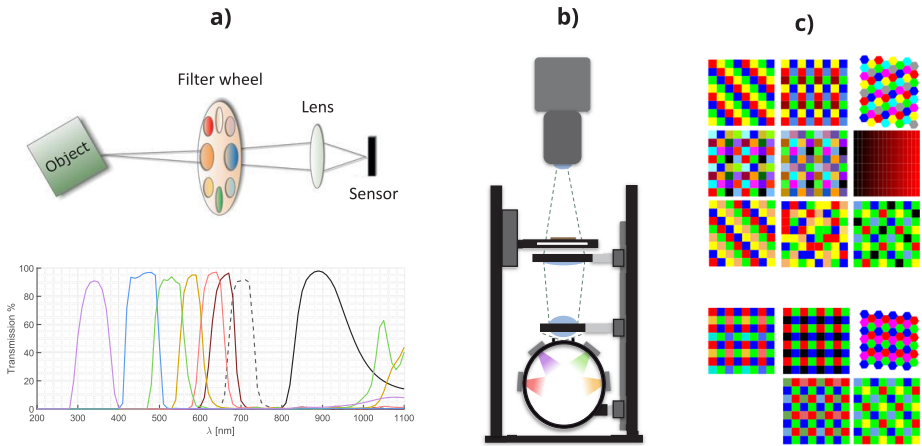


Figure 2.23: **a)** Filter wheel and typical transmission spectra of wide band-pass filters from UV to NIR range. Upper figure source: [126]. Data source for the plot: Midwest Optical Systems, INC [127]. **b)** Schematization of a MSI system for the scanning of film in transmission mode deploying multiple LED sources in an integrating sphere. Figure adapted from [128] with permission from authors. **c)** Typical SFA patterns showcasing different shapes and numbers of deployed filters. Figure source: [126]

wheel, but allows to achievement of good levels of image quality by carefully selecting individual integration times for each filter. Similarly, a monochrome or RGB camera can be deployed in combination with various illumination sources with more or less narrow emission spectra, depending on the application [128, 131].

Having to rotate between filters and illumination sources can sometimes represent an unsurmountable obstacle in an image capture pipeline. Multispectral high-speed quality control and video are challenging to put into practice with sequential acquisition unless snapshot solutions like Spectral Filter Array (SFA) [126], closely related to the already discussed color filter array, can be deployed. The main research questions that were asked during the development of such technologies concerned the maximum number of applicable filters, their arrangement on the SFA, and the interpolation methodology.

Depending on the number of implemented channels and their inherent properties such as bandwidth and relative distribution, a MSI system allows its user to reconstruct a reflectance spectrum more easily than a characterized color camera [132]. This property, in conjunction with the relative ease of designing an imaging set-up and the relatively inexpensive costs, made MSI an attractive imaging technique for

the study of CH artifacts all around the world [133], even in smaller institutions. MSI can be used as a preliminary observation technique, especially when one of the deployed channels works in the UV range, thus allowing the investigation of the varnish layer of paintings [134].

2.5.5 Spectroscopy Principles

The word spectrum has different meanings, also in an optical context. It can be used interchangeably to define the range of wavelengths (or frequencies) in which electromagnetic radiation oscillates, or it can define the continuous distribution of material properties, such as reflection and transmission, over an interval of wavelengths.

Continuous quantities are common in daily life but pose computational challenges. Quantization, as discussed with light detection, converts continuous data into discrete signals for processing. Integrals are often simplified using summations. Similarly, it is better to represent continuous spectra by discrete data points. But then, what makes a spectrum, a spectrum?

A series of non-negative data points can be described by the spanned interval of wavelengths and by the number of recorded data points. Taken individually, these two properties do not promptly define a spectrum, but their ratio, also called spectral resolution, is arguably one of the most important features of a spectrum. In a uniformly sampled continuous spectrum, the spectral resolution is the difference in wavelength between two consecutive data points. A spectroscopic system can be thought of as a measurement device that samples a continuous quantity into discrete data points. Sampling, whether it takes place by selecting a monochromatic light source or by filtering the incoming radiation, cannot happen punctually on the spectrum. As observed in Figure 2.23a, the transmission of a filter tends to approximate a Gaussian function. When a narrow-band filter is deployed, the central wavelength defines the measured point on the spectrum (the nominal wavelength), whereas the Full Width at Half Maximum (FWHM) is typically adopted to characterize the spectral resolution of systems with uniformly distributed selective functions [135].

The Nyquist-Shannon theorem of sampling states that in order to reconstruct an aliasing-free signal, the sampling rate must be at least twice the highest frequency present in the signal [136]. Thus, the number of points needed to reconstruct a spectrum depends on the shape of the signal. The Nyquist-Shannon theorem is useful to evaluate if a spectroscopy system is suitable to measure certain spectra, but manufacturers prefer to define their spectral resolutions a priori. In the context of measuring reflectances in the visible and infrared ranges, typical spectromet-

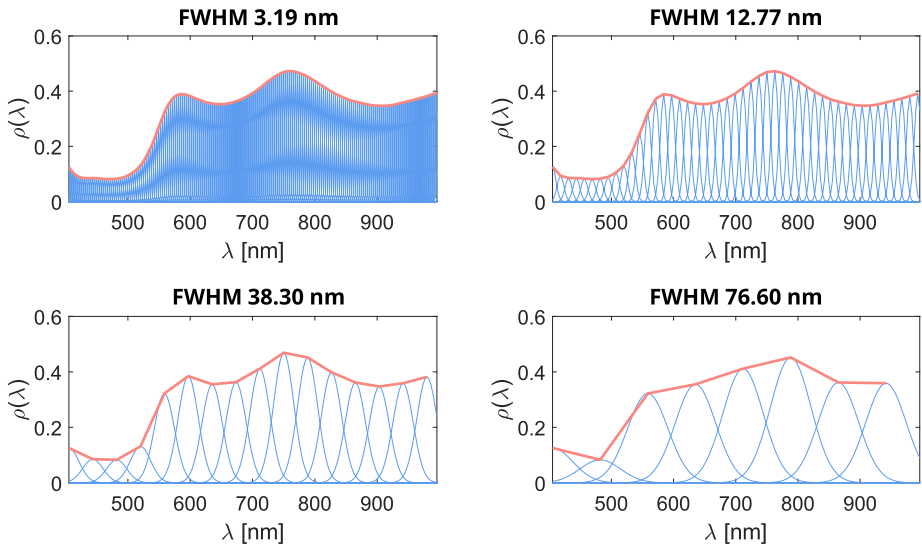


Figure 2.24: Sampling a recorded spectrum with different FWHM (spanning from 3.19 nm, as originally measured by the device to 77.60 nm) leads to loss of information.

ers display a spectral resolution that varies from 0.5 nm to 10 nm. The effect of deploying different FWHM is illustrated in Figure 2.24.

Most reflectance curves extracted from natural materials in the visible range display a certain degree of smoothness [137], whereas artificially manufactured materials can sometimes showcase spiky features in their spectra. With this notion in mind, smoothness is a characteristic of spectra that is visually investigated to judge the quality of the measurement. Local spikes in a reflectance spectrum are normally associated with noise and are usually corrected with smoothing techniques such as Savitzky-Golay filtering [138] (Figure 2.25).

2.5.6 Collection of spectra

Punctual spectral measurements in the visible and infrared (NIR and SWIR) can be carried out with three distinctive techniques, namely filtering, dispersion, and interferometry. When a measurement is defined as punctual, it is typically intended that the measurement area is relatively small and the output is a 1D vector. A typical punctual device such as a spectroradiometer usually involves the usage of an integrating sphere or fiber to minimize the degree of polarization of the incoming radiation and thus record an unpolarized measurement. When the sensor is a 2D matrix designed to capture images, this implementation cannot be performed.

Filtering methods for the collection of reflectance spectra involve the selection

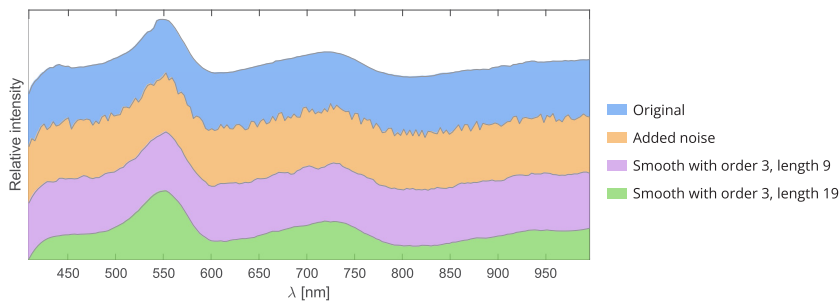


Figure 2.25: Example showing a spectrum to which noise is added in a first step, then Savitzky-Golay filtering with different parameters is applied in an attempt of denoising. The resulting spectra resemble the original one, although some differences can still be appreciated.

process either at the illumination source or at the sensor. The sensor can be constituted by a single pixel, as the filtering of different frequencies takes place sequentially. The most deployed systems of this kind are based on monochromators [139] (for the filtering of the illumination) and on Liquid Crystal Tunable Filters [140] (LCTF, for the filtering of radiation at the sensor level). Filtering methods offer the possibility to tune optimal acquisition parameters for each wavelength, but at the same time, they can be slow and not ideal for the sequential collection of data points.

Some of the most diffused spectroscopy systems rely on the presence of a dispersive element that splits the incoming radiation into the different constituting wavelengths [141] typically by exploiting optical properties of material like the refractive index, or by enforcing the generation of diffraction by etching a series of narrow slits on an appropriate substrate. The prism (Figure 2.26a) is a classic example of optical dispersion as it exploits the wavelength-dependency of the refractive index. The diffraction grating (Figure 2.26b) represents a valid alternative to dispersion by refractive index. When light passes through a diffraction grating with hundreds of equally-spaced and equally-sized slits per millimeter, multiple diffraction patterns defined by *orders* can be observed. Each pattern splits the radiation and thus, an image sensor can be aligned to the first order (the most intense in terms of brightness) to measure the contribution of each spectral component.

Prisms have some inherent efficiency-related advantages over gratings since all the refracted radiation can be measured and not only the one related to the first diffraction order. Moreover, it is possible that the diffraction order signals can be mixed, thus generating the so-called second-order errors [143]. On the other hand, systems that adopt diffraction gratings allow the observation of larger spectral in-

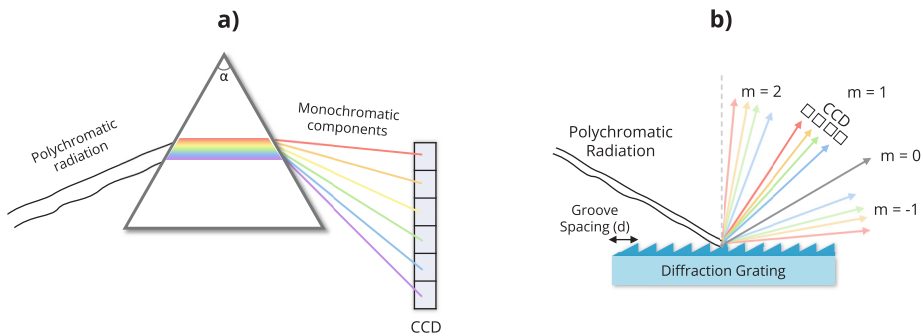


Figure 2.26: **a)** Dispersion of radiation by use of a prism. **b)** Schematization of the working principle of a diffraction grating. The orders of diffraction must not mix to accurately measure the spectral components at the edge of the operating range. Image source: Meet-Optics [142].

tervals and are more compact as they do not require large angles or the introduction of additional optics for realignments.

A novel methodology for the collection of reflectance spectra exploits the fact that the temporal profile of a signal and its spectrum in the wavelength domain are linked by the Fourier Transform [144]. An interferometer can then be deployed to generate delayed replicas of a signal using a beam splitter. The resulting interferogram obtained by applying different delays can then be transformed into a spectrum with the Fourier Transform. This methodology for the acquisition of spectra is promising since it does not require the presence of a dispersive or filtering element, and it can thus operate in response to low light levels.

2.5.7 Spectral Metrics

Once spectra are measured and their quality assessed, they are usually stored in *spectral libraries*. Spectral libraries represent a vital resource to researchers in conservation science and various application fields as they store the *fingerprints* of materials. A typical application when investigating a new artifact is the comparison of the retrieved spectra with those spectra stored in the spectral library. In the ideal scenario, if a high similarity is found then it is likely that the material under examination is known and conservation treatments can begin. In the following, we are going to explore the most common ways of comparing spectra.

When two spectra are plotted at the same time, it is evident for the visual observer to spot differences in magnitude and shape qualitatively. Quantitative measures of spectral difference can be computed with metrics and distances. These two terms are sometimes used interchangeably, but they refer to slightly different entities.

Indeed, all metrics are distances, but not all distances are metrics. A metric is such if it respects five fundamental properties [145]:

- ⤵ Reflexivity: $d(x, x) = 0$
- ⤵ Non-negativity: $d(x, x) \geq 0$
- ⤵ Symmetry: $d(x, y) = d(y, x)$
- ⤵ Identity of indiscernible: $d(x, x) = 0 \leftrightarrow x = y$
- ⤵ Triangular inequality: $d(x, y) \leq d(x, z) + d(y, z)$

whereas a distance must comply with reflexivity, non-negativity, and symmetry.

Most distance functions originate from the general definition of the Minkowski formula (Equation 2.38) for two generic spectra x and y defined over the spectral interval with N data points.

$$M(x, y) = \left(\sum_{i=1}^N |x_i - y_i|^p \right)^{1/p} \quad (2.38)$$

in which p represents the Minkowski order. By varying the order value, it is possible to obtain a series of distances, some of which respect all the properties to be labeled as metrics. In the case of $p = 0.5$, for instance, the triangular inequality is not respected, whereas Manhattan ($p = 1$), Euclidean ($p = 2$), and Chebychev ($p = \infty$) distances are all metrics. The Root Mean Square Error (RMSE, Equation 2.39) is a weighted extension of the Euclidean distance and is one of the most commonly deployed metrics [146].

$$RMSE(x, y) = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2} \quad (2.39)$$

A spectrum can be considered as a vector in the N -dimensional space generated by its sampled wavelengths. In this environment, the difference between two spectra can be quantified by the angle between them. The cosine distance is usually adopted, in a form denominated Spectral Angle (SA, Equation 2.40) [147].

$$SA(x, y) = \cos^{-1} \left(\frac{\sum_{i=1}^N x_i \cdot y_i}{\sqrt{\sum_{i=1}^N x_i^2} \sqrt{\sum_{i=1}^N y_i^2}} \right) \quad (2.40)$$

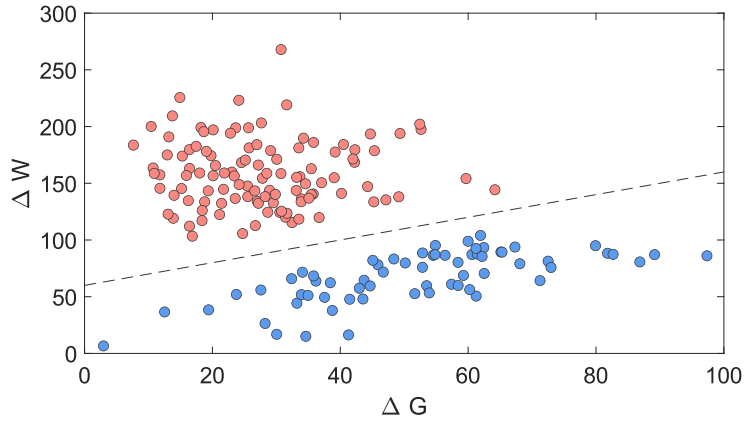


Figure 2.27: Example of a spectral library projected onto the $\Delta G/\Delta W$ space. Two macro-clusters can be easily identified.

The concept of spectral angle is tightly connected to the concept of spectral shape. Indeed, SA measures the shape similarity between two spectra, neglecting any magnitude difference. For this reason, the property of identity of indiscernible is not respected, as $SA(x, y) = 0$ can be obtained $\forall x = \alpha \cdot y$, in which α represents a scalar multiplier.

A way to simultaneously assess magnitude and shape differences between spectra is provided by the Kullback–Leibler Pseudo Divergence (KLPD) [148]. This metric is constituted by two components, an intensity difference ΔW and a spectral shape difference ΔG . Moreover, KLPD is a user-friendly visualization tool since it allows to plot spectra in a 2D space, the $\Delta G/\Delta W$ projection (Figure 2.27). Here, each data point is a spectrum with ΔW and ΔG coordinates computed with respect to an arbitrarily selected spectrum. This space can be used to cluster similar spectra and assess the distributions present in a spectral library.

When spectra are considered as probability density functions, it is possible to estimate the level of correlation. The Pearson coefficient has been adapted to yield the Spectral Correlation (SC, Equation 2.41) [149].

$$SC(x, y) = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (2.41)$$

Similarly to SA, SC evaluates the shape similarity between two spectra, and therefore cannot be considered a metric. However, SC is considered to be an improve-

ment over SA because $SC(x, y) = 1$ (the maximum similarity) can be obtained $\forall x = \alpha \cdot y$ and $\forall x = \beta + y$, in which β represents a constant magnitude shift.

The distances and metrics introduced so far produce a numerical output that is supposed to provide insights regarding the global comparison between two spectra. If from a point of view of quantification this is a desirable outcome, a single number cannot explain from *where* in the spectra the possible similarities or dissimilarities come from. To this purpose, *local* distances can help in visualizing the spectral region responsible for a high/low similarity score. All distances introduced so far can be transformed into their *rolling* counterpart, where a spectral window progressively slides across the spectra and computes the distance locally [150]. The output is then a new spectrum-like plot that enables different levels of interpretation. Equation 2.42 illustrates the local spectral correlation using k to represent the spectral bands and w as the window width.

$$LSC_k(x, y) = \frac{\sum_{i=k-w/2}^{k+w/2} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=k-w/2}^{k+w/2} (x_i - \bar{x})^2} \sqrt{\sum_{i=k-w/2}^{k+w/2} (y_i - \bar{y})^2}} \quad (2.42)$$

2.5.8 Hyperspectral Imaging a.k.a. Reflectance Imaging Spectroscopy

It took some introductory pages, but we now have all the instruments to be able to properly define Reflectance Imaging Spectroscopy as the *marriage* between imaging and spectroscopy. With this powerful imaging technique, each pixel can pixel associated with a reflectance spectrum, thus enabling the simultaneous investigation of spatial and spectral properties of a scene.

Reflectance Imaging Spectroscopy originated in the field of remote sensing, with NASA being one of the first major investors in the acquisition of satellite images [151]. Around the same time, the term hyperspectral imaging started to circulate and was used when discussing imaging spectroscopy. This term is not wrong, but according to the Working Group IEEE P4001 [152], which sets its goals on standardization of terminologies and quantities, the term RIS should be preferred. In this dissertation, we try to refer to the technique as RIS, whereas in the articles, in order to reach a broader audience, the HSI nomenclature was sometimes adopted. In this dissertation, the resulting images are termed *hyperspectral*, as an alternative terminology is not yet provided in the standard. Another term that requires certain care when used is *image cube*. A hyperspectral image can be thought of as a 3D structure with the spatial coordinates x, y and spectral coordinate z . The term *cube* is attributed as a convention since the numbers of rows, columns, and spectral bands almost never match. The more appropriate term would be *image*

parallelepiped, which is frankly a mouthful.

The boundary between MSI and RIS seems to lie on a fuzzy continuum of definitions that take into account the number of bands and the spectral resolution of the sensors. However, given the usual range of operation of spectral systems, an unofficial boundary is set on the spectral resolution for color imaging, classifying a system with a spectral resolution lower than 10 nm as belonging to the spectroscopic domain [153].

2.5.9 Methods for capturing hyperspectral images

The available image sensors are notably 2D structures, thus forcing the acquisition of hyperspectral images to resort to some sort of sequential data collection, be it in the spectral, spatial, or temporal domains.

Before delving into the different sequences adopted, let us discuss snapshot RIS. Although it would be a great commodity to record dense 3D information (meaning high spatial and spectral resolution), the task at hand is an extremely challenging one. Spectral Filter Arrays are at the moment the technology that approximates the most this ideal scenario, but at the current state-of-the-art, these systems can only be deemed to be multispectral [154].

Hyperspectral images can be acquired by sequentially collecting information along the spectral dimension deploying *staring* systems. In this configuration, which is similar to the way multispectral images are collected, the preferred solution is to deploy LCTFs [140], given their high selectivity when it comes to the filtering of incoming radiation.

Another methodology that has already been mentioned in this dissertation, spectroscopy by interferometry and Fourier Transform, can be extended to collect hyperspectral images in the temporal domain [155]. Given the absence of dispersive and filtering elements, the potentiality of such devices in RIS is promising for the acquisition in static conditions like in the CH field. On the other hand, devices with this working principle cannot be mounted on Unmanned Aerial Vehicles or satellites for Earth Observation purposes.

The most commercially diffused systems for RIS deploy dispersion elements (in particular diffraction gratings) and a sequential acquisition along the spatial dimensions. The *whiskbroom* configuration is comparable to a sequential acquisition of single-point spectroscopy measurements [141], like in the case of Fiber Optics Reflectance Spectroscopy (FORS) [156]. Differently from FORS, which is typically a hand-held device, whiskbroom cameras are mounted on a support connected to a motorized stage moving along the two spatial directions (x, y) . The

spatial resolution can then be decided by the user and is limited by the sensitivity of the motorized stage. On the other hand, selecting a high spatial resolution will proportionally increase the time needed to scan a certain area.

While whiskbroom systems only need an image sensor constituted by a stripe of pixels, *pushbroom* devices make use of a whole 2D pixel matrix that can be filled by sequentially scanning the spectral profiles of a spatial line of the scene [141]. The hyperspectral image is then composed by stacking the collected spectral lines. Pushbroom systems are ideal for the acquisition of hyperspectral images in the context of Earth Observation, where the camera moves relatively to the scene, and when translational stages are deployed so that the scene moves relatively to the camera. In this configuration, two fundamental directions are defined. The *along-track* direction corresponds to the direction of movement of the camera or scene. The *across-track* direction is perpendicular to the along-track one and is where the spatial lines are sampled from the scene.

In this dissertation, the deployed hyperspectral systems are of the pushbroom type coupled with a translational stage [157, 158], thus the assessment of quality and calibration procedures will be specific to the case at hand. When a pushbroom system is coupled with a translational stage, the illumination source and the camera are kept fixed while the scene moves in the along-track direction.

2.5.10 The spatio-spectral trade-off

A well-known trade-off in the field of RIS concerns the number of pixels that an image sensor can allocate to the detection of spatial or spectral information. In other words, spatial resolution and spectral resolution coexist in a conflict of one over the other [159]. The reason for this boils down to the inherent pixel response and its SNR.

When the pixel pitch is reduced and more pixels are introduced in an image sensor, the spatial resolution increases but the SNR decreases, as the noise of the individual pixels remains unchanged while the signal must be divided amongst more units. A strategy to increase the SNR back can involve the widening of the bandwidth, which in turn leads to a decrease in spectral resolution. Similarly, increasing the number of pixels in the spectral direction by having narrower FWHMs will make the SNR decrease, and acceptable levels of SNR can be restored by increasing the pixel pitch, thus decreasing the spatial resolution.

RIS sensors are then built considering this fine balance between the spatial and spectral resolution, while at the same time, research efforts are poured into developing sensors with lower electrical and thermal noise to push the boundaries of the tradeoff forward.

2.5.11 Quality of RIS systems

The term *quality* in an imaging context is related to a broad field of research that makes use of notions from vision science, psychology, and signal processing [160]. The high dimensionality of hyperspectral images makes their visualization a challenging task, and thus evaluating image quality via standard methodologies is not feasible.

Here, we discuss a series of parameters that can be investigated to evaluate the performances of a hyperspectral imager. Such parameters are usually evaluated at an early stage of the life of a RIS system and should be tested periodically but not on an everyday basis since their evaluation is not straightforward.

The SNR (Equation 2.43), probably the most indicative parameter of a RIS system, can be measured by recording a Lambertian-like surface uniformly illuminated (in controlled laboratory conditions this is achieved through the usage of an integrating sphere). The SNR is then computed spectrally as the ratio between the mean value of the band and its standard deviation:

$$SNR_{\lambda} = \frac{\mu(k_{\lambda})}{\sigma(k_{\lambda})} \quad (2.43)$$

Since SNR and quantum efficiency are closely related, the SNR curve tends to mimic the behavior of the quantum efficiency of the sensor. The classic shape that falls toward the extremities of the spectral interval is also typically used to determine the operating range in the first place [161].

When we think of a hyperspectral image as a 3D geometrical entity, the fact that the 2D images that compose the stack are co-aligned is often taken for granted. Misregistration of a pixel in adjacent bands in a hyperspectral image is quantified with *keystone* [162], which is an optical distortion that sprouts from misalignments in the composite system made of the slit (aperture) - diffraction grating - image sensor. Keystone can be evaluated by collimating a broadband point light source and by recording the center of gravity of the PSF band-wise (Figure 2.28a). If the deviation is in the order of a fraction of a pixel, then the system can be considered keystone-corrected.

Similarly, optical distortions can take place along the spatial dimension when the system is exposed to monochromatic inputs (for example by recording an integrating sphere in controlled laboratory conditions). In this instance, the distortion takes the name of *smile* [162] because of the typical bent shape of a spectral line across the field of view (Figure 2.28b). Similarly to keystone, if the deviation is in the order of a fraction of a spectral band, then the system can be considered

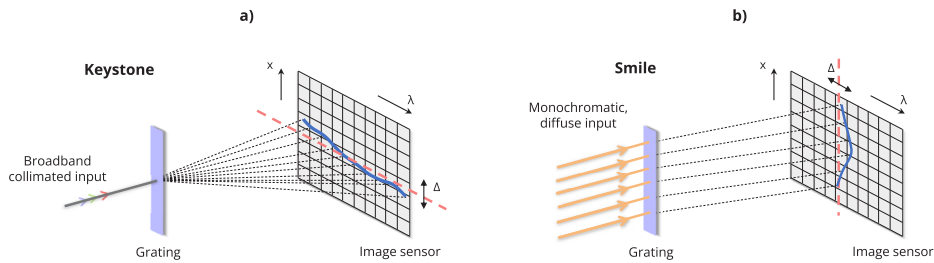


Figure 2.28: Schematization of keystone **a)** and smile **b)** in a RIS system. In the presented instances, both smile and keystone exceed the distortion limit of 1 pixel.

smile-corrected.

The pixels distributed across-track devoted to the capturing of spatial information are subject to optical distortions due to the fact that each pixel possesses a field of view that varies according to the position of the pixel with respect to the central spatial pixel. This information is usually encapsulated in a sensor model that can be used to correct image artifacts that squeeze and expands pixels in the across-track direction [163].

2.5.12 Acquisition pipeline

In a classic imaging setup, the four main components are an object of interest, an illumination source that sends photons to it, an imaging device able to capture and record the reflected photons, and a reference target with a known reflectance factor. We have already decided what our imaging device will be, so now we need to formulate some requirements that can help us in obtaining good-quality hyperspectral images. For the moment, we will consider the case of single-object (a relevant artifact), single-light source, and single-pushbroom RIS device.

Pushbroom RIS systems represent a convenient solution for the capturing of hyperspectral images because of their flexibility when it comes to scanning objects of varying dimensions. On the other hand, typical systems are built with a fore optic that only allows them to be focused at a specific acquisition distance. This leads to the generation of a constraint regarding the topography of the artifacts that can be scanned. Flat objects are privileged, but this is a mostly ideal condition. Indeed, most artifacts like paintings, manuscripts, and textiles, for one reason or the other, present topographical reliefs. As we already observed, paintings are produced sometimes with peculiar pictorial techniques such as impasto to confer a perception of depth, whereas parchments and textiles can lose their original shapes when exposed to non-ideal levels of temperature and relative humidity [27]. The

depth of field of the lens can thus be used as a parameter to decide if an artifact is suitable for scanning or not, but as we will see when discussing Flat Fielding (Section 2.5.13), objects with complex topographies and marked reliefs can have serious implications on the calibration process.

The illumination source must be selected considering two criteria of benefit. Firstly, a light source must not induce any irreversible change to the artifact and thus has to be selected respecting recommendations regarding light dosage and operating temperature [31]. Once these factors are validated, a source with a certain power can be placed at a suitable, harmless distance. At the same time, the illumination must be chosen so that it is beneficial for the data collection. For RIS, it is important to deploy a broadband light source (or a combination of several light sources) that has enough emission, especially in those wavelengths where the quantum efficiency of the sensor is lower. Once a decision is taken, the illumination should be as uniform as possible on the across-track direction in correspondence with the aperture slit. For historical artifacts, it is also recommended that the light source acts only on this line, so as to not illuminate parts of the artifact that are not imaged.

When it comes to the adjustments that can be made at the stage in which the imaging setup has already been built, every fine-tuning should be performed with the goal of increasing the SNR, while at the same time protecting the artifact. Unfortunately for us, high SNR and preservation do not go hand in hand, as the methodologies to increase SNR involve increasing the integration time as much as saturation allows it and frame averaging, i.e. the multiple scanning of a line designed to increase the SNR by a factor \sqrt{N} where N is the number of scans. Both procedures to increase the SNR increase the artifact exposure as well, so every instance must be carefully designed to meet the specific needs of preservation.

Finally, the presence of a reference target in the scene is crucial for the computation of the reflectance factors of the scene [164]. Typical targets are made of Spectralon, a material developed by Spherelab that approximates a Lambertian surface [165]. Spectralons come at different global reflection levels and with more or less stable reflectance curves in the range between 250 nm and 2500 nm. In a line-scanning environment, these targets are usually placed before or after the object of interest, at the same camera distance as the artifact, so that the whole scene receives the same amount of irradiance.

2.5.13 Calibration pipeline

In the previous sections, we have reviewed how the collection of photons triggers the generation of electric currents that, if transformed into digital signals can then be stored as digital numbers in a file accessible by means of various software. At

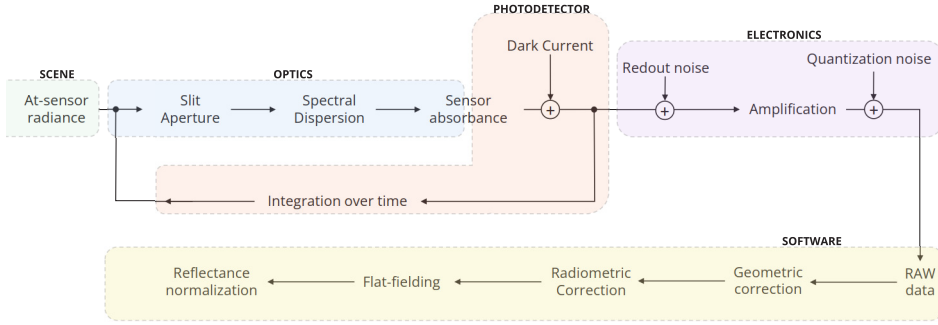


Figure 2.29: Full pipeline of data collection, reading, and calibration. Figure inspired from [161]. The different color shades serve to segment the processing steps into semantic areas: green for scene, blue for optics, orange for photodetector, purple for electronics, and yellow for software.

this stage, we refer to the values as being *raw*, meaning that they are the total product of interaction between light source, object, and factors related to the intrinsic properties of the imaging device as well as parameters introduced by the user when the images were captured. The goal of a calibration pipeline is to progressively discard all these effects until only the material-related information (the reflectance in the case of RIS) is left. The pipeline adopted in the articles presented in this dissertation can be found in Figure 2.29.

Geometric Correction

The first step of the pipeline can be performed on RAW data and involves accounting for the optical distortions introduced on the across-track direction by the variation of individual pixel fields of view. Moreover, a second distortion can be introduced in pushbroom system if the camera slit aperture and the movement of the scene are not exactly perpendicular [163]. The general correction model can be expressed as indicated in Equation 2.44.

$$\forall (x, y) \in [M] \times [N], I(x, y) = D[h(x, y), v(x, y)] \quad (2.44)$$

in which the transformation h accounting for the across-track distortion can be derived from the sensor model of the camera, whereas the transformation v related to the along-track distortion depends on the way the setup is built and must be obtained experimentally.

Radiometric Correction

The effect of camera-dependent and user-dependent parameters are discarded with a radiometric correction [166]. The goal of this correction is to perform a transformation of data from RAW to relative radiance (RAD), hence the name. The formulation of the correction, reported in Equation 2.45, is system-dependent and is related to an acquisition model that is inverted.

$$L(x, \lambda) = \frac{[RAW(x, \lambda) - DC(x, \lambda)] \cdot h \cdot c}{\eta(\lambda) \cdot RE(x, \lambda) \cdot SF \cdot A \cdot t \cdot \omega \cdot \Delta\lambda(\lambda) \cdot \lambda} \quad (2.45)$$

The camera-dependent parameters are intrinsic to the system and cannot be changed by an external user. They are: quantum efficiency η , gain matrix RE , aperture A , pixel field of view ω , scaling factor SF , spectral bandwidth (FWHM - $\Delta\lambda$), and central wavelength λ . The user can modify the integration time t , thus influencing the dark current signal DC .

Flat-Fielding

As already mentioned when discussing light sources, the illumination should be uniform on the acquisition line of a pushbroom camera. However, intensity fluctuations can still take place and can be observed when imaging a homogeneous achromatic material. In most instances, the Spectralon already present in the scene can fulfill this purpose, but a standardized material is not strictly necessary. The whole scene can then be flat-fielded by arithmetically dividing each spatial line by the line acquired on the homogenous material.

The selected wording, *flat-fielding*, already tells a lot of the expected use that one should do. Indeed, this procedure is designed to work on strictly flat materials, but what should be the best solution for those objects that have features that can be scanned by hyperspectral systems because their reliefs do not exceed the depth of field limits? A non-flat-fielding should be adopted, but the surface topography must be estimated a-priori or within the same system [167].

A pushbroom system with a fixed camera position and fixed illumination can be considered advantageous when performing a flat-fielding correction if we consider that typical diffuse materials generate diffusive surface scattering thanks to their rugosity. If the spatial resolution of the system is able to capture these local variations, then flat-fielding would not work correctly because of the introduced noise of an individual scanned line. However, it is possible to capture a series of adjacent lines that once averaged can result in an appropriate approximation of the light field arriving on the acquisition line.

Reflectance Normalization

The flat-fielded radiance values of a pixel in the scene are transformed into reflectance via a band-wise arithmetic division by the flat-fielded radiance values of the Spectralon target and consequent multiplication by the reference reflectance curve, which is usually provided by the manufacturer of Spectralon. Equation 2.46 encompasses the transformation from relative radiance values $L(x, \lambda)$ to reflectance $\rho(x, \lambda)$ encompassing both flat-fielding and reflectance normalization. Here, $\zeta(x, \lambda)$ represents a spatio-spectral radiance line extracted from a flat-fielding target, whereas $s(\lambda)$ is the flat-fielded spectrum extracted from a potential Spectralon target and $r(\lambda)$ is the reference reflectance provided a priori. It is important to point out that Equation 2.13 inverts the radiative transfer model following the assumption that the observed object is flat, matte, and diffuse.

$$\rho(x, \lambda) = \frac{L(x, \lambda)}{\zeta(x, \lambda)} \cdot \frac{s(\lambda)}{r(\lambda)} \quad (2.46)$$

2.5.14 Multivariate analysis for hyperspectral images

A possible definition for a good-quality hyperspectral image that captures a scene illuminated by a broadband source (not an LED or a Mercury lamp for example) could be of a spatially sharp image content with smooth spectral features.

If a spectrum varies smoothly, the difference between adjacent spectral bands is likely to be small. This leads to two fundamental conclusions that are usually both debated as positive and negative aspects when discussing hyperspectral images: band correlation and redundancy [168]. If on one hand having similar adjacent bands can simplify computation operators, on the other, the redundancy can hinder them and represents a challenge for data management, especially for problems related to storage and data compression [169]. In the end, spectral smoothness is a blessing and a curse of RIS.

The high spectral redundancy makes hyperspectral images optimal targets to apply dimensionality reduction techniques that can transform a dataset with hundreds of channels into a reduced but similarly meaningful dataset with a handful of channels that do not necessarily refer to spectral information [170]. Reducing the dimensionality of hyperspectral data can help in reducing noise, generating new features, and visualizing data highlighting underlying patterns that would not be so evident in a more standard visualization scenario.

Multivariate analysis can be applied to examine hyperspectral images by considering the pixels to be individual observations, whereas their reflectance acquired at

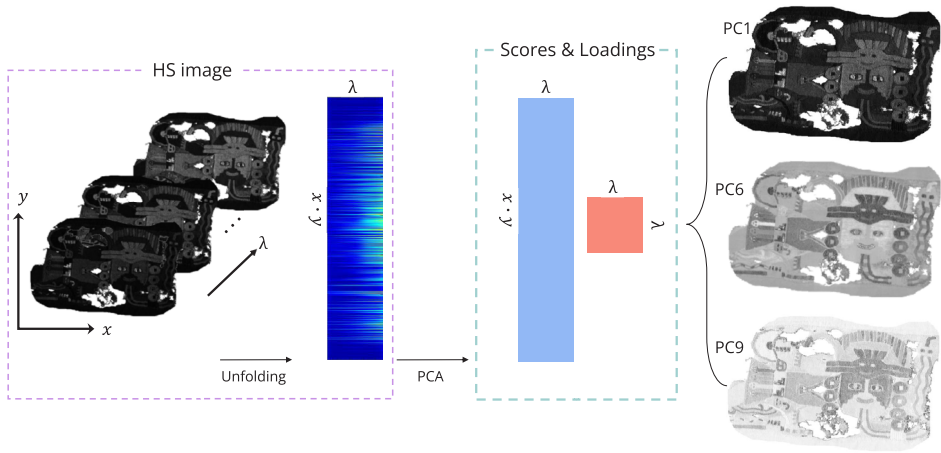


Figure 2.30: Example of PCA applied to a hyperspectral image.

the sampled wavelengths can be considered the variable of the dataset. Multivariate analysis techniques usually do not take into account the spatial arrangements of pixels and can be performed without any loss of information on apparently noisy and meaningless images whose pixels got randomly shuffled. This can sometimes be seen as a shortcoming, but nonetheless, the potential to extract meaningful information is still significant.

Principal Component Analysis (PCA) is a multivariate statistical technique that transforms the variables of hyperspectral images related to the observations, into new variables, denominated Principal Components (PCs), which are linearly uncorrelated between themselves [25]. The application of PCA to hyperspectral imaging involves constructing a covariance matrix from the data, followed by the computation of its eigenvectors and eigenvalues. The so-called first PC is then the linear combination of original variables that explains the most variance, whereas the subsequent PCs capture progressively less variance. The data is then projected into a new space of lower dimensionality by selecting a subset of the PCs that can highlight the most significant patterns. Figure 2.30 illustrates an example of PCA applied to the case of a hyperspectral image in which some patterns are not as visible in the original capturing as in the newly projected principal components.

In PCA terminology, the terms *scores* and *loadings* are often used, but they do not directly translate to eigenvalues and eigenvectors, although they are related. We can think of eigenvectors as the directions in the newly projected space in which the variance of the data is maximized, while the eigenvalues quantify the variance explained along each eigenvector. The scores are the coordinates of the newly

projected observations onto the eigenvectors, whereas the loadings indicate the relationship between each variable and the new eigenvectors (PCs). Thus, a high loading represents a variable that is highly linearly correlated with the observed PC.

PCA exploits Singular Value Decomposition (SVD, Equation 2.47), a matrix factorization technique that decomposes the observation matrix A into three new matrices: an orthogonal matrix U containing the *left* eigenvectors, a diagonal matrix Σ containing the singular values, and a transposed orthogonal matrix V^t containing the *right* eigenvectors. The role of SVD is to estimate the principal components of U from the covariance matrix computed from the original observations.

$$A = U \Sigma V^t \quad (2.47)$$

Other matrix factorization techniques are particularly effective on multi-band images. Independent Component Analysis (ICA) [171] works similarly to PCA, with the difference that each new component is found so that it is a linear combination of original variables statistically independent from all the other components. The main difference is then that the PCs are found sequentially following a variance explained-based ranking, whereas the ICs are computed all at once. ICA tries to find independent components in mixed signals and is then sometimes useful in decomposing images into different image sources, like in the example of an outdoor reflection on a window looking inside a room.

PCA and ICA can sometimes be challenging to interpret since the new spaces into which the features are projected are computed unconstrainedly. So, reflectance data are sometimes projected with negative values, which do not carry much physical meaning. Non-negative Matrix Factorization (NMF) [172] is a decomposition technique that overcomes this problem, providing a data reduction with values that remain strictly positive.

2.5.15 RIS for Cultural Heritage Analysis

One of the most fundamental tools dominating the RIS analysis scene is undoubtedly the spectral library, a collection of individual *spectral signatures* (reflectance, transmittance, absorbance, etc.) related to pure or compounded components that have found use in the production of historical artifacts. Numerous institutions across the world have been collecting spectral signatures with the goal of creating and sharing large databases for research and archiving purposes [173–175].

The optical characteristics of pigments in their mass-stone form (the powder),

binders, pigments diluted in different concentrations of binders, dyestuff, mordants, fibers, etc. are all subject to be stored in spectral libraries. Typically, spectral signatures are included in a library either by the direct extraction from artifacts or with the usage of mockups. In order to be able to include a signature from an artifact it is necessary to have an auxiliary analytical measurement that can pinpoint the composition of materials present in the measured area [176]. Producing mockups from scratch, on the other hand, allows researchers to accurately know all the materials deployed and thus does not require an auxiliary measurement to be involved [177].

At this point, a premise is necessary. In the context of Heritage Science, it is highly unlikely to encounter pure endmembers as such. Through the ages, artists have experimented with all kinds of materials to generate the colors, optical effects, paint consistencies, and textures that eventually ended up on the canvas. As a result, finding a match for all materials contained in a newly examined artifact is not something that can be taken for granted. If a spectral library is not available a priori, then it is possible to operate according to two alternatives that by no means intend the endmembers as *pure* materials, but rather a rough approximation that is used for labeling simplicity.

The first alternative requires the exploitation of prior knowledge, which can be in the form of expert knowledge regarding the history of the artifact, the historical context in which it was placed at the moment of its creation, and beliefs regarding specific regions of interest. Prior knowledge can also be intended as previous acquisition campaigns operated with different techniques such as XRF and FTIR in some areas of interest [176].

When prior knowledge cannot be exploited in any form, then it is possible to estimate the main spectral signatures directly from the scene. This practice is possible thanks to the development of Endmember Extraction Algorithms (EEAs) in the field of remote sensing [178]. Here, the term *endmember* is used to define a spectrum contained in the spectral library which represents a defined material in its purity. The concept of purity is however highly dependent on the context and on the criteria imposed by the user. In remote sensing, typical endmembers indicate water, grass, canopy, rocks, etc., but the definition is tightly connected to the available spatial resolution. When a low-altitude drone is deployed, the endmember -canopy- likely becomes not pure enough, and distinctions must be made between tree species.

Assuming that the observed scene contains a certain number of pure materials is a good starting point to begin the search for endmembers. Consequently, only a handful of pure pixels are present in the scene, whereas the nonpure ones (the

vast majority) are considered to be the result of a combination of the endmembers [179]. In the simplest case, the signal of a pixel $Y(\lambda)$ is assumed to be a linear combination of the endmembers $e(\lambda)$, thus implying a linear mixing model (Equation 2.48).

$$Y(\lambda) = \sum_{i=1}^q e_i(\lambda) \cdot \alpha_i + b(\lambda) \quad (2.48)$$

in which q is the number of endmembers, and α are the relative concentrations or abundances of the endmembers. An additional level of band-dependent noise $b(\lambda)$ must also be included. In this configuration, two constraints are typically applied to the concentration vector. The non-negativity constraint allows the abundances to be strictly zero or positive $\alpha_i \geq 0 \forall i \in \{1, \dots, q\}$, whereas the sum-to-one constraint ensures the energy conservation principle by not allowing the loss or creation of matter $\sum_{i=1}^q \alpha_i = 1$. In some applications, to allow a certain degree of tolerance, the sum-to-one constraint can be relaxed by a certain percentage.

One of the first EEAs to be developed, the Pixel Purity Index [180], seeks to geometrically find the endmembers as the vertices of the smallest simplex that contains the observed data in an N -dimensional space. Similarly, the N-FINDR [181] method deploys the complex of N -dimensional simplex, but the endmembers are found by iteratively growing the simplex from within the data. The stopping criterion is set so that the simplex connecting the purest pixels is larger than any simplex connecting the pixels found as combinations of others. The Vertex Component Analysis [182] draws from both predecessors and projects the pixel spectra in an N -dimensional space assuming that every spectrum is the product of a linear combination of a certain number of endmembers. The endmembers are then found iteratively as those linear combinations that include only small concentrations of the other presumed endmembers. Some EEAs have the disadvantage (or sometimes advantage) that a certain user proficiency is required, especially in determining the number of endmembers to be extracted, which is necessary information for the algorithms to initiate their search and has a significant impact on the performances [183]. This can be beneficial in contexts in which the palette of a painter is known to limit the extraction of meaningful signatures but also can leave a user clueless and having to repeat the procedure many times until a somewhat optimal solution is found. Recent developments of fully automated EEAs [151] have focused their efforts on considering the possibility that pure pixels may not exist within a scene [184, 185].

Endmembers are fundamental tools to identify materials, but it is natural to think that within a complex scene, a series of factors can concur to produce local ap-

pearance differences. Usually, these differences help humans infer shapes, illumination, texture, and material properties. The endmember -grass- related to a data collection conducted on a sunny day should still be able to recognize grass even when a mountainside casts a shadow on it or on a rainy day. A spectral system, however, is not likely to overcome these differences, unless it becomes aware of the concept of spectral variability [186]. In RIS for CH artifacts, the spectral variability of a spectral signature is dictated mostly by the interaction of the spatial resolution of the imaging system and the surface roughness, which can be summarized in the BRDF. Micro-shadows and partial specularities can also be considered instances in which an endmember changes its spectral signature. The main idea for conducting RIS analysis with spectral variability is to consider *bundles* of N endmembers rather than individual endmembers. This will produce more robust results, at the expense of an increase in computational requirements.

A carefully designed set of endmembers (or bundles of endmembers) is crucial to conduct the two most popular analysis techniques of artifacts by means of RIS: spectral mapping [187–192] and spectral unmixing [177, 179, 193–195]. In spectral mapping, each pixel spectrum is compared to each signature contained in the spectral library, and a similarity score is assigned. The pixel then receives a label corresponding to the element in the library with which it shares the highest similarity. The result is then a map of endmembers (Figure 2.31). There is no free-lunch rule for the choice of the similarity measure to adopt, although it has been proved that the Spectral Correlation represents a more robust version of the Spectral Angle [149]. The typical outcome assigns a label to each pixel, but by defining quality thresholds on the similarity scores it is possible to leave those pixels that do not achieve a good match unclassified. This is typically performed to avoid false positive cases, which can in turn have unwanted repercussions in the selection of wrong conservation treatments.

Spectral unmixing aims to destructure the spectrum of a pixel into a weighted combination of endmembers according to a mixing model $Y(\lambda) = f(E(\lambda, q), C(q))$ in which $E(\lambda, q)$ represents the spectral library and $C(q)$ is the concentration vector of individual abundances α_i . We have already seen how the linear model of Equation 2.48 is an example of a mixing model. However, when we consider typical mixing occurring in paintings and textiles, it is not easy to envision linear mixing at the microscopic level. Indeed, linear mixing is designed to resolve instances in which signal mixing takes place at the camera level because two or more signatures, spatially separated in reality, are represented within a single pixel value. Pigments and dyestuff, on the other hand, mix more *intimately* and thus a linear model is likely to be not suitable. Non-linear mixing models are for example based on subtractive mixing [196] (Equation 2.49) and on the more famous

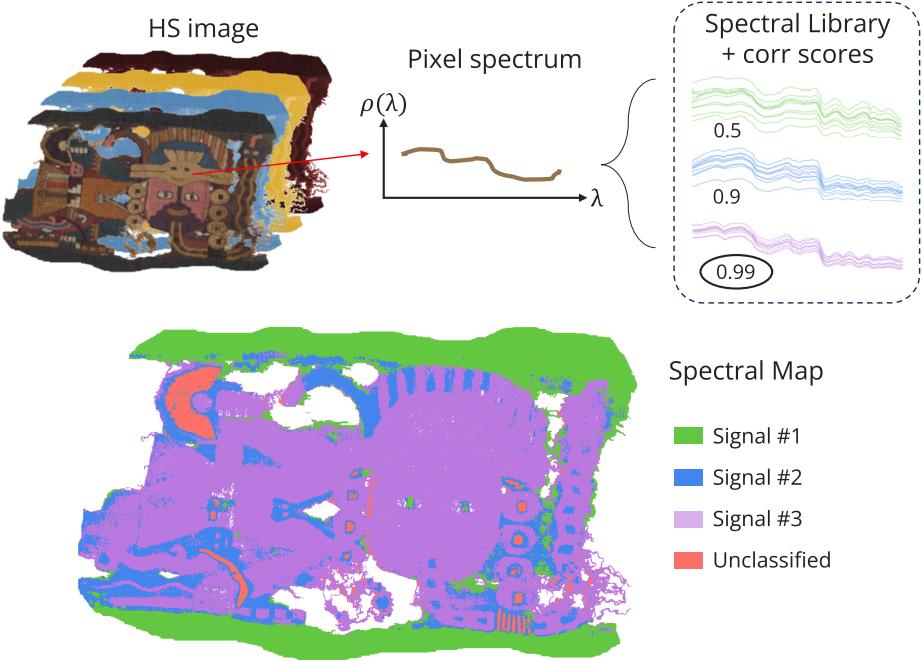


Figure 2.31: Example of spectral mapping with three endmembers bundles. A pixel receives a label corresponding to an endmember if the correlation score is higher than a certain threshold. Otherwise, it remains unclassified (pixels in red).

Kubelka-Munk [197, 198] model for opaque materials (Equation 2.50).

$$Y(\lambda) = \prod_{i=1}^q e_i(\lambda)^{\alpha_i} + b(\lambda) \quad (2.49)$$

$$Y_{\infty} = 1 + \frac{K}{S} - \sqrt{\left(\frac{K}{S}\right)^2 + 2\frac{K}{S}} \quad (2.50)$$

in which Y_{∞} represents the reflectance of an infinitely thick (opaque) surface, and K and S are its corresponding absorption and scattering coefficients. When mixing takes place, K and S are computed as a linear combination of the absorption and scattering coefficients of the individual endmembers (Equation 2.51).

$$K = \sum_{i=1}^q k_i \cdot \alpha_i \quad S = \sum_{i=1}^q s_i \cdot \alpha_i \quad (2.51)$$

Spectral unmixing boils down to inverting the mixing model in order to retrieve the concentration vector. This is usually carried out by a constrained optimization applying the non-negativity and sum-to-one constraints. Moreover, regularization can be enforced to limit the realm of possible solutions [199]. Lasso L_1 regularization has been proven to be quite efficient since it limits the number of endmembers that can be selected within a mixing. This is a convenient procedure to adopt since it reflects the reality in which colors were obtained by mixing no more than three to five pure pigments or dyes. Equation 2.52 reports a typical constrained cost function for an optimization that solves the unmixing problem.

$$J = \arg \min_C \frac{1}{2} \sum_{i=1}^N \left\| Y_i - \prod_{j=1}^q e_{i,j}^{\alpha_j} \right\|^2 + \lambda \|C\|_1 \quad (2.52)$$

$$s.t. \quad \alpha_i \geq 0 \forall i \in \{1, \dots, q\}, \quad \sum_{i=1}^q \alpha_i = 1 \quad (2.53)$$

The product of spectral unmixing is then a series of q concentration images in which the pixel values are included in the interval $[0, 1]$. Concentration images offer more information than a spectral map, but they encounter the same visualization problems as multispectral images, and the pixel-wise compositions can only be looked up in large tables.

2.6 Fusion of VNIR-SWIR data

The first step to fuse two sets of data is to understand how the production of the data takes place at the sensor level. In the previous sections, we examined how photodetectors are constructed and their working principles. Moreover, when introducing the sensors deployed for RIS, the trade-off between spatial and spectral resolution was highlighted.

The size of sensors and the number of pixels they can host along the spatial and spectral directions depend on a variety of factors including but not limited to the material of the photosensitive area, the type of sensor (CCD or CMOS), the desired spectral bandwidth, and the covered spectral range. Therefore, if we compare sensors for imaging in different spectral ranges, we will observe arrays of different sizes expressed as numbers of pixels. The most direct consequence of this is that the resulting images, acquired at the same acquisition distance, will display different spatial resolutions in terms of GSD. A combination of image registration [200] (Section 2.6.1) and sharpening [201] techniques (Section 2.6.3) can be deployed to harmonize the GSD of two different image datasets by spatially aligning one scene to the other.

If the goal is to form a unified spectral image that covers two or more spectral ranges originally separated, it is convenient that the unified spectrum is continuous across the two ranges, with a uniform sampling rate. However, when spectra are extracted from matching areas, it is possible to observe that the values do not connect smoothly (or do not match in instances in which the spectral intervals overlap). Therefore, a spectral alignment, or *splicing* [202] (Section 2.6.2), is needed to obtain smoothly connected spectra that resemble the true nature of the collected data.

2.6.1 Image Registration

The spatial alignment of images belonging to the same scene is a classic problem in Computer Vision, with solutions that span different scenarios and needs of accuracy [200]. Generally, the problem is described with two images, one *fixed* image, serving as a reference, and one *moving* image, also called the target. As the name suggests, the moving image is the one that is aligned with the reference. The solution of image registration boils down to finding the homography, i.e. the transformation matrix, that transforms the target image so that it appears as it was observed from the camera viewpoint used to capture the reference image. The general homography is a 3×3 matrix in which each element refers to a specific geometrical transformation.

The simple geometrical operations of scaling, shifting, and rotation are included in

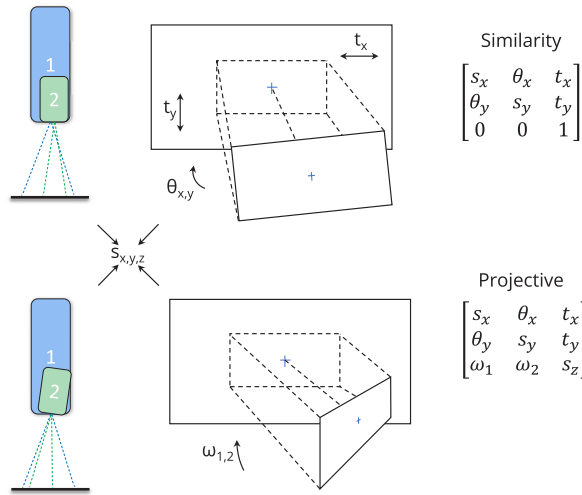


Figure 2.32: Schematization of a dual-imaging system with different spatial resolutions. When the camera objectives are aligned (top), the projective transformations are neglected, whereas they have a significant influence when the camera objectives are not parallel (bottom). Each element in the transformation matrix is responsible for the various operations of scaling, rotation, translation, and projection.

the transformation that is denominated similarity. A similarity transform is perhaps the most likely to take place in the case of two pushbroom hyperspectral cameras with parallel across-track directions. However, if there is no safety regarding the parallelism of the camera objective, a user might consider that a geometric transformation of shearing is present as well, and a projective transform might become then more suitable. A brief schematization of the different situations and related geometric transform can be found in Figure 2.32.

In the context of multi-band imaging, it can be assumed that each hyperspectral image is internally co-registered. Therefore, a unique geometric transform can be used to connect the two hyperspectral datasets.

Another distinction to be made when defining an image registration problem concerns the nature of the images that need to be aligned. Multimodal imaging [200] is a crucial practice in many application fields, including medical and CH imaging [203]. Multiple devices capturing different types of radiation can be deployed to extract large amounts of information from an artifact. Since the goal of multimodal imaging is usually to fuse information from different sources, the general tendency is to collect data in an overlapping field of view, trying to maximize the image similarity beforehand, at the experimental stage. On the other hand, unimodal imaging

deploys a single device, and therefore its registration problems tend to deal with captures at different viewpoints, either at different scales or camera projections.

The algorithms developed to learn the homography matrix are classified into four main categories: feature-based, area-based, hybrid approaches, and deep learning-based [200]. Feature-based approaches usually involve the deployment of sequential actions to compute a homography matrix. The first stage concerns the detection of keypoints in both fixed and moving images. Depending on the rationale of the method, keypoints can have different meanings. Amongst the others, Harris keypoints [204] compute local gradients in both x and y directions to identify corners, whereas Scale-Invariant Feature Transform (SIFT) [205] points are computed using a multi-resolution approach to ensure that a point is relevant regardless of the scale it is observed at. Each keypoint is then associated with a feature vector of variable (depending on the method) length. The way features are computed usually takes into account the gradients in a local neighborhood around a detected keypoint. According to carefully designed distances, the feature from the fixed and moving images are then matched, and a geometric transform is learned usually deploying methods such as RANSAC (RANdom SAMple Consensus) [206] to build a certain resistance to potential outliers in the matched keypoints.

For multi-band images produced with RIS, features can also consider spectral information and thus can help find more robust matches [207, 208].

Area-based methods focus on solving an optimization problem in which the cost function reflects the quality of the alignment of the images. In instances in which the two images are collected at a similar spatial resolution but with different modalities, e.g. an XRF image and a color image, area-based methods can be useful for finding matching structural patterns where the pixel intensity differences are too dissimilar. The similarity can be computed with statistical measures coming from information theory like Mutual Information (MI) and its normalized version (nMI) [209].

Feature-based and area-based methods can be used in combination to solve problems of multimodal image registration at different scales, exploiting the advantages of both methodologies [210]. A common framework can involve alternating between the two families of methods to finely tune the homography matrix at each iteration.

As in most Computer Vision tasks, the coming of deep learning represented a leap in image registration. The bottlenecks have been for years the need for large amounts of training data and the necessity of generalization to adopt trained neural networks in different application fields. Recently, the development of generative

networks has pushed the capabilities of deep learning even more forward [211].

Evaluating if a registration procedure has been successful is not an easy task, both from visual and computational standpoints. If we opt to visually evaluate an image alignment, it can be possible that the different geometrical transformations introduce local distortions that can be difficult to track and compare, especially when the size of the image increases. The main challenge encountered by computational methods is the lack of ground truth. By approximation, the reference image is usually deployed in this role. Classical intensity-based distances like RMSE can be used, but only if the images are acquired with the same conditions of illumination and instrumentation. The Peak Signal-to-Noise Ratio (PSNR) is one of the best indicators to evaluate image quality [160] when it comes to color imaging, but its usage to evaluate non-perceptual images can be sometimes not straightforward. A commonly accepted practice computes the so-called pixel displacement [212]. Once the moving image has been aligned, the process of keypoint detection and matching is performed again. By tracking the coordinates difference of the matched pixels it is possible to provide an average pixel displacement error and set an acceptance threshold. In the case of multi-modal imaging, where it is not possible to match pixels, the same distances or metrics used in the cost function are usually deployed.

2.6.2 Spectral Splicing

The word *splicing* is quite an unused term in the context of spectroscopy and RIS, but it represents a well-defined operation that is perhaps better described with a practical example.

Going back for a second to the domain of 1D spectroscopy, let us assume that a spectroradiometer is deployed to capture the reflectance of a material in a wide range spanning the visible and the infrared. Neglecting the internal structure of the instrument, which might include complex optical components, we have discussed how it is impossible to perform the measurement with a single detector. Let us assume then that visible radiation is captured with a silicon sensor, in the range from 400 nm to 900 nm, whereas IR radiation is captured with an InGaAs sensor sensitive in the range from 900 nm to 1800 nm. Once the measurement is performed, it is possible to observe that the resulting curves are *displaced* (Figure 2.33a). The *spectral jump* can be ascribed to various factors such as the decreasing SNR at the end of the sensitive spectral range, a possible different bandwidth deployed by the different detectors, and the operating temperature of the detector, which tends to affect more significantly silicon sensors that do not typically deploy strong cooling systems [213–215].

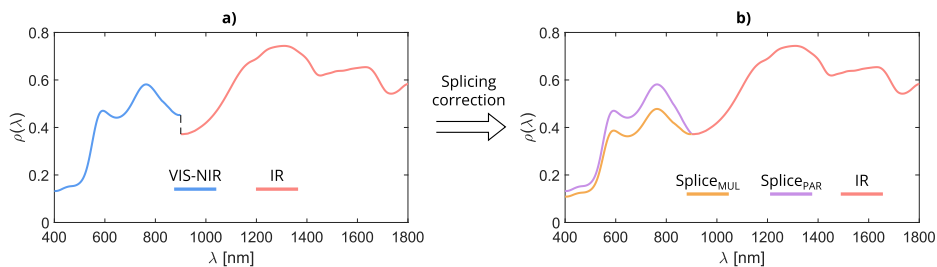


Figure 2.33: Spectral splicing correction example. In **a)** it is possible to observe the discrepancy in measured reflectance at the junction wavelength. In **b)**, two possible splicing corrections (parabolic and multiplicative) are applied on the spectrum related to the visible range, while the infrared spectrum remains unchanged and acts as a reference.

If the curves are simply connected, the resulting spectrum would have a shape that is highly unlikely for both natural and artificial materials. For this reason, a splicing correction is applied. Splicing, which is a term borrowed from biology in the process of RNA transcription¹, smoothly connects two spectral curves affected by some degree of discontinuity so that the resulting spectrum has a plausible shape (Figure 2.33b).

Simple splicing corrections can be performed by deciding which spectrum remains unchanged and consequently, which one is transformed. A global multiplicative factor or an addend can be applied [202]. However, these corrections apport a change even at those wavelengths in which the quantum efficiency of the detector is reliable. In most spectroscopic devices the deployed splicing correction makes use of a parabolic correction [217] that affects more significantly those data points found closer to the junction wavelength.

A splicing correction is therefore needed whenever spectral information is acquired in adjacent or overlapping spectral ranges since detectors behave according to their intrinsic properties and can produce slightly different measurements of the same quantity. We will see in Chapter 3 how the effect of spectral jumps is amplified in a RIS system and how a new ad-hoc splicing correction is necessary to consistently obtain smooth spectra.

2.6.3 Hypersharpenering

We have discussed at length how sensors deployed for the imaging of radiation belonging to different portions of the electromagnetic spectrum tend to have dif-

¹RNA (ribonucleic acid) splicing is the process of removing introns and joining together the remaining exons to create the mature RNA molecule that can be used by the cell to make proteins [216].

ferent sizes. In addition to the specific needs of a manufacturer and the construction limitations dictated by the spatio-spectral trade-off, this is also the product of economics, with expensive materials that are not easy to source and are used in limited quantities, thus leading to smaller sensors.

If the spatial quality of an image is dissatisfactory or it does not meet some pre-defined requirements, this can be changed computationally to some extent with image sharpening techniques. Before delving into the topic, it is necessary to make a premise. In its common meaning, the term *image sharpening* refers to techniques and algorithms that enhance the overall quality of an image while maintaining its original size. Examples of this are histogram equalization and contrast stretching [218]. The problem at hand, however, involves the fusion of images of different sizes, so it is necessary that while the image quality is improved, the spatial resolution increases. In this thesis, we will use the term -sharpening- according to the last definition, identifying the instance in which the resolution of an image is enhanced by exploiting the presence of another image of higher resolution (and captured with a different device) of the same scene.

Another premise regards the technique called Super-Resolution [219]. Super-Resolution algorithms try to enhance the resolution of an image following a pre-determined set of empiric rules (single-image Super-Resolution) or by exploiting the presence of multiple views of the same scene (multi-frame Super-Resolution, typically captured with the same device from different distances and angles). This thesis does not treat this subject.

In conditions of simultaneous data collection, two sensors of different sizes can be deployed at the same acquisition distance, thus leading to the resulting images having different Ground Sampling Distances (GSD). This is the typical case of remote sensing imaging, where spectral sensors are frequently used in combination with a monochrome or RGB camera with significantly better GSD. In this context, to *sharpen* an image means to increase the spatial resolution (overall quality and image size) of a lowly resolved image by injecting spatial details coming from a more highly resolved image [201].

Depending on what source is deployed as a high-resolution image, the task at hand takes on different names. A typical problem in remote sensing is *pansharpening* [220], in which a lowly resolved multispectral or hyperspectral image is sharpened using the spatial details of a panchromatic image generated by a monochrome broadband camera. In the instance in which two hyperspectral images are captured at the same time *hypersharping* [221] can be performed by injecting the spatial details of the more highly resolved image (typically the one deploying a silicon-based sensor).

The simplest sharpening technique, which not coincidentally is used as a starting point by more advanced techniques is a bicubic interpolation. Here, the new image of higher resolution is obtained by extrapolating the pixel values in the gaps created when the image is upsampled. This roughly upsampled image is then refined according to four main classes of pansharpening methods [201]: Component Substitution, Multiresolution Analysis, Variational Optimization, and the always-present class of deep-learning methods, for which here a description is not provided.

The first pansharpening methods to be developed belong to the class of Component Substitution techniques [201]. Here, the lowly resolved image is projected into a different space in which its spatial and spectral components are separated and the spatial component is replaced by the spatial structures present in the panchromatic image. The main challenge presented by these techniques is the potential introduction of spectral distortions induced by the new spatial structure which are almost blindly injected. To reduce this effect, new-generation Component Substitution algorithms usually consider a prior segmentation of the image and have changed the injection rule by adopting statistical modeling.

Much robust spectral quality can be achieved with Multiresolution Analysis approaches [201], which deploy multi-scale decomposition of the panchromatic image. The way in which the low-resolution version of the panchromatic image is obtained usually characterizes the methods belonging to this class. On the downside, the spatial quality that can be obtained is not usually at the same level as Component Substitution approaches. A significant improvement has been brought to Multiresolution Analysis techniques by including sensor-specific parameters (such as MTF) to estimate the involved images at resolutions different from the original.

A general improvement, coming at the same time with much higher computational complexity, is brought to pansharpening techniques by approaches based on Variational Optimization [201]. Here, the pansharpening problem is cast as an optimization of a sensor model that relates the originally acquired panchromatic and lowly resolved images to the desired output image of higher resolution. The problem cast as such is heavily ill-posed, meaning that the straightforward inversion of the model leads to noise amplification. Therefore, it is crucial to constrain the optimization by including prior beliefs regarding the desired highly resolved image.

The nature of pansharpening problems is an obvious landing field for the development of Deep Learning techniques [222]. Pre-trained neural networks represent now the standard in remote sensing applications since the generated high-resolution images are produced with very low spatial and spectral distortions. The

common challenges belonging to deep learning approaches apply to this set of problems as well: the need for large amounts of data, pre-labeling of training data, and the need for generalization to different targets.

Most of the techniques developed for pansharpening can be adapted to work in the case of hypersharpening problems with little effort. However, when we are presented with two hyperspectral sets, one of which needs to be sharpened, a simple question arises: What is the panchromatic image? [221] The hyperspectral set of higher resolution should produce a 2D image that has to work as a spatial reference for the lowly resolved hyperspectral set. It is however clear that given the internal variation of channels present in images that span large spectral intervals, the panchromatic image cannot be the same for all bands (this is better illustrated in Chapter 3). With this reasoning in mind, two approaches are proposed to tackle the problem of panchromatic image generation, producing a panchromatic image for each band of the lowly resolved set.

In the first criterion, denominated *band selection*, the most appropriate band is directly extracted from the set of highly resolved images. The suitability of an image can be subjected to different user-dependent choices, but the resulting panchromatic image is typically the band that minimizes a distance or maximizes a similarity. Typical approaches can involve the minimization of RMSE or the maximization of correlation, structural similarity, or mutual information.

The second criterion, denominated *band synthesis*, generates a panchromatic image from the highly resolved set, but without picking a specific spectral band. This methodology is more prone to user creativity, as the criteria for generation can abide by specific needs hardly generalizable to other instances. Examples of band synthesis include the average along the spectral dimension of the whole or partial set and an image related to a component of PCA, ICA, and NMF. A synthesized band can also be generated by an optimization problem that again can aim at minimizing a distance or maximizing a similarity. In this case, the optimization can learn weights that are assigned to each spectral band of the highly resolved set.

In the general formulation of a hypersharpening problem (Equation 2.54), the goal is to estimate band-wise the weights w that transform the highly resolved hyperspectral set H into its panchromatic version \hat{p}_i , based on the properties of the lowly-resolved image h_i . The hyperspectral set H has M bands, whereas the set h has N bands. For completeness, it is necessary to express that each band of h is affected by random noise b_i , although many methodologies assume it to be negligible.

$$\hat{p}_i = \sum_{k=1}^M w_{i,k} \cdot H_k + b_i \quad (2.54)$$

When the criterion of band selection is applied, the band \bar{m} that better satisfies the imposed conditions (Equation 2.55 illustrate correlation) is selected, and thus the weight $w_{\bar{m}}$ is set to 1 whereas the remaining weights are set to 0. The comparison is usually carried out between the band of h interpolated at the resolution of H , indicated by h_{exp} and a version of the original high-resolution image H that has been sequentially downgraded and interpolated back at its original resolution, indicated with \tilde{H}_{exp} .

$$\bar{m} = \arg \max_m \text{corr}(h_{n,exp}, \tilde{H}_{m,exp}) \rightarrow w_{i,k} = \begin{cases} 1, & \text{if } m = \bar{m} \\ 0, & \text{otherwise} \end{cases} \quad (2.55)$$

The application of band synthesis requires solving the optimization problem reported in Equation 2.56 to find the weights related to the i^{th} band of h :

$$\arg \min_{w_i} |h_{i,exp} - \tilde{H}_{exp} \cdot w_i|^2 \quad (2.56)$$

Evaluating the quality of hypersharpening

The ill-posed nature of the sharpening problem, which welcomes infinite solutions, poses a quite enigmatic riddle regarding the evaluation of the final sharpened products. Relating perceptual metrics to the case of multi-band imaging (sometimes referring to non-perceptual image domains like IR) is not a viable way. Moreover, if an image needs to have its resolution increased in the first place, it means that a high-resolution version is not available. This plainly translates to the fact that evaluating the results of pansharpening and hypersharpening is a complex matter because of the lack of ground truth.

The most popular way of evaluating the performances of pansharpening algorithms relies on the generation of the ground truth from the very dataset used. Indeed, the procedure known as Wald's protocol [223] treats the originally captured image of low resolution as the ground truth, while at the same time, both originally captured sets are degraded to lower resolutions. By doing so, it is expected that by sharpening the now degraded version of h the solution should approximate h itself. Because of the reduction of scale involved, the evaluation is said to be conducted at Reduced Resolution. A schematization of Wald's protocol is illustrated in Figure 2.34.

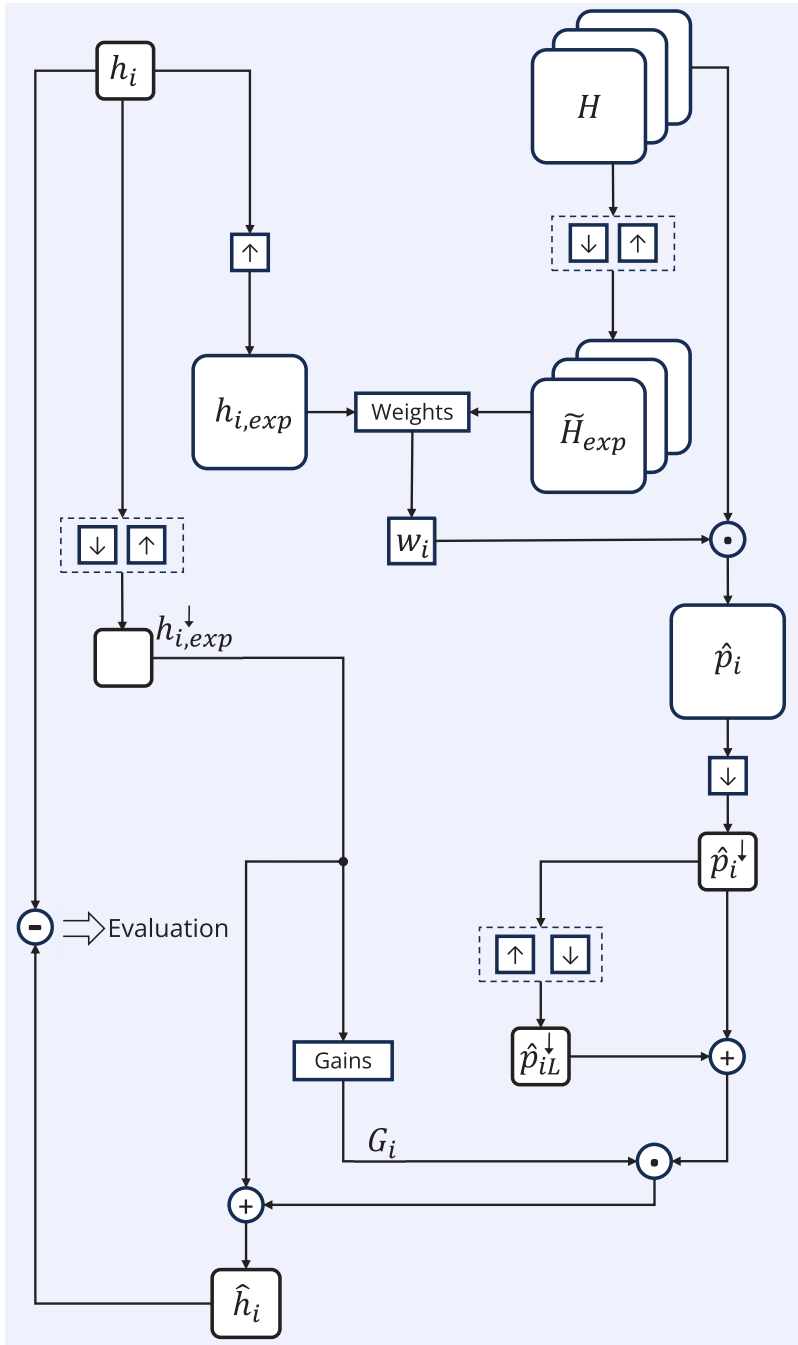


Figure 2.34: Example of hypersharpening the i^{th} band of a lowly resolved hyperspectral image, following Wald's protocol and a typical multiresolution analysis approach.

Wald's protocol is based on two fundamental properties: consistency and synthesis. Consistency is evaluated by degrading the sharpened image and spectrally comparing it to the originally captured h . Synthesis assumes that the sharpened image would be the result of acquiring h with a larger sensor. This last property leads to the protocol being based on a quite adventurous assumption, namely the scale-invariance of the methodology. Indeed, it is implied that if a series of methods can be ranked according to their performances at a certain resolution scale, then the same ranking will be maintained at different resolutions.

Over the years, effort was poured into developing strategies to evaluate the product of sharpening algorithms at Full Resolution, where there is no direct comparison with a reference image. Regardless of the lack of ground truth, the developed measures are designed in a way that the measured distortion would be zero if the ground truth was actually present. The most used Full Resolution evaluation measures are Quality with no Reference and Hybrid Quality with no Reference [224, 225].

2.7 Spectral and Polarization Imaging

The term polarization has come up several times during this dissertation when introducing different topics. We have previously defined the polarization of light in its classic wave-related meaning as the direction of oscillation of the electric field perpendicularly to the direction of propagation. We have also defined the status of polarization in the more intuitive cases of linear, circular, and elliptical polarization for single waves, whereas when bundles of waves with different polarization states move jointly, the resulting radiation is described as unpolarized.

Another instance in which we encountered polarization is when we discussed the Fresnel coefficients of reflectance and transmission that define the behavior of an optical interface in response to incoming light at different incident angles. Particularly, specular reflections with a defined polarization state can be generated by the interaction of unpolarized incident light at the Brewster angle. Finally, we have introduced how deploying carefully designed polarization filters in front of a camera can aid in reducing the presence of specular reflection in a scene.

Adding polarization information to a scene can be a valid asset to perform tasks such as image segmentation and classification of materials, which in the context of Heritage Science can help in possibly detecting materials used in undocumented conservation treatments, and allow material classification where specular reflections are detected and computationally removed.

The main component that enables the characterization of polarization information of a scene is a polarization filter that is placed in front of the camera, typ-

ically denominated a Polarization State Analyzer (PSA). When polarized light passes through a linear PSA, the registered intensity depends on the relative angle between the oscillation angle of polarization and the orientation of the nanowires that usually constitute the filter. When the main axis of the filter is aligned to the polarization of light, the recorded intensity I is the same as the input I_0 , whereas if the filter is consequently rotated by 90° , then no light passes. Equation 2.57 described the variation between minimum and maximum points following the Malus law.

$$I = I_0 \cos^2(\theta) \quad (2.57)$$

The three fundamental θ -dependent properties of linear polarizers are their transmittance, extinction ratio (often referred to as contrast ratio), and polarization efficiency. Transmittance can be defined as the ratio of light that passes through at a given rotation angle, thus defining T_1 for the maximum transmittance, and T_2 in its relative perpendicular direction. The contrast ratio is defined as T_1/T_2 , whereas the polarization efficiency can be thought of as a Michaelson contrast for polarizers since their expressions are similar: $(T_1 - T_2)/(T_1 + T_2)$.

Recalling the formulation of the dichromatic model (Equation 2.28), the recorded intensity corresponding to light reflected off a surface I_r can be split into an unpolarized diffusive component I_d , and a potentially polarized or partially polarized specular component. Thus, if a polarizer is rotated at an angle θ in front of a camera, the intensity measurement can be described as in Equation 2.58.

$$I_r(\theta, \lambda) = \frac{1}{2}I_d(\lambda) + I_{sp,c}(\lambda) + I_{sp,v}(\lambda) \cos 2(\theta - \phi) \quad (2.58)$$

Here, the specular component is further split into a constant part relative to the angle of the polarization filter $I_{sp,c}$, and a variable term representing the amplitude of a cosine function $I_{sp,v}$, which also depends on the polarization angle of the incoming light ϕ [226]. Figure 2.35 proposes a simulation of this measurement.

Polarimetric information can be extracted from a scene by capturing images while at the same time rotating a polarization filter in front of the camera. Stokes imaging can be performed to describe the polarization state of reflected light by estimating the first element of the Stokes vector [4] $S = [S_0 \ S_1 \ S_2 \ 0]^t$. In particular, S_0 describes the total power of the incoming light, while the S_1 and S_2 components express the difference between intensities measured through orthogonal directions of the polarizer.

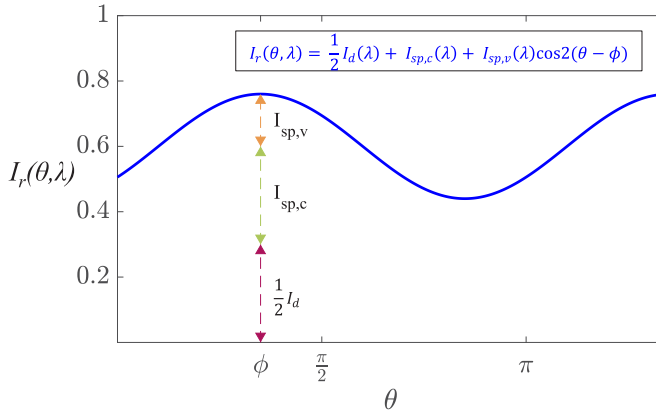


Figure 2.35: Simulation of a partially linearly polarized light passing through a linear polarizer after a surface reflection.

Equation 2.59 describes the pixel response of a single PSA to a particular input Stokes vector.

$$I = A S \quad (2.59)$$

where I is the pixel response, S is the polarization state of the input light, and $A = \begin{bmatrix} a_0 & a_1 & a_2 & 0 \end{bmatrix}$ is the analyzer vector which embeds the polarizer characteristics such as transmission, polarizing angle, and extinction ratio coefficient. If assuming ideal transmission and extinction ratio, the analyzer vector is only a function of the rotation angle, and Equation 2.59 can be rewritten as in Equation 2.60.

$$I(\theta) = \frac{1}{2} \begin{bmatrix} 1 & \cos 2\theta & \sin 2\theta & 0 \end{bmatrix} S \quad (2.60)$$

An important feature of Stokes imaging is that it does not necessarily require polarized incident radiation, although polarized light would allow us to infer depolarizing properties of the examined material. On the other hand, Mueller imaging, which can be thought of as the next step of Stokes imaging, compulsorily requires the deployment of polarized radiation. The result of Mueller imaging is a 4×4 matrix that not only describes the polarization properties of the reflected light but encompasses the polarization properties of the material. Moreover, Mueller imaging requires the presence of a Polarization State Generator which manipulates the polarization state of the incoming light, so that the characterization involves more than the linear states, but also circular and elliptical states.

Since it is known that the intensity follows a cosine-like law, it is not necessary to collect images of a scene rotating the PSA at a fine sampling. Indeed, it has

been demonstrated how only three measurements are needed to recover the Stokes vector efficiently and with little levels of noise [227]. What is crucial is to capture images at meaningful rotation angles of the PSA, which means spanning uniformly sampling from the period of the cosine function, between $[0, 180]^\circ$.

If for example, four measurements are performed, the intensity vector is defined by:

$$I = \begin{bmatrix} I_{\theta_1} \\ I_{\theta_2} \\ I_{\theta_3} \\ I_{\theta_4} \end{bmatrix} = \mathbf{W}S = \frac{1}{2} \begin{bmatrix} 1 & \cos(2\theta_1) & \sin(2\theta_1) & 0 \\ 1 & \cos(2\theta_2) & \sin(2\theta_2) & 0 \\ 1 & \cos(2\theta_3) & \sin(2\theta_3) & 0 \\ 1 & \cos(2\theta_4) & \sin(2\theta_4) & 0 \end{bmatrix} \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ 0 \end{bmatrix}$$

in which \mathbf{W} is the analysis matrix that combines the four analyzer vectors A and θ_{1-4} are the polarizer angles of the four PSA configurations. Equation 2.61 shows that the Stokes vector $\hat{S} = [\hat{S}_0 \ \hat{S}_1 \ \hat{S}_2 \ 0]^t$ can be computed from I for each pixel in the image.

$$\hat{S} = \hat{\mathbf{W}}^+ I. \quad (2.61)$$

in which $\hat{\mathbf{W}}^+$ is the pseudo-inverted PSA matrix estimation. The Stokes vector offers the possibility to assign intuitive parameters, namely the degree of polarization ρ and the angle of linear polarization ϕ (Equation 2.62).

$$\rho = \frac{\sqrt{S_1^2 + S_2^2}}{S_0} \quad \phi = \frac{1}{2} \arctan\left(\frac{S_2}{S_1}\right) \quad (2.62)$$

The capturing of purely polarimetric images has been made much more compact by the development of Polarization Filter Arrays (PFAs), following the principle of Spectral Filter Arrays in the context of MSI. A full polarimetric image can then be recovered by demosaicing. The extension of polarimetric imaging to embrace the spectral domain is not straightforward [228], as the various studies in combining it with MSI, and thus developing Spectral Polarization Filter Arrays, demonstrate. Fusing polarization and RIS is something for which we are not technologically ready if we consider the problem from a sensor development standpoint. The subsequent acquisition of hyperspectral images at different angles of a PSA, although time-consuming, is a valuable alternative to further characterize the reflective properties of materials [229].

When fusing polarization and RIS, it is crucial that the selected PSA is able to work within the spectral range of the imagers. By construction, linear polarizers tend to exhibit broadband properties more than their circular counterparts, which are typically built with anisotropic materials that introduce a phase delay between the polarization components. This requirement, which makes the capturing of spectral information related to circular polarization hard, at the same time highlights the challenges of building a Mueller spectral imaging setup.

Chapter 3

An overview of needs and limitations

After a long excursion in the theory that governs the capturing and analysis of RIS data, it is now the moment to delve into the more practical aspects that are treated in this dissertation. By doing so, it is perhaps beneficial to re-explore the research questions proposed in Chapter 1 under a more critical and knowledgeable lens. With this chapter, we want to dissect the aspects that work synergically toward the achievement of the ultimate goal of this thesis, e.g. to allow conservation science practitioners to deploy VNIR-SWIR RIS in the most advantageous way.

The ultimate goal of this thesis is to allow conservation science practitioners to deploy VNIR-SWIR RIS in the most advantageous way. The definition of this goal, in relation to the state-of-the-art knowledge at our disposition, leads us to further define five sub-goals related to different aspects:

To better keep the discussion level on track, we introduce an example in which two hyperspectral images of a scene representing a historical textile have been captured. The goal is to obtain a fused hyperspectral image \mathbf{F} defined in the image domain \mathcal{D} with spatial coordinates (x, y) and spectral sampling in correspondence of wavelengths $\lambda \in \Lambda$ defined in $[\lambda_{min}, \lambda_{max}]$. Equation 3.1 illustrates how $\mathbf{F}(x, y, \lambda)$ can be computed by applying an appropriate transform \mathcal{T} to the individually captured hyperspectral images $\mathbf{V}(x_v, y_v, \lambda_v)$ and $\mathbf{S}(x_s, y_s, \lambda_s)$, respectively associated with the VNIR and SWIR spectral ranges.

$$\mathbf{F}(x, y, \lambda) = \mathcal{T} [\mathbf{V}(x_v, y_v, \lambda_v), \mathbf{S}(x_s, y_s, \lambda_s)] \quad (3.1)$$

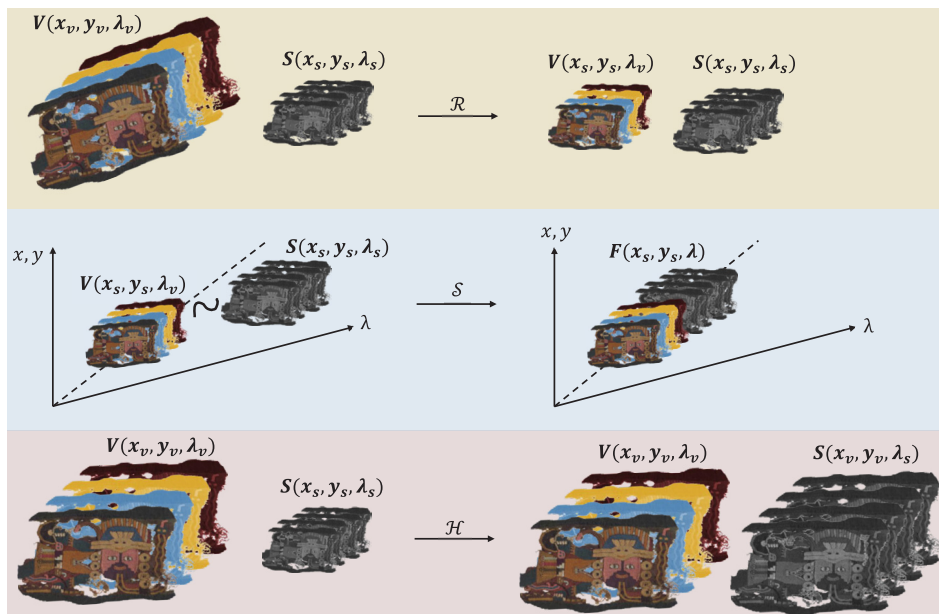


Figure 3.1: From top to bottom, schematization of image registration (\mathcal{R}), splicing (\mathcal{S}), and hypersharpening (\mathcal{H}) for a pair of hyperspectral images of a historical textile.

The role of \mathcal{T} is broad since it can represent the combination of various transformations such as image registration, spectral splicing, and hypersharpening. In the following, the individual transformations will be indicated with \mathcal{R} , \mathcal{S} , and \mathcal{H} respectively. Figure 3.1 summarizes the purposes of the three transformations.

RQ1a: What are the factors and decisions that influence the performance of spatio-spectrally aligning two hyperspectral images coming from two different sources and electromagnetic ranges?

Spatial alignment

When the concept of image registration was introduced, the key aspects were explained in terms of image source and camera viewpoint. When the nature of the two images to be aligned is the same, we have an instance of unimodal imaging, whereas multimodal imaging represents a case in which the two images are captured by different systems deploying diverse imaging models. Analyzing the issue at hand, it is possible to observe that aligning \mathbf{V} and \mathbf{S} presents at the same time features that can be traced back to instances of unimodal and multimodal imaging. It is unimodal because the nature of the image is the same, as both cameras

capture the scene irradiance that later is converted to reflectance. Moreover, the presence of overlap between the spectral ranges can be a factor of unimodality strengthening. On the other hand and with the due limitations, a dual RIS system can sometimes be thought of as a multimodal imaging example. Although \mathbf{V} and \mathbf{S} are captured following the same imaging principle, the resulting spatial content is sometimes not preserved when comparing single bands related to far regions of the electromagnetic spectrum extracted from \mathbf{V} and \mathbf{S} .

Based on these observations, it was decided to term this specific registration problem as *mildly* multimodal. The presence of a spectral overlap in which the two systems supposedly reach a measurement agreement (it is known that they do not) cannot always be taken for granted, and multimodal features tend to prevail when the reflectance properties are considered at distant wavelengths.

Multimodal image registration can be addressed straightforwardly by maximizing a similarity measure when the two images have a comparable resolution. When this is not the case, template matching techniques [230] can be deployed to find an area within the reference image that contains the target image. As the issue at hand is not strictly a template matching instance (the field of view of \mathbf{V} and \mathbf{S} is generally maximized experimentally), multimodal registration techniques must address the difference in resolution between moving and fixed images.

When feature-based methods are considered, it is important to point out that extracting spatio-spectral features can have questionable usability, given that the inherent spectral characteristics refer to different wavelengths. Traditional feature matching methods such as SIFT used for unimodal image registration are claimed to possess certain robustness to local changes in intensity but showcase a limited range of action when it comes to matching features extracted from images at different wavelengths. As illustrated in Figure 3.2, SIFT benefits from the usage of bands extracted at similar nominal wavelengths, as the number of matched SIFT points quickly decreases as the distance from the junction wavelengths is increased.

Operating around the ends of the available spectral ranges, as the SIFT method would suggest, is often considered a risky practice, given that the typical SNR curves of the sensors tend to quickly decrease due to a decrease in quantum efficiency and the increase of noise sources related to optical aberrations and second-order distortions, especially in RIS systems that deploy dispersive elements. In Section 2.5.14 we have introduced the concept of hyperspectral redundancy and how it affects negatively a handful of operations mostly related to computational efficiency and data storage/compression. In this case, however, it is possible to exploit redundancy in its positive role, as it can be used to generate a new image

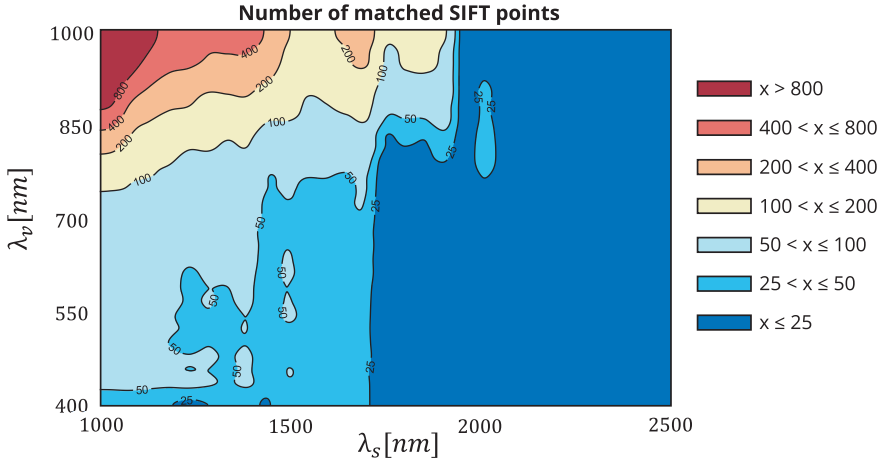


Figure 3.2: Number of matched SIFT points obtained by sampling the individual bands of \mathbf{V} and \mathbf{S} . As we approach a potential overlap or the junction wavelength, the number of obtained matches quickly increases.

with reduced noise.

The generation of new images can be traced back to the concept of image synthesis introduced in Section 2.6.3 when discussing hypersharpening. Indeed, the problem of registering \mathbf{V} and \mathbf{S} can be tackled with the same paradigms of band selection and band synthesis. However, before either selecting the two most correlated bands or synthesizing a suitable reference-target pair, it is crucial to decide the *arrival venue*, or in other words, the resolution at which the images are aligned.

When an artifact is digitized, the proposed guideline is usually to strive for the highest possible spatial resolution. Upscaling \mathbf{S} without the application of sharpening techniques, and thus using simple bicubic interpolation, can result in a scenario in which the evaluation of the final result is quite complicated, and the computational and storage costs increase drastically. On the other hand, downscaling \mathbf{V} can lead to a significant information loss, as the ratio of resolutions between sensors working in the visible and infrared is typically included between 4:1 and 8:1 [231].

When a hypersharpening problem is tackled following the band synthesis [221] approach, for each band of \mathbf{S} , the algorithm seeks a suitable linear combination of bands of \mathbf{V} to generate the panchromatic image, and thus the image pair is formed by an individual band of \mathbf{S} and a synthesized image from \mathbf{V} . This results in a wavelength-dependent transform \mathcal{H}_λ . As we have introduced in Section 2.6.1, given that the bands of \mathbf{V} and \mathbf{S} are assumed to be internally co-registered, it is not necessary to loop through the bands, as \mathcal{R} is wavelength-independent. Thus,

a single \mathcal{R} can be learned from a suitable pair of images synthesized from \mathbf{V} and \mathbf{S} . However, Equation 3.2 shows that synthesizing two images from \mathbf{V} and \mathbf{S} is a highly ill-posed problem:

$$\arg \min_{u,w} \left\| \sum_{i=1}^{N_{\lambda_v}} (u_i \cdot V_{\lambda_i}) - \sum_{j=1}^{N_{\lambda_s}} (w_j \cdot S_{\lambda_j}) \right\|_2^2 \quad (3.2)$$

The image registration problem cannot thus be addressed straightforwardly with a band synthesis approach. Equation 3.2 can be constrained in the case in which one of the two terms (the synthesized version of \mathbf{V} or \mathbf{S}) is generated through a pre-determined strategy (for instance band-averaging within a specific spectral interval), thus leaving room for uncertainties due to user decisions and variability.

Spectral alignment

Satellites for earth observations are often equipped with a series of detectors sensitive to different portions of the electromagnetic spectrum [232, 233]. In most cases, the captured spectral bands are adjacent or overlapping. It is then legitimate to wonder why with these contiguously sensitive sensors no effort was spent on performing splicing. The reason probably lies within the specific spectral region in which the sensor concatenation would take place. If we observe the responses of typical VNIR and SWIR sensors, their junction point is found at about 950 nm - 1000 nm. These particular wavelengths generally do not allow accurate retrieval of parameters from the Earth's surface due to the high absorption coefficient of water (vapor) [234], and thus it is often preferred to remove the bands and leave *spectral holes* behind.

In the field of spectroscopy, the state-of-the-art splicing correction is represented by the parabolic correction [217] described by Equation 3.3, in which the reflectance curve coming from the system working in the visible (not cooled, and therefore affected by warm-up noise) is connected to the reflectance of the infrared through the applications of coefficients learned as:

$$c_{\lambda_1:\lambda_2} = \frac{((\lambda_1 : \lambda_2) - \lambda_1)^2 \Delta \rho}{\rho_{\lambda_2} (\lambda_2 - \lambda_1)^2} + 1 \quad (3.3)$$

Here, the spectral range indicated by $\lambda_1 : \lambda_2$ is pre-defined based on the sensor characteristics, and thus the correction is able to handle only a limited series of instances. Indeed, the parabolic correction is deemed reliable up to cases in which the spectral difference $\Delta \rho$ is smaller than 6% in its relative terms [213]. For ap-

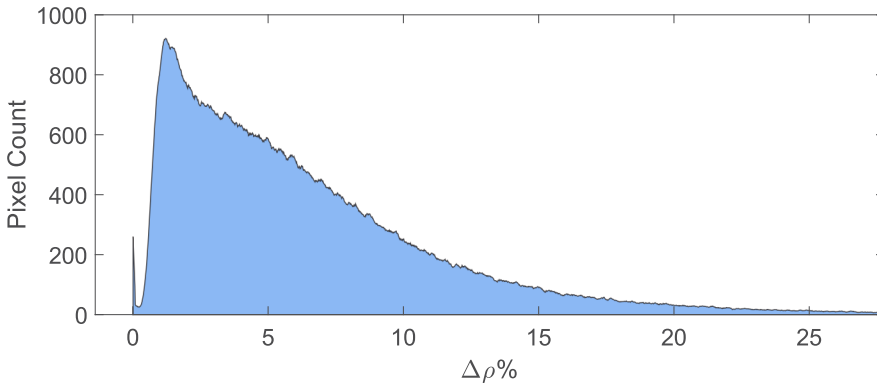


Figure 3.3: Relative spectral discrepancy between an instance of co-registered $\mathbf{V}(x_s, y_s, \lambda_v)$ and $\mathbf{S}(x_s, y_s, \lambda_s)$. More than 40% of pixels have a $\Delta\rho\%$ greater than 6% and would lead to unsatisfactory splicing results when corrected with the current state-of-the-art methods.

plications of spectroscopy, however, this limit is rarely exceeded and the parabolic correction can easily suffice the quality requirements.

As we have discussed in Section 2.6.2, when a spectroscopy system and a RIS system are compared, the sources of noise in the latter are more numerous and amplified. The acquisition geometry of punctual spectroscopic systems is seldom comparable to that of a hyperspectral camera, and this is translated into a higher BRDF variance among neighboring pixels.

When we consider the issue at hand more closely, the fact that the images to be spliced have a different spatial resolution can be considered an additional (and highly significant) source of noise. Indeed, if we consider the instance in which there exists a 4:1 resolution ratio between \mathbf{V} and \mathbf{S} , this means that 16 spatial pixels in \mathbf{V} are condensed into a single pixel in \mathbf{S} . Moreover, the image registration procedure is usually considered successful if the pixel displacement is smaller than the unit value, but even a fraction of a pixel is enough to introduce significant noise and amplify the discrepancy in reflectance values at the pixel level.

Thus, accounting for all the additional noise sources, it becomes highly likely that the relative spectral differences can exceed the limits set in a spectroscopy environment, as Figure 3.3 demonstrates. Consequently, this highlights the need for a dedicated splicing correction for RIS.

A splicing correction able to handle the smooth connection of hyperspectral images should obey the following requirements:

- **Minimal intervention.** Only the wavelengths that are more significantly affected by noise should experience the stronger part of the correction, and thus, spectral bands away from the noisy region should ideally remain unchanged.
- **General.** The correction should be applicable to different types of hyperspectral data (reflectance, transmittance, etc.), while at the same time, it should be able to handle various instances of spectral overlap, if any.
- **Adaptive.** The relative spectral discrepancy depends on the complexity of the scene and on the geometry of the experimental setup. The correction should be able to adapt to different levels of $\Delta\rho\%$.
- **Evaluation.** Two spectra can be connected in infinite ways and as a result, evaluating the product of a splicing correction is not straightforward. A quantitative evaluation paradigm needs to be developed to measure the similarity between computed and expected results. Moreover, it is possible that the classical spectral metrics introduced in Section 2.5.7 cannot fully encompass the properties that need to be evaluated in this specific instance.

RQ1b: What are the implications of sharpening techniques developed in remote sensing when translated to proximal sensing applications?

With the due distinctions, it is possible to draw a number of parallels between remote sensing imaging for Earth Observation and imaging of mostly planar artifacts in the context of CH imaging. Indeed, if today the field of Heritage Science has been enriched by non-invasive analysis through RIS, this is thanks to the developments that have been produced over the years in the field of remote sensing. In Section 2.5.15 we have observed how directly applying an analytical technique such as spectral unmixing without considering carefully the matter at hand can lead to results that are not entirely satisfactory. In particular, we have observed how it is crucial to change the mixing model depending on the application.

Similarly, it is legitimate to wonder if the techniques developed to sharpen lowly resolved satellite images need some level of adaptation prior to being applied in a context for which they have not been tested before. This question stems from the observation that historical artifacts tend to segment their information within a contained spectral range [235, 236]. This is particularly evident for textiles, which are prone to exhibit characteristic features of dyes in the visible range and fiber-related features in the infrared [189, 190]. From a global image standpoint, this is translated to a rather radical change in the appearance of textiles when single-band images are examined individually.

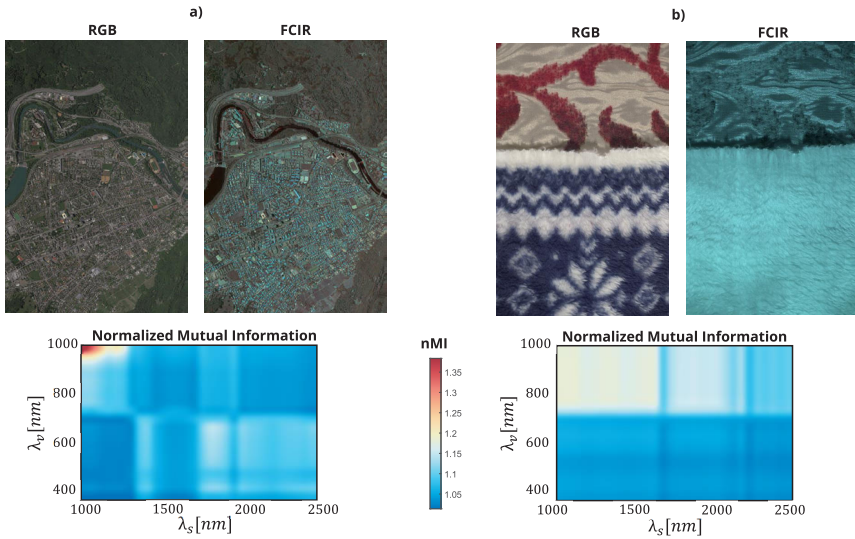


Figure 3.4: Comparison of remotely sensed image (a) and proximally sensed image of a textile (b). The normalized mutual information charts at the bottom show a rather clear information separation taking place at around 700 nm, especially for the textile scene. The comparison between a color rendering (RGB) and an infrared false color rendering (FCIR) of the scenes highlights the non-preservation of spatial patterns in the textile.

Figure 3.4a displays how for a remote sensing image (captured by the Airborn Prism Experiment APEX satellite [237]) the spatial features related to edges are typically preserved when switching from a visible-related visualization to an infrared-related one. On the other hand, in an image of a textile the dye information is not *transmitted* to the infrared range, and thus only the fiber information is preserved (Figure 3.4b).

Blindly applying a sharpening technique to increase the resolution of \mathbf{S} would lead to injecting untruthful spatial information extracted from \mathbf{V} , thus generating the possibility of false-positive detection in a potential classification task.

Besides the consideration of spatial pattern preservation, *blind* algorithms to sharpen images have been commonly accepted because the sharpening problems in remote sensing popularly deal with instances in which both image sources (highly and lowly resolved) belong to the same portion of the electromagnetic spectrum: panchromatic image and multispectral, multispectral and hyperspectral VNIR, etc.

Algorithms belonging to the Component Substitution family of sharpening techniques are highly prone to inject unwanted spatial details onto the lowly resolved image since they exploit the difference between the panchromatic image and a lin-

ear combination of the bands of the lowly resolved image to extract the spatial details to be transferred over [238]. Algorithms belonging to the Multi-resolution Analysis family can be made context-aware by implementing a local correlation measure to avoid the injection of spatial details that are unlikely to be present in the lowly resolved image [239].

As sharpening research is now highly profused by deep learning approaches, questions have been raised about the generality of the methodologies. Recently, some works highlighted the shortcomings of neural networks that had been trained on datasets of specific satellites when sharpening images coming from different detectors, and therefore research efforts were poured into making the networks more flexible [240]. However, what is still missing is an evaluation framework of such methodologies, not only for different sensors but for completely different domains.

RQ1c: Is it possible to develop a paradigm for the joint analysis of VNIR-SWIR RIS and polarimetric imaging?

From the standpoint of capturing radiometrically meaningful data, the ideal surface of an object should be flat. If we follow this reasoning, most historical artifacts would represent sub-optimal examples of objects to be studied by means of RIS. However, it is commonly accepted to allow small radiometric fluctuations as long as the surface is quasi-planar, with the limits typically set by the depth of field of the optics.

As we have observed in Section 2.3.1 and 2.7, the interaction of directional light with surfaces at particular angles of incidence can sometimes generate specular reflections with various degrees of polarization. The presence of specularities can hinder the material analysis and thus a polarimetric imaging system can help to detect and remove such image flaws thus allowing to uncover what is hidden underneath.

The fusion of spectral and polarization imaging systems is non-trivial and has been proposed for applications in food analysis [241], computer vision [242], and target detection [243], but never in a CH framework. Moreover, the application of spectral and polarization systems is usually limited to either multispectral systems or hyperspectral systems working in the visible range [228], and thus the extension to a VNIR-SWIR domain represents uncharted territory.

A key advantage of spectro-polarimetric imaging is that it allows the examination of the polarization-related properties of materials, which can then be used later on as features. An example of a relevant spectro-polarimetric feature that can be extracted from CH objects such as paintings is the correlation between the degree

of linear polarization and the amount of reflectance [244, 245].

The implementation of a polarimetric RIS VNIR-SWIR finds its challenges in the sourcing of appropriate optical interfaces able to provide meaningful polarimetric information in the whole considered spectral range. Indeed, not all polarization filters offer a broadband capability, due to the way the filters are manufactured. Circular polarizers, for example, are quite λ -selective, and thus, the capturing of circular polarization information is limited, preventing the complete implementation of Mueller imaging. Linear polarizers are more broadband and thus can enable the implementation of Stokes imaging. Integrating a polarimetric framework onto an existent RIS one presents challenges that are highly dependent on the available instrumentation, and so, there could be a high variability in the way the system is calibrated. With this question, we want to achieve an efficient spectro-polarimetric calibration by deploying the available imaging setup (introduced in Chapter 4).

RQ2: Can we improve the typical tasks conducted in RIS analysis for CH considering the full extension of the spectral data in the VNIR-SWIR ranges?

A limitation of recent works conducted deploying a dual VNIR-SWIR RIS system is that often the two datasets are analyzed independently. However, it seems that there is a clear reason for that. Historical artifacts like paintings and textiles have a peculiar response to radiation in the visible and infrared ranges. Backing up the separation theory is the fact that there is a strong belief that the visible range is strictly associated with the features that characterize pigments and dyes, whereas the infrared is typically strongly associated with information regarding the fibers of a textile or the layers found underneath the paint layer in a painting. Thus, it is plausible that very little information mixing takes place.

It is however arguable that by adding new features (reflectance values at different wavelengths) to the observations (the pixels), the *cross-talk* of data, intended in its positive meaning, can be beneficial to the detection of underlying patterns that could not be observed while keeping the spectral sets separated. On the other hand, while performing VNIR-SWIR fusion may introduce a series of practical benefits related to pixel correspondence, it is not mandatory that the processing is conducted on the whole extension of the spectral range, and analyses can be carried out on appropriate spectral subsets, depending on the needs and the nature of the studied artifacts.

Chapter 4

Summary of Articles

The first part of this chapter takes a wider overview of the articles that contribute to this thesis. Here, we shall look at a grouping based on their individual scopes, as well as introduce the targets on which the methodologies are demonstrated and evaluated. After that, the imaging setup common to all supporting articles shall be presented in detail. Finally, each supporting article is summarized.

4.1 Articles grouping

The five contributing articles of this dissertation consider three main categories of approaches:

- **Data fusion-oriented.** Articles **A1** and **A2** deal directly with matters related to the technicalities of fusing the RIS data coming from the two independent VNIR and SWIR cameras from the spatial and spectral perspectives, respectively. Article **A4** belongs to this category as well as it concerns the application of hypersharpening to hyperspectral images of CH. Article **A5** is based upon a pre-existent framework of hyperspectral data fusion and implements polarimetric imaging on top of that.
- **Material analysis-oriented.** Articles **A3** and **A5** investigate different material properties exploiting the consolidated fusion framework. The former applies classical RIS analytical techniques such as spectral mapping and spectral unmixing for the study of textiles, whereas the latter examines the spectro-polarimetric properties of a mockup painting.
- **Dissemination-oriented.** Article **A4** finally fits a category of its own, as one of its goals was to introduce the full fusion framework to an audience represented by

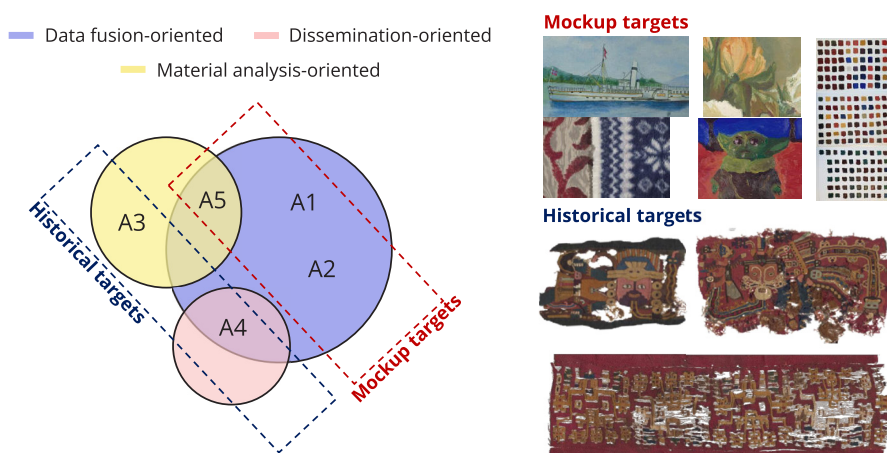


Figure 4.1: Grouping of the supporting articles by scope and by nature of the deployed targets, which are illustrated on the right-hand side of the figure.

the CH community.

An additional article grouping can be performed by considering the nature of the targets that are imaged. Articles **A1**, **A2**, and **A5** focus on works in their first steps of experimentation, so it was decided to not take any chances with potentially damaging historical artifacts. Therefore, these articles make use of mockup targets represented by a series of planar surfaces either fabricated with the goal of being mockups or collected from around the laboratory facilities because of their suitability to the experiments. Opposite to that, Articles **A3** and **A4** dealt with historical textiles. Figure 4.1 illustrates the article grouping and the deployed targets.

The mockup targets include a commercial postcard from Gjøvik (Norway) depicting *Skibladner*, the oldest paddle steamer still in service, a painting of flowers, a textile piece made of cotton from the HyTexiLa database [163], a series of mockup mixtures of seven pigments in different known concentration (data available from [177]), and an oil painting made using mixtures of those same pigments.

The historical targets are three Paracas [246] textiles belonging to a small collection housed at the Department of Collection Management, University of Oslo (UiO).

4.2 Imaging setup

The common base imaging setup for all supporting articles of this thesis is schematically reported in Figure 4.2. The imaging setup deployed in Article **A5** slightly

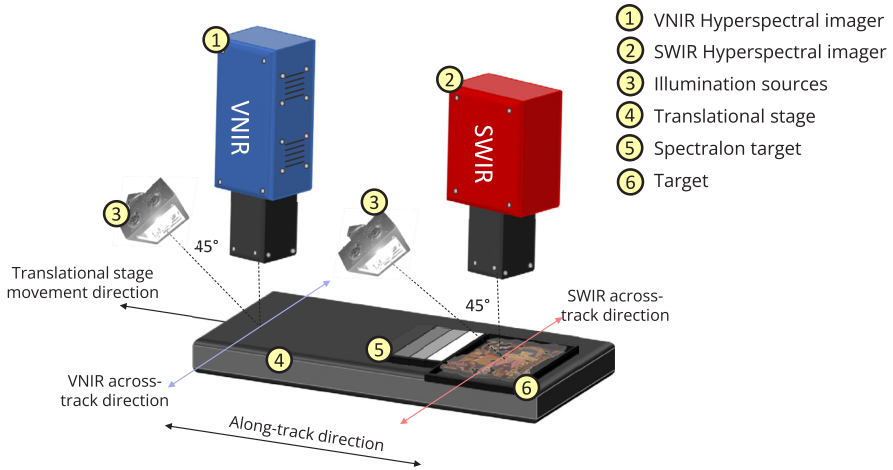


Figure 4.2: Schematization of the imaging setup adopted in the articles of this thesis.

differs from this depiction because of the inclusion of polarization elements and because the hyperspectral imagers are used sequentially and not simultaneously.

An intrinsic advantage brought by the pushbroom system is that the illumination source needs to shine only in a small area corresponding with the line field of view of the camera, thus avoiding unnecessary illumination of parts of the artifact that are not presently imaged. For this reason, it is advisable to use illumination sources that can focus their outputs on small areas. The deployed illumination sources were unpolarized halogen lights for both VNIR and SWIR scenes. These light sources have broadband SPD that can effectively cover the whole extension of the VNIR-SWIR spectrum, but come at the cost of producing heat on the surface due to their infrared content. For what concerns the geometry, the 45/0 imaging Standard is aimed at.

Both hyperspectral imagers are manufactured by Hypspec (Norsk Elektro Optikk, Norway), and their main features are reported in Table 4.1.

The scene laid out on the translational stage is constituted by the target and a Spectralon calibration reference. The translational stage then shifts the scene in the along-track direction at a speed that is synchronized with the framerate of the cameras. The framerate is computed from the selected integration time and potentially by the number of averaging frames. Frame averaging is one of the strengths of this dual RIS system as it allows to drastically increase the SNR by a factor \sqrt{N} for N averaged frames. In this particular case, it becomes advantageous because the high IR content of the halogen SPD typically allows the selection of very short

	Hypex VNIR-1800	Hypex SWIR-384
Sensor	CMOS (Silicon)	HgCdTe
Dispersive element	Grating	Grating
Cooling	NA	150 K
Spatial pixels	1800	384
Spectral bands	186	288
Spectral range	400-1000 nm	950-2500 nm
Overlap range	950-1000 nm	
Bands in overlap	16	9
FWHM	3.19 nm	5.45 nm
FOV (across-track)	17°	16°
GSD @ 30 cm	≈ 50 μm	≈ 200 μm
Noise floor	2.4 e-	150 e-

Table 4.1: Features of the two hyperspectral imagers deployed in this thesis.

integration times for the SWIR camera in comparison to the VNIR camera. So, while the speed of the stage is decided according to the longer integration time, it is possible to perform multiple acquisitions of the same SWIR line, and thus increase the SNR. Another way to experimentally increase the SNR is to accurately estimate the dark current, and this can be performed at the software level by selecting an appropriate number of *dark images* to be averaged before and after the measurement.

The main limitation of the system is represented by the low flexibility to adapt to different sizes of objects to scan, due to the fact that the fore optics only works with pre-designed lenses and no adjustable focus. This provides however an advantage to image quality when the system is used in its ideal conditions.

4.3 A1: Hyperspectral VNIR-SWIR image registration: Do not throw away those overlapping low SNR bands

This article tackles the problem of spatially registering two hyperspectral images **V** and **S** acquired with the setup depicted in Figure 4.2. In particular, rather than focusing on a method survey, the work looks into whether making decisions in a certain way affects the final result, and if it does, how the final performances change according to the decisions taken.

The registration problem is described as a *mild* multimodal instance characterized by a large difference in spatial resolution and small possible translations and ro-

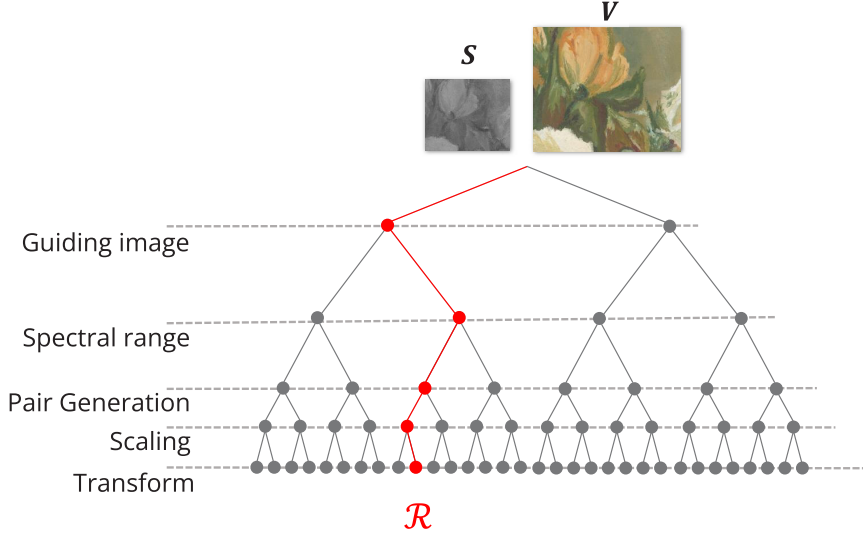


Figure 4.3: Image registration process described as a decision tree.

tations that can be due to the manual positioning of the cameras with respect to the target. The process that brings to learn the transform \mathcal{R} can be depicted as a decision tree in which each branch leads to a different solution to the problem, as illustrated in Figure 4.3. The structure of the tree is characterized by five decision layers, although different registration methods can slightly alter this since they require skipping some layers because of their working principles.

The identified decisional factors are the following:

- **Guiding Image.** The first decision determines the spatial resolution arrival venue, thus which spectral set is to be taken as the reference, and consequently which spectral set acts as the target.
- **Spectral range.** At this step the decision is whether the whole spectral range of \mathbf{V} and \mathbf{S} is adopted or if only the overlapping range is considered, bearing in mind that this range is notoriously affected by a significantly lower SNR for both cameras.
- **Pair Generation.** The moving-fixed image pair is generated at this step, either following the band correlation paradigm (selecting the two most correlated bands of \mathbf{V} and \mathbf{S}) or synthesizing the new bands using the first component image of PCA.
- **Scaling.** The large scale transformation can be learned intrinsically within the

image registration process or can be performed a priori to facilitate the application of area-based methods in a more straightforward manner.

✧ **Transform.** Finally, the last decision concerns what type of transform to apply to the target image. In the simplest scenario, an affine transform should describe the registration problem. If the objectives of the cameras are considered not parallel, then a projective transform is more appropriate.

The decision tree is explored in all its branches for four registration methods (feature-based, spectral feature-based, hybrid, and area-based), and four scenes representative of different complexities. At each iteration, a transform \mathcal{R} is learned and applied to register an image pair arbitrarily generated from the overlapping range. Five evaluation indexes are adopted, four represent image quality metrics measuring different characteristics: Peak Signal-to-Noise Ratio, Ground Sampling Distance Error, truncated structural similarity (structural similarity without luminance component [247]), and normalized Mutual Information. A multiband quality index (ERGAS [248]) is also estimated. The performance results are then investigated in a multivariate analysis framework (PCA).

The main takeaway of the work is that despite showcasing low levels of SNR, the overlapping bands can be considered a valuable asset to perform image registration in the specific instance. A clear correlation between scene complexity and method performance is observed, and among the methods, the one termed Maximum of Wavelet Transform Phase Cross-Correlation (MWTPXC) [210] turned out to be the more consistent. On the other hand, the performance of the technique based on the extraction of spectral features turned out to be poor in comparison to the other methods.

Limitations

In the article, the impact that the paradigm of band synthesis has on registration problems of this kind is most likely understated. This is due to the way the bands are synthesized in the article, i.e. by arbitrarily deciding to compute the PCA and extract the image related to the first principal component. However, the paradigm of band synthesis is more consistent and based on the computation of an appropriate image by solving an optimization problem. It is highly likely that by correcting this, the band synthesis decision would correlate with better performances.

In connection with the previous point, the evaluation is conducted through the usage of an image pair selected arbitrarily. It is plausible that by adopting a more rigorous criterion the analysis of results can be more robust.

Another important limitation is represented by the lack of knowledge of how gen-

eralizable this work is to other examples of VNIR-SWIR image registration that deploy different cameras. For example, a way of quantifying what is the lower limit of admissible noise in the overlapping range is missing.

4.4 A2: Logistic splicing correction for VNIR–SWIR reflectance imaging spectroscopy

This article deals with the spectral alignment that needs to be performed to smoothly connect two co-registered hyperspectral images \mathbf{V} and \mathbf{S} affected by spectral discontinuities or *jumps*.

A first attempt of splicing was performed in the supporting article **S1**: *Radiometric spectral fusion of VNIR and SWIR hyperspectral cameras*. Here, splicing was conducted by applying global multiplicative factors to both VNIR and SWIR curves solving an optimization problem. Although this method could not be generalized to different image contents (the optimization was conducted on standardized targets), analyzing its shortcomings helped in defining the desirable properties of a splicing correction.

Article **A2** starts by reviewing the state of the art in splicing corrections in spectroscopy and by raising awareness regarding the need for a dedicated splicing correction in RIS. This is due to the fact that stronger sensor mismatches take place in an imaging system because of a much higher local BRDF variance and image registration misalignments, as well as the intrinsic image sensor characteristics that differ from sensors deployed in spectroscopy.

A new logistic splicing correction for RIS is proposed. The correction deploys the presence of an overlapping range or a junction wavelength to smoothly connect the spectra at the pixel level. Following the guidelines exposed in Section 3, the new splicing correction aims to be adaptive to different levels of spectral discrepancy and to be respectful of the shape of the original spectra, introducing perturbations only in the spectral bands more affected by spectral jumps. It was found that learning the correcting coefficients in the form of a logistic function would suffice these requirements.

An important contribution of the article is the evaluation of the methodology, typically a tricky topic in cases in which a unique solution does not exist. The ground truth was measured with a Lambda1050 spectroradiometer [139] (Perkin Elmer inc.) that provided continuous spectral data in the overlap range of the two cameras. Using the external data, it was possible to quantify the perturbation brought by the newly proposed correction and compare it against the existing state of the art. The results show that the logistic correction tends to better preserve the ori-

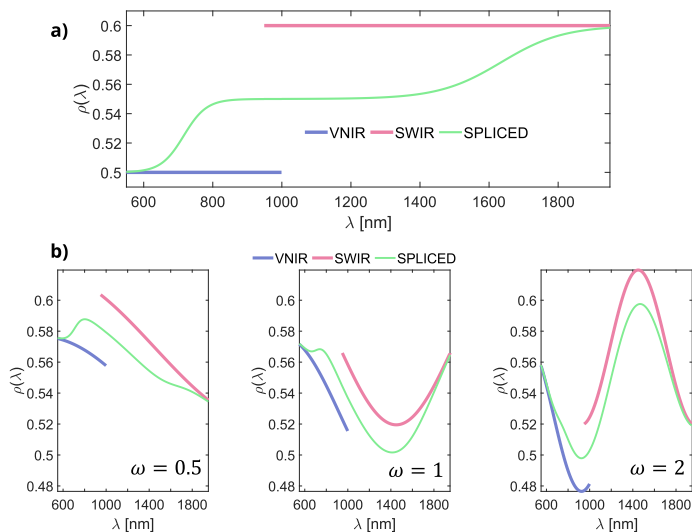


Figure 4.4: Extreme cases in which the proposed splicing correction seems to not perform correctly. **a)** With flat spectra the solution can lie in between, but based on how the correction is formulated, spectral shape differences are introduced to not change the whole extent of the original spectra. **b)** The proposed correction seems to benefit from the local complexity of the curves: as the angular frequency ω of the cosine function is increased, the proposed curves look like more realistic spectra.

ginal spectral shapes in instances of high spectral discrepancies, resulting in being the most robust method when compared to the traditional techniques.

The logistic splicing correction proposed in this article has been made publicly available within a GitHub repository that includes Matlab functions to splice hyperspectral images and single spectra. Link to the repository: <https://github.com/federigr/HyperspectralSplicingCorrection>.

Limitations

Upon testing the correction, a few instances in which it achieved unsatisfactory results were highlighted. When the spectra to be connected are mostly flat (Figure 4.4a), correcting the spectrum with a logistic function can introduce unnatural details. However, we should interrogate ourselves and wonder what the best result should be in this instance: a stepped connection or a flat average spectrum that perturbs the original spectra at all wavelengths? So, for the moment, it seems that the correction works efficiently on more complex signals where introducing small local variations does not provoke significant shape alterations (Figure 4.4b).

4.5 A3: Reflectance imaging spectroscopy for the study of archaeological pre-Columbian textiles

This is the first article that tackles material analysis reaping the benefits of the technical work conducted for articles **A1**, **S1**, and **A2**. In this article, two out of the five Paracas textiles housed at the Department of Collection Management of the University of Oslo are analyzed by means of RIS. The specific analysis is placed in a broader context in which other analytical techniques like Surface-Enhanced Raman Scattering spectroscopy are applied to the artifacts to reveal their chemical composition.

Once a unique hyperspectral image is obtained from the fusion of the individual VNIR and SWIR, several Region Of Interest (ROIs) are identified by an expert in conservation science and familiar with the textiles. Such ROIs are selected based on the belief that they might contain relevant spectral features that can be traced back to dyestuff and fiber identification. Recalling the notion that the surface of a textile is non-flat, it was decided to operate in a spectral variability framework, thus associating with each ROI a bundle of endmember spectra, rather than an individual endmember.

In a preliminary step, the endmembers contained in the spectral library underwent spectral unmixing, to potentially discover if possible dye-mixing took place and to possibly identify *purser* endmembers. Based on the unmixing result and on previous knowledge, a series of selected endmembers was produced. A second set of filtered endmembers was iteratively generated in an automatic manner by exploring the clustering properties of the projection of the endmember ROIs in the KLPD space.

The two selected endmember sets were then used to perform spectral mapping, first considering the whole spectral range, and then individually the visible range for dyestuff mapping, and the infrared range for fibers mapping.

If the results proposed by the analysis of the visible information were of interest, but expected, those coming from the infrared mapping have elicited some surprises. Indeed, by using the automatically selected endmembers it was possible to enhance the presence of an invisible (to the human eye) vertical pattern (Figure 4.5). This certainly represented an exciting result for which the causes are still to be clarified, with possible explanations tracing back to the manufacturing technique of the textile and the possibility that the textile had been folded for centuries.



Figure 4.5: Enhancing the vertical pattern resulting from the fibers mapping in the SWIR range on an RGB rendering of textile UEM37914.

Limitations

In hindsight, this work could have benefited from a technical adjustment, if we consider the context of this thesis. Indeed, the work does exploit the data fusion, but only to go back to considering separately the VNIR and SWIR ranges for the analysis of dyestuff and fibers. Although this was a decision dictated by the nature of the data and the scope of the article in its publication venue, the data-fusion aspect could have been exploited more, for example by subdividing the spectral range differently so as to include spectral bands originally belonging to the different cameras.

4.6 A4: Full VNIR-SWIR hyperspectral imaging workflow for the monitoring of archaeological textiles

This article was initially designed to be a cross-disciplinary exercise, so as to disseminate the work done to fuse hyperspectral images to a CH community represented by the attendance of the Archiving conference. The problem of sharpening proximally sensed images is also tackled, so to have a twofold goal with this work: present a full fusion workflow, and raise awareness towards the fact that not all sharpening methods are alike and lead to similar quality results.

In order to demonstrate the last point, two popular empiric sharpening methods

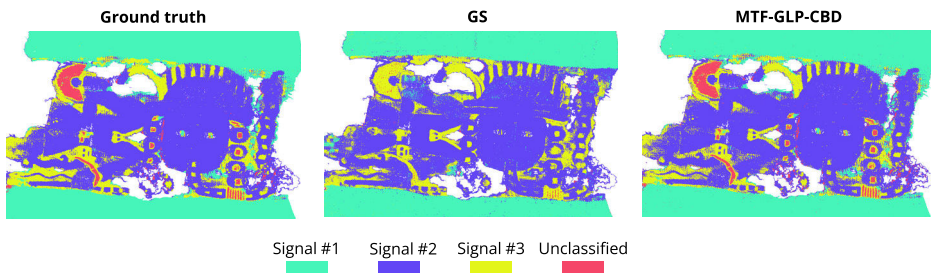


Figure 4.6: Unsupervised spectral mapping of the SWIR scene by extracting the most relevant spectral signatures. The ground truth image is obtained following Wald's protocol, thus allowing a comparison with the sharpened versions.

are selected: the Gram-Schmidt (GS) method, belonging to the Component Substitution family, and the Modulation Transfer Function - Generalized Laplacian Pyramid - Context-Based method, belonging to the Multiresolution Analysis family. Out of the two, only the latter is deemed to be context-aware, and thus it does not blindly inject spatial details from the spatially highly resolved image onto the lowly resolved one.

While both methods are able to generate a visually pleasing enhanced version of the SWIR scene, this should not mislead the user. Indeed, by performing an unsupervised spectral mapping, it was possible to display how the mapping was affected by the sharpening method (Figure 4.6), and how this could lead to a misinterpretation of the data, and to a poor conservation or preservation decision when the GS method is deployed.

Limitations

This work was designed to disseminate information and raise awareness regarding carefully selecting the sharpening methods to be deployed. By doing so, it targeted two selected sharpening techniques, and therefore its robustness could be increased by conducting a comprehensive review of existing methods for applications in CH, assessing strengths and weaknesses in generalizing solutions.

The presented workflow stops at the formulation of the hypothesis that a fused VNIR-SWIR hyperspectral image is a bulky dataset, without addressing topics related to data management and storing that are valuable in the field of heritage archiving. This certainly leaves ground for interesting future work to make hyperspectral imaging a more sustainable technique.

4.7 A5: Relationship between reflectance and degree of polarization in the VNIR-SWIR: A case study on art paintings with polarimetric reflectance imaging spectroscopy

The last article of the collection tackles the implementation of a polarimetric imaging framework on top of the already existing VNIR-SWIR RIS system. High importance is given to the process of spectro-polarimetric calibration since the imaging setup contains polarization elements that cannot be controlled with absolute precision.

Indeed, there is no possibility of knowing the exact angle of the polarizers mounted in front of the cameras, and therefore the first part of the polarimetric calibration is devoted to this experimental estimation. A set of linear polarizers with known orientation angles is placed in the scene as a reference, thus allowing the estimation of the analysis angles by inversion of the cosine law that models the intensity of light transmitted by the filters.

The joint spectro-polarimetric calibration allowed the pixel-wise computation of the spectral Stokes vectors and consequently, the degree of polarization was estimated.

The negative correlation between reflectance and degree of polarization previously studied in the visible range [244] was investigated, considering the full extension of the spectrum, and with a multi-resolution approach. In this way, it was possible to observe a significant influence of the surface topography on the measured correlation, hinting at the fact that observations made with spectroscopic systems might not hold in a much more complex imaging scenario. Moreover, while investigating the spectral behavior of the correlation, it was found that in general, the correlation is present and significant in the visible range, but then it completely disappears (complete decorrelation) in the range recognizable as the NIR, whereas in the last part of the spectrum (SWIR), a more fluctuating behavior was observed.

Limitations

The article focuses most of its efforts on implementing and describing the spectro-polarimetric calibration, which is achieved successfully. While the observations regarding the fluctuation of correlation through the VNIR-SWIR spectrum are interesting, there are still a few questions that need to be addressed. Amongst these questions, it is legitimate to wonder whether this behavior can only be associated with the examined target or if the observations can be generalized. Moreover, the question regarding the causes of the behavior remains unanswered, with only a few speculations in place.

Chapter 5

Discussion

In this chapter, we will take a step back to examine the overall picture drawn by the research described in this thesis. In Chapter 1, two main research questions (four overall questions considering the subdivision of **RQ1**) were defined to accomplish the previously set research goals. The purpose of the present chapter is to reflect upon the proposed RQs and examine how our knowledge regarding the treated topics has changed, thanks to both the contribution of the published articles and the insights gained in three years of research in which it is honestly admissible that more was learned from the mistakes rather than from the successes.

The second part of the chapter takes a further step back and reflects on the global contributions and impact that this thesis has/could have in the field of Reflectance Imaging Spectroscopy for the applications to Cultural Heritage artifacts.

5.1 Revisiting the Research Questions

RQ1: How to efficiently build a fused hyperspectral image starting from two hyperspectral images in the VNIR and SWIR?

The first RQ proposed in this thesis was tackled in practice from different angles. For this reason, it was decided to address this RQ by splitting it into three sub-questions that treat different topics that once interlaced all cooperate towards the resolution of **RQ1**. Thus, addressing the sub-questions of **RQ1** deals with considerations related to the imaging setup and processing. The fusion of VNIR and SWIR RIS takes place in five main steps, two related to image capturing (setup designing and inclusion of additional modalities), and three related to data processing (spatial registration, spectral splicing, and resolution sharpening). The identification of these steps alone can be considered a way to answer **RQ1**, but the

details of each individual block are explained in the following sections.

RQ1a: What are the factors and decisions that influence the performance of spatio-spectrally aligning two hyperspectral images coming from two different sources and electromagnetic ranges?

At the start of Chapter 3, the fusion problem (Equation 3.1) of the two hyperspectral images $\mathbf{V}(x_v, y_v, \lambda_v)$ and $\mathbf{S}(x_s, y_s, \lambda_s)$ was defined through the implementation of a generic transform \mathcal{T} that acts on both \mathbf{V} and \mathbf{S} to generate a fused hyperspectral image $\mathbf{F}(x, y, \lambda)$ that presents sharp spatial features, indicating good single-image quality, and smooth, discontinuity-free spectra that point at realistic spectral features for the examined objects. The transform \mathcal{T} can be split into the contributions of individual transforms, identified as image registration (spatial alignment, \mathcal{R}), spectral splicing (spectral alignment, \mathcal{S}), and hypersharpening (resolution enhancement, \mathcal{H}).

Learning the transform \mathcal{R} boils down to estimating the homography matrix that geometrically warps one spectral set onto the reference system (same point of view and same spatial resolution) of the spectral set assumed as reference. Since \mathbf{V} and \mathbf{S} are individually internally co-registered, a single homography is sufficient to spatially align all spectral bands involved.

In article **A1_c**, we have identified in detail a series of decisional factors that can alter the performances of image registration at the processing level. An additional factor that was identified was the impact of the image scene complexity, but that is in principle not controllable, although it can help assess if fused VNIR-SWIR RIS can be a suitable tool for an accurate digitization and material analysis.

The spectral redundancy that characterizes hyperspectral images can be considered an asset rather than an obstacle to processing since it helps us denoising those low SNR bands present at the extremes of the spectral ranges because of the manufacturing limits of the sensors. In the context of image registration, this represents an advantage to the deployment of feature-based techniques that are claimed to be robust to magnitude differences but are indeed facilitated by more similar responses, found indeed in correspondence with similar nominal wavelength bands.

Spectral alignment, or splicing, is needed due to a series of noise sources that result in the non-matching output of sensors measuring the same quantity, i.e. spectral reflectance at a given wavelength. The transform \mathcal{S} in this case is applied to both spectral sets \mathbf{V} and \mathbf{S} in order to first harmonize their spectral responses in the overlapping range (if present), and then to accordingly modifying the remaining of the spectra in a respectful way, finding a fine balance between maintaining the original spectral shapes and obtaining a plausible solution of the connected spec-

trum. The driving parameter in splicing is the relative spectral difference between the two responses, and thus a correction should be effective for a wide range of discrepancies that might take place in a RIS scenario.

Besides designing a suitable correction, our task is also to facilitate the processing at the experimental level by adopting an appropriate imaging setup whose goal is to maximize the BRDF similarity between the two independent imaging devices. By doing so, a beneficial effect is also brought to the image registration problem.

As introduced in Section 2.3.7, the BRDF describes the proportion of reflected light off of a surface as a function of illumination and observation angles (Equation 2.26). If we consider an imaging system constituted by a scene and a classic DLSR camera, we can approximate the camera objective as a single point, so that each pixel, translated into real-world points in space is defined by two observation angles, the azimuth angle (elevation) and the zenith angle. Throughout the scene, the two angles vary according to the position of the pixels. In a pushbroom system, the variation is experienced only for the angle that changes in the across-track direction (the zenith), since the relative position in the along-track direction of pixels is constant with respect to the sensor. For this reason, we can think of a line scanning system as a geometry simplifier. This can be considered an advantage for material analysis, but certainly constraints the usage of line scanners for the rendering of perceptually ultra-realistic images.

During the time in which this thesis was conducted, we have experimented with different imaging setups for VNIR-SWIR RIS, and we have been able to highlight the strengths and shortcomings of each one of them, thus allowing us to decide on the final image setup described in Section 4.2. A schematization of the four setups that we have analyzed is reported in Figure 5.1.

Figure 5.1 does not report the case in which the cameras are deployed subsequently leaving the acquisition geometry unchanged. This instance, although it is a valid option, involves a level of instrumentation handling that is normally discouraged.

The key features that can be used to assess the pros and cons of each setup are speed of acquisition, impact on the examined target, and similarity of the BRDF. The setup depicted in Figure 5.1a is the one adopted in all contributing articles. The same geometry allows the maximization of the BRDF as long as the alignments are respected at the practical level. This is not the fastest setup available, since it deploys only one illumination source per camera, which at the same time allows being more *gentle* on the artifact as the developed heat is more controlled. On the other hand, the setup of image Figure 5.1d could be deemed risky because of the amount of radiation incident on the examined artifacts, even if it is the fastest and

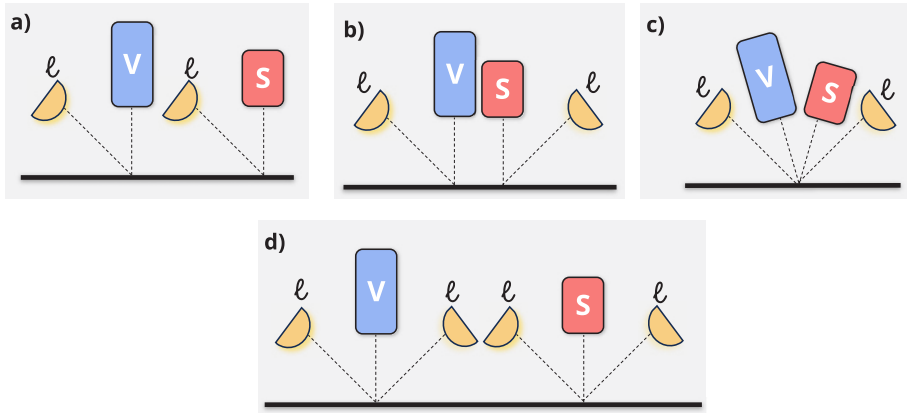


Figure 5.1: Schematization of four identified possible imaging setups for VNIR-SWIR RIS. The letters V and S indicate the VNIR and SWIR cameras respectively, while the letter ℓ indicates a light source. **a)** Single-light, same geometry (setup currently adopted). **b)** Single-light, symmetric geometry. Sometimes deployed in fixed frameworks. **c)** Double light, co-boresighted. **d)** Double light, same geometry.

allows the capturing of a highly similar BRDF.

Sometimes, because of practical needs, fixed frameworks can be deployed in which, for the sake of time and space constraints, the imaging setup resembles the illustration of Figure 5.1b. In this case, however, the captured BRDF of the VNIR and SWIR setup is not the same. Moreover, when the topography of the surface cannot be approximated to a flat plane, height reliefs cast shadows that are imaged symmetrically by the two cameras. This would then require an additional step in image pre-processing to carefully treat the shadow area before proceeding to analyze the pixels involved. It is worth pointing out that line-scanning imaging can introduce some differences in the way shadows are perceived when compared to a DSLR camera. When a standard camera is deployed to image a scene, the length of the shadows cast by the reliefs will depend on the height of the obstacles and on the distance between the light source and the obstacle. Moreover, the proportion of umbra and penumbra will also be affected by the same factors. When the same scene is scanned by shifting it across the field of view of a line scanner, the distance between obstacles and the light source will be the same for all reliefs. Thus, the resulting shadow distribution will be different than the one acquired by a staring system. Moreover, the ratio of umbra and penumbra can be considered constant for all reliefs. This last property makes the pushbroom system approach a system where the illumination source is at an infinite distance and can be a useful insight in studies in which the goal is to retrieve the topography of the examined scene.

The co-boresighted system depicted in Figure 5.1c can be considered the fastest among the four, although a risk assessment regarding incoming radiation should be performed. Moreover, it suffers from BRDF differences due to different observation directions and the need to assume a projective transform when registering the images, thus allowing an additional degree of freedom that can result in noisy homography learning.

As we have highlighted in article A1c, the scene complexity intended as surface topography has an impact on the performance of image registration, which in turn affects the spectral discrepancies that need to be corrected during splicing. Thus, if adopting different imaging setups changes the perception of the imaged surface, a careful assessment must be conducted when a specific artifact is presented. In the instance in which the surface is approximately flat, the setup depicted in Figure 5.1b can suffice the requisites if the assumption of an isotropic BRDF is in place. Otherwise, the other BRDF-preserving setups should be privileged.

Lastly, a note on the illumination source. In the setup illustrated in Figure 4.2, halogen lights are deployed individually on the fields of view of the VNIR and SWIR cameras. These lights are used for their IR content, but if the purpose is to extract spectral reflectance, nothing forbids the usage of a different light source for the VNIR range, as a lower IR emission would be beneficial to the heat development of the surface. It is important however to ensure geometry alignment between the two independent setups and always operate in a camera and illumination-independent space like reflectance, as the comparison of radiance or irradiance could be meaningless.

RQ1b: What are the implications of sharpening techniques developed in remote sensing when translated to proximal sensing applications?

In Chapter 3 we have observed the intrinsic differences that exist between a remotely sensed image of the Earth surface and that of a proximally sensed textile. In particular, we have shown how typically, the spatial patterns related to edges tend to be preserved in the former instance when passing from a visualization in the visible to the infrared, while in the latter this does not happen due the reflective properties of dyes and fibers. This lead us to the observation that the techniques developed for sharpening remotely sensed images could not be straightforwardly applied to the Cultural Heritage imaging domain.

However, a category of sharpening methods deemed context-based is claimed to be able to selectively neglect the spatial details that are injected onto a lowly resolved image. Upon testing of a sharpening technique belonging to this family (Modulation Transfer Function - Generalized Laplacian Pyramid - Context Based) in

article **A4_c**, it seems that the performances are satisfactory, both from a visual and analytical standpoint. Thus, the need for a dedicated hypersharpening technique for CH artifacts does not seem so urgent.

A few options can be worth exploring to improve the performances. For example, most empirical sharpening techniques are based on the inherent properties of specific sensors as they are developed on a case-specific basis. The MTF of a sensor is usually the leading characteristic to compute intermediate images deployed in sharpening, and thus a prior sensor characterization would be necessary before being able to work at full capacity. This can be hard to achieve without proper instruments and sensor manufacturers might not be so willing to share the information. Neural Networks, although they too would benefit from sensor models, could be a workaround to develop a CH-specific sharpening method. The usage of Neural Networks trained on satellite images for the sharpening of proximally sensed images is questionable, given that the sharpening performances already decay when a Neural Network is deployed on images acquired with a satellite imager not present in the training set.

RQ1c: Is it possible to develop a paradigm for the joint analysis of VNIR-SWIR RIS and polarimetric imaging?

To answer the question with brevity: yes. In article **A5_j** we have designed a system for the capturing of hyperspectral images in the VNIR and SWIR jointly with polarimetric information. Due to space limitations, the setup had to be designed in a way that allowed the sequential capturing of VNIR-polarimetric and then SWIR-polarimetric data. This is the first point where improvement can be sought after.

The imaging paradigm that we proposed can be adapted to the instrumentation and needs of other research institutions since it is highly dependent on the availability of specific components. However, the robustness of the spectro-polarimetric calibration was the main point that we wanted to showcase. Indeed, the usage of reference polarizers with known orientation angles in the scene proved to be an economic solution to estimate the angle of the analysis polarizer in front of the camera, there where the orientation of this last component cannot be known a priori due to its manual mounting and rotation. An analysis polarizer could be mounted on a rotational stage, but that would encounter a series of practical challenges that could end up degrading the overall image quality.

As we discussed in Chapter 3, the ultimate goal of studying material-related polarization properties would be to recover the full Mueller matrix, while at the moment, only the Stokes vectors can be recovered at the pixel level. However, this is already a good starting point that allows us to study the polarization properties

of the reflected light and can thus enable the speculation of hypotheses regarding the optical properties of the examined artifacts. The development of a Mueller hyperspectral-polarimetric system is something that in the foreseeable future could not be accessible, but starting from the study of similar multispectral systems [249] it is possible that such technology will be available in the years to come.

RQ1c finds its connection to the broader reach of this thesis in the way spectral splicing is inserted in the pipeline. If we examine the imaging paradigm, we have introduced how it is possible that the analysis polarizer in front of the camera is controlled manually. This means that it is highly unlikely that the rotation angle of the VNIR and SWIR analyzer matches, thus disabling the possibility of splicing the reflectance images as they are captured.

A requisite of splicing is indeed that the two spectral sets must capture the same quantity in the same (or as much as possible from an experimental standpoint) conditions. When discussing **RQ1a** we have introduced that it could be possible to deploy different illumination sources, reminding that spectral fusion takes place in a space that discards the influence of the illumination. Similarly, before applying splicing in a polarimetric context, it is necessary to discard the effects of the polarizing filters placed in front of the cameras. The Stokes space can represent such an environment, so the spectral Stokes vectors undergo splicing to obtain a unique fused version of spectral Stokes images between 400 nm and 2500 nm.

When we examine the Stokes vector of a pixel, the $S_0(\lambda)$ component is representative of its reflectance, while $S_1(\lambda)$ and $S_2(\lambda)$ indicate the difference between intensities measured through orthogonal directions of the polarizer. When the degree of polarization of a pixel is low, it is usually due to $S_1(\lambda)$ and $S_2(\lambda)$ being small and flat. These two characteristics tend to trigger challenging scenarios (high relative spectral discrepancy and *simple* spectral shape) for the proposed splicing correction, as we have introduced in Chapter 4. Thus, it could be worth exploring a dedicated spectral Stokes splicing correction.

RQ2: Can the performances of the typical tasks conducted in RIS analysis for CH be improved by considering the full extension of the spectral data in the VNIR-SWIR ranges?

In Chapter 3, we have highlighted a lack of joint spectral analysis when RIS is deployed in the VNIR and SWIR range for applications in Cultural Heritage, explaining that especially in the instance of textiles the two ranges are used separately to conduct dyestuff and fiber analysis, respectively. On the other hand, we have argued how adding additional features to a system of observations could lead to the enhancement of patterns that could not be observed when considering the feature

sets independently.

Considering the breadth of **RQ2** and the related state-of-the-art status, it was not possible to fully address this question directly. However, the analytical tasks conducted in article **A3_j** helped to gain some important insights that are hereby discussed. Our observations are based only on the task of spectral mapping, as spectral unmixing in article **A3_j** was performed only on the visible range, according to a prior belief that fiber mixing was not yet discovered as a practice by pre-Columbian civilizations in South America. The spectral mapping performed considering the whole extension of the spectral range from 400 nm to 2500 nm could segment meaningful areas, but in the end, was not as informative as the mapping conducted on the SWIR range. In this particular case, this could also be due to the selected endmembers, which in the case of the SWIR analysis were in inferior numbers.

So, to concisely answer **RQ2**, it does not seem that there is a clear advantage to using the full range in classic analytical tasks such as spectral mapping and unmixing. To elaborate on this conclusion, we have identified three main reasons:

1. **Wrong domain of application:** As we discussed at length, the spectral response of textiles tends to segregate the information related to colorants and fiber substrates. Colorants showcase relevant spectral features in the range 400 nm to 780 nm, while fibers, which appear mostly achromatic, tend to show characteristic peaks and valleys in the deeper infrared. The semantic nature of the information that can be extracted from historical textiles, however, seems to be endangered when a wider spectral range is considered, and thus the interpretation can become more challenging. For this reason, we label this particular application as *wrong*: not because VNIR-SWIR RIS should not be applied at all, but because the analysis of full-range spectra is not more advantageous than considering carefully selected spectral subsets.
2. **Curse of dimensionality:** Adding features to an already high-dimensional space is sometimes counterproductive. This is known in the field of data science as *curse of dimensionality* [250] and its first repercussion affects the modeling perspective of spaces. Indeed, as the space dimensionality increases, the data points become more spread out and the sparsity of the system increases, thus making it difficult to gather observations in those data-lacking areas to draw conclusions. Thus, to an increase in dimensionality, an increment in observations should follow. The *curse*, besides bringing a higher demand for computational complexity and data management in all instances, is particularly effective when feature engineering approaches (or in other words, when new features are created by already existing ones) are

adopted without prior considerations, or when the new measured features brought into the system do not possess a meaningful statistical relevance and thus contribute only to increase the noise in the system. With this question, we want to assess if joining spectra at the pixel level actually positively contributes to having statistically meaningful observations.

3. **Effect of increased dimensionality on spectral metrics:** When dimensions are added to two vectors, their quantitative comparison changes according to the nature of the newly introduced features and their interaction with the existing ones. An observed effect on historical textiles is the significant spreading of the KLPD projections when the full range is considered against the individual VNIR and SWIR ranges (Figure 5.2). This can be interpreted as a consequence of the curse of dimensionality, whose more pronounced effects are visible as a stretching along the ΔG (spectral shape) axes.

Indeed, spectral distances based on non-linear operators such as cosine distance (spectral angle) and correlation tend to exhibit a behavior that is quite challenging to predict and interpret when the number of elements of a spectrum increases. Thus, the deployment of such evaluation techniques needs to be carefully considered prior to usage. In Section 5.3, we consider an instance that potentially can help in increasing the interpretation of similarity measures of wide spectra.

4. **Mixing model validity:** Specifically for the case of spectral unmixing, the mixing model assumes a central role in the assessment of the performances. Given the typical complexity with which the painting layer and the dyestuff are applied on their respective substrates, it could be possible that the current state-of-the-art mixing models are not able to describe the phenomenon accurately across the whole spectral range. Perhaps a revision of the full approach is necessary, since typically the task at hand is to *unmix pigments*, whereas this is hardly achieved when considering spectral ranges that encompass multiple structures and signal sources.

5.2 On the order of \mathcal{R} , \mathcal{S} , and \mathcal{H}

In article A4_c, we presented a workflow that operated at low resolution in order to show the difference in classification performances of different sharpening methods. Here, the pipeline followed these steps: registration deploying the SWIR spectral set as a reference, splicing correction, downscaling following Wald's protocol for evaluation purposes, and hypersharpening at the original SWIR resolution. However, in a real scenario application, the desired output is produced at the

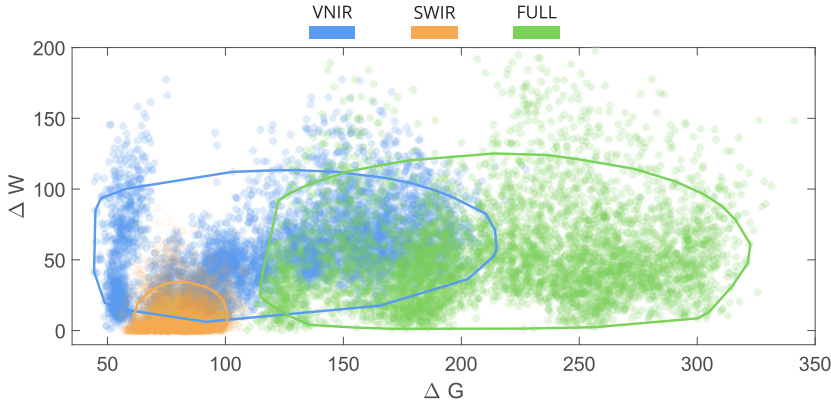


Figure 5.2: KLPD projections and corresponding convex hulls of spectra included in the spectral library of one of the Paracas textiles studied in article A3j. Fusing the VNIR and SWIR spectral information results in a spreading of the projections along the ΔG axes.

original spatial resolution of the VNIR spectral set. In this instance, following the above workflow is possible, but careful considerations are necessary.

The first observation is that both splicing and hypersharpening operations require the two spectral sets to be spatially aligned. Image registration is thus the first step, although it is important to point out that its evaluation can be trickier at full resolution compared to when it is performed at LR due to the presence of extrapolated pixels.

With this in mind, we are left with two alternatives:

1. $\mathcal{R} \rightarrow \mathcal{H} \rightarrow \mathcal{S}$
2. $\mathcal{R} \rightarrow \mathcal{S} \rightarrow \mathcal{H}$

Alternative 1 could find its challenges in the application of the hypersharpening step. Here, it is possible that without prior splicing, the pixel values of the un-spliced bands of the VNIR set would lead to the generation of sub-optimal pan-chromatic images. On the other hand, alternative 2 might encounter a noise amplification issue since the splicing step is applied on extrapolated pixels that did not get treated in the sharpening step. We have therefore an instance in which both splicing and sharpening need the other to work more efficiently. In this scenario, the development of a joint step that solves both problems iteratively could be beneficial.

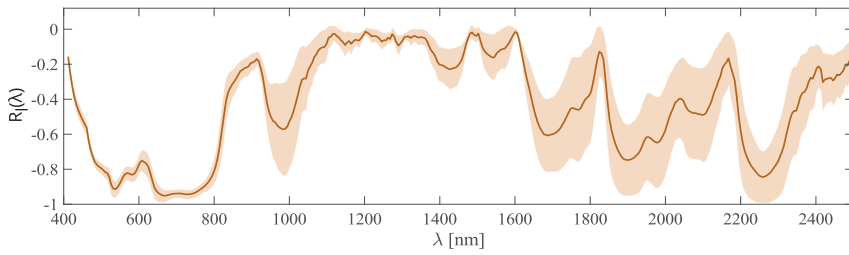


Figure 5.3: Rolling spectral correlation between the first component of the Stokes vector S_0 and the degree of polarization, extracted from article **A5_j**. In this instance, the rolling spectral window is approximately 104 nm wide (19 equally sampled bands).

5.3 On the usage of rolling metrics

Connecting to the last point in the above list, we want to reflect upon the interpretability of a single number as a result of the application of a spectral metric. In many instances, a single spectral angle or spectra correlation value is used to indicate the similarity of two spectra, but it is legitimate to wonder about the interpretability of this when the dimensionality of the data increases in a way like in the fusion of VNIR-SWIR RIS, where a feature doubling takes place.

To tackle this issue, in article **A5_j** we have tried to use a rolling spectral correlation (taking inspiration from the research on seismology) to identify regions in the spectra in which two vectors are more or less correlated. As we illustrate in Figure 5.3, which is extracted from **A5**, with this visualization it can become straightforward to highlight the spectral regions that are more or less correlated.

It is however arguable that the implementation of rolling metrics could represent a challenge in a computational environment (for example, what is the most appropriate spectral window width?), but likely a direction worth exploring to analyze wide spectra without recurring to dimensionality reduction techniques.

5.4 Visualization opportunities

As we have discussed the risk that the curse of dimensionality poses to fuse two already large datasets and how historical artifacts cast a *curse* of their own to the wide spectral analysis, we finally try to look for the silver linings offered by the definite information separation that takes place in textiles.

Over the recent years, museums and exhibitions from around the world have started expressing interest in fully digitizing their collections towards the production of digital twins [251]. Digital twins can include virtual 3D reconstructions of an artifact, long lists of metadata regarding its manufacturing, history, discovery, con-

servation treatments, and preservation conditions, as well as an array of complementary visualizations that enhance different structures. Among these, it is likely that in the instance of a historical textile, a simultaneous visualization of mapped dyes and fiber-related signals could be of high interest (Figure 5.4). The interest does not stop at the contribution to the digital twin constitution, but such visualizations (rendered with the appropriate procedures) could be used in the context of an exhibition to showcase exactly where different dyes and fibers concur.

Displaying a lot of information at once is both exciting and challenging. As observers, we only have a limited attentional span and the overcrowding of information can be too much to digest at times. Hyperspectral images are inherently high-dimensional entities and as such represent a difficult target for visualizations. The usage of visualization techniques that do not find a place in a manuscript like this thesis (videos, animations, interactive tools, etc.), can offer however a nice gateway to explore visualization possibilities in dissemination and exhibition contexts. Among them, the International Image Interoperability Framework (IIIF) [252] pioneers the integration of images represented in layers and their corresponding metadata, while web-based and offline viewers such as Mirador [253] and Butterfly Viewer [254] allow the interactive exploration of scenes represented with multimodality. The usage of animations and interactive tools have historically helped the visualization of data obtained with RIS, and many opportunities, limited only by our creativity, await.



Figure 5.4: Visualization of the simultaneous mapping of six dyes and four fibers on sample UEM37914 of the UiO Paracas collection examined in article A3_j. The dyes are color-coded, whereas the different fiber signals are highlighted by the overlaying of different textures. As previously discussed, the signal labeled as -Fiber #1- could be attributed to a mechanical event (folding) rather than an actually different fiber material.

Chapter 6

Conclusion

6.1 Breadth of the Research

The term -mononumerosis- has been informally coined to indicate those instances in which, in order to describe a spectral comparison, the whole evaluation is trusted upon the usage of a single value computed by a metric [255]. By doing so, it is possible that only one side of the story can be told, with the risk of having an incomplete picture of the observed phenomenon. Similarly, being prone to tunnel vision when trying to bring developments to an already established technology can be counter-productive. That is why this thesis work was always conducted with the primary goal of being at the service of Cultural Heritage, and thus a multidisciplinary, multi-approach layer helped keep the full picture always available. In doing this, the constant interface with experts in their sector was crucial, not only to reach a valuable publication but to organically grow shared knowledge.

This thesis tackled the fusion of two techniques of RIS (VNIR and SWIR) for applications in Cultural Heritage, with studies that ranged in various aspects of scientific research such as technical-oriented, application-oriented, and dissemination-oriented. Moreover, within the technical-oriented approaches, the problem at hand was addressed from different angles that concerned different aspects of imaging, starting from the initial setups to the implementation of processing steps and the addition of further layers of complexity represented by spectral image multimodality.

6.2 Contributions

The multi-approach featured in this thesis has the main result that the highlighted contributions belong to different aspects of research:

- **Understanding of imaging setup:** Although this particular aspect did not stem into a research publication, the insights and reflections that were gathered are extremely valuable. Putting in perspective the implications that each imaging setup has in combination with the complexity of the scene is crucial when assessing the capability of a RIS system, as it influences the quality of intermediate processing steps and can change the outcome of later analysis.
- **Spectral splicing :** A splicing correction dedicated to hyperspectral images was something that was lacking in the landscape of RIS processing. With the newly proposed logistic correction, we hope that more and more users can analyze VNIR-SWIR spectral data more efficiently and in different ways.
- **Polarimetric RIS:** The proposed practical implementation of a VNIR-SWIR polarimetric RIS system allowed for the retrieval of high-quality spectral and polarimetric data and can then represent a base framework for further analysis.
- **Sharpening for CH:** The consequences that specific sharpening methods developed for remote sensing, namely methods belonging to the Component Substitution family, have on the spectral of the final sharpened image are known. However, the application of such methods and the impact of these consequences on classification tasks performed on historical artifacts had not been showcased before. This contribution is indeed a natural consequence of observations performed in a different imaging domain, but a much-needed one in the Cultural Heritage domain, where new imaging techniques start to emerge.

6.3 Impact of the Research

Assessing the impact of research is not an easy task, especially in the early stages and in the period immediately after the work is concluded. As it is not possible to discuss how much the research presented in this thesis has impacted its field of application, we can offer a reflection on the tangible impact that we have observed so far and on the potentialities that lie ahead.

The contributing articles of this thesis are presented in an order that does not follow the chronology of publication but rather follows a semantic division that helps in having this dissertation flow more freely. However, it is clear how the first two articles **A1_c** and **A2_j** are instrumental to the production of the subsequent ones. Indeed, the lessons learned regarding image registration and the developed splicing correction are deployed to form the pre-processing basis in articles **A3_j**, **A4_c**, and **A5_j**.

The splicing correction proposed in article **A2_j** can be considered to have the potentially highest immediate resonance, as the concept of splicing is not all that

popular in RIS. One of its strengths is that it is not limited to Cultural Heritage applications, and it can be used to fuse spectral data of different natures, provided that the BRDF (or BTDF, in case of transmittance data for example), is matched between the deployed imaging systems.

With article **A5_j**, we have tackled the fusion of RIS with polarimetric imaging in the VNIR-SWIR, opening what it is possible to think of as a new framework to compute characteristic spectro-polarimetric features of materials commonly used for the constitution of historical artifacts at the imaging scale.

Another way of assessing research impact more practically is to analyze the publication venues of the different contributing articles. Article **A1_c** was published at the *Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing* (WHISPERS), an international conference organized yearly by IEEE that gathers the community working on RIS, mostly in the domain of remote sensing, but that in recent years has opened up to various proximal applications such as food analysis and Cultural Heritage. Participating in WHISPERS allows for gathering insights regarding the new trends and feedback from experts in both the academic and industrial research areas.

The decision to publish Article **A2_j** in *Optics Letters* (Optica group) stemmed from the need to provide a concise solution to the problem of spectral splicing. The journal publishes exclusively short communications (four pages) that can have an immediate resonance in the field they address, also thanks to the accumulated notoriety that the journal acquired over the years (2022 Impact Factor 3.6).

Article **A3_j** has been submitted with the goal of being published in a well-established venue in the field of Heritage Science.

Article **A4_c** was published in the context of the *Archiving Conference* organized by the Society for Imaging Science and Technology, which gathers different profiles of practitioners in preservation. Thus, it was thought of as an optimal venue to present and disseminate the conducted work, while at the same time raising awareness regarding the sometimes tricky applications of advanced image-enhancing techniques.

Finally, article **A5_j** has been submitted aiming for a publication in an established multidisciplinary venue, given the transversal topics that it addresses.

A more circumscribed vision of impact can be thought of as the collaboration between researchers from different fields. Through the development of this thesis, research groups have been brought together to interact, collaborate, and most importantly, transfer high-level knowledge. This kind of impact does not immediately

resonate at the community level, but it still cannot be overlooked.

6.4 Future directions and perspective

During this thesis, we have discussed in detail different research questions, and while trying to answer them, we have perhaps opened new paths leading to more questions that can be addressed in the future. Here, we try to summarize new directions and perspectives that are worth exploring to push the research in combined VNIR-SWIR RIS forward.

➤ **Band-synthesis from two spectral sets:** Image registration is highly dependent on the two images that lead the learning of the connecting transform matrix, and thus, an optimal way of generating two suitable images from the two spectral sets can be devised. The problem is ill-posed by nature, but the introduction of constraints and carefully designed regularization rules could help in simplifying the search for the appropriate image pair.

➤ **Shadow treatment:** When technical constraints are applied to the imaging setups and complex surfaces are imaged, it is possible to face a sub-optimal situation in which the BRDF does not match at the pixel level. If this situation cannot be avoided and analysis is still needed, then a methodology must be devised to treat those areas that show non-matching lit and shadow areas. A shadow detection and correction framework can potentially exploit the image difference at the spectral overlap similar to the case of splicing.

➤ **Joint splicing-sharpening:** In the previous Chapter, we have highlighted how splicing and sharpening need each other to work at full capacity. A processing step that implements both simultaneously is therefore in the realm of possibilities and worth exploring.

➤ **Dedicated sharpening for CH:** Although it does not seem that there is a need to develop a dedicated empiric sharpening technique for CH, it is possible that a leap forward can be achieved by deploying Neural Networks specifically trained on proximally sensed objects. Since it is demonstrated that Neural Networks struggle with data from different satellite sensors, it is foreseeable that an application of Neural Networks trained for remote sensing and deployed on historical artifacts would not perform as well as their empiric counterparts.

➤ **Mixing models in extended spectral ranges:** As much as the validity of commonly used mixing models such as the Kubelka-Munk theory is established in describing the way paints mix in the visible range, more investigation is necessary to explore the way spectral signals interact when multiple substrates are considered once paints and pigments stop having significant responses in the infrared

range. The definition of an appropriate mixing model, not necessarily behaving uniformly wavelength-wise, can be crucial to solving the unmixing problem in the whole VNIR-SWIR range.

➤ **Widening of users pool:** The equipment that enables hyperspectral imaging is typically a wish item at many institutions around the world, but it is rarely within budget. In the context of the current global crises, efforts should be spent on increasing accessibility to technology, which has already begun with the Iperion HS consortium [256] but can be followed by preventive actions in those areas where wars and climate disasters endanger people's lives and the sustenance of their identity through their heritage.

➤ **Spectro-polarimetric features on mockups:** Mockups have always represented a cornerstone for research in heritage science, and in this instance, they should be no less. The computation of the spectral Stokes vector achieved with VNIR-SWIR RIS can offer a venue for the design of features, and the correlation between reflectance and degree of linear polarization can already be considered one of them. It is however crucial to assess if there exists a feature variance that will allow the discrimination of materials otherwise indistinguishable.

Bibliography

- [1] UNESCO General Conference, *Draft Medium-term Plan, 1990-1995: General Conference, Twenty-fifth Session, Paris, 1989*. Paris, France: UNESCO, 1989.
- [2] ICOMOS, “ICOMOS glossary.” <https://www.icomos.org/en/2016-11-10-13-53-13/icomos-and-the-world-heritage-convention-4>, 2023. [Online; accessed 24-August-2023].
- [3] F. Lenzerini, “Intangible cultural heritage: The living culture of peoples,” *European Journal of International Law*, vol. 22, no. 1, pp. 101–120, 2011.
- [4] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Saint Louis: Elsevier Science & Technology, 1980.
- [5] P. C. Hobbs, *Building electro-optical systems: Making it all work*. John Wiley & Sons, 2009.
- [6] R. H. Kingston, *Detection of optical and infrared radiation*, vol. 10. Springer, 1978.
- [7] W. R. McCluney, *Introduction to radiometry and photometry*. Artech House, 2014.
- [8] K. Takahashi, A. Yoshikawa, and A. Sandhu, “Wide bandgap semiconductors,” *Verlag Berlin Heidelberg*, 2007.
- [9] R. G. Driggers, M. H. Friedman, and J. Nichols, *Introduction to infrared and electro-optical systems*. Artech House, 2012.

- [10] M. Strlič, “Heritage Science: A future-oriented cross-disciplinary field,” *Angewandte Chemie International Edition*, vol. 57, no. 25, pp. 7260–7261, 2018.
- [11] Technical Committee CEN/TC 346: Conservation of Cultural Heritage, *Conservation of cultural heritage - Conservation process - Decision making, planning and implementation*. European Committee for Standardization, 2017.
- [12] P. Merico, M. Faccoli, and G. Cornacchia, “Analysis of ancient slag inclusion-metal systems as a method to disclose processing thermochemical parameters: The case study of a medieval lombard steel bar from northern Italy,” *Metallurgical and Materials Transactions B*, vol. 54, no. 3, pp. 1408–1421, 2023.
- [13] A. Nanetti, “Defining heritage science: A consilience pathway to treasuring the complexity of inheritable human experiences through historical method, AI, and ML,” *Complexity*, vol. 2021, pp. 1–13, 2021.
- [14] B. Cornelis, A. Dooms, J. Cornelis, F. Leen, and P. Schelkens, “Digital painting analysis, at the cross section of engineering, mathematics and culture,” in *2011 19th European Signal Processing Conference*, pp. 1254–1258, IEEE, 2011.
- [15] American Institute for Conservation of Historic and Artistic Works, *Code of Ethics and Guidelines for Practice*. AIC, 1994.
- [16] Federal Agencies Digitization Guidelines Initiative, *Technical Guidelines for Digitizing Cultural Heritage Materials, Third Edition*. FADGI, 2023.
- [17] ISO/TR 19263-1:2017, *Photography - Archiving Systems - Part 1: Best Practices For Digital Image Capture Of Cultural Heritage Material*. ISO, 2017.
- [18] DS/ISO/TS 19264-1:2017, *Photography - Archiving Systems - Image Quality Analysis - Part 1: Reflective Originals*. ISO, 2017.
- [19] F. Di Cicco, M. W. Wijntjes, and S. C. Pont, “Understanding gloss perception through the lens of art: Combining perception, image analysis, and painting recipes of 17th century painted grapes,” *Journal of vision*, vol. 19, no. 3, pp. 7–7, 2019.
- [20] N. Proctor, “The Google Art Project: A new generation of museums on the web?,” *Curator: The Museum Journal*, vol. 54, no. 2, pp. 215–221, 2011.

-
- [21] M. Hughes-Hallett, C. Young, and P. Messier, "A review of RTI and an investigation into the applicability of micro-RTI as a tool for the documentation and conservation of modern and contemporary paintings," *Journal of the American Institute for Conservation*, vol. 60, no. 1, pp. 18–31, 2021.
- [22] L. Lanteri and C. Pelosi, "2D and 3D ultraviolet fluorescence applications on cultural heritage paintings and objects through a low-cost approach for diagnostics and documentation," in *Optics for Arts, Architecture, and Archaeology VIII*, vol. 11784, pp. 156–165, SPIE, 2021.
- [23] M. Alfeld, "MA-XRF for historical paintings: State of the art and perspective," *Microscopy and Microanalysis*, vol. 26, no. S2, pp. 72–75, 2020.
- [24] C. Margariti, "The application of FTIR microspectroscopy in a non-invasive and non-destructive way to the study and conservation of mineralised excavated textiles," *Heritage Science*, vol. 7, pp. 1–14, 2019.
- [25] C. Cucci, J. K. Delaney, and M. Picollo, "Reflectance hyperspectral imaging for investigation of works of art: Old Master paintings and illuminated manuscripts," *Accounts of chemical research*, vol. 49, no. 10, pp. 2070–2079, 2016.
- [26] R. Qureshi, M. Uzair, K. Khurshid, and H. Yan, "Hyperspectral document image processing: Applications, challenges and future prospects," *Pattern Recognition*, vol. 90, pp. 12–22, 2019.
- [27] A. Papanikolaou, D. Dzik-Kruszelnicka, S. Saha, and M. Kujawinska, "3D digital image correlation system for monitoring of changes induced by RH fluctuations on parchment," *Electronic Imaging*, vol. 2021, no. 18, pp. 65–1, 2021.
- [28] J. D. Cutajar, A. Babini, H. Deborah, J. Y. Hardeberg, E. Joseph, and T. Frøysaker, "Hyperspectral imaging analyses of cleaning tests on Edvard Munch's monumental Aula paintings," *Studies in Conservation*, vol. 67, no. sup1, pp. 59–68, 2022.
- [29] M. González-Cabrera, A. Domínguez-Vidal, and M. Ayora-Cañada, "Monitoring UV-accelerated alteration processes of paintings by means of hyperspectral micro-FTIR imaging and chemometrics," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 253, p. 119568, 2021.
- [30] V. Gonzalez, I. Fazlic, M. Cotte, F. Vanmeert, A. Gestels, S. De Meyer, F. Broers, J. Hermans, A. van Loon, K. Janssens, *et al.*, "Lead (II) formate

- in Rembrandt's Night Watch: Detection and distribution from the macro to the micro-scale," *Angewandte Chemie*, vol. 135, no. 16, 2023.
- [31] M. Strlič, D. Thickett, J. Taylor, and M. Cassar, "Damage functions in heritage science," *Studies in Conservation*, vol. 58, no. 2, pp. 80–87, 2013.
- [32] K. L. Pendergrass, W. Sampson, T. Walsh, and L. Alagna, "Toward environmentally sustainable digital preservation," *The American Archivist*, vol. 82, no. 1, pp. 165–206, 2019.
- [33] M. Stols-Witlox, *Historical recipes for preparatory layers for oil paintings in manuals, manuscripts and handbooks in North West Europe, 1550-1900: analysis and reconstructions*. Universiteit van Amsterdam Amsterdam, 2014.
- [34] I. C. A. Sandu, S. Schäfer, D. Magrini, S. Bracci, and C. A. Roque, "Cross-section and staining-based techniques for investigating organic materials in painted and polychrome works of art: A review," *Microscopy and Microanalysis*, vol. 18, no. 4, pp. 860–875, 2012.
- [35] J. Plesters, "Cross-sections and chemical analysis of paint samples," *Studies in conservation*, vol. 2, no. 3, pp. 110–157, 1956.
- [36] P. Kum-Essuon, "Painting supports: Significance, historical review, types and appropriate preparations for effective painting," *Arts and Designs Studies*, vol. 45, 2016.
- [37] I. Hutanu, I. Sandu, V. Vasilache, and L. Nica, "Studies on wood consolidation and completing gaps in panel paintings.," *Pro Ligno*, vol. 9, no. 4, 2013.
- [38] B. Wang, "An investigation of the grounds in Nineteenth-century Chinese export oil paintings," *Studies in Conservation*, pp. 1–13, 2022.
- [39] M. Strojnik, G. Paez, and A. Ortega, "Near IR diodes as illumination sources to remotely detect under-drawings on century-old paintings," in *22nd Congress of the International Commission for Optics: Light for the Development of the World*, vol. 8011, pp. 1931–1936, SPIE, 2011.
- [40] F. Presciutti, J. Perlo, F. Casanova, S. Glöggler, C. Miliani, B. Blümich, B. G. Brunetti, and A. Sgamellotti, "Noninvasive nuclear magnetic resonance profiling of painting layers," *Applied Physics Letters*, vol. 93, no. 3, 2008.

-
- [41] J. R. Barnett, S. Miller, and E. Pearce, "Colour and art: A brief history of pigments," *Optics & Laser Technology*, vol. 38, no. 4-6, pp. 445–453, 2006.
- [42] M. P. Colombini and F. Modugno, "Characterisation of proteinaceous binders in artistic paintings by chromatographic techniques," *Journal of Separation Science*, vol. 27, no. 3, pp. 147–160, 2004.
- [43] R. S. Berns and E. R. De la Rie, "The effect of the refractive index of a varnish on the appearance of oil paintings," *Studies in conservation*, vol. 48, no. 4, pp. 251–262, 2003.
- [44] S. V. Sgourev, "The alchemy of painting: How the technology of oil paint transmuted art," *Technology and Creativity: Production, Mediation and Evaluation in the Digital Age*, pp. 37–62, 2020.
- [45] B. H. Berrie, "Mining for color: new blues, yellows, and translucent paint," *Early science and medicine*, vol. 20, no. 4-6, pp. 308–334, 2015.
- [46] J. M. Adovasio and T. F. Lynch, "Preceramic textiles and cordage from Guitarrero Cave, Peru," *American Antiquity*, vol. 38, no. 1, pp. 84–90, 1973.
- [47] C. H. Fisher, "History of natural fibers," *Journal of Macromolecular Science—Chemistry*, vol. 15, no. 7, pp. 1345–1375, 1981.
- [48] A. Melelli, D. U. Shah, G. Hapsari, R. Cortopassi, S. Durand, O. Arnould, V. Placet, D. Benazeth, J. Beaugrand, F. Jamme, and A. Bourmaud, "Lessons on textile history and fibre durability from a 4000-year-old Egyptian flax yarn," *Nature Plants*, vol. 7, no. 9, pp. 1200–1206, 2021.
- [49] H. Yule and A. C. Burnell, *Hobson-Jobson: The Anglo-Indian Dictionary*. Wordsworth Editions, 1996.
- [50] E. S. Ferreira, A. N. Hulme, H. McNab, and A. Quye, "The natural constituents of historical textile dyes," *Chemical Society Reviews*, vol. 33, no. 6, pp. 329–336, 2004.
- [51] I. Degano, E. Ribechini, F. Modugno, and M. P. Colombini, "Analytical methods for the characterization of organic dyes in artworks and in historical textiles," *Applied Spectroscopy Reviews*, vol. 44, no. 5, pp. 363–410, 2009.
- [52] K. Prabhu, A. S. Bhute, *et al.*, "Plant based natural dyes and mordants: A review," *J. Nat. Prod. Plant Resour*, vol. 2, no. 6, pp. 649–664, 2012.

- [53] A. H. Compton, “A quantum theory of the scattering of X-rays by light elements,” *Physical review*, vol. 21, no. 5, p. 483, 1923.
- [54] A. Einstein, “Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt,” *Annalen der Physik*, vol. 322, pp. 132–148, Jan. 1905.
- [55] University of Central Florida, “Young’s Double Slit Experiment.” <https://pressbooks.online.ucf.edu/phy2053bc/chapter/youngs-double-slit-experiment/>, 2023. [Online; accessed 15-September-2023].
- [56] Jitender Singh, “Diffraction.” <https://www.concepts-of-physics.com/optics/diffraction.php>, 2023. [Online; accessed 15-September-2023].
- [57] LSPDD: Lamp Spectral Power Distribution Database, “Lamp Index.” <https://lspdd.org/app/en/lamps>, 2023. [Online; accessed 15-September-2023].
- [58] J. Dorsey, H. Rushmeier, and F. Sillion, *Digital modeling of material appearance*. Elsevier, 2010.
- [59] M. D. Fairchild, *Color appearance models*. John Wiley & Sons, 2013.
- [60] F. O. Bartell, E. L. Dereniak, and W. L. Wolfe, “The theory and measurement of bidirectional reflectance distribution function (BRDF) and bidirectional transmittance distribution function (BTDF),” in *Radiation scattering in optical systems*, vol. 257, pp. 154–160, SPIE, 1981.
- [61] D. Guarnera, G. Guarnera, A. Ghosh, C. Denk, and M. Glencross, “BRDF representation and acquisition,” *Computer Graphics Forum*, vol. 35, no. 2, pp. 625–650, 2016.
- [62] H. A. Khan, J.-B. Thomas, and J. Y. Hardeberg, “Analytical survey of highlight detection in color and spectral images,” in *Computational Color Imaging: 6th International Workshop, CCIW 2017, Milan, Italy, March 29-31, 2017, Proceedings 6*, pp. 197–208, Springer, 2017.
- [63] Q. Yang, J. Tang, and N. Ahuja, “Efficient and robust specular highlight removal,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 6, pp. 1304–1311, 2014.

-
- [64] A. Artusi, F. Banterle, and D. Chetverikov, "A survey of specular removal methods," in *Computer Graphics Forum*, vol. 30, pp. 2208–2230, Wiley Online Library, 2011.
- [65] W. J. Adams and J. H. Elder, "Effects of specular highlights on perceived surface convexity," *PLoS computational biology*, vol. 10, no. 5, p. e1003576, 2014.
- [66] S. A. Shafer, "Using color to separate reflection components," *Color Research & Application*, vol. 10, no. 4, pp. 210–218, 1985.
- [67] K. E. Torrance and E. M. Sparrow, "Theory for off-specular reflection from roughened surfaces," *Josa*, vol. 57, no. 9, pp. 1105–1114, 1967.
- [68] C. P. Huynh and A. Robles-Kelly, "A solution of the dichromatic model for multispectral photometric invariance," *International Journal of Computer Vision*, vol. 90, no. 1, pp. 1–27, 2010.
- [69] Y. Imai, Y. Kato, H. Kadoi, T. Horiuchi, and S. Tominaga, "Estimation of multiple illuminants based on specular highlight detection," in *Computational Color Imaging: Third International Workshop, CCIW 2011, Milan, Italy, April 20-21, 2011. Proceedings 3*, pp. 85–98, Springer, 2011.
- [70] J. Van De Weijer and S. Beigpour, "The dichromatic reflection model-future research directions and applications.," *VISAPP*, p. 011, 2011.
- [71] A. Woo, P. Poulin, and A. Fournier, "A survey of shadow algorithms," *IEEE Computer Graphics and Applications*, vol. 10, no. 6, pp. 13–32, 1990.
- [72] B. A. Maxwell, R. M. Friedhoff, and C. A. Smith, "A bi-illuminant dichromatic reflection model for understanding images," in *2008 IEEE conference on computer vision and pattern recognition*, pp. 1–8, IEEE, 2008.
- [73] R. Gershon, A. D. Jepson, and J. K. Tsotsos, "Ambient illumination and the determination of material changes," *JOSA A*, vol. 3, no. 10, pp. 1700–1707, 1986.
- [74] D. Liu, C. Long, H. Zhang, H. Yu, X. Dong, and C. Xiao, "ARShadowgan: Shadow generative adversarial network for augmented reality in single light scenes," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 8139–8148, 2020.
- [75] Z. Liu, H. Yin, Y. Mi, M. Pu, and S. Wang, "Shadow removal by a lightness-guided network with training on unpaired data," *IEEE Transactions on Image Processing*, vol. 30, pp. 1853–1865, 2021.

- [76] Wikipedia, “2020–2023 global chip shortage — Wikipedia, the free encyclopedia.” <http://en.wikipedia.org/w/index.php?title=2020%E2%80%932023%20global%20chip%20shortage&oldid=1169845210>, 2023. [Online; accessed 16-August-2023].
- [77] PV Education, “Optical Properties of Silicon.” <https://www.pveducation.org/pvcdrom/materials/optical-properties-of-silicon>, 2023. [Online; accessed 09-September-2023].
- [78] M. Dresselhaus and J. Tauc, *Optical properties of solids*. Part II, 1998.
- [79] H. Föll, “Semiconductors I.” https://www.tf.uni-kiel.de/matwis/amat/semi_en/index.html, 2021. [Online; accessed 09-September-2023].
- [80] G. Hansen and J. Schmit, “Calculation of intrinsic carrier concentration in $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$,” *Journal of Applied Physics*, vol. 54, no. 3, pp. 1639–1640, 1983.
- [81] J. Nakamura, *Image sensors and signal processing for digital still cameras*. CRC press, 2017.
- [82] S. Battiato, A. R. Bruna, G. Messina, and G. Puglisi, *Image processing for embedded devices*. Bentham Science Publishers, 2010.
- [83] D. P. D’Amato and R. M. Centamore, “Two applications for microlens arrays: detector fill-factor improvement and laser diode collimation,” in *Miniature and Micro-Optics: Fabrication and System Applications*, vol. 1544, pp. 166–177, SPIE, 1991.
- [84] R. C. Gonzalez, *Digital image processing*. Pearson, 2009.
- [85] J.-M. Belloir, J.-B. Lincelles, A. Pelamatti, C. Durnez, V. Goiffon, C. Virmondois, P. Paillet, P. Magnan, and O. Gilardx, “Dark current blooming in pinned photodiode CMOS image sensors,” *IEEE Transactions on Electron Devices*, vol. 64, no. 3, pp. 1161–1166, 2017.
- [86] E. Guni, J. Durst, B. Kreisler, T. Michel, G. Anton, M. Fiederle, A. Fauler, and A. Zwerger, “The influence of pixel pitch and electrode pad size on the spectroscopic performance of a photon counting pixel detector with CdTe sensor,” *IEEE Transactions on Nuclear Science*, vol. 58, no. 1, pp. 17–25, 2010.

-
- [87] K. Hirakawa, "Cross-talk explained," in *2008 15th IEEE International Conference on Image Processing*, pp. 677–680, IEEE, 2008.
- [88] N. J. Murray, D. J. Burt, A. D. Holland, K. D. Stefanov, J. P. Gow, C. McCormick, B. J. Dryer, and E. A. Allanwood, "Multi-level parallel clocking of CCDs for: improving charge transfer efficiency, clearing persistence, clocked anti-blooming, and generating low-noise backgrounds for pumping," in *UV/Optical/IR Space Telescopes and Instruments: Innovative Technologies and Concepts VI*, vol. 8860, pp. 160–168, SPIE, 2013.
- [89] G. E. Smith, "Nobel lecture: The invention and early history of the CCD," *Reviews of modern physics*, vol. 82, no. 3, p. 2307, 2010.
- [90] M. Bigas, E. Cabruja, J. Forest, and J. Salvi, "Review of CMOS image sensors," *Microelectronics journal*, vol. 37, no. 5, pp. 433–451, 2006.
- [91] P. Magnan, "Detection of visible photons in CCD and CMOS: A comparative view," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 504, no. 1-3, pp. 199–212, 2003.
- [92] D. L. Snyder, C. W. Helstrom, A. D. Lanterman, M. Faisal, and R. L. White, "Compensation for readout noise in CCD images," *JOSA A*, vol. 12, no. 2, pp. 272–283, 1995.
- [93] O. A. Skydan, F. Lilley, M. J. Lalor, and D. R. Burton, "Quantization error of CCD cameras and their influence on phase calculation in fringe pattern analysis," *Applied optics*, vol. 42, no. 26, pp. 5302–5307, 2003.
- [94] A. K. Boyat and B. K. Joshi, "A review paper: noise models in digital image processing," *arXiv preprint arXiv:1505.03489*, 2015.
- [95] B. Moomaw, "Camera technologies for low light imaging: overview and relative advantages," *Methods in cell biology*, vol. 114, pp. 243–283, 2013.
- [96] M. Freeman, *The DSLR Field Guide*. CRC Press, 2013.
- [97] J. Sasián, *Introduction to aberrations in optical imaging systems*. Cambridge University Press, 2013.
- [98] A. Chakrabarti and T. Zickler, "Depth and deblurring from a spectrally-varying depth-of-field," in *Computer Vision—ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part V 12*, pp. 648–661, Springer, 2012.

- [99] H. Zhu and P. Blackborow, “Etendue and optical throughput calculations,” *Energetiq Technologies Inc. Application Note ref*, pp. 002–2, 2011.
- [100] J. T. Enns and S. C. MacDonald, “The role of clarity and blur in guiding visual attention in photographs.,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 39, no. 2, p. 568, 2013.
- [101] A. Saha and Q. M. J. Wu, “High frequency content based framework for perceptual sharpness assessment in natural images,” in *2015 International Conference on Computers, Communications, and Systems (ICCCS)*, pp. 182–187, 2015.
- [102] C. Wittpahl, H. B. Zakour, M. Lehmann, and A. Braun, “Realistic image degradation with measured PSF,” *Electronic Imaging, Autonomous Vehicles and Machines 2018*, no. 17, 2018.
- [103] A. A. Michelson, *Studies in optics*. University of Chicago Press, 1927.
- [104] E. Manson, L. Bambara, R. Nyaaba, J. Amuasi, J. Flether, C. Schandorf, and A. Amable, “Comparison of modulation transfer function measurements for assessing the performance of imaging systems.,” *Medical Physics*, vol. 5, no. 2, 2017.
- [105] J. R. Kopacz, R. Herschitz, and J. Roney, “Small satellites an overview and assessment,” *Acta Astronautica*, vol. 170, pp. 93–105, 2020.
- [106] D. Gigilashvili, J. B. Thomas, J. Y. Hardeberg, and M. Pedersen, “On the nature of perceptual translucency,” *Workshop on Material Appearance Modeling*, 2020.
- [107] M. Gabriel, *Why the world does not exist*. John Wiley & Sons, 2015.
- [108] J. Y. Hardeberg, *Acquisition and reproduction of color images: Colorimetric and multispectral approaches*. Universal-Publishers, 2001.
- [109] G. Thierry, P. Athanasopoulos, A. Wiggett, B. Dering, and J.-R. Kuipers, “Unconscious effects of language-specific terminology on preattentive color perception,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 11, pp. 4567–4570, 2009.
- [110] M. Jossierand, E. Meeussen, A. Majid, and D. Dediu, “Environment and culture shape both the colour lexicon and the genetics of colour perception,” *Scientific Reports*, vol. 11, no. 1, p. 19095, 2021.

-
- [111] M. P. Keating, *Geometric, physical, and visual optics*. Elsevier Health Sciences, 1988.
- [112] H.-C. Lee, *Introduction to color imaging science*. Cambridge University Press, 2005.
- [113] J. Hirsch and C. A. Curcio, “The spatial resolution capacity of human foveal retina,” *Vision research*, vol. 29, no. 9, pp. 1095–1101, 1989.
- [114] R. Shapley and M. Hawken, “Neural mechanisms for color perception in the primary visual cortex,” *Current opinion in neurobiology*, vol. 12, no. 4, pp. 426–432, 2002.
- [115] E. N. Johnson, M. J. Hawken, and R. Shapley, “The orientation selectivity of color-responsive neurons in macaque V1,” *Journal of Neuroscience*, vol. 28, no. 32, pp. 8096–8106, 2008.
- [116] H. S. Fairman, M. H. Brill, and H. Hemmendinger, “How the CIE 1931 color-matching functions were derived from Wright-Guild data,” *Color Research & Application*, vol. 22, no. 1, pp. 11–23, 1997.
- [117] A. D. Broadbent, “A critical review of the development of the CIE1931 RGB color-matching functions,” *Color Research & Application*, vol. 29, no. 4, pp. 267–272, 2004.
- [118] S. Y. Kahu, R. B. Raut, and K. M. Bhurchandi, “Review and evaluation of color spaces for image/video compression,” *Color Research & Application*, vol. 44, no. 1, pp. 8–33, 2019.
- [119] M. Berthier and E. Provenzi, “From Riemannian trichromacy to quantum color opponency via hyperbolicity,” *Journal of Mathematical Imaging and Vision*, vol. 63, no. 6, pp. 681–688, 2021.
- [120] B. Bayer, “Color imaging array,” *United States Patent*, no. 3971065, 1976.
- [121] D. Menon and G. Calvagno, “Color image demosaicking: An overview,” *Signal Processing: Image Communication*, vol. 26, no. 8-9, pp. 518–533, 2011.
- [122] D. Akkaynak, T. Treibitz, B. Xiao, U. A. Gürkan, J. J. Allen, U. Demirci, and R. T. Hanlon, “Use of commercial off-the-shelf digital cameras for scientific data acquisition and scene-specific color calibration,” *JOSA A*, vol. 31, no. 2, pp. 312–321, 2014.

- [123] J. Zhu, X. Xie, N. Liao, Z. Zhang, W. Wu, and L. Lv, “Spectral sensitivity estimation of trichromatic camera based on orthogonal test and window filtering,” *Optics Express*, vol. 28, no. 19, pp. 28085–28100, 2020.
- [124] Y.-T. Lin and G. D. Finlayson, “Physically plausible spectral reconstruction,” *Sensors*, vol. 20, no. 21, p. 6399, 2020.
- [125] R. S. Berns, *Billmeyer and Saltzman’s principles of color technology*. John Wiley & Sons, 2019.
- [126] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton, “Multispectral filter arrays: Recent advances and practical implementation,” *Sensors*, vol. 14, no. 11, pp. 21626–21659, 2014.
- [127] Midwest Optical Systems, INC, “Bandpass Filters.” <https://midopt.com/filters/bandpass/>, 2023. [Online; accessed 09-September-2023].
- [128] G. Trumpy, J. Y. Hardeberg, S. George, and B. Flueckiger, “A multispectral design for a new generation of film scanners,” in *Optics for Arts, Architecture, and Archaeology VIII*, vol. 11784, pp. 138–146, SPIE, 2021.
- [129] Allied Vision, “Alvium 1800 C-234.” <https://www.alliedvision.com/en/products/alvium-configurator/alvium-1800-c/234/>, 2023. [Online; accessed 17-August-2023].
- [130] G. Bianco, F. Bruno, and M. Muzzupappa, “Multispectral data cube acquisition of aligned images for document analysis by means of a filter-wheel camera provided with focus control,” *Journal of Cultural Heritage*, vol. 14, no. 3, pp. 190–200, 2013.
- [131] R. S. Berns, “Theory and practice of dual-RGB imaging,” *Rochester: Studio for Scientific Imaging and Archiving of Cultural Heritage, Rochester Institute of Technology*, 2016.
- [132] R. Wu, Y. Li, X. Xie, and Z. Lin, “Optimized multi-spectral filter arrays for spectral reconstruction,” *Sensors*, vol. 19, no. 13, p. 2905, 2019.
- [133] C. Fischer and I. Kakoulli, “Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications,” *Studies in Conservation*, vol. 51, no. sup1, pp. 3–16, 2006.
- [134] Y. Castro, A. Siatou, M. Rossé, H. Chatoux, R. Luxman, G. Le Goïc, and A. Mansouri, “Extended framework for multispectral RTI,” in *Archiving*

- Conference, Society for Imaging Science and Technology Publisher Location: IS&T 7003 Kilworth Lane, Springfield, VA 22151 USA*, vol. 19, pp. 56–61, 2022.
- [135] E. F. Prentice, M. B. Henriksen, T. A. Johansen, F. N. Medina, and A. G. San Juan, “Characterizing spectral response in thermal environments, the HYPISO-1 Hyperspectral Imager,” in *2022 IEEE Aerospace Conference (AERO)*, pp. 1–10, IEEE, 2022.
 - [136] C. E. Shannon, “Communication in the presence of noise,” *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.
 - [137] S. Westland, J. Shaw, and H. Owens, “Colour statistics of natural and man-made surfaces,” *Sensor Review*, vol. 20, no. 1, pp. 50–55, 2000.
 - [138] H. Yang, D. Zhang, W. Huang, Z. Gao, X. Yang, C. Li, and J. Wang, “Application and evaluation of wavelet-based denoising method in hyperspectral imagery data,” in *Computer and Computing Technologies in Agriculture V: 5th IFIP TC 5/SIG 5.1 Conference, CCTA 2011, Beijing, China, October 29-31, 2011, Proceedings, Part II 5*, pp. 461–469, Springer, 2012.
 - [139] Perkin Elmer, “LAMBDA 1050+ UV/Vis/NIR Spectrophotometer.” <https://www.perkinelmer.com/uk/product/lambda-1050-2d-base-inst-no-sw-16020055>, 2023. [Online; accessed 17-August-2023].
 - [140] X. Wang, Y. Zhang, X. Ma, T. Xu, and G. R. Arce, “Compressive spectral imaging system based on liquid crystal tunable filter,” *Optics Express*, vol. 26, no. 19, pp. 25226–25243, 2018.
 - [141] J. M. Amigo, *Hyperspectral imaging*. Elsevier, 2019.
 - [142] Meet Optics, “Groove Density.” <https://www.meetoptics.com/academy/groove-density>, 2023. [Online; accessed 09-September-2023].
 - [143] Z. Xu and M. H. Brill, “Correction of second-order-diffraction errors in spectrophotometry,” *Color Research & Application*, vol. 42, no. 2, pp. 189–192, 2017.
 - [144] A. Oriana, J. Réhault, F. Preda, D. Polli, and G. Cerullo, “Scanning fourier transform spectrometer in the visible range based on birefringent wedges,” *JOSA A*, vol. 33, no. 7, pp. 1415–1420, 2016.

- [145] H. Deborah, N. Richard, and J. Y. Hardeberg, “A comprehensive evaluation of spectral distance functions and metrics for hyperspectral image processing,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 3224–3234, 2015.
- [146] A. B. López-Baldomero, M. Martínez-Domingo, E. M. Valero, R. Fernández-Gualda, A. López-Montes, R. Blanc-García, and T. Espejo, “Selection of optimal spectral metrics for classification of inks in historical documents using hyperspectral imaging data,” in *Optics for Arts, Architecture, and Archaeology (O3A) IX*, vol. 12620, pp. 99–111, SPIE, 2023.
- [147] F. Kruse, A. Lefkoff, and J. Dietz, “Expert system-based mineral mapping in northern Death Valley, California/Nevada, using the airborne visible/infrared imaging spectrometer (AVIRIS),” *Remote Sensing of Environment*, vol. 44, no. 2-3, pp. 309–336, 1993.
- [148] N. Richard, D. Helbert, C. Olivier, and M. Tamisier, “Pseudo-divergence and bidimensional histogram of spectral differences for hyperspectral image processing,” *Journal of Imaging Science and Technology*, vol. 60, no. 5, pp. 504021–5040213, 2016.
- [149] O. A. De Carvalho and P. R. Meneses, “Spectral correlation mapper (SCM): an improvement on the spectral angle mapper (SAM),” in *Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18*, vol. 9, p. 2, JPL publication Pasadena, CA, USA, 2000.
- [150] G. Liu, S. Fomel, L. Jin, and X. Chen, “Stacking seismic data using local correlation,” *Geophysics*, vol. 74, no. 3, pp. V43–V48, 2009.
- [151] S. Peyghambari and Y. Zhang, “Hyperspectral remote sensing in lithological mapping, mineral exploration, and environmental geology: an updated review,” *Journal of Applied Remote Sensing*, vol. 15, no. 3, pp. 031501–031501, 2021.
- [152] C. N. Durell, “An overview of the emerging IEEE P4001 hyperspectral standard,” in *Imaging Spectrometry XXV: Applications, Sensors, and Processing*, vol. 12235, p. 122350F, SPIE, 2022.
- [153] H. Trusseli and M. S. Kulkarni, “Sampling and processing of color signals,” *IEEE Transactions on image processing*, vol. 5, no. 4, pp. 677–681, 1996.
- [154] N. Hagen and M. W. Kudenov, “Review of snapshot spectral imaging technologies,” *Optical Engineering*, vol. 52, no. 9, 2013.

-
- [155] A. Perri, B. N. De Faria, D. T. Ferreira, D. Comelli, G. Valentini, F. Preda, D. Polli, A. De Paula, G. Cerullo, and C. Manzoni, “Hyperspectral imaging with a TWINS birefringent interferometer,” *Optics Express*, vol. 27, no. 11, pp. 15956–15967, 2019.
- [156] M. Bacci, M. Picollo, G. Trumpy, M. Tsukada, and D. Kunzelman, “Non-invasive identification of white pigments on 20th-century oil paintings by using fiber optic reflectance spectroscopy,” *Journal of the American Institute for Conservation*, vol. 46, no. 1, pp. 27–37, 2007.
- [157] Hypspec, “Hypspec VNIR-1800.” <https://www.hypspec.com/hypspec-products/hypspec-classic/hypspec-vnir-1800/>, 2023. [Online; accessed 17-August-2023].
- [158] Hypspec, “Hypspec SWIR-384.” <https://www.hypspec.com/hypspec-products/hypspec-classic/hypspec-swir-384/>, 2023. [Online; accessed 17-August-2023].
- [159] J. Jia, J. Chen, X. Zheng, Y. Wang, S. Guo, H. Sun, C. Jiang, M. Karjalainen, K. Karila, Z. Duan, *et al.*, “Tradeoffs in the spatial and spectral resolution of airborne hyperspectral imaging systems: A crop identification case study,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–18, 2021.
- [160] M. Pedersen, J. Y. Hardeberg, *et al.*, “Full-reference image quality metrics: Classification and evaluation,” *Foundations and Trends in Computer Graphics and Vision*, vol. 7, no. 1, pp. 1–80, 2012.
- [161] T. Skauli, “Sensor noise informed representation of hyperspectral data, with benefits for image storage and processing,” *Optics Express*, vol. 19, no. 14, pp. 13031–13046, 2011.
- [162] K. Lenhard, A. Baumgartner, and T. Schwarzmaier, “Independent laboratory characterization of NEO Hypspec imaging spectrometers VNIR-1600 and SWIR-320m-e,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 1828–1841, 2014.
- [163] H. A. Khan, S. Mihoubi, B. Mathon, J.-B. Thomas, and J. Y. Hardeberg, “HyTexiLa: High resolution visible and near infrared hyperspectral texture images,” *Sensors*, vol. 18, no. 7, p. 2045, 2018.
- [164] K. A. Dooley, S. Lomax, J. G. Zeibel, C. Miliani, P. Ricciardi, A. Hoenigswald, M. Loew, and J. K. Delaney, “Mapping of egg yolk and animal skin

- glue paint binders in Early Renaissance paintings using near infrared reflectance imaging spectroscopy,” *Analyst*, vol. 138, no. 17, pp. 4838–4848, 2013.
- [165] Labsphere, “Spectralon diffuse reflectance standards.” <https://www.labsphere.com/product/spectralon-diffuse-reflectance-standards/>, 2023. [Online; accessed 16-August-2023].
- [166] T. Skauli, “Specifying radiometric performance of hyperspectral and conventional cameras: a minimal set of independent characteristics,” in *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imagery XXVI*, vol. 11392, pp. 86–97, SPIE, 2020.
- [167] H. Liang, A. Lucian, R. Lange, C. S. Cheung, and B. Su, “Remote spectral imaging with simultaneous extraction of 3D topography for historical wall paintings,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 95, pp. 13–22, 2014.
- [168] C. Cucci, A. Casini, M. Picollo, M. Poggesi, and L. Stefani, “Open issues in hyperspectral imaging for diagnostics on paintings: When high-spectral and spatial resolution turns into data redundancy,” in *O3A: Optics for arts, architecture, and archaeology III*, vol. 8084, pp. 57–66, SPIE, 2011.
- [169] Y. Dua, V. Kumar, and R. S. Singh, “Comprehensive review of hyperspectral image compression algorithms,” *Optical Engineering*, vol. 59, no. 9, 2020.
- [170] B. M. Devassy and S. George, “Dimensionality reduction and visualisation of hyperspectral ink data using t-SNE,” *Forensic science international*, vol. 311, p. 110194, 2020.
- [171] A. Villa, J. Chanussot, C. Jutten, J. A. Benediktsson, and S. Moussaoui, “On the use of ICA for hyperspectral image analysis,” in *2009 IEEE International Geoscience and Remote Sensing Symposium*, vol. 4, pp. IV–97, IEEE, 2009.
- [172] X.-R. Feng, H.-C. Li, R. Wang, Q. Du, X. Jia, and A. Plaza, “Hyperspectral unmixing based on nonnegative matrix factorization: A comprehensive review,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 4414–4436, 2022.
- [173] A. Cosentino, “FORS spectral database of historical pigments in different binders,” *E Conserv. J*, vol. 2, pp. 54–65, 2014.

-
- [174] Institute of Applied Physics (IFAC), “Fiber optics reflectance spectra (FORS) of pictorial materials in the 270–1700 nm range.” <http://fors.ifac.cnr.it/index.php>, 2011. [Online; accessed 24-August-2023].
- [175] H. Deborah, “Hyperspectral pigment dataset,” in *2022 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1–5, IEEE, 2022.
- [176] B. Singer, T. E. Aslaksby, B. Topalova-Casadiegos, and E. S. Tveit, “Investigation of materials used by Edvard Munch,” *Studies in conservation*, vol. 55, no. 4, pp. 274–292, 2010.
- [177] F. Grillini, J.-B. Thomas, and S. George, “Comparison of imaging models for spectral unmixing in oil painting,” *Sensors*, vol. 21, no. 7, p. 2471, 2021.
- [178] A. Plaza, G. Martín, J. Plaza, M. Zortea, and S. Sánchez, “Recent developments in endmember extraction and spectral unmixing,” *Optical Remote Sensing: Advances in Signal Processing and Exploitation Techniques*, pp. 235–267, 2011.
- [179] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, “Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches,” *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 5, no. 2, pp. 354–379, 2012.
- [180] J. W. Boardman, “Automating spectral unmixing of AVIRIS data using convex geometry concepts,” in *JPL, Summaries of the 4th Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop*, 1993.
- [181] M. E. Winter, “N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data,” in *Imaging Spectrometry V*, vol. 3753, pp. 266–275, SPIE, 1999.
- [182] J. M. Nascimento and J. M. Dias, “Vertex component analysis: A fast algorithm to unmix hyperspectral data,” *IEEE transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, 2005.
- [183] A. Plaza and C.-I. Chang, “Impact of initialization on design of endmember extraction algorithms,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3397–3407, 2006.

- [184] J. Li and J. M. Bioucas-Dias, “Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data,” in *IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium*, vol. 3, pp. III–250, IEEE, 2008.
- [185] J. Li, A. Agathos, D. Zaharie, J. M. Bioucas-Dias, A. Plaza, and X. Li, “Minimum volume simplex analysis: A fast algorithm for linear hyperspectral unmixing,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 9, pp. 5067–5082, 2015.
- [186] R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez, C. Richard, J. Chanussot, L. Drumetz, J.-Y. Tourneret, A. Zare, and C. Jutten, “Spectral variability in hyperspectral data unmixing: A comprehensive review,” *IEEE geoscience and remote sensing magazine*, vol. 9, no. 4, pp. 223–270, 2021.
- [187] J. Huang, H. He, R. Lv, G. Zhang, Z. Zhou, and X. Wang, “Non-destructive detection and classification of textile fibres based on hyperspectral imaging and 1D-CNN,” *Analytica Chimica Acta*, vol. 1224, p. 340238, 2022.
- [188] Z. Jianxin, Z. Kangping, W. Junkai, and H. Xudong, “Color segmentation and extraction of yarn-dyed fabric based on a hyperspectral imaging system,” *Textile Research Journal*, vol. 91, no. 7-8, pp. 729–742, 2021.
- [189] J. K. Delaney, P. Ricciardi, L. Glinsman, M. Palmer, and J. Burke, “Use of near infrared reflectance imaging spectroscopy to map wool and silk fibres in historic tapestries,” *Analytical Methods*, vol. 8, no. 44, pp. 7886–7890, 2016.
- [190] C. Vlachou-Mogire, J. Danskin, J. R. Gilchrist, and K. Hallett, “Mapping materials and dyes on historic tapestries using hyperspectral imaging,” *Heritage*, vol. 6, no. 3, pp. 3159–3182, 2023.
- [191] H. Deborah, S. George, and J. Y. Hardeberg, “Pigment mapping of The Scream (1893) based on hyperspectral imaging,” in *Image and Signal Processing: 6th International Conference, ICISP 2014, Cherbourg, France, June 30–July 2, 2014. Proceedings 6*, pp. 247–256, Springer, 2014.
- [192] J. K. Delaney, K. A. Dooley, A. Van Loon, and A. Vandivere, “Mapping the pigment distribution of Vermeer’s Girl with a Pearl Earring,” *Heritage Science*, vol. 8, pp. 1–16, 2020.
- [193] N. Rohani, E. Pouyet, M. Walton, O. Cossairt, and A. K. Katsaggelos, “Pigment unmixing of hyperspectral images of paintings using deep neural net-

- works,” in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3217–3221, IEEE, 2019.
- [194] D. Bai, D. W. Messinger, and D. Howell, “A hyperspectral imaging spectral unmixing and classification approach to pigment mapping in the Gough & Selden Maps,” *Journal of the American Institute for Conservation*, vol. 58, no. 1-2, pp. 69–89, 2019.
- [195] B. Grabowski, W. Masarczyk, P. Głomb, and A. Mendys, “Automatic pigment identification from hyperspectral data,” *Journal of Cultural Heritage*, vol. 31, pp. 1–12, 2018.
- [196] S. A. Burns, “Subtractive color mixture computation,” *arXiv preprint arXiv:1710.06364*, 2017.
- [197] P. Kubelka, “Ein beitrag zur optik der farbanstriche (Contribution to the optic of paint),” *Z. tech. Phys*, vol. 12, pp. 593–601, 1931.
- [198] R. S. Berns and M. Mohammadi, “Single-constant simplification of Kubelka-Munk turbid-media theory for paint systems—a review,” *Color Research & Application*, vol. 32, no. 3, pp. 201–207, 2007.
- [199] H. Deborah, M. O. Ulfarsson, and J. Sigurdsson, “Fully constrained least squares linear spectral unmixing of The Scream (verso, 1893),” in *2021 11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1–5, IEEE, 2021.
- [200] A. A. Goshtasby, *Image registration: Principles, tools and methods*. Springer Science & Business Media, 2012.
- [201] G. Vivone, M. Dalla Mura, A. Garzelli, R. Restaino, G. Scarpa, M. O. Ulfarsson, L. Alparone, and J. Chanussot, “A new benchmark based on recent advances in multispectral pansharpening: Revisiting pansharpening with classical and emerging pansharpening methods,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, no. 1, pp. 53–81, 2020.
- [202] W. Dorigo, M. Bachmann, and W. Heldens, “AS toolbox and processing of field spectra,” *User’s manual, German Aerospace Center (DLR), Oberpfaffenhofen*, 2006.
- [203] A. L. de Queiroz Baddini, J. L. V. de Paula Santos, R. R. Tavares, L. S. de Paula, H. da Costa Araújo Filho, and R. P. Freitas, “PLS-DA and data fusion of visible reflectance, XRF and FTIR spectroscopy in the classification of mixed historical pigments,” *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 265, p. 120384, 2022.

- [204] C. Harris and M. Stephens, “A combined corner and edge detector,” in *Alvey vision conference*, vol. 15, pp. 10–5244, Citeseer, 1988.
- [205] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International journal of computer vision*, vol. 60, pp. 91–110, 2004.
- [206] H. Cantzler, *Random sample consensus (RANSAC)*. Institute for Perception, Action and Behaviour, Division of Informatics, University of Edinburgh, 1981.
- [207] S. L. Al-Khafaji, J. Zhou, A. Zia, and A. W.-C. Liew, “Spectral-spatial scale invariant feature transform for hyperspectral images,” *IEEE Transactions on Image Processing*, vol. 27, no. 2, pp. 837–850, 2017.
- [208] H. Chatoux, N. Richard, F. Lecellier, and C. Fernandez-Maloigne, “Gradient in spectral and color images: from the Di Zenzo initial construction to a generic proposition,” *JOSA A*, vol. 36, no. 11, pp. C154–C165, 2019.
- [209] A. F. McDaid, D. Greene, and N. Hurley, “Normalized mutual information to evaluate overlapping community finding algorithms,” *arXiv preprint arXiv:1110.2515*, 2011.
- [210] D. M. Conover, J. K. Delaney, and M. H. Loew, “Automatic registration and mosaicking of technical images of Old Master paintings,” *Applied Physics A*, vol. 119, pp. 1567–1575, 2015.
- [211] K. Kuppala, S. Banda, and T. R. Barige, “An overview of deep learning methods for image registration with focus on feature-based approaches,” *International Journal of Image and Data Fusion*, vol. 11, no. 2, pp. 113–135, 2020.
- [212] O. Lisický, S. Avril, B. Eydan, B. Pierrat, and J. Burša, “Evaluation of image registration for measuring deformation fields in soft tissue mechanics,” *Strain*, vol. 58, no. 4, p. e12424, 2022.
- [213] A. Hueni and A. Bialek, “Cause, effect, and correction of field spectroradiometer interchannel radiometric steps,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 4, pp. 1542–1551, 2017.
- [214] T. H. Hemmer and T. L. Westphal, “Lessons learned in the postprocessing of field spectroradiometric data covering the 0.4–2.5- μ m wavelength region,” in *Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI*, vol. 4049, pp. 249–260, SPIE, 2000.

-
- [215] M. Danner, M. Locherer, T. Hank, and K. Richter, "Spectral sampling with the ASD FieldSpec4," *GFZ Data Services*, 2015.
- [216] Wikipedia, "RNA splicing — Wikipedia, the free encyclopedia." https://en.wikipedia.org/wiki/RNA_splicing, 2023. [Online; accessed 18-September-2023].
- [217] D. Beal and M. Eamon, "Preliminary results of testing and a proposal for radiometric error correction using dynamic, parabolic linear transformations of "stepped" data," *Analytical Spectral Devices, Inc*, 2009.
- [218] G. Singh and A. Mittal, "Various image enhancement techniques: A critical review," *International Journal of Innovation and Scientific Research*, vol. 10, no. 2, pp. 267–274, 2014.
- [219] D. C. Lepcha, B. Goyal, A. Dogra, and V. Goyal, "Image super-resolution: A comprehensive review, recent trends, challenges and applications," *Information Fusion*, 2022.
- [220] T. R. Peery and D. W. Messinger, "Panchromatic sharpening enabling low-intensity imaging of cultural heritage documents," in *Image Sensing Technologies: Materials, Devices, Systems, and Applications VI*, vol. 10980, p. 1098004, SPIE, 2019.
- [221] M. Selva, B. Aiazzi, F. Butera, L. Chiarantini, and S. Baronti, "Hyper-sharpening: A first approach on SIM-GA data," *IEEE Journal of selected topics in applied earth observations and remote sensing*, vol. 8, no. 6, pp. 3008–3024, 2015.
- [222] L. Chen, G. Vivone, J. Qin, J. Chanussot, and X. Yang, "Spectral-spatial transformer for hyperspectral image sharpening," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.
- [223] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?," in *Third conference" Fusion of Earth data: merging point measurements, raster maps and remotely sensed images"*, pp. 99–103, SEE/URISCA, 2000.
- [224] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva, "Multispectral and panchromatic data fusion assessment without reference," *Photogrammetric Engineering & Remote Sensing*, vol. 74, no. 2, pp. 193–200, 2008.

- [225] B. Aiazzi, L. Alparone, S. Baronti, R. Carlà, A. Garzelli, and L. Santurri, “Full-scale assessment of pansharpening methods and data products,” in *Image and Signal Processing for Remote Sensing XX*, vol. 9244, p. 924402, SPIE, 2014.
- [226] J. Riviere, I. Reshetouski, L. Filipi, and A. Ghosh, “Polarization imaging reflectometry in the wild,” *ACM Trans. Graph.*, vol. 36, no. 6, pp. 1–14, 2017.
- [227] J. S. Tyo, “Design of optimal polarimeters: maximization of signal-to-noise ratio and minimization of systematic error,” *Appl. Opt.*, vol. 41, pp. 619–630, Feb 2002.
- [228] S. Sattar, P.-J. Lapray, A. Foulonneau, and L. Bigué, “Review of spectral and polarization imaging systems,” in *Unconventional Optical Imaging II*, vol. 11351, pp. 191–203, SPIE, 2020.
- [229] B. D. Bartlett, A. Schlamm, C. Salvaggio, and D. W. Messinger, “Anomaly detection of man-made objects using spectropolarimetric imagery,” in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVII*, vol. 8048, pp. 109–115, SPIE, 2011.
- [230] X. Zhang, C. Leng, Y. Hong, Z. Pei, I. Cheng, and A. Basu, “Multimodal remote sensing image registration methods and advancements: A survey,” *Remote Sensing*, vol. 13, no. 24, p. 5128, 2021.
- [231] G. Vivone, M. Dalla Mura, A. Garzelli, and F. Pacifici, “A benchmarking protocol for pansharpening: Dataset, preprocessing, and quality assessment,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 6102–6118, 2021.
- [232] S. Pignatti, A. Palombo, S. Pascucci, F. Romano, F. Santini, T. Simoniello, A. Umberto, C. Vincenzo, N. Acito, M. Diani, *et al.*, “The PRISMA hyperspectral mission: Science activities and opportunities for agriculture and land monitoring,” in *2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS*, pp. 4558–4561, IEEE, 2013.
- [233] J. Irizar, M. Melf, P. Bartsch, J. Koehler, S. Weiss, R. Greinacher, M. Erdmann, V. Kirschner, A. P. Albinana, and D. Martin, “Sentinel-5/UVNS,” in *International Conference on Space Optics—ICSO 2018*, vol. 11180, pp. 41–58, SPIE, 2019.
- [234] J. A. Curcio and C. C. Petty, “The near infrared absorption spectrum of liquid water,” *JOSA*, vol. 41, no. 5, pp. 302–304, 1951.

-
- [235] C. Cucci, M. Picollo, L. Stefani, and R. Jiménez, “The man who became a Blue Glass. Reflectance hyperspectral imaging discloses a hidden Picasso painting of the Blue Period,” *Journal of Cultural Heritage*, vol. 62, pp. 484–492, 2023.
- [236] F. Gabrieli, J. K. Delaney, R. G. Erdmann, V. Gonzalez, A. van Loon, P. Smulders, R. Berkeveld, R. van Langh, and K. Keune, “Reflectance imaging spectroscopy (RIS) for Operation Night Watch: Challenges and achievements of imaging Rembrandt’s masterpiece in the glass chamber at the Rijksmuseum,” *Sensors*, vol. 21, no. 20, p. 6855, 2021.
- [237] M. E. Schaepman, M. Jehle, A. Hueni, P. D’Odorico, A. Damm, J. Weyermann, F. D. Schneider, V. Laurent, C. Popp, F. C. Seidel, *et al.*, “Advanced radiometry measurements and earth science applications with the airborne prism experiment (APEX),” *Remote Sensing of Environment*, vol. 158, pp. 207–219, 2015.
- [238] R. Restaino, M. Dalla Mura, G. Vivone, and J. Chanussot, “Context-adaptive pansharpening based on image segmentation,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 753–766, 2016.
- [239] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, “Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis,” *IEEE Transactions on geoscience and remote sensing*, vol. 40, no. 10, pp. 2300–2312, 2002.
- [240] L. Chen, G. Vivone, Z. Nie, J. Chanussot, and X. Yang, “Spatial data augmentation: Improving the generalization of neural networks for pansharpening,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–11, 2023.
- [241] S. Ono, “Snapshot multispectral imaging using a pixel-wise polarization color image sensor,” *Optics Express*, vol. 28, no. 23, pp. 34536–34573, 2020.
- [242] M. Á. Martínez-Domingo, E. M. Valero, J. Hernández-Andrés, S. Tominaga, T. Horiuchi, and K. Hirai, “Image processing pipeline for segmentation and material classification based on multispectral high dynamic range polarimetric images,” *Optics Express*, vol. 25, no. 24, pp. 30073–30090, 2017.
- [243] W. Hubbard, G. Bishop, T. Gowen, D. Hayter, and G. Innes, “Multispectral-polarimetric sensing for detection of difficult targets,” in *Electro-Optical and Infrared Systems: Technology and Applications V*, vol. 7113, pp. 149–155, SPIE, 2008.

- [244] L. Le Hors, P. Hartemann, and S. Breugnot, “Multispectral polarization active imager in the visible band,” *Proc. SPIE*, vol. 4035, pp. 380 – 389, 2000.
- [245] L. Le Hors, P. Hartemann, D. Dolfi, and S. Breugnot, “Phenomenological model of paints for multispectral polarimetric imaging,” *Proc. SPIE*, vol. 4370, pp. 94 – 105, 2001.
- [246] H. E. M. Leon, *La cultura Paracas [in Spanish]*. Universidad Nacional De Trujillo, 2019.
- [247] Z. Wang, E. P. Simoncelli, and A. C. Bovik, “Multiscale structural similarity for image quality assessment,” in *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, vol. 2, pp. 1398–1402, IEEE, 2003.
- [248] L. Wald, *Data fusion: definitions and architectures: fusion of images of different spatial resolutions*. Presses des MINES, 2002.
- [249] T. Novikova, J. Rehbinder, S. Deby, H. Haddad, J. Vizet, A. Pierangelo, P. Validire, A. Benali, B. Gayet, B. Teig, *et al.*, “Multi-spectral Mueller matrix imaging polarimetry for studies of human tissues,” in *Clinical and Translational Biophotonics*, Optica Publishing Group, 2016.
- [250] M. Köppen, “The curse of dimensionality,” in *5th online world conference on soft computing in industrial applications (WSC5)*, vol. 1, pp. 4–8, 2000.
- [251] W. Luther, N. Baloian, D. Biella, and D. Sacher, “Digital twins and enabling technologies in museums and cultural heritage: An overview,” *Sensors*, vol. 23, no. 3, p. 1583, 2023.
- [252] International Image Interoperability Framework. <https://iiif.io/>, 2023. [Online; accessed 12-November-2023].
- [253] Mirador. <https://projectmirador.org/>, 2023. [Online; accessed 12-November-2023].
- [254] Lars Maxfield, “Butterfly Viewer.” https://olive-groves.github.io/butterfly_viewer/butterfly_viewer.html, 2023. [Online; accessed 12-November-2023].
- [255] F. H. Imai, M. R. Rosen, and R. S. Berns, “Comparative study of metrics for spectral match quality,” in *Conference on colour in graphics, imaging, and vision*, no. 1, pp. 492–496, Society for Imaging Science and Technology, 2002.

- [256] IPERION HS, “Integrating Platforms for the European Research Infrastructure ON Heritage Science.” <https://www.iperionhs.eu/>, 2023. [Online; accessed 12-November-2023].

Part II

Original Articles

Hyperspectral VNIR-SWIR image registration: Do not throw away those overlapping low SNR bands

Article A1_c

This article is a reprint of the publication:

Grillini, F., Thomas, J-B., George, S. (2022). *Hyperspectral VNIR-SWIR image registration: Do not throw away those overlapping low SNR bands*. 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (IEEE WHISPERS). DOI: <https://doi.org/10.1109/WHISPERS56178.2022.9955080>.

HYPERSPECTRAL VNIR - SWIR IMAGE REGISTRATION: DO NOT THROW AWAY THOSE OVERLAPPING LOW SNR BANDS

Federico Grillini, Jean-Baptiste Thomas, Sony George

Colourlab, Department of Computer Science, NTNU, Norway

ABSTRACT

We study the influence from a series of factors on the quality of the registration of VNIR and SWIR hyperspectral images. We specifically consider the registration of VNIR and SWIR images of different spatial resolutions acquired in controlled laboratory conditions on different historical artefacts. This registration problem is defined by a large scaling difference and small translation and rotation. We compare four methods and demonstrate that the largest effects on the quality metrics are due to the image contents and to the deployment of specific ranges of spectral bands. In particular, we demonstrate that using bands of similar nominal wavelengths gives significant advantage. While those bands have typically low values of Signal-to-Noise ratio and are frequently discarded, for this application, they can be treated as a valuable asset.

Index Terms— Hyperspectral imaging, Registration, VNIR-SWIR, Cultural Heritage

1. INTRODUCTION

In the field of Cultural Heritage, imaging spectroscopy is encountering a growing trend in popularity which makes the technique more accessible to many facilities and laboratories around the world. As a non-invasive, non-destructive technique, it represents a reliable tool for documentation, monitoring, and analysis of historical artefacts [1]. Another trend that this field of research is experiencing is the deployment of spectral data in different modalities and ranges of the electromagnetic spectrum. However, such modalities are usually analyzed and processed independently, with conclusions and findings that are eventually inferred by combining the individual results.

Image registration is a preliminary step that allows the combined analysis of separate datasets that refer to the same scene. Data alignment is a compulsory requirement for pixel-based analysis such as pigment mapping [2, 3] and spectral unmixing [4] or further fusion processing like pansharpening and hypersharpening [5, 6]. Therefore, registered images should be as close as possible to sub-pixel precision, in order to provide accurate inferences regarding the material properties. The literature in remote sensing offers a plethora of articles that address the problem of hyperspectral images reg-

istration, also in the case of different hyperspectral imaging modalities.

Image registration techniques can be broadly divided into two families: feature-based and intensity-based approaches [7]. Feature-based methods require the detection of control points in both fixed (or reference) and moving images; which at a later stage are matched according to the similarity of their constructed feature vectors. Examples of this family are the SIFT [8] and SURF [9]. Intensity-based methods rely on the optimization of a cost function that reflects the final quality of alignment between fixed and moving images. This family is often adopted to solve multimodal registration problems maximizing metrics such as correlation and Mutual Information (MI). Normally, both approaches are performed on a predefined fixed-moving image pair. However, in the case of hyperspectral imaging in two different modalities, these *roles* are not defined, i.e., the decision of which spectral set is the reference depends on the application. In addition, both fixed and moving datasets are represented by a series of images, which means that the spectral information can be exploited to build richer feature vectors, as proposed in several registration attempts [10, 11, 12]. The abundance of images can also be considered an asset in generating the most suitable image pair that leads the registration, in the same fashion of the band selection and band synthesis approaches adopted in hypersharpening [6]. In the field of Cultural Heritage imaging, multimodal registration represents a crucial task, since historical artefacts are often studied in cross-disciplinary frameworks that require the capturing of images carrying different information (topography maps, XRF, imaging spectroscopy, FTIR, etc.). A popular approach was proposed in [13], where the maximum of the modulus of the wavelet transform is deployed to identify control points and then the local maximum of normalized cross correlation between phase images helps in identifying potential matched pairs between fixed and moving images. The method, referred here as to MWTPXC (Max Wavelet Transform Phase Cross-Correlation), was developed mainly for applications on historical paintings.

In this article we tackle the specific problem of registering spectral data with two hyperspectral imagers working in complementary ranges of the electromagnetic spectrum: Visible-Near-Infrared (VNIR) and Shortwave-Infrared (SWIR) by evaluating the influence of selected factors on the quality of

the final aligned result. The capturing of hyperspectral data for historical artefacts usually takes place in laboratory or in-situ, with the latter being more challenging as several aspects - such as the illumination - cannot be controlled as well as in laboratory conditions. In this work, we adapt and compare several registration methods to the case of VNIR-SWIR images. We perform an experiment on images of objects that mock-up cultural heritage artefacts (paintings, textiles, drawings), which images were acquired in controlled laboratory conditions. We show that the methods benefiting from prior scaling and overlapping bands perform the best.

2. IMAGING SET-UP AND PROBLEM DESCRIPTION

The capturing set-up is illustrated in Fig.1 and it shows the two HySpex hyperspectral imagers manufactured by NEO (Norsk Elektro Optikk) - VNIR1800 and SWIR384 - in the pushbroom configuration. In this system, the translational stage shifts the scene across the fields of view of the cameras which synchronously acquire the full spectrum of a single spatial line per time of exposure. The VNIR camera deploys 186 spectral bands in the range 400-1000 nm, with 1800 pixels on the acquisition line. On the other hand, the SWIR camera has 288 bands in the range 950-2500 nm, with only 384 pixels on the acquisition line. Thus, a scaling spatial resolution ratio of approximately 4.6 exists between the two datasets, while the difference in field of view for the two imagers - 17° for VNIR and 16° for SWIR - introduces shifts in the registration problem. The illumination geometry is designed in a way that the irradiance impinging the target is not harmful and does not exceed the light dosage limits for historical artefacts [14]. The relative angles between camera, target, and lights are the same for both VNIR and SWIR sub-sets in order to avoid significant differences in intensity due to the Bidirectional Reflectance Distribution Function (BRDF) of the object surface. The deployed halogen lights emit continuously in the range 400-2500 nm, but although they are of the same model, the reading of illuminance on the target in lux does not match. For this reason, working with spectral reflectance factors instead of radiance data will facilitate the registration, especially for feature-based methods.

The registration of VNIR and SWIR spectral data can be regarded as a *mild* multimodal registration problem, with the multimodality that arises from the fact that the information acquired comes from different regions of the electromagnetic spectrum. However, the two imagers capture the same type of information (spectral reflectance), and the multimodality becomes less accentuated if we consider that there exists a narrow interval (950-1000 nm) in which the two datasets share overlapping nominal wavelengths. At a closer inspection of the data in the narrow shared region it is possible to observe that the spectral values do not concatenate accurately. This is due to a series of factors such as the decrease in signal-to-

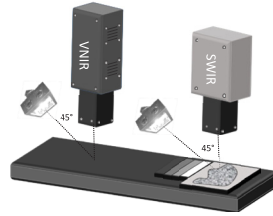


Fig. 1. Schematic representation of the dual camera imaging set-up.

noise ratio (SNR) at the extremities of the two spectral ranges [15] (as illustrated in Fig.2), the change in bandwidth between the two sensors (CMOS and MCT), and the differences in BRDF that exist at the pixel level, due to spurious differences in angles between target, camera, and illumination.

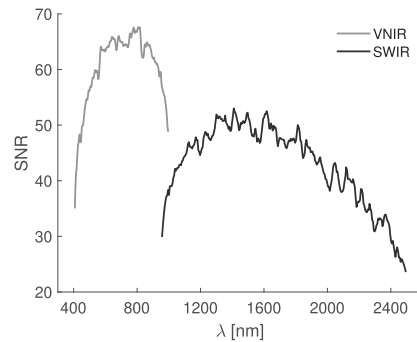


Fig. 2. Signal-to-Noise Ratio of the two hyperspectral imagers as a function of wavelength. The curves are obtained capturing a standardized uniform 99% reflective Spectralon tile and applying the ratio between the mean and standard deviation for each spectral band.

Since the spectral sets are internally co-registered, it is assumed that is necessary to learn only one homography between a fixed-moving image pair that can then be applied to the rest of the moving bands.

3. MATERIAL AND METHODS

In this section we describe the selected factors that potentially affect the results of the registration performances. In order to evaluate the quality of registration, five metrics are selected. Three of them are commonly used to evaluate image similarity: normalized Mutual Information (nMI) [16], Peak Signal to Noise Ratio (PSNR), and Relative Dimensionless Global Error (ERGAS) [17]. Two quality metrics are adapted to suit the evaluation of VNIR-SWIR alignment: the Structural

Similarity Index Measure (SSIM) [18] is modified to account only for its contrast and structure components and renamed truncated-SSIM (tSSIM):

$$tSSIM(X, Y) = \frac{2\sigma_X\sigma_Y + c}{\sigma_X^2 + \sigma_Y^2 + c} \cdot \frac{\sigma_{XY} + c/2}{\sigma_X\sigma_Y + c/2} \quad (1)$$

In which X and Y are the examined images, σ^2 the variance of the image, σ_{XY} the covariance, and $c = 0.03L^2$ (with L being the dynamic range of the images). In order to compare pixel distances across images with significantly different spatial resolution, the commonly used pixel displacement is modified to account for the real-world dimension of a pixel, in a measure here called Ground Sampling Distance Error (GSDE):

$$GSDE(X, Y) = \Delta p(X, Y) \cdot g(X) \quad [\mu m] \quad (2)$$

In which Δp is the classic pixel displacement computed with the euclidean distance between two matched points in images X and Y , and g is the ground sampling distance of the fixed image X (g is approximately $50 \mu m$ for VNIR and $200 \mu m$ for SWIR, computed from a 30 cm distance between camera and target).

The selected factors are of three types: registration methods, image contents, and related to decisions. The selected registration methods include feature-based methods like SIFT [8] and spatio-spectral SIFT [12] (SS SIFT), and a combination of feature and area based methods such as SIFT followed by maximization of nMI (SIFT OPT) and MWTPXC [13]. The image contents, which mock-up historical artefacts and are illustrated in Fig.3, are a drawing on paper (D), a flat painting (FP), a painting with textural reliefs (RP), and a piece of textile (T). The factors related to decisions regard: selection of reference image (VNIR band or SWIR band), considered spectral range (full or overlap), modality of fixed and moving images generation (band selection or band synthesis), scaling (a priori or intrinsic in the homography matrix), and type of homography (affine or projective).

In this study we perform the image registration task covering all the combinations of the reported factors. In total, 352 tasks are performed. Some of the factors do not apply to all the methods, such as SS SIFT that does not require the generation of a single band and operates only on same spectral ranges. Similarly, MWTPXC requires the two images to be always scaled to the same size a priori.

When the overlap range is considered, only the last 13 bands of VNIR and first 8 bands of SWIR constitute the available spectral sets. For the generation of single fixed and moving image pairs from the spectral sets it was decided to choose the approaches of band selection and band synthesis. Band selection picks the two most correlated spectral bands between VNIR and SWIR sets, while band synthesis generates the bands using the first component of Principal Component Analysis (PCA) performed on the images.

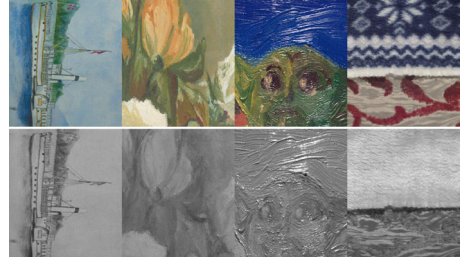


Fig. 3. Image contents in sRGB and IR (band-averaging). From left to right: Drawing on paper (D), Flat painting (FP), Painting with textural reliefs (RP), Textile (T).

4. RESULTS

The single-image quality metrics are computed after registering a pair of test bands that are located in the overlap range of the two spectral imagers, while ERGAS is computed on the spectral datasets of the overlapping range, up-sampling the SWIR data with linear interpolation to match the nominal wavelengths of VNIR.

Since the highlighted potential factors are several, a Principal Component Analysis was performed to inspect possible correlations and interactions. Fig.4 reports the loadings plot of the evaluated conditions along the two first components, which roughly explain 30% of the total variance. It is worth to point out that this percentage is typically a low value. However, by inspecting the next components (not shown) we can corroborate similar conclusions. The performance metric scores are highly correlated, taking into account that ERGAS and GSDE good performance exhibit a low score and that nMI, tSSIM and PSNR good performance exhibit a high score.

From this plot it appears that the most influencing factors are the selected methods, two of the image contents (D and RP), and factors like scaling, selected wavelength range, and band generation. The positions of the loadings with respect to the metrics suggest that points that lie close to the metric cluster (tSSIM, nMI, PSNR) positively influences the results, and conversely what lies further away provides a negative contribution. We also conclude that for further examination (and for simplicity) it is possible to refer to a single quality metric.

To evaluate the influence of the registration methods and image contents Fig.5 reports the statistics related to normalized Mutual Information in the registration tasks. SIFT OPT and MWTPXC come out as the best performing approaches, while SIFT and SS SIFT probably suffer from the intensity differences that exist between the images. Amongst the image contents, it is clear how the presence of textural reliefs and specularities negatively affects the results, as in the case of RP. Quite surprisingly, the textile image T did not pose a

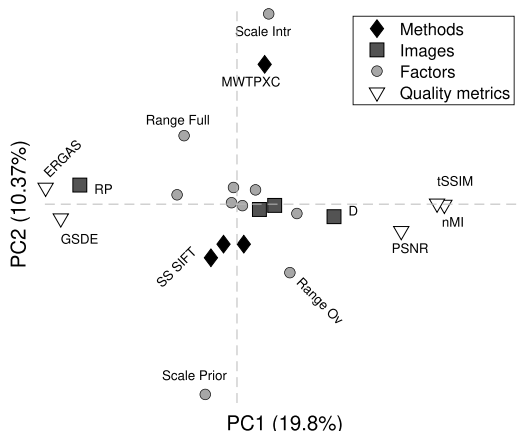


Fig. 4. Loadings of first 2 components of PCA. Only the names of the most influential variables reported for clarity.

big challenge, as consistently good results were achieved.

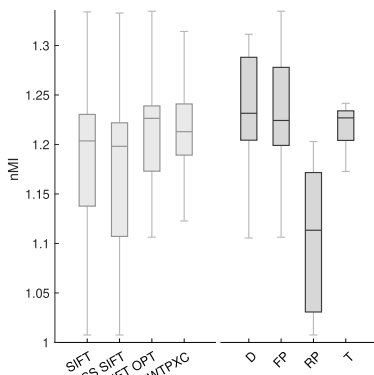


Fig. 5. Statistics of nMI for groups of registration runs divided into methods (left) and image contents (right).

To better appreciate the influence of the selected factors on the registration performances it is possible to examine the image contents individually. We quantified a factor influence by regressing a linear model that relates the conditions of observation with the normalized Mutual Information of the registered images. By using this analytical technique and appropriately normalizing the data, the sign of the regression coefficients will provide information regarding the nature of the influence: either positive or negative. By consequence, the absolute value of the coefficients will give indications on the degree of influence. A visualization of this analysis is provided in Fig.6 in which it is possible to observe that the

exploitation of the overlapping range (950-1000 nm) brings a positive contribution to nMI. Other factors of positive contribution are the selection of the SWIR image as reference, and the prior scaling (a negative contribution from learning the scale intrinsically is highlighted). This analysis does not solve the question on whether it is suggested to proceed with band selection or band synthesis. Running the same analysis excluding the observations where SS SIFT is deployed (i.e., when band generation is not a factor) it would appear that band selection works better for the image contents of D, FP and T, while band synthesis is preferred for RP.

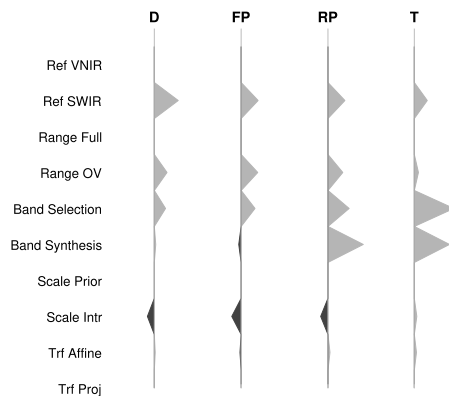


Fig. 6. Linear model coefficients to describe the factors influence on nMI. Light-gray triangles pointing to the right denote a positive influence on the registration result. Conversely, darker triangles pointing to the left highlight negative effects.

To highlight the positive effects of considering only the overlapping range of spectral bands, Table 1 reports the values of the quality metrics of full range against overlap for two image contents - FP and D - when applying SIFT OPT, using SWIR as reference, with band synthesis generation, prior scaling, and affine transformation.

Image		nMI	tSSIM	ERGAS	GSDE (μm)
	Ideal	2	1	0	0
D	Start	1.13	0.70	1.97	380
	Full λ	1.19	0.85	1.66	150
	Ov λ	1.29	0.98	1.30	29.1
FP	Start	1.13	0.45	4.51	822
	Full λ	1.11	0.45	4.75	996
	Ov λ	1.31	0.92	4.58	55.0

Table 1. Registration of some of the quality metrics comparing cases of Full spectral range against Overlap.

5. CONCLUSION

Obtaining sub-pixel precision when registering hyperspectral images coming from different modalities enables the further processing and accurate study of historical artefacts. Knowing which factors improve or degrade the performances of alignment is crucial to obtain the best possible results. With the aid of multivariate techniques we highlighted that in the case of VNIR-SWIR image registration a prior scaling, the deployment of SWIR as reference, and the usage of the bands shared by both imagers contribute positively in achieving good results. In particular, we advise against discarding spectral bands in the overlapping wavelength range that upon a first inspection have a lower signal-to-noise ratio, since they can actually be used as a valuable asset.

6. REFERENCES

- [1] Costanza Cucci, John K Delaney, and Marcello Piccolo, "Reflectance hyperspectral imaging for investigation of works of art: old master paintings and illuminated manuscripts," *Accounts of chemical research*, vol. 49, no. 10, pp. 2070–2079, 2016.
- [2] Hilda Deborah, Sony George, and Jon Yngve Hardeberg, "Pigment mapping of The Scream (1893) based on hyperspectral imaging," in *International Conference on Image and Signal processing*. Springer, 2014, pp. 247–256.
- [3] John K Delaney, Kathryn A Dooley, Annelies Van Loon, and Abbie Vandivere, "Mapping the pigment distribution of Vermeer's Girl with a Pearl Earring," *Heritage Science*, vol. 8, no. 1, pp. 1–16, 2020.
- [4] Chen Shi and Le Wang, "Incorporating spatial information in spectral unmixing: A review," *Remote Sensing of Environment*, vol. 149, pp. 70–87, 2014.
- [5] Gemine Vivone, Mauro Dalla Mura, Andrea Garzelli, and Fabio Pacifici, "A benchmarking protocol for pan-sharpening: Dataset, preprocessing, and quality assessment," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 6102–6118, 2021.
- [6] Massimo Selva, Bruno Aiazzi, Francesco Butera, Leandro Chiarantini, and Stefano Baronti, "Hyper-sharpening of hyperspectral data: A first approach," in *2014 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. IEEE, 2014, pp. 1–4.
- [7] Arthur Ardeshir Goshtasby, *2-D and 3-D image registration: for medical, remote sensing, and industrial applications*, John Wiley & Sons, 2005.
- [8] David G Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [9] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool, "Speeded-up robust features (SURF)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [10] Suhad Lateef Al-Khafaji, Jun Zhou, Ali Zia, and Alan Wee-Chung Liew, "Spectral-spatial scale invariant feature transform for hyperspectral images," *IEEE Transactions on Image Processing*, vol. 27, no. 2, pp. 837–850, 2017.
- [11] Yanshan Li, Qingteng Li, Yan Liu, and Weixin Xie, "A spatial-spectral SIFT for hyperspectral image matching and classification," *Pattern Recognition Letters*, vol. 127, pp. 18–26, 2019.
- [12] Yang Yu, Yong Ma, Xiaoguang Mei, Fan Fan, Jun Huang, and Jiayi Ma, "A spatial-spectral feature descriptor for hyperspectral image matching," *Remote Sensing*, vol. 13, no. 23, pp. 4912, 2021.
- [13] Damon M Conover, John K Delaney, and Murray H Loew, "Automatic registration and mosaicking of technical images of Old Master paintings," *Applied Physics A*, vol. 119, no. 4, pp. 1567–1575, 2015.
- [14] Matija Strlič, David Thickett, Joel Taylor, and May Cas-sar, "Damage functions in heritage science," *Studies in Conservation*, vol. 58, no. 2, pp. 80–87, 2013.
- [15] Unal Okyay and Shuhab D Khan, "Spatial co-registration and spectral concatenation of panoramic ground-based hyperspectral images," *Photogrammetric Engineering & Remote Sensing*, vol. 84, no. 12, pp. 781–790, 2018.
- [16] Aaron F McDaid, Derek Greene, and Neil Hurley, "Normalized mutual information to evaluate overlapping community finding algorithms," *arXiv preprint arXiv:1110.2515*, 2011.
- [17] Lucien Wald, *Data fusion: definitions and architectures: fusion of images of different spatial resolutions*, Presses des MINES, 2002.
- [18] Zhou Wang, Eero P Simoncelli, and Alan C Bovik, "Multiscale structural similarity for image quality assessment," in *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, 2003. Ieee, 2003, vol. 2, pp. 1398–1402.

Logistic splicing correction for VNIR–SWIR reflectance imaging spectroscopy

Article A2j

This article is a reprint of the publication:

Grillini, F., Thomas, J-B., George, S. (2023). *Logistic splicing correction for VNIR–SWIR reflectance imaging spectroscopy*. Optics Letters, vol. 48, pg. 403–406 (Optica Publishing Group). DOI: <https://doi.org/10.1364/OL.478691>

This paper is not included due to copyright restrictions.

Reflectance imaging spectroscopy for the study of archaeological pre-Columbian textiles

Article A3j

This article is a reprint of the publication:

Grillini, F., De Ferri, L., Pantos, G.A., George, S., Veseth, M. (2023).
*Reflectance imaging spectroscopy for the study of archaeological pre-
Columbian textiles.* (Under review)

This paper is under review for publication and is therefore not included.

Full VNIR-SWIR hyperspectral imaging workflow for the monitoring of archaeological textiles

Article A4_c

This article is a reprint of the publication:

Grillini, F., Thomas, J-B., George, S. (2023). *Full VNIR-SWIR hyperspectral imaging workflow for the monitoring of archaeological textiles*. Archiving Conference, vol. 20, pg 192–197 (Society for Imaging Science and Technology). DOI: <https://doi.org/10.2352/issn.2168-3204.2023.20.1.39>

This paper is not included due to copyright restrictions.

Relationship between reflectance and degree of polarization in the VNIR-SWIR: A case study on art paintings with polarimetric reflectance imaging spectroscopy

Article A5j

This article is a reprint of the publication:

Grillini, F., Aksas, L., Lapray, P-J., Foulonneau, A., Thomas, J-B., George, S., Bigué, L. (2023). *Relationship between reflectance and degree of polarization in the VNIR-SWIR: A case study on art paintings with polarimetric reflectance imaging spectroscopy*. (Under review)

This paper is under review for publication and is therefore not included.

ISBN 978-82-326-7546-3 (printed ver.)
ISBN 978-82-326-7545-6 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)



NTNU

Norwegian University of
Science and Technology